

CATARINA DE BRITO CAMACHO PEREIRA COUTINHO

**IDENTIFICAÇÃO DE NÍVEIS DE ATIVIDADE FÍSICA AO
LONGO DE 12 MESES, COM *MACHINE LEARNING*, E A SUA
ASSOCIAÇÃO COM VARIÁVEIS MOTIVACIONAIS, EM
ADULTOS EM PROCESSO DE MANUTENÇÃO DE PERDA DE
PESO NO PROJETO *NOHOW***

Orientador: **Prof. Doutor António João Labisa da Silva Palmeira**

**UNIVERSIDADE LUSÓFONA DE HUMANIDADES E TECNOLOGIAS
FACULDADE DE EDUCAÇÃO FÍSICA E DESPORTO
LISBOA**

2020

CATARINA DE BRITO CAMACHO PEREIRA COUTINHO

**IDENTIFICAÇÃO DE NÍVEIS DE ATIVIDADE FÍSICA AO
LONGO DE 12 MESES, COM *MACHINE LEARNING*, E A SUA
ASSOCIAÇÃO COM VARIÁVEIS MOTIVACIONAIS, EM
ADULTOS EM PROCESSO DE MANUTENÇÃO DE PERDA DE
PESO NO PROJETO *NOHOW***

Dissertação defendida em provas públicas para a obtenção do grau de Mestre em Exercício e Bem-Estar, no Curso de Mestrado em Exercício e Bem-Estar com especialidade em Exercício, Nutrição e Saúde, conferido pela Universidade Lusófona de Humanidades e Tecnologias, no dia 05/03/2020, com o Despacho de Nomeação de Júri nº 56/2020, com a seguinte composição:

Presidente: Prof.^a Doutora Sofia Cristina Carreiras Fonseca

Arguente: Prof. Doutor Diogo dos Santos Teixeira

Orientador: Prof. Doutor António João Labisa da Silva
Palmeira

**UNIVERSIDADE LUSÓFONA DE HUMANIDADES E TECNOLOGIAS
FACULDADE DE EDUCAÇÃO FÍSICA E DESPORTO
LISBOA
2020**

Agradecimentos

Vale a pena lembrar o trabalho e dedicação necessários para concluir uma tese de mestrado, são muitas horas, dias e meses de entrega, de obstáculos, de dificuldades, mas também de metas alcançadas, de resultados e de animação, que culminam num momento de total satisfação, materializada nesta dissertação.

Agradeço, muito sinceramente, o apoio na elaboração desta tese:

- ◇ Ao Professor Doutor António Palmeira pelo desafio que me colocou, pela confiança que depositou em mim e pela audácia de acreditar que é possível fazer mais e melhor, pela partilha da informação do projeto *NoHoW*, que permitiu que este trabalho fosse feito, pelo constante e precioso apoio e orientação no processo de elaboração da tese, através de sempre pertinentes comentários que certamente enriqueceram este trabalho. Sem o seu suporte, este trabalho não teria sido concretizado;
- ◇ À equipa do projeto *NoHoW* por permitir a utilização dos dados recolhidos com o objetivo de obter informação com mais valor;
- ◇ A toda a minha família pelo apoio constante e incondicional, por me motivarem a ser melhor todos os dias, acreditarem em mim e me encorajarem a nunca desistir e, sobretudo, contribuírem positivamente e diariamente para que a minha vida seja mais feliz;
- ◇ Aos meus amigos e colegas que, de alguma forma, ouviram “a história que tinha para contar” e que reflete a abordagem deste trabalho, que colaboraram com ideias e com críticas.

Lista de Abreviaturas

OMS – Organização Mundial de Saúde

DGS – Direção Geral de Saúde

AF – Atividade Física

WL – *Weight Loss* / Perda de peso

AI ou IA – *Artificial Intelligence* / Inteligência Artificial

ML – *Machine Learning*

ANN – *Artificial Neural Networks* / Redes Neurais Artificiais

PRISMA- Preferred Reporting Items for Systematic Reviews and Meta-analyses

BAA – *Behavioral Analytics Algorithm*

RCT – *Randomized Controlled Trial* / Ensaio Aleatório Controlado

RSL – Revisão Sistemática da Literatura

MET- Equivalente Metabólico da Tarefa

TAD – Teoria da Autodeterminação

FC – Frequência Cardíaca

bpm – batimentos por minuto

IMC – Índice de Massa Corporal

Ativ – Ativos

I-Ativ – Insuficientemente Ativos

N.D. – Não Disponível

N.A. – Não Aplicável

DK – Dinamarca

PT – Portugal

UK – Reino Unido

BPN – *Basic Psychological Needs*

BREQ-3 – *Behavioral Regulation in Exercise Questionnaire (versão 3)*

RAI – Índice de Autonomia Relativa do BREQ-3

Índice Geral

Resumo.....	9
Abstract.....	10
CAPÍTULO I.....	11
INTRODUÇÃO GERAL.....	11
<i>Perda de peso e manutenção do peso perdido</i>	<i>11</i>
<i>Inteligência Artificial e Métodos de Machine Learning</i>	<i>11</i>
<i>Machine Learning aplicado à prática de atividade física e gestão do peso</i>	<i>12</i>
CAPÍTULO II	15
ANÁLISE DA UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NO CONTEXTO DA ADEÇÃO À ATIVIDADE FÍSICA E GESTÃO DO PESO - REVISÃO SISTEMÁTICA DA LITERATURA RÁPIDA.....	15
Resumo.....	16
Abstract.....	17
II.1 Introdução	18
<i>Racional teórico – influência do perfil motivacional na prática de atividade física, no processo de manutenção do peso perdido</i>	<i>18</i>
<i>Racional teórico – como a inteligência artificial e os métodos de Machine Learning podem ser úteis em intervenções no contexto da atividade física e gestão do peso</i>	<i>18</i>
II.2 Método	19
<i>Protocolo</i>	<i>19</i>
<i>Estratégia de Pesquisa.....</i>	<i>19</i>
<i>Critérios de Exclusão</i>	<i>19</i>
<i>Seleção dos Estudos.....</i>	<i>19</i>
<i>Extração de Dados.....</i>	<i>20</i>
<i>Qualidade dos Estudos</i>	<i>20</i>
II.3 Resultados.....	21
<i>Pesquisa da Literatura</i>	<i>21</i>
<i>Características dos Estudos Incluídos</i>	<i>23</i>
<i>Instrumentos</i>	<i>23</i>
<i>Principais Resultados</i>	<i>25</i>
<i>Qualidade dos Estudos</i>	<i>39</i>
<i>Discussão dos Resultados.....</i>	<i>43</i>
<i>Limitações</i>	<i>43</i>
<i>Conclusões</i>	<i>44</i>
II.4 Referências Bibliográficas.....	45

CAPÍTULO III.....	48
IDENTIFICAÇÃO DE NÍVEIS DE ATIVIDADE FÍSICA AO LONGO DE 12 MESES, COM MACHINE LEARNING, E A SUA ASSOCIAÇÃO COM VARIÁVEIS MOTIVACIONAIS, EM ADULTOS EM PROCESSO DE MANUTENÇÃO DE PERDA DE PESO NO PROJETO NOHOW - UM ESTUDO DE COHORT PROSPETIVO.....	48
Resumo.....	49
Abstract.....	50
III.1 Introdução.....	51
<i>Racional teórico.....</i>	<i>51</i>
<i>Estudos efetuados – Estilo de Vida e Mudanças Comportamentais.....</i>	<i>51</i>
<i>Estudos efetuados – Estratégias baseadas na teoria da auto-determinação</i>	<i>54</i>
<i>Considerações acerca da abordagem metodológica.....</i>	<i>56</i>
<i>Introdução aos modelos de Machine Learning.....</i>	<i>56</i>
<i>Estudos efetuados com modelos de Machine Learning</i>	<i>57</i>
<i>Objetivo e GAP</i>	<i>58</i>
III.2 Método.....	59
III.2.a) Desenho do Estudo	60
III.2.b) Amostra.....	60
III.2.c) Medidas/Instrumentos.....	62
III.2.d) Procedimentos	64
III.2.e) Análise dos dados.....	64
(1) Validação das variáveis	65
(2) Caracterização dos participantes com base nas métricas da FitBit.....	66
i) Limpeza dos dados.....	66
ii) Engenharia das características relevantes	67
(3) Clustering dos participantes	68
i) Análise em Componentes Principais	69
ii) Algoritmo de aprendizagem não supervisionada – K-Means Clustering.....	69
iii) Avaliação do output através da aplicação de Testes de Hipóteses	70
(4) Cruzamento com as respostas aos inquéritos motivacionais	70
III.3 Resultados	72
III.3.a) Resultados da aplicação do algoritmo de <i>Machine Learning – K-Means</i>	72
III.3.b) Comparação dos indicadores de motivação dos Ativ face aos I-Ativ	77
III.3.c) Evolução dos indicadores de motivação dos indivíduos Ativ	79
III.3.d) Evolução dos indicadores de motivação dos indivíduos I-Ativ	80
III.3.e) Comparação dos indicadores de motivação em ambos os grupos.....	82
III.4 Discussão dos Resultados.....	84
III.5 Conclusão e Limitações.....	88
III.6 Referências Bibliográficas	89
CAPÍTULO IV	94
CONCLUSÃO GERAL.....	94

IV.1 Discussão geral.....	95
IV.2 Conclusões finais.....	96
IV.3 Referências Bibliográficas	97
CAPÍTULO V.....	99
ANEXOS.....	99
V.1 Website do projeto <i>NoHoW</i>.....	99
V.2 Questionários motivacionais – Questionário de Regulação Comportamental para o Exercício (BREQ-3)	100
V.3 Questionários motivacionais – Escala de Necessidades Psicológicas Básicas (BPN)	101
V.4 <i>Machine Learning</i> – Código Fonte.....	102
V.5 Análises Estatísticas – Testes de Hipóteses (output do SPSS).....	123

Índice de Figuras, Gráficos e Tabelas

CAPÍTULO II - Análise da utilização de Inteligência Artificial no contexto da adesão à atividade física e gestão do peso – Revisão Sistemática da Literatura rápida

Figura 1 – Diagrama do processo de seleção dos estudos para a RSL rápida.....	22
Tabela 1 – Categorização dos instrumentos.....	23
Tabela 2 – Principais resultados obtidos – estudos tipo 1.....	25
Tabela 3 – Principais resultados obtidos – estudos tipo 2.....	31
Tabela 4 – Resultados da análise da qualidade dos estudos - Estudos observacionais, transversais ou de <i>Cohort</i>	39
Tabela 5 – Resultados da análise da qualidade dos estudos – RSL rápida.....	42

CAPÍTULO III - Identificação de níveis de atividade física ao longo de 12 meses, com *Machine Learning*, e a sua associação com variáveis motivacionais, em adultos em processo de manutenção de perda de peso no projeto *NoHoW* - um estudo de *Cohort* prospectivo

Tabela 1 – Distribuição percentual dos participantes por característica.....	61
Tabela 2 – Valores médios da amostra total e por género.....	61
Tabela 3 – Seleção das variáveis.....	66
Gráfico 1 – Variância explicada pelas diferentes componentes principais.....	72
Gráfico 2 – Aplicação do <i>K-Means</i> às componentes principais (2 e 3 clusters)	73
Gráfico 3 – Projecção das 2 componentes principais para 2 clusters.....	74

Tabela 4 – Valores médio das variáveis da <i>FitBit</i> por <i>cluster</i>	75
Tabela 5 – Resultados obtidos com a aplicação do teste t- <i>Student</i>	76
Tabela 6 –Variação dos indicadores de motivação dos Ativ face aos I-Ativ, no <i>Baseline</i> , 6 e 12 meses.....	77
Tabela 7 – Indicadores motivacionais - valores mais elevados nos Ativ.....	78
Tabela 8 – Indicadores motivacionais - valores mais elevados nos I-Ativ.....	78
Tabela 9 – Variação dos valores médios dos indicadores motivacionais dos Ativ – variações estatisticamente significativas.....	79
Tabela 10 – Variação dos valores médios dos indicadores motivacionais dos I-Ativ – variações estatisticamente significativas.....	80
Gráfico 4 – Comparação das variações dos indicadores motivacionais para os Ativ e os I-Ativ (entre o <i>Baseline</i> e os 6 meses)	82
Gráfico 5 – Comparação das variações dos indicadores motivacionais para os Ativ e os I-Ativ (entre os 6 e os 12 meses)	82
Gráfico 6 – Comparação das variações dos indicadores motivacionais para os Ativ e os I-Ativ (entre o <i>Baseline</i> e os 12 meses)	83

Resumo

Objetivo: Esta dissertação teve como objetivo principal a identificação de níveis de atividade física ao longo de 12 meses, com *Machine Learning*, e a sua associação com variáveis motivacionais, em adultos em processo de manutenção de perda de peso no projeto *NoHoW*.

Método: O trabalho foi dividido em duas partes, em primeiro lugar foi efetuada uma Revisão Sistemática da Literatura rápida, na qual se procurou analisar a utilização de inteligência artificial no contexto da adesão à atividade física e gestão do peso e seguidamente foi realizado um estudo de *Cohort* prospetivo, com uma amostra de 1627 indivíduos. Utilizou-se a informação recolhida de *FiBit* (24/7) durante 12 meses e respostas a questionários motivacionais aplicados no *Baseline*, 6 e 12 meses.

Resultados: Na Revisão Sistemática da Literatura rápida verificou-se que a literatura aponta para que intervenções personalizadas são essenciais para o sucesso individual da gestão do peso. Relativamente aos estudos focados em modelos preditivos, os resultados mostraram que a prática de atividade física se ajusta aos padrões de vida atuais e que os preditores para a prática de atividade física parecem estar associados a fatores psicológicos e comportamentais dos indivíduos. No estudo de *Cohort* prospetivo identificaram-se 2 grupos com diferentes níveis de atividade física, os ativos (Ativ) e os insuficientemente ativos (I-Ativ), com perfis motivacionais distintos. No grupo Ativ, as variáveis RAI, Motivação Autónoma, Regulações Intrínseca, Integrada e Identificada têm valores mais elevados, nos períodos em análise, em comparação com os valores dos indivíduos I-Ativ. Os indicadores motivacionais variam ao longo do tempo, com tendência crescente no início do projeto, verificando-se uma estagnação a partir dos 6 meses e redução em ambos os grupos no final dos 12 meses.

Conclusões: No que respeita a Revisão Sistemática da Literatura rápida, concluiu-se que a inteligência artificial e os métodos de *Machine Learning* aplicados à gestão do peso e à promoção da prática de atividade física podem ser uma enorme vantagem e poderão apoiar na redução das taxas de obesidade. No estudo de *Cohort* prospetivo foi possível a segregação de indivíduos por padrões de atividade física, com métodos de *Machine Learning*. São níveis elevados de Motivação Autónoma, Regulação Intrínseca, Integrada e Identificada ao longo dos 12 meses que parecem permitir que os indivíduos pratiquem ativamente atividade física. A Autonomia, no contexto da gestão do peso, aos 12 meses deverá ser avaliada, pela sua diminuição face ao nível no *Baseline*.

Palavras chave: Atividade Física, *Machine Learning*, Inteligência Artificial, Motivação, Regulação Comportamental, Gestão de peso, Perda de peso, Manutenção do peso perdido

Abstract

Objective: This dissertation main objective was to identify physical activity levels over 12 months, with *Machine Learning*, and its association with motivational variables, in adults undergoing weight loss maintenance in the *NoHoW* project.

Method: The work was divided in two parts, first a Fast-Systematic Review of the Literature was carried out to analyze the use of artificial intelligence in the context of adherence to physical activity and weight management and second a Cohort's prospective study was conducted, we used a sample of 1627 individuals, FiBit information (24/7) for 12 months and motivational questionnaires output applied at Baseline, 6 and 12 months.

Results: In the Fast-Systematic Review of the Literature, studies indicated that personalized interventions are essential for individual weight management success. Concerning studies focused on predictive models, the results showed that physical activity practice adapts to current living standards and that predictors for physical activity practice seem to be associated with psychological and behavioral factors. In the prospective Cohort study, 2 groups with different levels of physical activity were identified: active (Activ) and somewhat active (I-Activ), with different motivational profiles. In the Activ group, the variables RAI, Autonomous Motivation, Intrinsic, Integrated and Identified Regulations have higher values when compared to the values of I-Activ group. Motivational indicators vary over time, with an upward trend at the beginning of the project, stagnation starting at 6 months and decreasing in both groups at the end of 12 months.

Conclusions: The Fast-Systematic Review of the rapid literature made us conclude that artificial intelligence and Machine Learning methods applied to weight management and the enhancement of physical activity practice are a huge advantage and may help to reduce obesity rates. In the prospective Cohort study, with Machine Learning methods, it is possible to segregate individuals by patterns of physical activity. High levels of Autonomous Motivation, Intrinsic Regulation, Integrated and Identified over the 12 months seem to allow individuals to actively engage in physical activity. Autonomy in the context of weight management at 12 months should be assessed.

Keywords: Physical Activity, Machine Learning, Artificial Intelligence, Motivation, Behavioral Regulation, Weight Management, Weight Loss, Lost Weight Maintenance

CAPÍTULO I

INTRODUÇÃO GERAL

Perda de peso e manutenção do peso perdido

O aumento da obesidade a nível mundial é, hoje em dia, um fenómeno identificado e reconhecido. Muitas são as formas e os métodos que advogam conseguir colmatar este flagelo. Desde suplementos alimentares a intervenções cirúrgicas, surgem-nos diariamente inúmeras soluções (supostamente) eficazes para perder peso, nos diferentes meios de comunicação. Mas, na verdade, a solução está ainda longe de ter sido encontrada, e sobretudo uma solução sustentada que permita não só perder peso como manter um peso e uma vida saudáveis ao longo do tempo.

Na avaliação das estratégias de dieta eficazes para perda de peso verificou-se que, para que uma dieta tenha sucesso e seja segura é necessário que ocorra um deficit calórico atingido sobretudo pela redução do consumo de gordura. O aumento de ingestão de fibra, a inclusão de atividade física e treino comportamental, e a auto monitorização, surgem como catalisadores da manutenção do peso (Ramage, Farmer, Eccles, & McCargar, 2014).

A perda de peso é um desafio difícil e a manutenção do peso perdido um problema efetivo. Na era das novas tecnologias, as intervenções web que permitem a auto monitorização, seja da atividade física seja da dieta alimentar, aparecem como forma de ajudar a população a ter consciência do seu comportamento na gestão de peso, no dia-a-dia. Estas intervenções parecem estar associadas a uma maior eficácia quando comparadas com intervenções mais tradicionais e menos tecnológicas, mas ainda é necessária mais pesquisa nesta área (Sorgente et al., 2017). Um estudo de Painter et al. (2017) efetuou uma análise retrospectiva com 2113 participantes entre 2011 e 2015 e concluiu que comportamentos de auto monitorização do peso, dos passos diários, da atividade física de alta intensidade e o registo do diário alimentar apresentaram-se como preditores da perda de peso numa intervenção com duração de 6 meses.

Assim, é pertinente afirmar que as novas tecnologias parecem ter um papel fundamental no que respeita as intervenções de perda de peso e manutenção de peso perdido, tanto na sua análise como na forma de as tornar mais eficazes e sustentadas.

Inteligência Artificial e Métodos de Machine Learning

Segundo Hamet e Tremblay (2017), quando pensamos em inteligência artificial, pensamos na modelação da inteligência humana ou de comportamento inteligente através da utilização de tecnologia, e, olhando atrás no tempo, remete-nos para a criação de robôs. Apesar do termo inteligência artificial

poder remontar a épocas mais longínquas, o surgimento da inteligência artificial como ciência e engenharia surge em 1956. Hoje em dia aplica-se um enorme leque de áreas, pelo potencial que aporta à análise da informação e o apoio à tomada de decisão, e consequente retorno económico.

O processo de aprendizagem, no que respeita inteligência artificial, é implementado através da utilização de métodos matemáticos, estatísticos, lógicos e de programação computacional e permite que o modelo de inteligência artificial seja treinado com dados de forma iterativa com ajustamento de parâmetros via tentativa/erro. Uma abordagem da inteligência artificial é o *Machine Learning*, que se baseia no treino de novos dados utilizando um método adaptativo como a rede neuronal sem a programação explícita de novas regras (como noutros tipos de abordagens), os coeficientes (pesos) do modelo são ajustados iterativamente, com representações repetidas dos dados, até que o índice de desempenho seja otimizado. Um exemplo é o modelo de aprendizagem não supervisionada, como o algoritmo de *clustering K-Means*, que tem como objetivo encontrar agrupamentos ou definir características nos dados (Benke & Benke, 2018).

Nos últimos anos, o *Machine Learning* tornou-se uma "*tecnologia de uso geral*", na medida em que é difundida, pode ser aprimorada ao longo do tempo e tem potencial para criar inovação (Panch, Szolovits, & Atun, 2018), é por isso muito utilizada nas mais diversas áreas, designadamente na área da saúde e atividade física e bem-estar. As técnicas de *Machine Learning* criam processos executados através da combinação de esforços humanos e tecnológicos, permitindo ultrapassar a normal capacidade humana de gerar hipóteses. Estas abordagens permitem a análise de grandes volumes de dados para encontrar associações com capacidade preditiva (Panch, Szolovits, & Atun, 2018). A utilização destas técnicas só se afigurou exequível devido aos avanços na capacidade de processamento, na memória, no armazenamento de informação e na quantidade de informação que hoje está disponível. Só assim se tornou possível transferir um conjunto de complexas tarefas para computadores, que as executam com enorme sucesso. Empresas e instituições diversas têm vindo a aplicar análises complexas aos seus dados, muitas na área da saúde, no entanto esta área ainda não tem beneficiado das contribuições das abordagens de *Machine Learning* como poderia, pois de alguma forma, estas abordagens põem em causa a forma como se pratica medicina atualmente (Deo, 2015).

Se pensarmos em áreas da saúde, mas menos sensíveis, e nas quais os resultados da aplicação de algoritmos de *Machine Learning* não sejam tão impactantes ou controversos, podemos olhar para a área do bem-estar e da prática de atividade física e como é que estes algoritmos nos podem ajudar a que a população mundial seja mais ativa no seu dia-a-dia.

Machine Learning aplicado à prática de atividade física e gestão do peso

Já existem diversos estudos científicos que aplicam abordagens de Inteligência Artificial ou *Machine Learning* a problemas relacionados com a prática de atividade física e com a gestão de peso. São assuntos complexos que não derivam apenas de questões físicas, mas relacionam-se com características psicológicas e motivacionais dos indivíduos. Nesse sentido, surgem alguns estudos que

tentam facilitar e incentivar a prática de atividade física, através da ajuda da tecnologia, e dessa forma melhorar a capacidade individual de gestão do peso.

No estudo de Willetts, Hollowell, Aslett, Holmes e Doherty (2018) foram utilizadas técnicas de *Machine Learning* no contexto da prática de exercício. Como foi referido, as atuais diretrizes existentes referentes à prática de atividade física e à duração do sono estão limitadas a evidência subjetiva e auto reportada. Ao contrário, neste estudo foi possível detectar, através da utilização de monitores de pulso de atividade física, vários níveis de atividade física e comportamentos do sono. Um modelo de *Machine Learning* criado a partir de uma amostra de informação recolhida em 132 adultos pôde ser utilizado posteriormente para inferir padrões de atividade num universo de 96.220 indivíduos. Por exemplo, verificou-se que os homens passaram mais tempo em atividades ou de baixa ou de alta intensidade, enquanto as mulheres apresentaram comportamentos mistos, as caminhadas foram mais frequentes na primavera e no verão, o número de horas de sono diminuiu.

Outro exemplo de técnicas de *Machine Learning* é o utilizado no estudo de Kerr et al. (2017), onde foi efetuada a comparação de três técnicas de estimação de atividade física através da utilização de um acelerómetro, em diferentes localizações e colocações. A amostra era constituída por 321 mulheres de idade média ou avançada que utilizaram durante 7 dias um acelerómetro GT3X+, no pulso e anca. Foi possível concluir que as diferentes técnicas produzem resultados diferentes e que a comparação de atividade física proveniente de contextos e instrumentos diferentes deve ser efetuada com cautela.

De notar que, essencialmente, foram utilizadas abordagens de *Machine Learning* para detecção ou previsão de padrões ou de comportamentos ligados à prática de atividade física, mas também à criação de *coaching* virtual para a promoção e incentivo à prática de atividade física. Por exemplo, nos estudos de Alley, Jennings, Plotnikoff e Vandelanotte (2014, 2016), nos quais se advogava que as intervenções web para promover a prática de atividade física melhorariam o comprometimento e comportamento individual dos participantes, verificou-se que uma opção mista com vídeo-*coaching* e *coaching* virtual foi efetivo na promoção da prática de atividade física.

Outra abordagem que poderia fazer sentido nesta área é a aplicação de técnicas de classificação de tipologias de atividade física, e categorização dos respetivos comportamentos de atividade física. Por exemplo, como é que é possível identificar grupos de pessoas idênticas no que respeita a prática de atividade física. Será possível identificar praticantes de atividade física num conjunto de indivíduos? E será que este processo identifica indivíduos que apresentam características que se sabem estar associadas a diferentes níveis de AF, como sejam, por exemplo, níveis de motivação diferenciados? É nesta tipologia de abordagens que se tornam mais escassos os estudos científicos publicados e é nesse sentido que este trabalho pretende seguir. Dessa forma, o presente estudo foi dividido em duas partes: uma na qual se efetuou uma revisão sistemática da literatura, onde se verificou a aplicação de inteligência artificial no contexto da adesão à atividade física e gestão do peso e uma segunda parte na qual se procurou através de um estudo de *cohort* prospetivo, efetuar a identificação de níveis de

atividade física em indivíduos adultos, através de informação acedida continuamente durante 12 meses através de *wearables*, num processo de manutenção do peso perdido, com uma abordagem de *Machine Learning*, utilizando informação fornecida pelo Projeto *NoHoW*.

CAPÍTULO II

ANÁLISE DA UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NO CONTEXTO DA ADESÃO À ATIVIDADE FÍSICA E GESTÃO DO PESO - REVISÃO SISTEMÁTICA DA LITERATURA RÁPIDA

Resumo

Objetivo: O principal objetivo desta análise foi aferir qual a informação existente na literatura no que respeita a utilização de inteligência artificial no contexto da adesão à atividade física e gestão do peso.

Método: Foi efetuada uma Revisão Sistemática da Literatura rápida, na qual se procurou analisar os estudos efetuados com recurso a inteligência artificial, no contexto da adesão à atividade física e gestão do peso.

Resultados: Verificou-se que todos os estudos apontaram para que intervenções personalizadas foram essenciais para o sucesso individual da gestão do peso. No que respeita os estudos focados em modelos preditivos, os resultados mostraram que a prática de atividade física se ajustava aos padrões de vida atuais, menos tempo de prática, mas com maior intensidade, e que os preditores para a prática de atividade física estavam mais associados a fatores psicológicos e comportamentais.

Conclusões: Concluiu-se que a inteligência artificial e os métodos de *Machine Learning* aplicados à gestão do peso e à promoção da prática de atividade física foram uma enorme vantagem e poderiam apoiar na redução das taxas de obesidade.

Palavras chave: Atividade Física, *Machine Learning*, Inteligência Artificial, *Data Mining*, *Neural Network*, Gestão de peso, Perda de peso, Manutenção do peso perdido

Abstract

Objective: The main objective of this analysis was to find the existing information in the literature regarding the use of artificial intelligence in the context of adherence to physical activity and weight management.

Method: A Fast-Systematic Review of the Literature was carried out to analyze the use of artificial intelligence in the context of adherence to physical activity and weight management.

Results: All studies indicated that personalized interventions were essential for individual weight management success. Concerning studies focused on predictive models, the results showed that physical activity practice adapted to current living standards, less time but higher intensity, and that predictors for physical activity practice seemed to be associated with psychological and behavioral factors.

Conclusions: Artificial intelligence and Machine Learning methods applied to weight management and the promotion of physical activity practice were a huge advantage and may be useful to reduce obesity rates.

Keywords: Physical Activity, Machine Learning, Artificial Intelligence, Data Mining, Neural Network, Weight Management, Weight Loss, Lost Weight Maintenance

II.1 Introdução

Racional teórico – influência do perfil motivacional na prática de atividade física, no processo de manutenção do peso perdido

Segundo a OMS (2016), a prevalência do excesso de peso e da obesidade tem vindo a aumentar de forma preocupante nas últimas décadas. Ainda assim, enquanto que indivíduos com obesidade conseguem, habitualmente, ter sucesso na perda de peso, a manutenção desse peso perdido revela-se extremamente complexa. Verifica-se, ainda hoje, um escasso entendimento do fenómeno de manutenção do peso perdido (Poulimeneas, Yannakoulia, Anastasiou, & Scarmeas, 2018) e, continua a permanecer a opinião generalizada que é muito complicado manter o peso perdido a longo prazo, e que poucos conseguem fazê-lo eficazmente. Existem estudos que mostram que apenas 20% dos indivíduos com excesso de peso conseguem ter sucesso na manutenção do peso perdido a longo prazo, considerando uma perda de peso de pelo menos 10% do peso inicial, e o período de manutenção do peso perdido de pelo menos 1 ano (Wing & Phelan, 2005).

Racional teórico – como a inteligência artificial e os métodos de Machine Learning podem ser úteis em intervenções no contexto da atividade física e gestão do peso

Existem já diversos estudos, no contexto da atividade física e gestão do peso, que fazem uso da inteligência artificial. A inteligência artificial, apesar de diferente assemelha-se de alguma forma à inteligência humana, é largamente utilizada para resolução de problemas. Como exemplo, avaliou-se o estudo de Novatchkov e Baca (2013) que implementava métodos de reconhecimento de padrões para a avaliação de exercícios realizados em máquinas de treino. Os resultados obtidos mostraram bons resultados de desempenho do modelo de previsão, indicando a viabilidade e a capacidade das técnicas de inteligência artificial na avaliação automática do desempenho em equipamentos de musculação e no aconselhamento imediato dos atletas.

Segundo Aswani, Kaminsky, Mintz, Flowers e Fukuoka (2019) é conhecido que as intervenções de perda de peso através do aumento da atividade física e da modificação da dieta têm sucesso na redução do peso dos indivíduos, no entanto torna-se difícil manter esses programas por longos períodos de tempo, devido à falta de adesão dos indivíduos a esse estilo de vida. Uma abordagem promissora para aumentar a adesão dos indivíduos foi através da personalização dos programas alimentares e de atividade física, e isso tornou-se possível alcançar através da utilização de modelos preditivos, como evidenciado neste artigo.

Assim, e não se conhecendo revisões sistemáticas da literatura (RSL) que trabalhem a aplicação de algoritmos de *Machine Learning* para prever a adesão à atividade física em programas de manutenção de peso perdido e gestão do peso, o objetivo desta RSL rápida foi analisar a literatura

existente, na qual houve recurso à utilização de técnicas de Inteligência Artificial e/ou algoritmos de *Machine Learning* nas intervenções efetuadas, no sentido de tentar obter soluções mais robustas no contexto da promoção da prática de atividade física, na melhoria da capacidade de gestão do peso e numa perspetiva de previsão de comportamentos de adesão ou desistência da prática de atividade física.

II.2 Método

Protocolo

Para a RSL rápida deste trabalho foi utilizado PRISMA-P 2015 *checklist* (Moher et al., 2009). Através da utilização deste protocolo foram definidos os itens relevantes para a elaboração da mesma. Adicionalmente foram também utilizados dois documentos de apoio (Bento, 2014) e (Dobbins, 2017).

Estratégia de Pesquisa

Para a pesquisa de informação foi utilizada a base de dados *PubMed*, e os termos de pesquisa foram, em primeiro lugar (*exercise OR physical activity*) AND (*artificial intelligence OR neural network OR data mining OR Big Data OR machine learning*) e posteriormente (*weight maintenance or exercise or physical activity or weight loss or weight management or weight*) AND (*artificial intelligence OR neural network OR data mining OR Big Data OR machine learning*).

CrITÉrios de Exclusão

Adicionalmente, foram definidos os critérios de exclusão de artigos, ou seja, excluÍram-se os artigos que não estavam diretamente relacionados com a previsão de adesão à atividade física, e que não utilizavam técnicas estatísticas avançadas, técnicas de inteligência artificial, redes neuronais, *Big Data* ou *Machine Learning* na sua análise. Esses critérios conduziram então à exclusão de artigos cujas análises utilizavam metodologias estatísticas tradicionais. Finalmente foram também eliminados um conjunto de artigos, que não se enquadravam nos objetivos do nosso estudo, relacionados com:

- 1- Identificação de padrões da passada/caminhada;
- 2- Identificação de riscos de doenças ou focados na prevenção de doenças;
- 3- Reconhecimento de imagens.

Seleção dos Estudos

Após a aplicação dos critérios de exclusão à pesquisa efetuada foi efetuado um trabalho no qual foram avaliados os restantes os artigos, para identificar se poderiam ser incluídos nesta RSL rápida. Para esses, foi feita a leitura dos títulos de cada um deles, e foram novamente selecionados apenas os artigos cujo título tinha alguma das palavras utilizadas na estratégia de pesquisa. Finalmente, e para os

artigos que restaram, foi lido o respetivo resumo e apenas selecionados os artigos cujo resumo poderia identificar-se com a análise desta revisão sistemática da literatura.

Extração de Dados

Para cada um dos estudos selecionados foi efetuada uma tabela, na qual foi inserido um conjunto de informação, de modo a sistematizar o processo de pesquisa. Foi extraída a seguinte informação:

- Autor(es) e ano de publicação do artigo
- Amostra
- Desenho do estudo e intervenções efetuadas no estudo (caso existam)
- Objetivo do estudo
- Instrumentos utilizados no estudo
- Algoritmos utilizados (de inteligência artificial, ou outros)
- Principais resultados e conclusões obtidos

Após esta sistematização da informação foram eliminados um conjunto de artigos relacionados com reconhecimento de comportamentos sedentários ou reconhecimento de padrões de atividade física, artigos referentes apenas a perdas de peso (sem inclusão de exercício) com recurso em soluções web e revisões baseadas na utilização de *Big Data*. Estes artigos foram removidos pois não se enquadravam com o estudo que se pretendia efetuar. Finalmente, foram ainda incluídos nesta análise, artigos recolhidos da leitura da bibliografia dos artigos selecionados até esta fase, através da leitura do respetivo título.

Qualidade dos Estudos

Para avaliar a qualidade dos estudos selecionados nesta RSL rápida foi utilizado o “*Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies*” e “*Quality Assessment of Systematic Reviews and Meta-Analyses*” obtido no website do “*the NIH Quality Assessment*”. Cada um dos estudos foi avaliado nas diversas dimensões existentes.

Além desta avaliação da RSL rápida foi também efetuada uma revisão por um outro autor, através de uma análise cruzada dos estudos, de modo a melhorar a qualidade da avaliação.

II.3 Resultados

Pesquisa da Literatura

Foram efetuadas duas pesquisas com formulações e palavras chave ligeiramente diferentes e cuja publicação estava disponível até dia 24 de fevereiro de 2019, e que resultaram no seguinte:

Pesquisa 1: utilizou-se a seguinte instrução de pesquisa: *(exercise OR physical activity) AND (artificial intelligence OR neural network OR data mining OR Big Data OR machine learning)*

- Artigos obtidos após pesquisa na *PubMed*: 3651
- Artigos excluídos: 3648
- Artigos selecionados como elegíveis: 3
- Artigos selecionados após leitura do resumo: 3

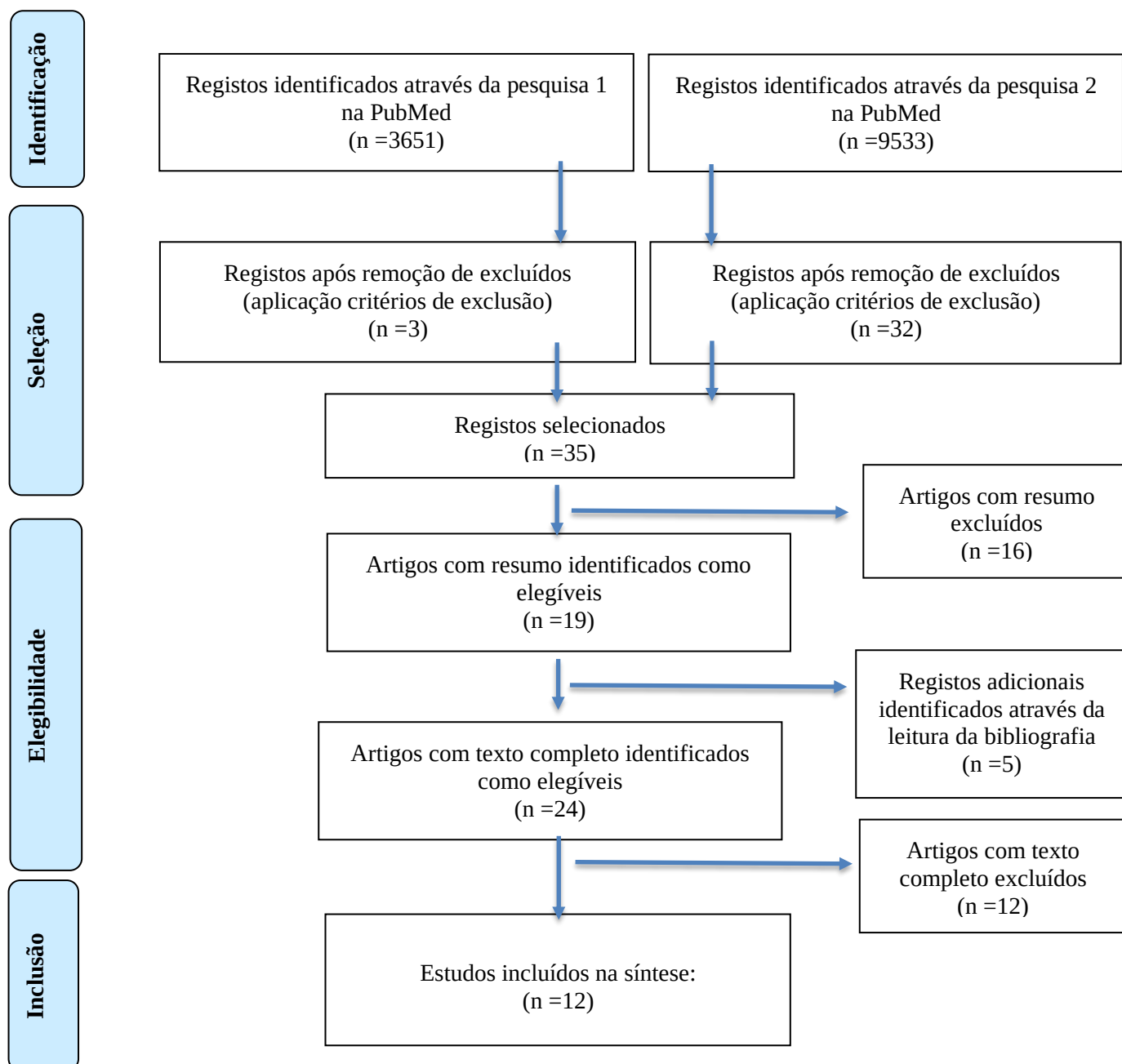
Pesquisa 2: utilizou-se a seguinte instrução de pesquisa: *(weight maintenance or exercise or physical activity or weight loss or weight management or weight) AND (artificial intelligence OR neural network OR data mining OR Big Data OR machine learning)*

- Artigos obtidos após pesquisa na *PubMed*: 9533
- Artigos excluídos: 9501
- Artigos selecionados como elegíveis: 32
- Artigos selecionados após leitura do resumo: 16

Ao total dos 19 artigos selecionados nas 2 pesquisas foram incluídos mais 5, após a leitura dos títulos dos artigos constantes na bibliografia dos 19 artigos selecionados. Do total dos 24 artigos, foram selecionados apenas 12 artigos, já que foram aplicadas as exclusões referenciadas acima, ou seja, foram removidos artigos relacionados com reconhecimento de comportamentos sedentários ou reconhecimento de padrões de atividade física, artigos referentes apenas a perdas de peso baseadas em soluções web e revisões baseadas na utilização de *Big Data*.

Resumindo a pesquisa, através da adaptação do diagrama PRISMA-P, obtivemos o seguinte:

Figura 1 – Diagrama do processo de seleção dos estudos para a RSL rápida



Características dos Estudos Incluídos

Os 12 artigos selecionados, com base nos critérios acima definidos, foram divididos em 2 categorias:

- **Tipo 1** - Artigos que utilizavam inteligência artificial, de modo a elaborar um treinador virtual adequado ao perfil específico de cada indivíduo e assim, o apoiasse na adesão efetiva à atividade física (*Health Coach*)
- **Tipo 2** – Artigos que utilizavam algoritmos de *Machine Learning* numa perspetiva de previsão. Neste caso incluíram-se previsão de adesão à atividade física, ou previsão de falha na capacidade de manutenção do peso perdido, ou de previsão de risco de obesidade ou de alterações de peso.

Considerou-se que ambas as tipologias de artigos se enquadravam na análise a efetuar, pois abrangem 2 áreas de aplicação de metodologia estatística avançada: um área relacionada com a dificuldade da gestão de peso e de prática de atividade física (aqui pretendeu-se entender como este tipo de metodologias avançadas pode permitir que um maior número de pessoas consiga de maneira eficaz manter um peso saudável e aderir à atividade física de forma consistente) e uma outra área que permitiu de alguma forma perceber e prever os respetivos mecanismos associados.

Instrumentos

Os instrumentos utilizados nos estudos do Tipo 1 e Tipo 2 foram divididos em 2 categorias: informação existente a priori, antes do estudo ser efetuado, ou seja, fontes de informação já existentes recolhidas no início do estudo, ou a posteriori, neste caso informação recolhida automaticamente durante o estudo, através de aplicação móvel e/ou de algum dispositivo em utilização pelo utilizador, de dados inseridos na aplicação pelo próprio utilizador e ainda de questionários respondidos pelos utilizadores da aplicação. Os instrumentos agregam-se da seguinte forma:

Tabela 1 – Categorização dos instrumentos

Instrumentos	Informação existente a priori	Informação recolhida durante o estudo ou a posteriori	
	Medições / Planos definidos / Bases de Dados	Informação recolhida automaticamente	Dados inseridos ou questionário feitos utilizador
Tipo 1 - N° de Estudos	3	3	4
Tipo 2 - N° de Estudos	6	1	1

Verificou-se que no que respeita os estudos de Tipo 1, os tipos de fontes de dados eram diversos, ou seja, a informação provinha de diferentes tipos de fontes e usualmente, existiam mais de que uma fonte de dados para a análise. No caso dos estudos classificados como Tipo 2, as fontes de dados foram maioritariamente provenientes de base de dados já existentes, ou seja, não foram recolhidas durante o estudo, mas obtidas a priori.

Principais Resultados

Tipo 1 - Health Coach - Neste subcapítulo incluíram-se 5 artigos que utilizavam inteligência artificial de modo a elaborar um treinador virtual adequado ao perfil específico de cada indivíduo para apoio na adesão à atividade física (*Health Coach*)

Tabela 2 – Principais resultados obtidos – estudos tipo 1

Título: “A Fully Automated Conversational Artificial Intelligence for Weight Loss: Longitudinal Observational Study Among Overweight and Obese Adults” (Stein & Brooks, 2017)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Natalie Stein; Kevin Brooks (2017)	70 indivíduos com excesso de peso ou obesidade (IMC $\geq$ 25 kg/m ²)	- Estudo longitudinal observacional. - Os participantes utilizaram HCAI (<i>Lark Weight Loss Health Coach AI</i>) de forma a encorajar a perda de peso e escolhas saudáveis de dieta através de elementos da terapia comportamental cognitiva.	Avaliar a perda de peso, alterações na qualidade das refeições, e aceitação da aplicação móvel HCAI por parte dos seus utilizadores, de modo a que, de forma abrangente, seja possível aumentar o acesso a cuidados de saúde através de serviços de saúde móvel.	- Dados acerca da perda de peso, qualidade das refeições, atividade física e sono foram recolhidos através do input do utilizador e relativamente ao sono e atividade física, houve também recolha automática através do telemóvel do utilizador. - Dados referentes ao comprometimento do utilizador foram definidos através da duração e quantidade de utilização da mesma. - Dados referentes à usabilidade e aceitação da aplicação móvel foram obtidos através da resposta a um questionário com 4 questões, disponível na própria aplicação.	Utilização de <i>Conversational Artificial Intelligence</i>	Resultados: Este estudo mostrou que a utilização de um treinador criado através de inteligência artificial está associada a uma perda de peso comparável com intervenções personalizadas no estilo de vida. Pode também encorajar modificações comportamentais e tem grande aceitação dos utilizadores. Conclusões: Pesquisa na área da Inteligência Artificial e as suas aplicações à telemedicina deve continuar através de ensaios clínicos que investiguem o comprometimento e aceitação dos utilizadores, os efeitos no peso e nos comportamentos de saúde.

Título: “ <i>Personalized Physical Activity Coaching: A Machine Learning Approach</i> ” (Dijkhuis, Blaauw, van Ittersum, Velthuisen, & Aiello, 2018)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Talko B. Dijkhuis, Frank J. Blaauw, Miriam W. van Ittersum, Hugo Velthuisen and Marco Aiello (2018)	48 participantes, divididos em 2 grupos: um grupo de intervenção e um grupo de controlo	- Estudo experimental - Contagem dos passos dados permitia a elaboração de um plano de treino quinzenal	Investigar a possibilidade de automatizar parte do procedimento de treino de atividade física, através do fornecimento de feedback personalizado e diário, do progresso dos participantes no sentido de atingir o objetivo pessoal de número de passos.	- Contagem de passos recolhidos através de <i>Fitbit Flex</i> . - Vários testes de performance e preenchimento de questionários para avaliar a eficácia da intervenção.	A contagem do nº de passos foi utilizada para treinar 8 algoritmos diferentes de <i>Machine Learning</i> , de modo a estimar a probabilidade de os utilizadores atingirem o objetivo diário de nº de passos	Resultados: Desenvolvimento de uma prova de conceito através da qual a aplicação móvel indica de forma personalizada se é expectável ou não o utilizador atingir o nº de passos diários. <u>Conclusões:</u> A utilização de <i>Machine Learning</i> poderá tornar-se essencial no processo de automatização de treino personalizado. Os algoritmos individualizados permitem prever o nível de atividade física diária e fornecem a possibilidade de intervir no momento certo.

Título: “Evaluating Machine Learning-Based Automated Personalized Daily Step Goals Delivered Through a Mobile Phone App: Randomized Controlled Trial” (Zhou et al., 2018)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Mo Zhou; Yoshimi Fukuoka; Yonatan Mintz; Ken Goldberg; Philip Kaminsky; Elena Flowers; Anil Aswani (2018)	64 participantes recrutados via e-mail. Foi requerida uma sessão inicial presencial	- Randomized controlled trial (RCT) de 10 semanas - O grupo de intervenção recebeu objetivos diários automáticos, adaptáveis e personalizados da meta de passos, enquanto que o grupo de controlo apenas recebeu a meta constante de 10.000 passos diários.	Avaliar a eficácia de uma intervenção automatizada, baseada numa aplicação móvel, personalizada e com objetivos adaptáveis e personalizados, de modo a comparar resultados pessoais com uma meta diária de 10.000 passos	- Contagem do número de passos diários foi obtido através da medição efetuada pela aplicação móvel. - Na 1ª sessão foi instalada a aplicação móvel e foram medidos peso, altura, e calculado o IMC dos participantes. - Na 2ª sessão foi removida a aplicação móvel e participantes responderam a 2 questionários: “ <i>Barriers to Being Active Quiz</i> ” e “ <i>The international physical activity questionnaire</i> ” (versão reduzida)	<i>Behavioral analytics algorithm</i> (BAA); usa <i>Machine Learning</i> para construir um modelo quantitativo preditivo para cada participante com base nos dados históricos de passos e objetivos e, em seguida, utiliza o modelo para gerar objetivos de nº de passos desafiadores, mas realísticos de forma adaptável, escolhendo-os com base na estimativa do modelo, de forma a maximizar a atividade física futura	Resultados: Os participantes do grupo de intervenção tiveram uma diminuição na contagem média de passos diários de 390 entre o início e as 10 semanas, em comparação com uma diminuição de 1350 passos dos participantes do grupo de controle. (960 passos de diferença). Conclusões: Os resultados mostraram a eficácia a curto prazo desta intervenção, que deve ser formalmente avaliada através de um RCT com um período de acompanhamento mais longo.

Título: “Mining Health App Data to Find More and Less Successful Weight Loss Subgroups” (Serrano, Yu, Coa, Collins, & Atienza, 2016)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Katrina Serrano, Mandi Yu, Kisha Coa, Linda Collins, and Audie Atienza (2016)	Uma população de N=12.427.196 dados (aplicação móvel <i>Lose It!</i>), divididos aleatoriamente em 24 subamostras, tendo sido utilizadas 3 delas num total de n=972.687 dados	Estudo transversal observacional	Analisar os dados de uma aplicação móvel comercial (<i>Lose It!</i>) de forma a identificar subgrupos nos quais houve perda de peso bem-sucedida, verificando a estabilidade dos resultados	- Criação de um plano de perda de peso com base na altura, no peso, no nível de exercício, no objetivo de peso desejado e na perda de peso semanal desejada. - Cálculo de um valor estimado de calorias para atingir um deficit de energia necessário para atingir o objetivo de perda de peso. - O plano de perda de peso consiste em registar a dieta, o exercício e o peso - A aplicação móvel oferece ferramentas de motivação e suporte.	Árvores de classificação e regressão, com o objetivo agrupar numa subamostra, indivíduos em subgrupos de idênticas características.	<u>Resultados:</u> Na subamostra na qual 14,96% dos utilizadores perderam 5% ou mais de peso, foram definidos 3 subgrupos: - “utilizadores ocasionais”, com menor proporção de perda de peso, - “utilizadores básicos”, perda de peso intermédia - “utilizadores <i>power</i>”, com maior capacidade de perda de peso Foram os fatores comportamentais que permitiram a caracterização dos indivíduos nos subgrupos <u>Conclusões:</u> Através destes estudos, verifica-se que a segmentação e adaptação das informações fornecidas a subgrupos específicos de utilizadores das aplicações móveis poderia aumentar o sucesso da perda de peso.

Título: “Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity”

(Chin et al., 2016)

Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Sang Ouk Chin, Changwon Keum, Junghoon Woo, Jehwan Park, Hyung Jin Choi, Jeong-taek Woo & Sang Youl Rhee (2016)	Um total de 35.921 participantes dos quais, 77,9% reportaram diminuição do peso corporal enquanto utilizavam a aplicação móvel	Um estudo <i>Cohort</i> retrospectivo	Investigar a eficácia de uma aplicação móvel na redução de peso em indivíduos com obesidade ou com excesso de peso	- Dados clínicos e dados registados pelos utilizadores da aplicação móvel <i>Noom Coach</i> entre outubro de 2012 e abril de 2014. - Informação incluindo: género, idade, altura, peso, variáveis relacionadas com a dieta (frequência de input da informação e calorias das refeições), variáveis relacionadas com exercício (frequência e calorias gastas) variáveis relacionadas com alterações do peso (frequência de input da informação e IMC).	Regressão Linear	Resultados: Este estudo demonstrou a utilidade clínica de uma aplicação móvel para uma redução de peso bem-sucedida para a maioria dos utilizadores; os efeitos foram mais significativos para os indivíduos que monitorizaram o seu peso e a respetiva dieta com maior frequência. Conclusões: Fica demonstrada a utilidade clínica de aplicações móveis para redução de peso com sucesso, mais significativa para indivíduos que monitorizaram o seu peso e dieta mais frequentemente. Deverão ser feitos mais estudos para desenvolvimento de métodos para encorajar a adesão de aplicações de Auto monitorização.

Analisando os resultados obtidos dos 5 estudos de Tipo 1 incluídos nesta RSL rápida verificou-se o seguinte:

- 3 estudos mostraram que a utilização de inteligência artificial permitia melhores resultados no que respeita a perda de peso ou gestão do peso quando comparado com a sua não utilização, já que é uma ferramenta que se assemelha à utilização de intervenções personalizadas;
- 2 estudos apresentaram modelos ou provas de conceito que permitiam que as aplicações móveis, de forma individual, alertassem os utilizadores para uma determinada probabilidade de uma ocorrência, ou seja, considerando as metas definidas (a nível da prática de exercício, por exemplo), o dispositivo identificava se eram ou não atingíveis e dava indicações de como agir em tempo real (*real-time*) para ajudar ao cumprimento das mesmas.

Da avaliação destes estudos resulta sempre algo de positivo, já que todos os estudos apontaram para que a aplicação de inteligência artificial era eficaz no que respeita apoiar a gestão do peso, através também da prática de atividade física.

Tipo 2 - Aplicação de Machine Learning numa perspetiva de previsão - Neste subcapítulo incluem-se 7 artigos que utilizaram algoritmos de *Machine Learning* numa perspetiva de previsão. Neste caso incluíram-se previsão de adesão à atividade física, ou previsão de falha na capacidade de manutenção do peso perdido, ou previsão de risco de obesidade ou de alterações de peso.

Tabela 3 – Principais resultados obtidos – estudos tipo 2

Título: “Patterns and correlates of physical activity in adult Norwegians: a forecasted evolution up to 2025 based on machine learning approach” (Rossi & Calogiuri, 2018)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Alessio Rossi and Giovanna Calogiuri (2018)	52.477 indivíduos adultos (noruegueses)	Estudo transversal observacional	Investigar padrões e correlatos da atividade física utilizando uma abordagem de <i>Machine Learning</i> , assim como construção de um modelo preditivo de futura atividade física.	- Dados retirados do questionário “ <i>Norsk Monitor survey</i> ”, que consiste em cerca de 3000 itens recolhidos entre 1985 e 2013, de características individuais e fatores socioculturais.	- Método de maximização de informação condicional mútua - Modelo autorregressivo para a previsão de atividade física	Resultados: A análise dos padrões de atividade física mostrou: um aumento progressivo da frequência, enquanto que a duração e intensidade de AF estabilizaram. O modelo autorregressivo previu um incremento geral da frequência da e intensidade da atividade física até 2025, enquanto se prevê uma redução da duração. Diferentes padrões surgiram nos diferentes subgrupos, sugerindo menores aumentos de atividade física em homens e indivíduos mais velhos, em comparação com mulheres e indivíduos mais jovens. Conclusões: Os resultados deste estudo podem ajudar a aumentar os níveis de atividade física, direcionando abordagens específicas a populações-alvo diferentes. A abordagem de <i>Machine Learning</i> é proposta como uma ferramenta útil para a monitorização e garantia da saúde pública.

Título: “<i>Application of Machine Learning to Predict Dietary Lapses During Weight Loss</i>” (Goldstein et al., 2018)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Stephanie P. Goldstein, Fengqing Zhang, John G. Thomas, Meghan L. Butryn, James D. Herbert, and Evan M. Forman	12 indivíduos	Participantes tinham que seguir uma dieta de controlo de peso durante 6 semanas	Utilização de técnicas de <i>Machine Learning</i> para prever falhas e avaliar a utilidade de combinar dados individuais e de grupo para realçar a predição de falhas.	- Medidas retiradas de “<i>ecological momentary assessment</i> (via <i>smartphone</i>) referente a falhas na dieta e <i>triggers</i> relevantes. - Participantes tinham que responder a 8 questões de forma quase aleatória.	O estudo treinou e testou um algoritmo de <i>Machine Learning</i> (árvores de decisão WEKA) capaz de prever falhas na dieta num programa de perda de peso comportamental com adultos com excesso de peso/obesidade.	Resultados: Previsões para o grupo com elevada precisão, no entanto o que não aconteceu com a sua generalização para cada indivíduo. Necessários dados individuais recolhidos durante 4 semanas, para otimizar as previsões. Conclusões: O algoritmo de preditivo poderá ser utilizado para fornecer intervenções em tempo real para prevenir falhas na dieta e dessa forma potenciar perda de peso.

Título: “<i>Machine Learning Techniques for Prediction of Early Childhood Obesity</i>” (Dugan, Mukhopadhyay, Carroll, & Downs, 2015)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
T. M. Dugan; S. Mukhopadhyay; A. Carroll; S. Downs (2015)	7519 crianças com idades entre os 2 e os 10 anos	Um estudo Cohort	Previsão a obesidade infantil depois dos 2 anos de idade, utilizando dados recolhidos até ao 2º aniversário por um sistema clínico de apoio à decisão chamado CHICA	- Informação recolhida do Sistema CHICA: questionários dinâmicos (idade e dados demográficos) e historial clínico	Análise de 6 algoritmos de <i>Machine Learning</i> diferentes: <i>RandomTree</i> , <i>RandomForest</i> , J48, ID3, <i>Naïve Bayes</i> , e <i>Bayes</i> .	Resultados: Dos algoritmos analisados, o modelo ID3 mostrou ser aquele com melhor performance, com maior precisão e sensibilidade. Os preditores de obesidade definidos pelo modelo correspondem aos identificados na literatura, suportando assim a validade do modelo, estes referem que a probabilidade de a criança ter excesso de peso ou obesidade depois dos 2 anos é influenciada se a criança o tiver sido antes dos 24 meses ou entre os 12 e os 24 meses. Conclusões: Este estudo mostra que informação recolhida de sistemas de decisão pode ser utilizada Construir modelos de <i>Machine Learning</i> para prever a obesidade em crianças depois dos 2 anos de idade.

Título: “A Data Mining Approach for Examining Predictors of Physical Activity Among Urban Older Adults”

(Yoon, Suero-Tejeda, & Bakken, 2015)

Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Sunmoo Yoon, Niurka Suero- Tejeda, and Suzanne Bakken (2015)	2514 adultos (mais de 55 anos), sobretudo de origem hispânica	Estudo transversal observacional	Desenvolver modelos preditivos para 2 aspetos da atividade física (transporte ativo e tempo de écran)	- Base de dados com respostas a questionários efetuado na comunidade, com informação física objetiva, informação comportamental e informação clínica	<i>Weka</i> <i>software:</i> algoritmo de <i>data mining</i>	<u>Resultados:</u> Principais preditores identificados neste estudo para: 1 - Transporte ativo: estatuto de imigrante elevado nível de ansiedade, ter um local para atividade física e vontade de encontrar tempo para fazer atividade física. 2- Tempo de écran: vontade de encontrar tempo para fazer atividade física, ter um local para atividade física, idade, disponibilidade de apoio familiar para encontrar informação de saúde na internet. Este estudo apoia a necessidade de ocorrer uma mudança nas metas de atividade física para reduzir o tempo de sedentarismo e promover atividades leves entre os idosos, o que pode ser mais realista do que o exercício intenso. <u>Conclusões:</u> Métodos de <i>data mining</i> mostram-se úteis na identificação de populações-alvo para aplicação de intervenção, assim como na definição personalizada dessas mesmas intervenções.

Título: “ <i>Artificial neural networks to predict activity type and energy expenditure in youth</i> ” (Trost, Wong, Pfeiffer, & Zheng, 2012)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Stewart G. Trost, Weng-Keen Wong, Karen A. Pfeiffer, and Yonglei Zheng (2012)	100 participantes com idades compreendidas entre os 5 e os 15 anos	- 12 ensaios de atividade que foram caracterizadas em 5 tipos de atividade física: sedentário, caminhada, corrida, atividades domésticas leves ou jogos, e jogos ou desportos de intensidade moderada a vigorosa	Desenvolver e testar redes neurais artificiais (ANN) para prever tipos de atividade física e dispêndio energético (MET) de dados recolhidos por acelerómetros a crianças e adolescentes.	-Dados recolhidos de acelerómetros colocados na cintura	Redes neurais artificiais (ANN)	<u>Resultados:</u> ANN categorizaram de forma correta tipo de atividade física, assim como para o dispêndio energético (exceto em atividades sedentárias). <u>Conclusões:</u> ANN podem ser utilizadas para prever tipo de atividade física e dispêndio energético em crianças e adolescentes, utilizando dados recolhidos de acelerómetros.

Título: “A Data Mining Approach for Exploring Correlates of Self-Reported Comparative Physical Activity Levels of Urban Latinos” (Yoon, Co, Suero-Tejeda, & Bakken, 2016)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Sunmoo Yoon, Manuel C. CO, MPhila, Niurka Suero-Tejeda, and Suzanne Bakken (2016)	5429 Latinos	Estudo transversal observacional	Construir modelos preditivos para obter informação acerca de níveis de atividade física com base para futuras intervenções junto da população urbana de origem latina.	- Base de dados comportamentais comunitária (WICER), com dados referentes a fatores psicológicos, ambientais, comportamentais, autorrelatos, e sócio demográficos	Técnicas aplicadas de <i>data mining</i> , designadamente 6 algoritmos de <i>Machine Learning</i>	Resultados: Depressão e fadiga (fatores psicológicos), toma de bebidas açucaradas e distúrbios do sono (fatores comportamentais) estão mais fortemente correlacionados com os níveis de atividade física que idade ou educação (fatores demográficos). Conclusões: Métodos de <i>data mining</i> foram úteis na construção de modelos preditivos de forma a obter uma visão acerca da percepção do comportamento de atividade física.

Título: “A Simple Model Predicting Individual Weight Change in Humans” (Thomas et al., 2011)						
Autor (Ano)	Amostra	Desenho do Estudo e Intervenções	Objetivo	Instrumentos	Algoritmos utilizados	Principais resultados / conclusões
Diana Thomas and Corby Martin and Steven Heymsfield and Leanne Redman and Dale Schoeller and James Levine (2011)	Restrição Calórica - Estudo 1: 36 indivíduos; Estudo 2: 23 mulheres com obesidade; <i>Overfeeding</i> : 22 indivíduos	Estudo transversal observacional	Entender os mecanismos e prever resultados de alteração e gestão do peso.	- Dados individuais de 3 estudos: 2 estudos de restrição calórica e 1 estudo de <i>overfeeding</i>	Modelo matemático (modelo com equação diferencial de uma dimensão) construído de modo a de forma exata prever alterações individuais de peso, percebendo como os indivíduos ganham e perdem peso	Resultados: O modelo forneceu melhores predições do peso individual, para os estudos avaliados. Conclusões: O modelo apresentado pode ser utilizado em ambiente clínico, para o propósito definido, assim como para objetivamente medir a adesão a prescrições calóricas.

Analisando os resultados obtidos dos 7 estudos de Tipo 2 incluídos nesta RSL rápida, verificou-se o seguinte:

- 2 estudos foram estudos de categorização, ou seja, a aplicação de *Machine Learning* teve como objetivo categorizar um fenómeno, para que este seja melhor compreendido e assim construir-se intervenções melhor direcionadas. Neste caso, ambos os estudos categorizaram a atividade física, respetivos padrões de intensidade e frequência, um deles junto de indivíduos adultos e o outro junto de uma população mais jovem.
- 5 estudos focaram-se em identificar preditores de diferentes fenómenos ligados à gestão do peso:
 - 1 estudo identificou preditores da falha na dieta durante a perda de peso;
 - 2 estudos identificaram preditores da obesidade / alterações de peso;
 - 2 estudos identificaram preditores da prática de atividade física.

Todos estes estudos tiveram como objetivo comum através da identificação de preditores, a sua manipulação, ou então, a atuação atempada nos preditores para que não ocorram os fenómenos daí decorrentes como: falhar a dieta, aumentar o peso ou não praticar atividade física. Estes estudos concluíram que um maior e mais profundo conhecimento destes preditores poderá ajudar a melhorar a personalização das intervenções na área da gestão do peso.

Qualidade dos Estudos

Foi efetuada a análise de qualidade dos estudos selecionados e categorizados como estudos observacionais, transversais ou de *Cohort*, obtida do “*Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies*” (questionário completo em <https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools>).

Tabela 4 – Resultados da análise da qualidade dos estudos - Estudos observacionais, transversais ou de *Cohort*

CrITÉRIOS de Avaliação	(Stein & Brooks, 2017)	(Dijkhuis et al., 2018)	(Serrano et al., 2016)	(Chin et al., 2016)	(Rossi & Calogiuri, 2018)	(Goldstein et al., 2018)	(Dugan et al., 2015)	(Yoon et al., 2015)	(Trost et al., 2012)	(Yoon et al., 2016)	(Thomas et al., 2011)
1. Clareza do Objetivo	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim
2. Definição da população	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim
3. Taxa de participação	não	N.D.	não	sim	sim	não	N.A.	sim	N.D.	sim	N.D.
4. Recrutamento	sim	sim	sim	sim	sim	sim	sim	sim	N.D.	sim	sim
5. Justificação da amostra	sim	não	sim	não	não	não	não	não	não	não	N.D.
6. Medição previa das variáveis independentes	sim	sim	N.A.	sim	sim	sim	sim	sim	sim	sim	sim

7. Espaço temporal	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim
8. Variáveis independentes medidas ao longo do tempo ou como variável contínua	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim
9. Definição das variáveis independentes	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim
10. Variáveis independentes medidas ao longo do tempo	não	sim	sim	sim	sim	sim	sim	não	sim	sim	sim
11. Definição das variáveis dependentes	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim
12. Resultados mascarados para participantes	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim	sim
13. Follow-up	sim	sim	N.D.	N.D.	N.D.	sim	N.A.	N.D.	N.D.	N.D.	N.D.

14. Variáveis medidas e ajustadas estatisticamente	N.D.	N.D.	N.D.	N.D.	sim	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Resultado Final	11/14	11/14	10/13	11/14	12/14	11/14	10/12	10/14	9/14	11/14	10/14
Classificação	Razoável	Razoável	Razoável	Razoável	Razoável	Razoável	Razoável	Razoável	Razoável	Razoável	Razoável

N.D. – Não disponível

N.A. – Não aplicável

Para a classificação dos estudos foi utilizada a metodologia definida em <https://www.nhlbi.nih.gov/node/80102> – “*Significance of the quality ratings of good, fair, or poor*”, que define a classificação dos estudos através do risco de viés dos mesmos. Dessa forma, os estudos foram categorizados como “bom”, “razoável” ou “pobre”. Todos os estudos apresentaram algum risco de viés, ou por não cumprirem com os critérios definidos ou por omitirem informação.

Assim, considerou-se a classificação generalizada de “razoável” para o conjunto de todos os estudos avaliados, ou seja, um nível de qualidade abaixo de “bom”, pois alguns dos itens não foram possíveis de ser classificados ou não foram considerados durante a elaboração do estudo.

Para o único estudo selecionado e categorizado como RSL, que se decidiu incluir nesta RSL rápida por se tratar de uma área de estudo recente e respetiva escassez de estudos existentes e pela necessidade de incluir toda a evidência possível, apresenta-se a respetiva análise da qualidade na seguinte tabela resumo obtida do “*Quality Assessment of Systematic Reviews and Meta-Analyses*” (questionário completo em <https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools>).

Tabela 5 – Resultados da análise da qualidade dos estudos – RCT

Critérios de Avaliação	(Zhou et al., 2018)
1. Estudo RCT	sim
2. Adequação do método de randomização	sim
3. Alocação do tratamento foi ocultada	sim
4. Atribuição do grupo de intervenção não conhecida	sim
5. Atribuição dos participantes aos grupos não conhecida	sim
6. Grupos similares no <i>Baseline</i>	sim
7. Taxa de <i>drop-out</i>	sim
8. Diferencial taxa de <i>drop-out</i>	sim
9. Alta adesão ao protocolo de intervenção	sim
10. Outras intervenções evitadas	sim
11. Resultados obtidos através de medidas confiáveis	sim
12. Tamanho da amostra suficiente	sim
13. Resultados reportados	N.D.
14. Participantes analisados no respetivo grupo	sim
Resultado Final	13/14
Classificação	Bom

N.D. – Não disponível

O único RCT considerado nesta RSL rápida apresentou um nível de qualidade elevado, com um risco de viés reduzido, revelando-se o único estudo com boa qualidade.

Discussão dos Resultados

Considerando o âmbito desta RSL rápida, que avaliou a aplicação de *Machine Learning* no contexto da gestão do peso e da adesão á atividade física, e considerando Deo (2015), verificou-se que é através deste tipo de técnicas, que combinam a possibilidade de aprender relações existentes nos dados com a capacidade computacional, que novas formas de interpretar e compreender informação surgem. Naturalmente, seria de esperar que os artigos que aplicam estas técnicas tivessem resultados positivos, pois são já amplamente conhecidas na literatura as vantagens da aplicação de métodos de *Machine Learning* com provas dadas em áreas tão dispares como finanças ou física.

No contexto da saúde, a aplicação deste tipo de técnicas reveste-se de cuidados adicionais, e é menos intensiva, naturalmente pelo possível risco e impacto que um output errado fornecido por uma máquina possa ter num qualquer paciente, daí que a intervenção humana no processo é sempre essencial. Assim, é perceptível o porquê da maioria dos artigos que foram encontrados nesta revisão utilizarem dados já existentes, pois dessa forma, as conclusões retiradas das análises já não têm qualquer impacto na respetiva população, e são meramente exercícios académicos.

Finalmente, e sumando os principais resultados obtidos dos estudos avaliados, verificou-se que todos apontaram para que intervenções personalizadas, com inclusão de atividade física regular, eram essenciais para o sucesso individual da gestão do peso, facto também já evidenciado na literatura (Ramage et al., 2014). No que respeita os estudos focados em modelos preditivos, os resultados mais interessantes foram a perceção que a prática de atividade física se ajustava aos padrões de vida atuais, ou seja, tem vindo a ocorrer uma redução da duração da prática, mas aumento da sua intensidade e, que os preditores para a prática de atividade física parecem estar mais associados a fatores psicológicos e comportamentais do que a fatores sociodemográficos.

Limitações

Para os artigos selecionados, que se focaram na utilização de técnicas de inteligência artificial, designadamente *Machine Learning*, no contexto da atividade física e gestão do peso, a principal limitação encontrada foi a fraca experiência na validação destas técnicas, assim como a existência de pouca literatura neste âmbito. Praticamente todos os estudos referiram a necessidade de validação dos modelos e de mais investigação, ainda que os resultados obtidos parecessem bastante promissores.

Por outro lado, os estudos que procuraram prever fenómenos ligados à gestão do peso indicaram algumas pistas que poderiam ser alvo de análise, mas que estavam bastante condicionadas pelas amostras dos estudos, que eram pouco abrangentes e dificilmente poderiam ser transpostas de forma

direta para a população geral. Não foi perceptível se funcionariam quando aplicadas em grupos mais abrangentes e mais diversificados de indivíduos.

Outra limitação prendeu-se com o tipo de estudos avaliados, já que foram encontrados mais estudos observacionais do que experimentais ou RCT, isto deveu-se também ao facto das técnicas de *Machine Learning* utilizarem grandes quantidades de informação e esta estar muitas vezes disponível em bases de dados já existentes. Possivelmente, no futuro, os modelos de *Machine Learning* poderão vir a ser alimentados de forma permanente com informação recolhida de forma ininterrupta, e fornecerem inteligência acionável em tempo real.

Finalmente, importa mencionar que se tratou de uma RSL rápida com apenas uma base de dados pesquisada e com apenas um autor a realizar a maior parte dos processos de pesquisa e avaliação dos estudos, o que se traduz por uma menor quantidade e precisão da evidência reportada.

Conclusões

Podemos concluir da análise efetuada que a inteligência artificial e as metodologias estatísticas mais avançadas são algo de inevitável, essencial e obrigatório, no que respeita a análise da informação. Hoje em dia, a quantidade de dados existente só pode ser transformada em informação de valor, quando se utilizam métodos capazes de a processar em tempo útil. O processamento da informação recolhida por *wearables* ou aplicações móveis e acelerómetros, entre outros, irá permitir que o fenómeno de gestão do peso e adesão a atividade física sejam compreendidos de forma mais clara, e dessa forma, sejam construídas intervenções totalmente personalizadas e adaptáveis no tempo, com base nas alterações detetadas ou reportadas, sejam elas de ordem física ou psicológica, motivacional ou sociodemográfica.

É muito claro que a inteligência artificial e os métodos de *Machine Learning* aplicados à gestão do peso e à promoção da prática de atividade física são uma enorme vantagem e poderão, desde já, apoiar na redução do flagelo já amplamente conhecido, que é a obesidade.

II.4 Referências Bibliográficas

- Aswani, A., Kaminsky, P., Mintz, Y., Flowers, E., & Fukuoka, Y. (2019). Behavioral Modeling in Weight Loss Interventions. *Eur J Oper Res*, 272(3), 1058-1072.
doi:10.1016/j.ejor.2018.07.011
- Bento, T. (2014) Revisões sistemáticas em desporto e saúde: Orientações para o planeamento, elaboração, redação e avaliação. *Motri*. [online]. 2014, vol.10, n.2, pp.107-123. ISSN 1646-107X. [http://dx.doi.org/10.6063/motricidade.10\(2\).3699](http://dx.doi.org/10.6063/motricidade.10(2).3699).
- Chin, S. O., Keum, C., Woo, J., Park, J., Choi, H. J., Woo, J. T., & Rhee, S. Y. (2016). Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity. *Sci Rep*, 6, 34563. doi:10.1038/srep34563
- Deo, R. C. (2015). Machine Learning in Medicine. *Circulation*, 132(20), 1920-1930.
doi:10.1161/CIRCULATIONAHA.115.001593
- Dijkhuis, T. B., Blaauw, F. J., van Ittersum, M. W., Velthuisen, H., & Aiello, M. (2018). Personalized Physical Activity Coaching: A Machine Learning Approach. *Sensors (Basel)*, 18(2). doi:10.3390/s18020623
- Dobbins, M. (2017). Rapid Review Guidebook. *Hamilton, ON*: National Collaborating Centre for Methods and Tools.
- Domingos, P. (2012). A few useful things to know about machine learning. *Commun. ACM*. 55(10):78–87.
- Dugan, T. M., Mukhopadhyay, S., Carroll, A., & Downs, S. (2015). Machine Learning Techniques for Prediction of Early Childhood Obesity. *Appl Clin Inform*, 6(3), 506-520.
doi:10.4338/ACI-2015-03-RA-0036.
- Ganann et al. (2010), Expediting systematic reviews: methods and implications of rapid reviews *Implementation Science*, 5:56. doi: 10.1186/1748-5908-5-56.
- Goldstein, S. P., Zhang, F., Thomas, J. G., Butryn, M. L., Herbert, J. D., & Forman, E. M. (2018). Application of Machine Learning to Predict Dietary Lapses During Weight Loss. *J Diabetes Sci Technol*, 12(5), 1045-1052. doi:10.1177/1932296818775757
- Markland D, Tobin V. A modification to the behavioural regulation in exercise questionnaire to include an assessment of Amotivation. *J Sport Exer Psychol* 2004;26:191–6.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097. Retrieved from <https://dx.plos.org/10.1371/journal.pmed.1000097>
- Novatchkov, H., & Baca, A. (2013). Artificial intelligence in sports on the example of weight training. *J Sports Sci Med*, 12(1), 27-37.

- Poulimeneas, D., Yannakoulia, M., Anastasiou, C. A., & Scarmeas, N. (2018). Weight Loss Maintenance: Have We Missed the Brain? *Brain Sci*, 8(9). doi:10.3390/brainsci8090174
- Ramage, S., Farmer, A., Eccles, K. A., & McCargar, L. (2014). Healthy strategies for successful weight loss and weight maintenance: a systematic review. *Appl Physiol Nutr Metab*, 39(1), 1-20. doi:10.1139/apnm-2013-0026
- Rossi, A., & Calogiuri, G. (2018). Patterns and correlates of physical activity in adult Norwegians: a forecasted evolution up to 2025 based on machine learning approach. *BMC Public Health*, 18(1), 913. doi:10.1186/s12889-018-5854-2
- Serrano, K. J., Yu, M., Coa, K. I., Collins, L. M., & Atienza, A. A. (2016). Mining Health App Data to Find More and Less Successful Weight Loss Subgroups. *J Med Internet Res*, 18(6), e154. doi:10.2196/jmir.5473
- Shamseer L., Moher D., Clarke M., Ghersi D., Liberati A., Petticrew M., Shekelle P., Stewart L. (2015). PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*, 349 (jan02 1):g7647
- Stein, N., & Brooks, K. (2017). A Fully Automated Conversational Artificial Intelligence for Weight Loss: Longitudinal Observational Study Among Overweight and Obese Adults. *JMIR Diabetes*, 2(2), e28. doi:10.2196/diabetes.8590
- Study Quality Assessment Tools. Nacional Heart, Lung and Blood Institute. Retrieved from: <https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools>
- Thomas, D. M., Martin, C. K., Heymsfield, S., Redman, L. M., Schoeller, D. A., & Levine, J. A. (2011). A Simple Model Predicting Individual Weight Change in Humans. *J Biol Dyn*, 5(6), 579-599. doi:10.1080/17513758.2010.508541
- Trost, S. G., Wong, W. K., Pfeiffer, K. A., & Zheng, Y. (2012). Artificial neural networks to predict activity type and energy expenditure in youth. *Med Sci Sports Exerc*, 44(9), 1801-1809. doi:10.1249/MSS.0b013e318258ac11
- Wing, R. R., & Phelan, S. (2005). Long-term weight loss maintenance. *Am J Clin Nutr*, 82(1 Suppl), 222S-225S.
- World Health Organization. Health topics. Obesity. Retrieved from: <http://www.who.int/topics/obesity/en/>
- World Health Organization. (2016). Obesity and overweight, Key facts. *Fact sheet* - Reviewed February 2018.
- Yoon, S., Co, M. C., Suero-Tejeda, N., & Bakken, S. (2016). A Data Mining Approach for Exploring Correlates of Self-Reported Comparative Physical Activity Levels of Urban Latinos. *Stud Health Technol Inform*, 225, 553-557.
- Yoon, S., Suero-Tejeda, N., & Bakken, S. (2015). A Data Mining Approach for Examining Predictors of Physical Activity Among Urban Older Adults. *J Gerontol Nurs*, 41(7), 14-20. doi:10.3928/00989134-20150420-01

Zhou, M., Fukuoka, Y., Mintz, Y., Goldberg, K., Kaminsky, P., Flowers, E., & Aswani, A. (2018). Evaluating Machine Learning-Based Automated Personalized Daily Step Goals Delivered Through a Mobile Phone App: Randomized Controlled Trial. *JMIR Mhealth Uhealth*, 6(1), e28. doi:10.2196/mhealth.9117

CAPÍTULO III

IDENTIFICAÇÃO DE NÍVEIS DE ATIVIDADE FÍSICA AO LONGO DE 12 MESES, COM *MACHINE LEARNING*, E A SUA ASSOCIAÇÃO COM VARIÁVEIS MOTIVACIONAIS, EM ADULTOS EM PROCESSO DE MANUTENÇÃO DE PERDA DE PESO NO PROJETO *NOHOW* - UM ESTUDO DE *COHORT* PROSPETIVO

Resumo

Objetivo: Identificar que características motivacionais estariam associadas a indivíduos adultos em processo de manutenção do peso perdido que mantiveram níveis mais elevados e níveis menos elevados de prática de atividade física a prática de atividade física ao longo de 12 meses.

Método: Estudo de *Cohort* prospetivo. O estudo contou com a participação de 1627 indivíduos, de 3 países europeus (Portugal, Reino Unido, e Dinamarca), com idade superior a 18 anos, com índice de massa corporal (IMC) acima dos 25 kg/m² nos 12 meses anteriores ao projeto *NoHoW*, com pelo menos uma tentativa de perda de peso durante a vida e com uma tentativa de perda de peso (completada ou ainda a decorrer) nos 12 meses anteriores. Foi utilizada informação de *FiBit* (24/7), e após limpeza e engenharia dos dados, reteve-se a informação de 1139 indivíduos, na qual foi aplicado um método de *clustering* de *Machine Learning*, o *K-Means*, para identificação de quais os participantes que praticavam atividade física de forma ativa “Ativos” (Ativ) e os outros, “Insuficientemente Ativos” (I-Ativ). Verificou-se com questionários motivacionais aplicados no *Baseline*, 6 e 12 meses, o respetivo perfil motivacional. Foram aplicados testes de hipóteses para comprovar as diferenças verificadas entre ambos os grupos.

Resultados: Verificou-se que é possível diferenciar os Ativ dos outros indivíduos, I-Ativ. O perfil motivacional dos Ativ inclui altos níveis de Motivação Autónoma, Regulação Intrínseca, Integrada e Identificada sobretudo nos primeiros 6 meses de avaliação, fase em que que estes índices aumentam. Estes indivíduos apresentaram também uma característica interessante, o seu nível de Autonomia, no contexto da gestão do peso, diminuiu entre os 6 e 12 meses, apresentando valor mais baixo no final dos 12 meses.

Conclusões: Os algoritmos de *Machine Learning* foram métodos eficazes na segregação de indivíduos por padrão de atividade física. Verificou-se que são níveis elevados de Motivação Autónoma, Regulação Intrínseca, Integrada e Identificada ao longo dos 12 meses que parecem permitir que os indivíduos façam e permaneçam empenhados na prática diária de atividade física. A Autonomia, no contexto da gestão do peso, aos 12 meses deverá ser avaliada, pela sua diminuição face ao nível no *Baseline*. Este estudo permitiu validar a informação já existente na literatura, que suporta que ao intervirmos a nível motivacional, promovemos a capacidade individual de adotar um comportamento em prol da prática de atividade física.

Palavras chave: Atividade Física, *Machine Learning*, Motivação, Regulação Comportamental

Abstract

Objective: Identify which motivational characteristics were associated with adult individuals in the process of weight loss maintenance, who were engaged in physical activity practice in higher and lower levels for 12 months.

Method: Prospective Cohort Study. The study involved 1627 individuals from 3 European countries (Portugal, the United Kingdom, and Denmark), aged over 18 years, with body mass index (BMI) above 25 kg/m² in the 12 months prior to *NoHoW* project, with at least one lifetime weight loss attempt and one weight loss attempt (completed or ongoing) in the previous 12 months. Information from *FiBit* (24/7) was used, and after data cleaning and engineering, information from 1139 individuals was retained, in which a *Machine Learning clustering* method, *K-Means*, was applied to identify which participants practice exercise actively, named “Active” (Ativ), and the others, named “Somewhat Active” (I-Ativ). With motivational questionnaires applied at Baseline, 6 and 12 months, it was verified their motivational profile. Hypothesis tests were applied to prove the differences between both groups.

Results: It was possible to differentiate Ativ from I-Ativ. The motivational profile of Ativ included high levels of Autonomous Motivation and Intrinsic, Integrated and Identified Regulation, mainly in the first 6 months of assessment, where we observed higher values. These individuals exhibited an interesting feature, their level of autonomy in weight management decreases between 6 and 12 months, presenting the lowest value at the end of the 12 months.

Conclusions: *Machine Learning* algorithms were effective methods when we aimed to segregate individuals by physical activity patterns. High levels of Autonomous Motivation and Intrinsic, Integrated and Identified Regulation over 12 months, seemed to allow individuals to maintain and remain engaged in daily practice of exercise. Autonomy in weight management context decreased at 12 months and should be assessed further. This study allowed us to validate the information already available in the literature, which supports that when we intervene at a motivational level, we enhance the ability of people to adopt physical activity behavior.

Keywords: Physical Activity, *Machine Learning*, Motivation, Behavioral Regulation

III.1 Introdução

Racional teórico

Segundo a DGS (2017), a obesidade “é um problema de saúde pública e uma doença crónica. É definida como uma patologia em que o excesso de gordura corporal acumulada pode afetar a saúde. Com origem em diversos fatores, a obesidade requer esforços continuados para ser controlada. Constitui uma ameaça grave para a saúde e um importante fator de risco para o desenvolvimento e agravamento de outras doenças”. É na tentativa de reverter este fenómeno, que se recomenda a prática ativa de atividade física de forma consistente e continuada, entre outras recomendações, essencial num estilo de vida saudável. A questão que se impõe é o que significa prática ativa de atividade física, como identificar padrões de atividade física que correspondam a uma prática ativa de exercício, de forma a ajudar a população a mitigar o sedentarismo e a reverter a tendência existente de aumento de peso da população.

Estudos efetuados – Estilo de Vida e Mudanças Comportamentais

Numa tentativa de reverter o fenómeno da obesidade tem-se investido em intervenções no estilo de vida de pessoas com excesso de peso e obesidade, e essas têm-se mostrado eficazes para a perda de peso. Contudo, falham quando se trata da manutenção do peso perdido, verificando-se frequentemente a recuperação do mesmo. Assim, tem-se assistido a uma mudança de paradigma nos investigadores, de modo a que não só a perda de peso seja motivo de análise, mas também a manutenção e a prevenção da recuperação do peso perdido. Parece existir uma falha na capacidade de desenhar intervenções assertivas, a nível de mudança do estilo de vida e da prática de atividade física, que resultem na efetiva manutenção do peso perdido (Hall & Jordan, 2008).

Adicionalmente, e segundo um estudo baseado na revisão sistemática de 26 artigos do *PubMed* e *Scopus*, utilizando palavras chave como: excesso de peso, obesidade, manutenção de peso, recuperação de peso e prescrição de dieta, verificou-se que estratégias como a substituição de refeições, a dieta de baixo índice glicémico, a alta ingestão de proteínas, o consumo moderado de gordura e uma dieta específica para controlar a hipertensão (DASH – *Dietary Approach to Stop Hypertension*) apresentaram alguns efeitos positivos na manutenção do peso, ou seja, sugerindo o papel benéfico da redução da ingestão calórica, presente em todas estas estratégias. No entanto, os resultados foram controversos e verificou-se a necessidade de mais estudos. A nível de mudanças do estilo de vida, alguns comportamentos foram associados a menor recuperação de peso, como “não ficar acordado até tarde de noite”, “beber menor quantidade de bebidas açucaradas” e “seguir um padrão de vida saudável”. Alguns alimentos (caféina e chá verde) foram sugeridos como auxiliares na manutenção de peso, mas sem validação ou evidência científica (Soeliman & Azadbakht, 2014).

Uma outra revisão sistemática da literatura (RSL) em base de dados como *CAB Abstracts*, *Central Register of Controlled Trials*, *EMBASE*, *MEDLINE*, *Food Science and Technology Abstracts*, e *Web of Knowledge* examinou a evidência científica existente nas estratégias eficazes para a perda de peso e manutenção de peso perdido, em adultos com excesso de peso e obesidade e com idades compreendidas entre os 18 e os 65 anos. Os 67 estudos analisados resultaram numa seleção de artigos, nos quais as intervenções aplicadas foram consideradas bem-sucedidas, ou seja, os participantes conseguiram, desde um momento marcado como ponto de partida (*Baseline*), perder 5% ou mais de peso e mantiveram essa perda de peso desde a *Baseline* até aos 12 meses. A nível global, e no que respeita a perda de peso, esta foi conseguida com sucesso, através do défice de ingestão calórica atingido através da redução de ingestão de gordura, e também do aumento da ingestão de fibra, que se revelou uma componente eficaz em 21% das intervenções. 88% das intervenções bem-sucedidas incluíram atividade física e 92% o treino comportamental, como a auto monitorização. A combinação de restrição calórica e de gordura, atividade física regular e estratégias comportamentais foi necessária na fase de manutenção do peso perdido. Assim, o estudo concluiu que uma abordagem abrangente fundamentada por evidência científica foi eficaz na perda de peso e manutenção do peso perdido, em adultos (Ramage et al., 2014). Por outro lado, tem vindo a ser assumido, sem qualquer evidência que o suporte, que o excesso de peso percebido é também um importante pré-requisito para uma eficaz manutenção do peso. O propósito do estudo de Haynes, Kersbergen, Sutin, Daly e Robinson (2018) foi encontrar evidência, através de RSL, da existência de relações entre tentativas de perda de peso, estratégias (saúdáveis e não saúdáveis) de controlo de peso, comportamentos relacionados com o peso (atividade física e hábitos alimentares), distúrbios alimentares e mudanças de peso. Foram reunidas evidências de 78 estudos e os resultados indicaram que o excesso de peso percebido estava associado, em alguns grupos, ao aumento da probabilidade de utilização de estratégias de controlo de peso (saúdáveis e não saúdáveis) com o objetivo de perda de peso. No entanto, o excesso de peso percebido não foi associado de forma inequívoca à prática de atividade física ou à escolha de uma alimentação saudável, foi ao contrário, associado a distúrbios alimentares. Ao invés de estar associado a uma melhor gestão do peso, houve evidências consistentes de que o excesso de peso percebido era preditivo de aumento da recuperação do peso perdido, ao longo do tempo. Apesar dos indivíduos que percebem que estão com excesso de peso serem mais propensos a efetuar tentativas de perda de peso, ao longo do tempo, estes acabam por voltar a ganhar peso (Haynes et al., 2018).

Adicionalmente, há evidência científica que corrobora que, quando se trata de perder peso, alterações nas rotinas diárias de alimentação e atividade física são suficientes para que os resultados ocorram. No entanto, no que respeita a manutenção do peso perdido, apenas alterações de comportamento, relacionadas com intervenções direcionadas à motivação dos indivíduos conseguem promover uma efetiva manutenção do peso a longo prazo (Santos, Silva, & Teixeira, 2016). Aponta-se então para a componente do comportamento humano, esse fenómeno tão complexo, mas tão essencial nesta análise. Coloca-se a necessidade de responder à seguinte questão: qual ou quais as melhores

técnicas de mudança de comportamento, no contexto da manutenção do peso perdido? E para responder a esta questão, o caminho parece ser perceber também, além dos caminhos abordados pelos estudos mais focados na nutrição, qual ou quais as técnicas comportamentais utilizadas pelos indivíduos que sucedem nesta tarefa de manutenção do peso.

Abordou-se, portanto um estudo qualitativo acerca do papel da regulação emocional na manutenção do peso perdido e das suas interações na autorregulação motivacional e comportamental. Neste estudo, que se enquadra no projeto *NoHoW* e que será igualmente a base do estudo aqui apresentado, foram efetuadas entrevistas qualitativas a 44 indivíduos que anteriormente tiveram obesidade (25 provenientes do Reino Unido e 19 provenientes da Dinamarca) e que perderam mais de 5% do seu peso corporal com sucesso após uma tentativa deliberada de perda de peso, e que estariam atualmente a tentar manter essa perda de peso. As conclusões mostraram que é muito importante considerar a regulação emocional individual, quando se tenta compreender o fenómeno de manutenção do peso perdido. A motivação e a autorregulação podem ficar comprometidas por fortes emoções ou por estratégias de *coping* ou enfrentamento emocional, definido como “esforços cognitivos e comportamentais para lidar com situações de dano, de ameaça ou de desafio quando não está disponível uma rotina ou uma resposta automática”, e isso deve ser considerado na definição de intervenções e no apoio para evitar a recuperação de peso perdido (Sainsbury et al., 2018).

Analisando um estudo muito recente, também publicado no âmbito do projeto *NoHoW*, com o propósito de compreender o processo de manutenção de peso perdido, associou-se a compulsão alimentar a alguns fenómenos subjacentes à manutenção do peso perdido. Esta desregulação emocional poderia constituir-se como uma barreira às intervenções a longo prazo para a prevenção e redução da obesidade. Assim, o objetivo do estudo foi determinar a relação entre as dificuldades na regulação das emoções e a recuperação do peso perdido em adultos que perderam peso e depois voltaram a ganhá-lo, e examinar as características associadas a essas dificuldades. O método utilizado nos 2000 participantes deste projeto foi a aplicação de um questionário online, para aferir a perda de peso e posterior recuperação do peso, na sua mais recente tentativa de perda de peso. Adicionalmente, os participantes preencheram um questionário para identificar a possível presença de compulsão alimentar, e caso tenham recuperado o peso perdido, era perguntado se considerariam que isso se devia a fatores emocionais, ou seja, ligando assim o fenómeno de recuperação de peso à dificuldade na regulação de emoções. Foram utilizados métodos estatísticos que incluíram o coeficiente de correlação de *Spearman* e a regressão logística, de modo a encontrar associações entre a regulação de emoções, a recuperação de peso e a estratégia utilizada, tendo os resultados apontado para uma associação positiva entre a dificuldade na regulação de emoções e uma maior recuperação do peso perdido (no total 1594 participantes que perderam peso e recuperaram o peso perdido). A atribuição da falha na manutenção do peso perdido a razões emocionais foi sugerida maioritariamente por participantes jovens do sexo feminino, com perda de controlo e compulsão alimentar identificada, com menor perceção de sucesso na manutenção do peso perdido, com utilização de mais estratégias de dieta e de autorregulação na

perda de peso e menores estratégias dietéticas na manutenção do peso perdido. Em conclusão, este estudo apontou para a ligação entre as dificuldades na regulação emocional relacionadas com o peso e uma maior recuperação do peso perdido, nos indivíduos que voltaram a ganhar peso após o terem perdido. Os indivíduos que perderam peso utilizaram estratégias comportamentais durante a fase de perda de peso, mas posteriormente não foram capazes de o fazer, de forma consistente, na fase de manutenção do peso perdido. Incentivar a utilização dessas estratégias na fase de manutenção de peso perdido, sem efetivamente ensinar as pessoas a desenvolver competências na regulação de emoções pode não ser o meio mais eficaz para garantir o sucesso da manutenção de peso, e evitar a recuperação do peso perdido (Sainsbury et al., 2018).

Estudos efetuados – Estratégias baseadas na teoria da auto-determinação

Estreitando a análise na linha das estratégias comportamentais no que concerne a manutenção do peso, importa abordar a teoria da autodeterminação (TAD) desenvolvida por Ryan e Deci (2000) baseada no conceito de motivação, sendo esta fundamental na mudança comportamental. A TAD é aqui abordada como uma possível e eficaz estratégia para ajudar na adesão à prática de atividade física, e consequentemente facilitar a manutenção do peso perdido.

Na TAD, a motivação é o conceito central na modificação comportamental, esta teoria é revista por Flannery (2017) demonstrando o ênfase na motivação intrínseca, já que esta está associada com melhoria no bem-estar e mudanças comportamentais consistentes. Aplicando este conceito à prática de atividade física, será então essencial que para a manutenção regular da prática de atividade física, um indivíduo precisará de incorrer num processo de modificação comportamental para adquirir motivação intrínseca.

A TAD baseia-se na satisfação de três necessidades básicas do ser humano: a competência, a autonomia e o relacionamento positivo, já que estas promovem a motivação intrínseca, que está associada à melhoria do bem-estar e à mudança comportamental sustentada. Ou seja, quando estas três necessidades básicas são satisfeitas, e segundo a TAD, é criado o contexto para que seja possível a um indivíduo sentir-se realizado, e consequentemente manter um determinado comportamento.

Considerando a análise efetuada no estudo de Teixeira, Silva, Mata, Palmeira e Markland (2012) verificou-se que as intervenções que têm como objetivo o controlo de peso são menos bem-sucedidas quando se trata da manutenção do peso perdido, isto porque negligenciam, em parte, as dimensões associadas à motivação, tais como a autonomia percebida, resultando numa desconexão entre perda de peso e comportamentos relacionados com o controlo do peso. É sugerido que se os indivíduos se comprometerem totalmente com os objetivos comportamentais relacionados com a perda de peso e se se sentirem competentes e autónomos no seu cumprimento, como sugerido na TAD, os seus esforços resultarão numa modificação comportamental sustentada.

Adicionalmente, foi considerada a RSL de Rodrigues et al. (2018) na qual foi analisado se o comportamento interpessoal influenciava a persistência e adesão à atividade física em indivíduos

adultos, utilizando então como base teórica a TAD. Foram avaliados um total de 35 estudos empíricos publicados entre 1985 e 2018. Verificou-se que o suporte da autonomia promovido pelos profissionais de atividade física revelou-se fundamental e consequentemente contribuiu de forma significativa na satisfação das necessidades psicológicas básicas dos indivíduos, mas menores regulações autodeterminadas pareceram afetar negativamente a prática de exercício. Assim, de acordo com esta RSL, as intervenções baseadas na TAD foram fundamentais quando se trata de promover a prática de atividade física a longo prazo.

Um outro estudo de Silva et al. (2008) referente a um RCT, abordou também a TAD na adesão à atividade física e controlo de peso. Considerando que adotar e manter um estilo de vida fisicamente ativo é claramente a estratégia mais eficaz para evitar a obesidade, sobretudo a longo prazo, foi considerada uma intervenção comportamental com a duração de um ano e follow-up nos 2 anos seguintes. O estudo incluía mulheres entre os 25 e 50 anos com excesso de peso ou com obesidade moderada. A intervenção foi definida em conformidade com a TAD e com as *guidelines* da entrevista motivacional. Os resultados deste RCT teve como objetivo entender como as características motivacionais, particularmente as relacionadas com a autorregulação comportamental na prática de atividade física, influenciavam o sucesso da intervenção e consequentemente como a SDT podia ser útil na promoção da modificação comportamental na saúde, num contexto de obesidade. Os resultados deste RCT encontram-se explanados no estudo posterior de Silva et al. (2011), onde foram consideradas 221 mulheres. Verificou-se que nem todos os tipos de motivação preveem os resultados comportamentais a longo prazo, e que exercício moderado e vigoroso medeiam a modificação do peso a longo prazo. Parece existir evidência forte na ligação entre motivação autónoma, exercício e manutenção do peso perdido a longo prazo, sendo, portanto, notória a importância de construir intervenções que se foquem na internalização da regulação comportamental na atividade física, tornando assim os momentos de atividade física em algo positivo e em experiências de valor, em vez de visar a mudança imediata de comportamento.

Importa referir a RSL de Teixeira, Carraça, Markland, Silva e Ryan (2012), na qual foi indicado que a motivação é um fator crítico para a manutenção sustentada da prática de atividade física e daí ser inquestionável a importância da TAD, onde foram analisados 66 estudos empíricos. Verificou-se que existe uma relação positiva entre formas de motivação autónoma e exercício. Interessante verificar que a regulação identificada previu mais eficazmente a iniciação da prática de atividade física, enquanto no caso da adesão à atividade física a longo prazo, esta ocorreu graças às componentes de motivação intrínseca e satisfação da competência. No entanto, na literatura disponível, parecem existir ainda algumas inconsistências no que respeita a ligação entre exercício e constructos específicos da TAD.

Finalmente e num estudo longitudinal com 94 participantes, no qual se utilizaram 2 tipos de questionários, um para perceber quais as expectativas associadas à prática de atividade física (IMPRAF-54) e um outro para avaliar o nível de motivação (BREQ-2), verificou-se que a expectativa de interação social foi um fator positivo na predição da adesão à atividade física e que a atividade física não era um

fim, mas um meio para atingir autonomia. Ou seja, a associação da prática de atividade física com atividades sociais pôde ajudar a que a população fosse fisicamente ativa e seria a chave para a adesão a programas de exercício (Rosa et al., 2015).

Considerações acerca da abordagem metodológica

Com base nos estudos que têm vindo a ser efetuados, e que aqui se apresentam, insurgem algumas pistas, mas também algumas incertezas e indefinições acerca das variáveis motivacionais que ajudariam a prever a manutenção do peso perdido. A incapacidade de definir estratégias eficazes de manutenção do peso perdido parece estar associada à multiplicidade de fatores que impedem ou facilitam a manutenção do peso perdido, e à dificuldade em identificar quais desses fatores são mais relevantes e determinantes para o sucesso individual.

Assim, parecem existir diversos fatores associados à eficácia na manutenção do peso perdido, que abrangem desde os comportamentos dos próprios indivíduos, aos seus estilos de vida, à prática de atividade física, à percepção que têm do seu peso, à sua motivação e à autorregulação emocional. É essa pluralidade de dimensões comportamentais, físicas, sociais e ambientais, e respetivas combinações possíveis, que faz com que as abordagens preditivas tradicionais, utilizando métodos de regressão que assumem linearidade entre variáveis, falhem. A variedade e complexidade das variáveis que podem explicar o sucesso da manutenção do peso perdido, não pode ser estudada apenas com recurso aos métodos estatísticos tradicionais de modelação. Em conclusão, é necessário perceber como novas técnicas de exploração de dados mais avançadas e com recurso a mais informação, designadamente *Machine Learning*, podem ajudar na compreensão deste fenómeno.

Introdução aos modelos de Machine Learning

Primeiramente, pode definir-se de forma muito genérica os sistemas de *Machine Learning* como “programas que automaticamente aprendem a partir de dados, sendo por isso uma alternativa muito atraente quando comparados com os programas resultantes de construção manual” (Domingos, 2012).

Segundo Géron (2017), os métodos de *Machine Learning* utilizam algoritmos para analisar grandes fontes de dados, e efetuar previsões de forma muito mais rápida que os métodos tradicionais, o que ajuda a encontrar soluções de forma muito mais eficaz. Importa especificar que *Big Data* não só implica um grande conjunto de informação como também inclui noções de complexidade e profundidade dos dados, e tudo isto, em tempo real, sendo também muito relevante a importância do processamento da informação. Métodos de *Machine Learning* têm como base métodos estatísticos, no entanto as premissas e os objetivos são diferentes, designadamente:

- Nos métodos estatísticos tradicionais trabalha-se com amostras considerando uma determinada distribuição teórica, enquanto em *Machine Learning* trabalha-se com *Big Data* proveniente de diversas fontes e com diferentes formatos, e não existe uma distribuição definida para os dados;

- A nível de resultados, em *Machine Learning* a probabilidade é utilizada para perceber qual a melhor decisão enquanto nas análises estatísticas tradicionais, o output tem em conta a variabilidade e incerteza dos parâmetros do modelo;
- Finalmente em *Machine Learning* os investigadores aprendem com os dados enquanto nos métodos estatísticos tradicionais assume-se uma determinada teoria e tenta-se prová-la.

Assim, é essencial avaliar como métodos inovadores de análise de dados podem ajudar a definir o perfil motivacional dos indivíduos que conseguem manter o peso perdido, através da prática sustentada de atividade física.

Estudos efetuados com modelos de Machine Learning

Explorando a utilização de técnicas estatísticas avançadas para prever os fenómenos ligados à perda de peso e consequente manutenção do peso perdido, avaliou-se o estudo de Batterham, Tapsell, Charlton, O'Shea e Thorne (2017). Neste estudo foi investigada a relação entre variáveis demográficas e de perda de peso para prever a eficácia da perda de peso. Foram utilizados métodos de *Data Mining* e de regressão logística para prever a eficácia da perda de peso (definida como superior ou igual a 5%) ao fim de 12 meses de intervenção alimentar consoante índice de massa corporal, género e idade. Os resultados do estudo identificaram, através da utilização de árvores de decisão, que o melhor preditor da perda de peso foi a percentagem de peso perdido ao fim de 1 mês.

Outro estudo avaliado que utilizou técnicas de *Machine Learning*, teve como objetivo prever a obesidade infantil, após os dois anos de idade, com base em informação histórica recolhida antes do 2º aniversário, de um sistema de apoio à decisão apelidado de CHICA (*Child Health Improvement through Computer Automation*). Foram aplicados 6 diferentes modelos, e avaliada a respetiva precisão, considerou-se o ID3 como o modelo mais bem ajustado à informação. Este modelo é um método de classificação baseado em aprendizagem através de árvores de decisão. A nível de resultados, o modelo ID3 mostrou ser aquele com maior precisão e performance, e retornou preditores para a obesidade infantil previamente identificados pela literatura existente. O preditor de obesidade infantil mais forte identificado pelo modelo ID3 foi excesso de peso antes dos 24 meses de idade. Na comunidade pediátrica não existe evidência clara que ter obesidade antes dos 2 anos de idade esteja correlacionado com o estado tardio de obesidade. No entanto, através da aplicação do modelo ID3 e da análise da árvore de decisão construída verificou-se por um lado que uma criança tem 2 vezes mais probabilidade de ter obesidade após os 2 anos de idade se tiver tido obesidade nos 24 meses anteriores, e que por outro lado, uma criança ter excesso de peso entre os 12 e os 24 meses de idade é fator chave para que tenha obesidade após o 2º ano de idade, sendo este risco acrescido se a criança não tiver excesso de peso antes dos 12 meses de idade. Este último resultado indica que o aumento de peso de forma muito acelerada, ou seja, entre os 12 e os 24 meses de idade, parece constituir-se como um risco. Em conclusão, este estudo demonstra que a informação recolhida por sistemas clínicos de apoio à decisão, que usualmente contém informação não estruturada, errônea e incompleta, pode ser utilizada para a construção de

modelos de *Machine Learning* precisos e válidos, para prever a obesidade infantil. Uma recomendação deste estudo foi a incorporação de um *checkup* pediátrico aos 2 anos de idade, para verificar se a criança estaria em risco de vir a desenvolver obesidade. Estamos perante um cenário em que estudos científicos providenciam claramente, aos médicos pediatras e outros profissionais de saúde, uma estratégia preventiva para intervir de modo a que seja reduzida a chance da criança vir a ter obesidade no futuro e assim proteger, possivelmente, as crianças de serem englobadas nesta epidemia (Dugan et al., 2015).

Objetivo e GAP

Retomando o âmbito deste trabalho, os seus focos principais foram a importância da atividade física na saúde e bem-estar das pessoas e nos mecanismos de manutenção do peso perdido. Então, sabendo que a prática regular de atividade física promove um peso saudável, julga-se que deverá ser incorporada no processo de manutenção do peso perdido, de forma voluntária e regular, e de modo a que se torne num hábito prazeroso.

Assim, este estudo teve como objetivo identificar o perfil motivacional da prática de atividade física, no processo de manutenção do peso perdido em adultos, ou seja, encontrar que variáveis motivacionais estariam associadas ao perfil de um indivíduo adulto que pratica de forma ativa, regular e intencional atividade física. O perfil motivacional é, portanto, aqui definido através de um conjunto de variáveis motivacionais que permitem identificar a prática ativa de atividade física e que agem como estimuladoras ou mesmo causadoras da prática. Se se verificar a prática ativa de atividade física, e se for possível identificar como é que esta prática ocorreu, então será mais fácil desenhar as intervenções que promovam a manutenção do peso. Na mesma linha dos dois estudos de predição abordados nos parágrafos anteriores, a investigação deste trabalho pretendeu adotar modelos de *Machine Learning*, mas agora com o propósito de aplicação numa área que não foi ainda explorada com esse tipo de técnicas, que é a identificação de diferentes padrões de atividade física e respetiva associação com variáveis motivacionais, no processo de manutenção do peso perdido.

Este trabalho teve por base a aplicação de modelos de *Machine Learning* à informação de um projeto europeu em curso, o projeto *NoHoW*. Este projeto inclui aproximadamente 1600 participantes de 3 países europeus: o Reino Unido, Portugal e a Dinamarca, que perderam peso e que estão em fase de manutenção do peso perdido. A grande diferença, face a indivíduos que tentam manter o peso perdido de forma autónoma e desacompanhada, é que, neste projeto, durante a fase de manutenção do peso perdido, os participantes do projeto *NoHoW* tiveram disponível uma *FitBit* – um instrumento (*wearable*) que mede a atividade física entre outras variáveis - para recolha de dados pessoais em tempo real, e esta informação seria utilizada para melhorar a compreensão do fenómeno da manutenção do peso perdido e eventualmente perceber se a informação recolhida poderia dar alguma indicação da ocorrência da prática de atividade física de forma ativa. Este fator revela-se extremamente importante no sucesso da manutenção do peso perdido, objetivo final dos participantes do projeto *NoHoW*.

As questões que se colocaram no âmbito do projeto *NoHoW* foram: Se já se consegue elaborar estratégias de perda de peso eficazes com resultados efetivos, por que razão a manutenção do peso perdido é tão difícil? Que fenómenos poderão explicar ou antever que alguém que perdeu peso e que está a tentar mantê-lo, possa não voltar a recuperá-lo através da incorporação regular da prática de atividade física? Que técnicas motivacionais poderão contribuir mais fortemente e eficazmente para a manutenção do peso perdido e porquê? Como identificar a falha na manutenção de peso perdido e como definir estratégias de intervenção eficazes para ajudar as pessoas a manter o peso a longo prazo? Como identificar padrões de prática de atividade física? Que variáveis motivacionais caracterizam indivíduos que praticam de forma ativa atividade física? (site *NoHoW*)

Propõe-se então, com este trabalho, complementar os estudos no âmbito do projeto *NoHoW*, não só explorando os mecanismos que possam explicar o fenómeno da manutenção do peso perdido, mas também e neste caso, tentando identificar níveis de prática de atividade física, já que esta é uma componente fundamental para que a manutenção do peso seja bem-sucedida. Para tal, foram aplicados algoritmos de *Machine Learning* à informação recolhida pelas *FitBit* distribuídas no âmbito do projeto *NoHoW* correlacionando-a depois com informação discreta obtida em questionários comportamentais e motivacionais. A abordagem proposta a este problema implicou a análise de um conjunto alargado e abrangente de informação individual, de adultos que perderam peso, de modo a identificar o respetivo padrão de prática ativa de atividade física no processo de manutenção do peso perdido. Não foram encontradas na literatura, este tipo de análises estatísticas mais avançadas e com recurso a *Big Data*, para explorar, compreender e explicar o fenómeno de manutenção do peso perdido através de padrões de prática ativa de atividade física, correlacionada com informação motivacional.

Idealmente, os resultados deste estudo permitiriam identificar os mecanismos motivacionais responsáveis para evitar a redução dos níveis de atividade física e do aumento de probabilidade de recuperação de peso perdido, ajudando na sua manutenção a longo prazo.

III.2 Método

Este trabalho utilizou informação do projeto *NoHoW*, e foi nesses dados que foram aplicados métodos de *Machine Learning*, para a identificação do perfil motivacional na prática de atividade física, no processo de manutenção do peso perdido. Grande parte da descrição do método foi baseada na informação disponibilizada e definida no âmbito do projeto *NoHoW*.

Este projeto teve como um dos principais objetivos avaliar se uma intervenção de modificação comportamental digital era eficaz na manutenção do peso perdido através da melhoria da autorregulação e da motivação, da regulação emocional, da combinação destes vários fatores e do controlo. De interesse para o presente estudo, parte da intervenção incluiu a monitorização da atividade física, através do

equipamento *FitBit*, um monitor de atividade física. Foi colocada como hipótese que a intervenção combinada de vários fatores seria mais eficaz do que intervenções individualizadas.

Outros objetivos passaram por determinar como é que estas intervenções poderiam afetar marcadores de saúde e a composição corporal, assim como ter impacto na atividade física, no sono, no aporte calórico, na depressão, na ansiedade, no stress, na qualidade de vida e no bem-estar e adicionalmente examinar os mediadores da manutenção do peso perdido, entre outros. De referir que o projeto *NoHoW* teve a duração de 18 meses, entre março 2017 e setembro de 2019 (Scott et al., 2019), mas no caso deste estudo a informação utilizada foi apenas dos primeiros 12 meses.

III.2.a) Desenho do Estudo

Embora o *NoHoW* seja um estudo randomizado controlado, na presente investigação fez-se uma análise em conjunto com todos os braços da intervenção, pelo que se considera que é um estudo de *Cohort* prospetivo. Foi seguido um conjunto de indivíduos ao longo do tempo, para testar a hipótese de que esse mesmo conjunto de indivíduos acaba por se distinguir em 2 grupos diferentes. Esses 2 grupos diferem no que diz respeito ao nível de prática de atividade física, ou seja, um grupo com indivíduos que praticam atividade física de forma ativa versus indivíduos que não o fazem. Para ser possível essa distinção, e durante o período de tempo de 12 meses foi utilizada a informação da *FitBit* utilizada pelos participantes e dados motivacionais obtidos de questionários efetuados no início do projeto (*Baseline*), após 6 e 12 meses.

III.2.b) Amostra

O estudo contou com a participação de 1627 indivíduos, de 3 países europeus (Portugal, Reino Unido, e Dinamarca), com idade superior a 18 anos, com índice de massa corporal (IMC) acima dos 25 kg/m² nos 12 meses anteriores ao projeto *NoHoW*, com pelo menos uma tentativa de perda de peso durante a vida e com uma tentativa de perda de peso (completada ou ainda a decorrer) nos 12 meses anteriores. Foi desenvolvido um questionário online, de modo a recrutar e avaliar a representatividade da amostra (por idade, género e localização) para cada um dos 3 países. O âmbito deste estudo foi a utilização total da informação, analisada dentro de cada um dos países participantes, mas não passou por efetuar comparações entre os participantes dos diferentes países.

Os potenciais participantes, ou seja, que demonstraram interesse em participar no estudo foram contactados por email, com a indicação de uma hiperligação de modo a completarem um questionário. Os participantes que após esse processo se tornaram elegíveis (seguindo uma lista de critérios de inclusão) foram direcionados a completar um questionário principal, recebendo pontos pela sua finalização. O recrutamento dos participantes ficou finalizado, tendo sido obtido para este estudo o seguinte número de participantes no *Baseline*: 536 participantes para Portugal, 555 Dinamarca e 536 para o Reino Unido. (Scott et al., 2019).

Em primeiro lugar foi efetuada a caracterização genérica da amostra. A amostra utilizada para este estudo compreendia inicialmente 1627 participantes, incluídos no projeto *NoHow*. A amostra globalmente pôde ser caracterizada da seguinte forma:

Tabela 1 – Distribuição percentual dos participantes por característica

Género	Feminino – 68,7%		Masculino – 31,3%
Grupo Etário (anos)	[0,30] – 12,6%	[30,50] – 54,7%	[50,90] – 32,7%
Grupo de perda de peso (kg)	(0,10] – 43,6%	(10,20] – 46,6%	(20,50] – 9,8%
Grupo de IMC* (kg/m ²)	(18,30] – 60,5%	(30,40] – 35,0%	(40,65] – 4,5%
Centro**	DK – 34,1%	PT – 32,9%	UK – 32,9%

*IMC – Índice de Massa Corporal

**Centro - DK- Dinamarca; PT- Portugal; UK- Reino Unido

Verifica-se que a amostra era constituída por mais mulheres que homens, sensivelmente na proporção de 1/3 de homens e 2/3 de mulheres. Relativamente ao grupo etário predominante, os indivíduos com idades compreendidas entre os 30 e os 49 anos eram mais de metade da amostra, indivíduos com mais de 50 anos eram 32,7% da amostra e os restantes 12,6% eram indivíduos com menos de 30 anos. No que respeita as variáveis ligadas ao peso, como peso perdido e o IMC, verificou-se que aproximadamente 90% dos indivíduos perderam entre 0 e 20kg e os restantes perderam mais de 20 kg, com um limite de 50 kg. O IMC variou entre 18 e 65 kg/m², com a seguinte distribuição: 60,5% tinham IMC superior a 18 e até 30 kg/m²; 35% tinham IMC superior a 30 kg/m² e até 40 kg/m² e os restantes 4,5% tinham IMC superior a 40 kg/m² e até 65 kg/m². Finalmente, os participantes estavam repartidos de forma quase idêntica em 3 centros diferentes: Dinamarca, Portugal e Reino Unido (aproximadamente 1/3 para cada um deles).

Foi possível caracterizar também a amostra segundo algumas características demográficas, que se apresentam na tabela infra:

Tabela 2 – Valores médios da amostra total e por género

	Total		Feminino		Masculino	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
Idade (anos)	44,0	11,9	44,7	12,2	42,6	10,9
IMC (kg/m ²)	29,7	5,3	29,7	5,6	29,5	4,8
Perda de peso (kg)	12,1	5,6	12,5	5,8	11,3	5,1

No que respeita a análise dos valores médios das variáveis da amostra, verificou-se que a média de idade se situava nos 44 anos, o IMC apresentou um valor médio de 29,7 kg/m² e o valor médio da perda de peso situava-se nos 12,1 kg.

III.2.c) Medidas/Instrumentos

Os instrumentos utilizados neste estudo dividiram-se em três categorias: o conjunto de informação individual recolhida através das *FitBit* entregues aos participantes, questionários e outras variáveis de caracterização dos participantes, recolhidas no âmbito do projeto *NoHoW*.

As variáveis da *FitBit* incluíram um conjunto de métricas recolhidas por minuto. Foi pedido aos participantes do projeto para utilizarem de forma permanente uma *Fitbit Charge 2* de modo a monitorizar a atividade física, indicadores associados e o sono. Segundo a informação obtida do protocolo do projeto *NoHoW*, estes dados foram registados de forma anónima, num ficheiro de texto e exportados para um ficheiro .csv posteriormente comprimido. A recolha dessa informação para utilização neste projeto foi feita através de um serviço de armazenamento de ficheiros (*file hosting service*). O total da informação recolhida para todos os participantes totalizava 4,86 GB de informação, sendo que para cada participante foram adquiridos aproximadamente mais de 600 mil registos (Scott et al., 2019).

Foram selecionados dois questionários neste estudo (da totalidade existente), considerando aqueles que eram mais pertinentes segundo os objetivos deste trabalho e por já terem sido abordados no âmbito deste mestrado, designadamente o questionário de satisfação das necessidades psicológicas básicas da teoria da autodeterminação: autonomia, competência e relacionamento positivo, BPN - *Basic Psychological Needs* (Ng, Lonsdale & Hodge, 2011) e (Ng, Ntoumanis, Thøgersen-Ntoumani, Stott, & Hindle, 2013). Foi feita uma adaptação a este questionário para o contexto da gestão do peso resultando numa abordagem que inclui a definição, para cada indivíduo, do seu nível no que respeita relações de pertença, necessidades psicológicas básicas, competência e autonomia, assim como um índice global. Este questionário tem originalmente 21 itens e a sua escala varia entre 1 e 7, em que 7 corresponde a “concordo totalmente” e 1 “discordo totalmente” (Moutão, Serra, Alves, Leitão, & Vlachopoulos, 2012). No âmbito do projeto *NoHoW* foi utilizada uma versão com 16 itens (<http://nohow.fmh.ulisboa.pt/>). Para aferir, por exemplo, o nível de autonomia eram apresentadas as seguintes afirmações: “É de livre vontade que faço esforços para manter o meu peso”; “Sinto que fazer esforços para manter o meu peso é uma decisão minha”; “Sinto que estou a tentar atingir os meus próprios objetivos”; entre outras. Estas afirmações foram contabilizadas de forma direta ou com escala inversa no caso de serem positivas ou negativas. Outros índices incluíam, por exemplo, relações de pertença medido através de “Sinto-me próximo/a de outras pessoas” e competência medido através de “Sinto-me compreendido/a” (<http://nohow.fmh.ulisboa.pt/>).

Outro questionário utilizado foi o que mede formas de regulação (externa, introjetada, identificada e integrada) no comportamento de exercício (BREQ-3) (Markland, & Tobin, 2004). Este questionário tem um conjunto de 24 itens e 6 sub-escalas, cada uma delas com 4 itens, e está desenhado em particular para aferir as regulações no âmbito do exercício (Markland, & Tobin, 2004). A escala varia entre 0 e 7, em que 0 significa “Não é verdade para mim”, 4 significa “Algumas vezes é verdade para mim” e 7 significa “Muitas vezes é verdade para mim”. No âmbito do projeto *NoHoW* utilizou-se uma versão com 18 itens (<http://nohow.fmh.ulisboa.pt/>). Para definir a amotivação, por exemplo, foram utilizados, por exemplo, os seguintes itens do questionário: “É importante para mim fazer exercício regularmente”; “Dou valor aos benefícios/vantagens do exercício”; “Penso que vale a pena o esforço de me exercitar”. Relativamente aos restantes índices: o índice da regulação externa foi medido através de “Não percebo porque é que tenho de fazer exercício”, por exemplo; o índice da regulação introjetada foi medido através de “Eu faço exercício porque é divertido”, por exemplo; o índice da regulação identificada foi medido através de “Sinto-me fracassado/a quando não faço exercício durante algum tempo”, por exemplo; o índice da regulação integrada foi medido através de “Eu faço exercício porque faz parte da minha identidade”, por exemplo (<http://nohow.fmh.ulisboa.pt/>). De referir que um dos indicadores obtido da aplicação do BREQ3 foi o RAI (índice de autonomia relativo). O RAI é uma medida obtida através das sub-escalas do BREQ3 que permite medir a autodeterminação. Este índice é obtido através de uma média ponderada de um conjunto de sub-escalas. As ponderações e respetivas sub-escalas são as seguintes: amotivação (-3), regulação externa (-2), regulação introjetada (-1), regulação identificada (+1), regulação integrada (+2), regulação intrínseca (+3). Quanto maior for o valor de RAI, maior é a autonomia, valores mais baixos ou negativos indicam maior regulação controlada (Cid et al., 2018).

Finalmente foram recolhidas outras variáveis, de modo a caracterizar os participantes, que incluíram género (feminino ou masculino), idade (em anos), o peso dos participantes utilizando roupa leve (em kg), utilizando o instrumento Seca 704s (SECA, Alemanha) obtido no *Baseline* e 12 meses após *Baseline*. Adicionalmente, foi solicitado aos participantes que se pesassem 2 vezes por semana de manhã em jejum utilizando as escalas da *FitBit* durante os 18 meses (esta informação não foi utilizada no âmbito deste trabalho). Foi também medida a altura dos participantes utilizando o instrumento Seca 704s (SECA, Alemanha) e consequentemente calculado o respetivo IMC (em kg/m²). Finalmente foi identificado o país de origem do participante, designado de Centro (Dinamarca; Portugal; Reino Unido) (Scott et al., 2019).

Resumindo, as variáveis consideradas neste estudo podem ser sumarizadas da seguinte forma:

- *FitBit* (dados por minuto): os dados recolhidos durante 12 meses da *FitBit*, designadamente Número de passos, Calorias despendidas, Distância percorrida (km), Nível de atividade física (foram definidos 4 níveis desde sedentário a muito ativo) e Frequência cardíaca (batimentos por minuto - bpm);

- Questionários: os indicadores obtidos de 2 questionários motivacionais aplicados no *Baseline*, aos 6 e aos 12 meses: Formas de regulação (externa, introjetada, identificada e integrada) no comportamento de exercício (BREQ-3) e Satisfação das necessidades psicológicas básicas da teoria da autodeterminação: autonomia, competência e relacionamento positivo (BPN - *Basic psychological needs*), no contexto da gestão do peso;
- Outras variáveis: Género, Idade (em anos), Peso (em kg), Índice de Massa Corporal (em kg/m²) e Centro (país de origem: Dinamarca; Portugal; Reino Unido).

De referir que no âmbito do projeto *NoHoW* foram utilizados outros instrumentos e outras medidas que não foram considerados no âmbito deste trabalho.

III.2.d) Procedimentos

O estudo do projeto *NoHoW* iniciou-se em novembro de 2016, mas a intervenção *NoHoW* iniciou-se oficialmente em março de 2017 e teve duração total de 5 anos, tendo sido financiado por fundos comunitários. Em cada um dos países participantes neste projeto, houve a aprovação ética do estudo nos respetivos comités (Universidades de Leeds (17-0082; 27 fevereiro 2017), de Lisboa (17/2016; 20 fevereiro 2017) e da Capital da região da Dinamarca (H-16030495, 8 março 2017). Foi obtido consentimento informado de todos os participantes, incluindo dos potenciais participantes (Scott et al., 2019).

Os dados utilizados para este trabalho foram obtidos através de questionários online, assim como através das *FitBit*, que foram entregues aos participantes no âmbito do projeto *NoHoW*. Todos os dados recolhidos cumpriram critérios de confidencialidade da identidade dos participantes através da anonimização da respetiva identidade, e foram armazenados de forma segura numa base de dados. Este trabalho utilizou a informação já anonimizada recolhida pelas *FitBit*, assim como a informação motivacional obtida pelos questionários.

III.2.e) Análise dos dados

A análise da informação efetuou-se recorrendo a técnicas de *Machine Learning*. Para tal, foram efetuados um conjunto de procedimentos seguindo as melhores práticas para este tipo de análises (Géron, 2017).

O primeiro passo foi a definição explícita do problema e de todos os pressupostos assumidos à partida. No âmbito deste trabalho, o problema a resolver através das técnicas de *Machine Learning* foi a segmentação dos participantes do projeto *NoHow* em dois (ou mais) grupos diferentes consoante o nível de atividade física.

Os pressupostos, que se assumiram, foram que através das variáveis recolhidas pela *FitBit*, seria possível identificar diferentes níveis de atividade física. Assumiu-se também que poderiam existir outras dimensões que não foram medidas ou recolhidas e que poderiam também elas contribuir para

essa caracterização, pelo que os resultados obtidos através desta análise, poderão ter que ser sujeitos a validações posteriores para se tornarem mais robustos e fidedignos.

De referir que para a análise da informação foram utilizados os seguintes softwares: *IBM SPSS Statistics* versão 26 e o *Visual Studio Code Version: 1.37.1*, com recurso à linguagem de programação *Python* (código fonte em anexo).

(1) Validação das variáveis

As variáveis a utilizar nesta análise foram divididas em 2 categorias.

Utilizaram-se por um lado variáveis individuais recolhidas em tempo real, por minuto, através das *FitBit* utilizadas pelos participantes do projeto *NoHow*. Estas variáveis foram úteis no sentido de caracterizar comportamentos dos participantes no que respeita a prática de atividade física no dia-a-dia. Foram recolhidos dados pela *FitBit* durante os 12 primeiros meses do projeto numa periodicidade 24/7, tendo sido registados os dados por minuto numa base de dados, e em seguida utilizados para aplicação dos algoritmos de *Machine Learning*. Ainda assim, não era clara a perceção de quais os valores dessas variáveis que indicariam que um determinado indivíduo pratica de forma ativa exercício, essa informação só seria obtida após a aplicação dos algoritmos de *Machine Learning*.

Paralelamente, existiam as variáveis motivacionais que foram recolhidas no início do projeto *NoHow* (*Baseline*) e passados 6 e 12 meses e, foi utilizado um subconjunto dessas variáveis no âmbito deste trabalho, de modo a identificar as variáveis motivacionais dos indivíduos que praticam ativamente atividade física ao longo dos 12 meses de análise. Importa ainda referir que no âmbito do projeto *NoHow*, e durante os primeiros 6 meses de intervenção, os participantes recebiam semanalmente e-mails a encorajar a adesão à atividade física (Scott et al., 2019), pelo que houve uma preocupação inicial em apoiar e motivar os participantes neste projeto.

Caracterizando a informação recolhida pelas *FitBit*, existiam um total de 1627 ficheiros nos quais foram registadas as métricas das *FitBit*. O conjunto de métricas de cada um dos participantes encontrava-se guardado num ficheiro distinto, existiam, portanto, um total de 1627 participantes e 1627 ficheiros. Em cada um dos ficheiros existia informação por minuto das métricas recolhidas pela *FitBit* durante os 12 meses iniciais do projeto *NoHow*. O trabalho preparatório da informação implicou um conjunto de passos de exploração dos dados, designadamente o estudo de cada uma das variáveis e das respetivas características. Foi recolhida a sua designação, o tipo da variável, a percentagem de “*missing values*” (valores em falta), a existência de “*outliers*” (valores anormais) e a identificação dos mesmos, a adequabilidade da variável para o estudo em causa e a distribuição da variável. Para essa exploração, foi necessário efetuar a visualização dos dados e um estudo da correlação entre as variáveis (Géron, 2017).

No que respeita o estudo das variáveis, das 14 disponíveis foram selecionadas apenas 8 para aplicação do modelo de *Machine Learning*, já que as restantes foram consideradas redundantes ou

irrelevantes para o estudo em causa. O resultado da seleção efetuada encontra-se espelhada no quadro seguinte:

Tabela 3 – Seleção das variáveis

Nome da variável	Tipo	Seleção da variável
Identificação do participante (<i>Participant_id</i>)	Float	Selecionada para o modelo
Data / Hora / Minuto (<i>Date_time</i>)	Datetime	Selecionada para o modelo
Passos (<i>Steps</i>)	Float	Selecionada para o modelo
Distância percorrida (<i>Distance</i>)	Float	Selecionada para o modelo
Nível (<i>Level</i>) (0-sedentary;1-lightly_active;2-fairly_active;3-very_active)	Float	Selecionada para o modelo
Calorias da atividade (<i>Calories_activity</i>)	Float	Não selecionada para o modelo*
Total de calorias consumidas (<i>Calories_overall</i>)	Float	Selecionada para o estudo
Sedentário (<i>Sedentary</i>)	Float / Binary	Não selecionada para o modelo**
Pouco ativo (<i>Lightly_active</i>)	Float / Binary	Não selecionada para o modelo**
Ativo (<i>Fairly_active</i>)	Float / Binary	Não selecionada para o modelo**
Muito ativo (<i>Very_active</i>)	Float / Binary	Não selecionada para o modelo**
Frequência cardíaca (<i>HR</i>)	Float	Selecionada para o modelo
Dia (<i>Fbday</i>)	Datetime	Selecionada para o modelo
Dia da semana (<i>Wkday</i>) (1-Mon; 2-Tue; 3-Wed; 4-Thu; 5-Fri; 6-Sat; 7-Sun)	Float	Não selecionada para o modelo***

* variável a 0 para todos os participantes

** variável redundante (pois tem informação idêntica à variável “Level”)

*** variável sem relevância para o estudo

(2) Caracterização dos participantes com base nas métricas da FitBit

Após a definição das variáveis a utilizar, o objetivo traçado foi caracterizar os participantes deste estudo através das mesmas. Estas métricas foram preparadas, de modo a que fosse facilitada a aplicação dos algoritmos de *Machine Learning* e que fossem perceptíveis padrões. Os procedimentos efetuados são descritos em seguida.

i) Limpeza dos dados

Primeiramente, houve necessidade de fazer uma limpeza dos dados. Um dos principais problemas em bases de dados é a sua integridade, a integridade da informação pode ficar comprometida quando existem valores em falta que podem comprometer a sua análise global ao enviesar os resultados

(Géron, 2017). Naturalmente nem todos os dados são recolhidos, nem toda a informação fica guardada, e muitas vezes existem omissões e falhas que resultam informação incompleta. Existem, por isso, um conjunto de procedimentos usuais que podem ser implementados para lidar com esta situação: apagar os registos que incluem dados omissos ou preencher os dados omissos com um determinado valor (por exemplo, o valor médio dos mesmos). A opção a tomar deve ter em conta o número de registos que são identificados nesta situação, já que bases de dados com muitos dados omissos não podem, por razões óbvias, ser alvo de remoção de registos (Géron, 2017).

No caso do nosso estudo, foram identificados 5 ficheiros correspondentes a 5 participantes sem informação, correspondendo a aproximadamente 0,3% do total dos participantes, pelo que estes foram descartados.

ii) Engenharia das características relevantes

Em seguida, foram aplicadas um conjunto de técnicas de engenharia dos dados, de modo a efetuar e tornar as características relevantes dos participantes utilizáveis no modelo de *Machine Learning* (Géron, 2017).

O primeiro passo foi efetuar o cálculo de métricas agregadas para cada dia e para cada participante, da seguinte forma, resultando na criação de 9 variáveis:

- Soma dos passos efetuados
- Soma da distância percorrida
- Média do nível de atividade física
- Soma das calorias consumidas
- Média, desvio padrão, máximo, mínimo e moda da frequência cardíaca

O resultado obtido foi a redução da dimensionalidade da informação, inicialmente tínhamos dados por minuto e após esta transformação, passámos a ter dados por dia. Assim, obteve-se um ficheiro para cada participante, com o número de registos igual ao número de dias de utilização da *FitBit* (obtido da informação inicial por minuto). De referir que o número de registos difere consoante o participante tenha utilizado mais ou menos dias a *FitBit*. Para cada um dos participantes obtivemos os valores para as 9 variáveis definidas acima.

Após esta fase, foi efetuada uma nova limpeza dos dados, designadamente:

- Foram removidos os registos (dias) com valores a 0, que correspondiam a dias em que não foi utilizada a *FitBit*;
- Foram excluídos os ficheiros (participantes) com menos de 30 dias (inclusive) de utilização da *FitBit*;
- Foram apenas selecionados os participantes para os quais a taxa de utilização da *FitBit* (variáveis com valores diferentes de 0) fosse superior a 50% do total de informação (ou seja, o número de dias em que o participante teve a *FitBit* colocada e a recolher informação teve que

ser superior a 50% do total de dias, considerando a totalidade de dias de participação no programa).

Estas situações resultaram em critérios de exclusão de participantes, pois foi considerado que se tratava de situações de fraca ou muito fraca utilização da *FitBit*, e assim sendo, não seria possível caracterizar o participante com tão pouca informação. Todos os critérios definidos acima resultaram das várias interações do grupo de trabalho do *NoHow*, assim como o cumprimento das recomendações e boas práticas definidas para a aplicação de algoritmos de *Machine Learning* (Géron, 2017).

Nesta fase da análise, houve necessidade de voltar a reduzir a dimensionalidade dos dados pois a aplicação dos algoritmos de *Machine Learning* não é exequível com a informação tão desagregada (Géron, 2017). O passo seguinte foi então uma nova redução do número de registos, mas simultaneamente, com incremento no número de variáveis, ou seja, a criação de um ficheiro com todos os participantes, em que cada registo corresponde a um participante, e para o qual foram calculados o valor médio e o desvio padrão de:

- Soma dos passos efetuados
- Soma da distância percorrida
- Média do nível
- Soma das calorias consumidas
- Média, desvio padrão, máximo, mínimo, moda da frequência cardíaca

Obtivemos um total de 18 variáveis (2 novas variáveis para cada uma das 9 existentes) e foi desta forma, que se definiu a caracterização genérica do participante quanto à sua utilização da *FitBit* e consequentemente quanto ao seu padrão de prática de atividade física.

Após este trabalho, o total de registos (participantes) obtido para aplicação do modelo de *Machine Learning* foi de 1139.

(3) Clustering dos participantes

Após a caracterização dos participantes com base na informação recolhida pela *FitBit*, ou seja, para cada um dos 1139 participantes selecionados foi possível ter disponível um conjunto de métricas, o passo seguinte teve como objetivo tentar agrupar participantes semelhantes. O racional desta ideia foi tentar perceber se existiam padrões de comportamento de atividade física semelhantes, identificáveis por algoritmos de *Machine Learning*, de modo a que indivíduos que praticam atividade física de forma ativar podessem ficar destacados no grupo total de participantes. Para tal, optou-se pela utilização de um algoritmo de *clustering*, o *K-Means*, por ser um dos mais utilizados para este propósito, e que tem como principal função agregar componentes com características idênticas, o que neste caso, se traduz por participantes com o mesmo comportamento face à prática de atividade física. O *clustering* dos participantes foi efetuado em 3 passos que se enumeram de seguida, a saber, primeiramente aplicou-se a análise em componentes principais para redução da dimensionalidade, em seguida aplicou-se o

algoritmo *K-Means* para distinguir a amostra de participantes em grupos distintos e finalmente aplicou-se teste de hipóteses *t-Student* para validar que os dois grupos eram estatisticamente diferentes.

i) Análise em Componentes Principais

Neste trabalho e antes da aplicação do algoritmo *K-Means* foi aplicada a análise em componentes principais, esta técnica serve para permitir a visualização dos dados multivariados disponíveis no âmbito deste projeto, em subespaços de menor dimensão (Gray, 2017), e é uma das técnicas base da análise multivariada de dados, que permite a redução da sua dimensionalidade (Gortler, Spinner, Streeb, Weiskopf, & Deussen, 2019). É um dos métodos mais utilizados e o mais eficaz quando os dados estão correlacionados, que é o que acontece com a informação deste estudo, em particular. Na análise em componentes principais é esta propriedade dos dados, o facto de estarem correlacionados, que permite a utilização de uma transformação ortogonal para converter os mesmos em variáveis linearmente não correlacionadas, as componentes principais. O número de componentes principais é sempre menor ou igual ao número de variáveis existentes na base de dados. Com esta análise é possível reduzir a dimensionalidade dos dados através de projeções interpretáveis, e obter-se o conjunto de informação que melhor explica a variância dos dados. Adicionalmente, é possível com as componentes principais calculadas passar de um espaço multidimensional para um subespaço de menor dimensão, e assim interpretar mais facilmente a informação disponível (Jolliffe, 2002). Após a aplicação da análise em componentes principais, é às componentes principais obtidas que é aplicado o algoritmo *K-Means*, esta opção procedimental foi tomada por ser, de forma consensual, uma boa prática a aplicar neste tipo de análises.

ii) Algoritmo de aprendizagem não supervisionada – K-Means Clustering

Considerando que não foi possível definir na literatura nenhum critério perfeitamente claro e evidente de como caracterizar padrões atividade física (excetuando-se a já conhecida recomendação de efetuar 10.000 passos por dia), optou-se pela aplicação de algoritmos de aprendizagem não-supervisionada. Esta técnica pressupõe que é o próprio algoritmo de *Machine Learning* que descobre, com base nos dados fornecidos, qual a segmentação mais adequada para os dados. Não existia, portanto, à partida uma ideia pré-definida acerca de qual a melhor forma de agrupar a informação, ou seja, de classificar os participantes do estudo, consoante características idênticas ou semelhantes. A ideia passou por descobrir características idênticas em participantes que praticam atividade física ou então características que separam os indivíduos que praticam atividade física de forma ativa dos que a praticam de forma insuficiente.

O método a aplicar foi um método heurístico *K-Means Clustering*, que é configurado através da definição prévia do número *k* de *clusters* (grupos) que se pretende obter e representa o número de centróides do modelo. Teoricamente, o método *K-Means Clustering* é um método de otimização que

pretende minimizar uma função objetivo correspondente à soma quadrática dos erros (ou variâncias *intra-cluster*).

Função objetivo: Considerando que: $\|x_i^{(j)} - c_j\|^2$ é a distância Euclidiana entre uma observação $x_i^{(j)}$ e o centroide do cluster c_j , k corresponde ao número de *clusters* e n corresponde ao número de observações:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2$$

que representa a distância das n observações aos centróides dos respetivos clusters (MacKay, 2003).

Assim, o algoritmo coloca cada uma das observações (participante do estudo) num determinado *cluster*, com base na sua similaridade às observações já existentes nesse mesmo *cluster*, ou seja, uma observação é assignada ao *cluster* se as suas características forem semelhantes às características das observações já contidas nesse cluster (quando comparadas com as dos outros clusters). Iterativamente, o método vai ajustando o número de *clusters* k definido previamente, com base no conhecimento que vai adquirindo das observações da amostra que processa, sendo que o número final de *clusters* é aquele que o método considera mais adequado ao problema. Assim, participantes com características idênticas (neste estudo, métricas da *FitBit*) ficarão agrupados no mesmo *cluster*. Participantes agrupados em *clusters* diferentes terão características distintas (MacKay, 2003).

Os resultados desta análise revelam-se como pistas para o entendimento do perfil dos participantes que integram cada um dos grupos definidos, e ajudam a perceber quais as características dos indivíduos que praticam atividade física de forma ativa, já que idealmente estes indivíduos ficam agrupados num mesmo *cluster*. Possivelmente, conseguimos obter conjuntos de variáveis que caracterizem este fenómeno complexo e multifatorial que é o processo da prática ativa de atividade física. Com recurso ao conhecimento existente no âmbito das teorias de modificação comportamental, os resultados desta análise são afinados de modo a definir as variáveis motivacionais, e fazer corresponder estes resultados a um modelo sustentado numa base teórica comportamental válida e sólida (MacKay, 2003).

iii) Avaliação do output através da aplicação de Testes de Hipóteses

Após a obtenção dos resultados da aplicação do algoritmo *K-Means*, e de forma a validar teoricamente (e não apenas através da interpretação das estatísticas obtidas por grupo) que ambos os grupos eram distintos, decidiu-se aplicar um teste de hipótese *t-Student*.

Esta análise permitiu assegurar a validade da classificação obtida pela análise de clusters e reforçar o pressuposto de que há grupos de participantes diferentes, consoante diferentes níveis de atividade física.

(4) Cruzamento com as respostas aos inquéritos motivacionais

Nesta fase da análise, obtivemos dois conjuntos de dados, a informação que serviu para a aplicação do modelo de *Machine Learning*, designadamente *K-Means*, com 1139 registos definidos em grupos e a base de inquéritos motivacionais com 1627 registos (1636 registos iniciais, sendo que 3 participantes tinham registos em triplicado, pelo que foram removidos 9 registos).

O objetivo com esta última análise foi tentar definir o perfil motivacional dos indivíduos classificados em cada um dos grupos. Perceber, nesses indivíduos, que tipo de motivação possuíam, o que os motivava para a prática de atividade física, com o intuito de no futuro, podermos trabalhar, promover, fomentar essas mesmas motivações em indivíduos que á partida não as possuem. Esperou-se obter com este trabalho, pistas para ajudar as pessoas a praticar, mais e de forma regular e consistente, atividade física.

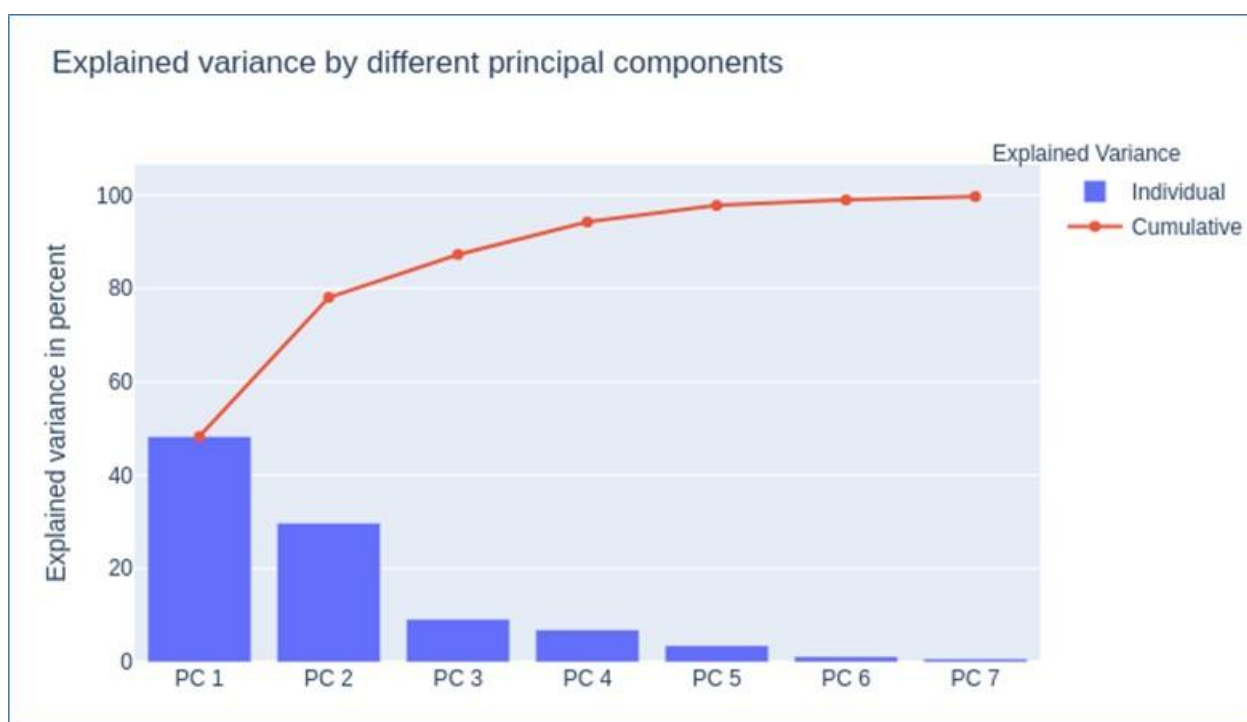
III.3 Resultados

III.3.a) Resultados da aplicação do algoritmo de *Machine Learning* – *K-Means*

Após a caracterização da amostra, avançou-se para a aplicação dos algoritmos de *Machine Learning*, de modo a classificar os indivíduos desta amostra em grupos distintos de prática de atividade física, para tal utilizou-se informação recolhida pela *FitBit*, por minuto, durante 12 meses.

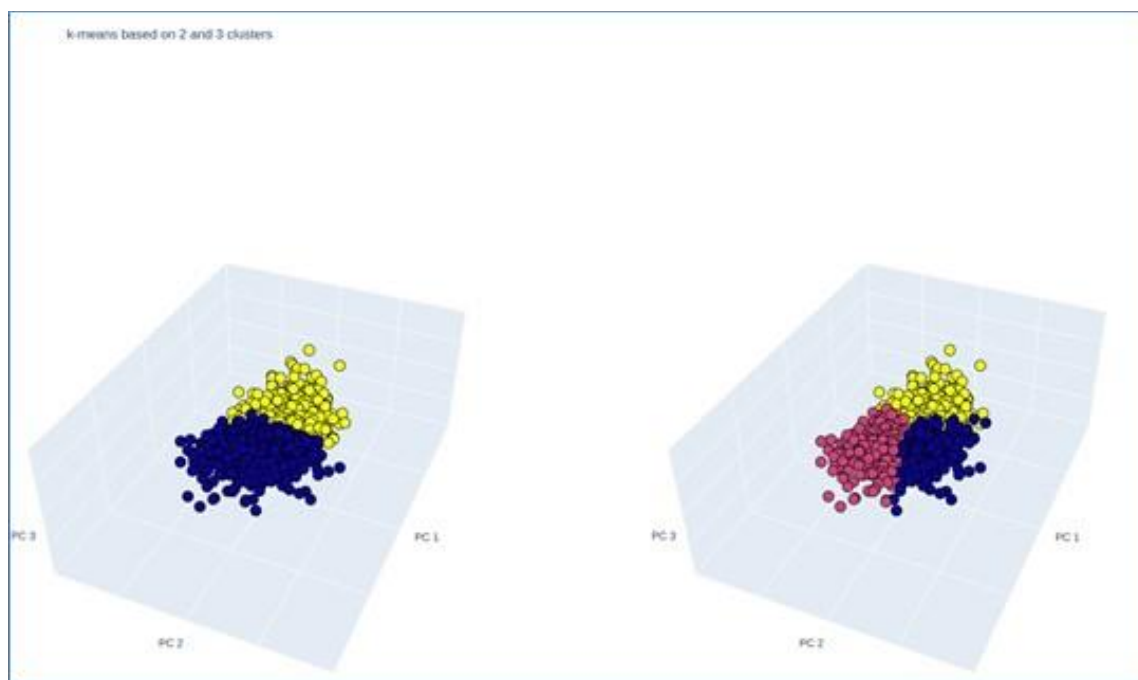
Foram efetuados os procedimentos identificados no capítulo anterior e selecionado o algoritmo de *Machine Learning* para *clustering* dos indivíduos, *K-Means*. Procedeu-se à aplicação da análise em componentes principais e apresentam-se os resultados através dos seguintes gráficos:

Gráfico 1 – Variância explicada pelas diferentes componentes principais



Tipicamente, aquando da utilização da análise em componentes principais, importa verificar a variância da informação que é explicada através das componentes principais obtidas. Através do gráfico 1 foi possível demonstrar que as duas primeiras componentes principais explicam aproximadamente 80% da variância das variáveis, e que o incremento de mais componentes principais se tornava quase redundante a partir da 4ª componente principal.

Gráfico 2 – Aplicação do *K-Means* às componentes principais (2 e 3 clusters)



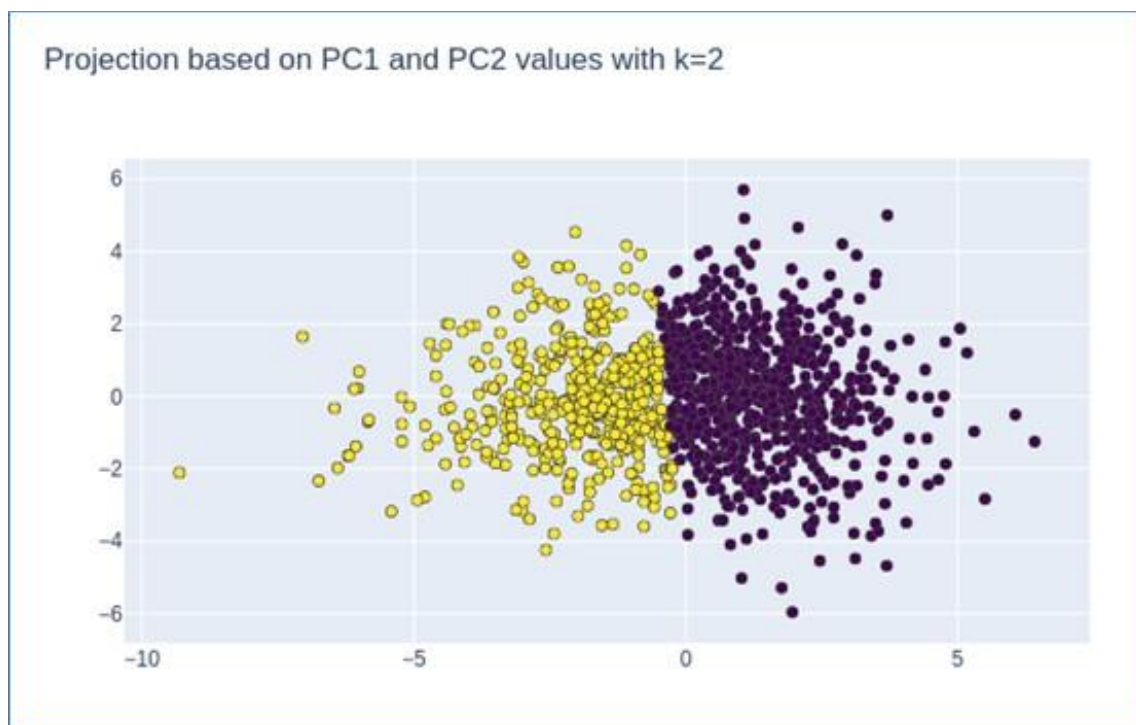
Optou-se por testar a aplicação do algoritmo *K-Means* às componentes principais, com 2 e 3 *clusters*, verificando-se na projeção 3D acima utilizando as 3 primeiras componentes principais, que existe uma clara distinção dos dados em 2 grupos diferentes (gráfico 2), e possivelmente a distinção em 3 clusters (gráfico 2).

Considerando que o âmbito deste trabalho passa por distinguir 2 grupos de indivíduos: os que têm um nível elevado de atividade física e os que têm um nível médio ou mais baixo de atividade física diária, e efetuando a análise visual dos 2 gráficos apresentados na figura acima, optámos por escolher os 2 grupos e prosseguir a análise com essa premissa. De notar que a escolha do número de *clusters* é sempre subjetiva, não existindo uma solução ótima.

De modo a validar a escolha do número de grupos, optou-se por efetuar 10 iterações do algoritmo *K-Means* com 2 e 3 *clusters* e verificou-se que à 10ª iteração, quando escolhidos 2 clusters, o resultado mantinha-se estável em 2 *clusters* distintos, no entanto quando selecionados 3, à 10ª iteração, 2 dos clusters passam a ter centroídes muito similares, o que indica que um dos clusters se torna redundante. Verificou-se então que a opção tomada de escolher 2 *clusters* foi a que pareceu ser a mais adequada (MacKay, 2003). Poderim ter sido aplicados outros métodos de otimização do número de *clusters*, como por exemplo *Elbow Method* ou *Average Silhouette Method*, no entanto não se considerou como essencial a implementação destes métodos no âmbito deste trabalho.

Eventualmente e numa outra análise, poderá ser alvo de averiguação como é que os participantes neste estudo podem ser diferenciados em 3 grupos.

Gráfico 3 – Projeção das 2 componentes principais para 2 clusters



Como resultado da identificação de 2 grupos distintos de indivíduos nesta amostra, foi construído o gráfico 3, no qual foi efetuada uma projeção 2D das duas primeiras componentes principais. A conclusão que se retirou deste gráfico foi que efetivamente foi possível a diferenciação em 2 grupos, com um corte vertical. Considerando a análise gráfica dos resultados da aplicação do *K-Means* foram selecionados 2 *clusters*. Para identificação dos 2 *clusters*, optou-se pela nomenclatura definida por Tudor-Locke (2011) e a nível do resultado do *clustering* identificaram-se 453 (27,8%) indivíduos no *cluster* 1, esse *cluster* foi apelidado de grupo de participantes com um nível elevado de atividade física – “Ativos” (*Active*), abreviado para “Ativ” e 686 (42,2%) indivíduos no *cluster* 2, esse *cluster* foi apelidado de grupo de participantes com um nível médio ou baixo de atividade física – “Insuficientemente Ativos” (*Somewhat Active*), abreviado para “I-Ativ”. De referir que os restantes 488 (30%) indivíduos já tinham sido excluídos pelos critérios identificados anteriormente (capítulo anterior).

De modo a avaliar a análise de *clustering* obtida através da aplicação do algoritmo *K-Means* foram calculados os valores médios das variáveis recolhidas pela *FitBit* por *cluster*, e adicionalmente juntaram-se os valores médios referentes ao peso (não utilizado na construção do modelo de *Machine Learning*) mas recolhidos no âmbito do projeto *NoHoW*, apenas para verificar se existiam diferenças do peso e da sua variação, nos 2 *clusters* definidos:

Tabela 4 – Valores médio das variáveis da *FitBit* por *cluster*

Valores médios	Ativ	I-Ativ	Total
Soma de passos diários	12.086,8	7.665,0	9.423,6
Média da frequência cardíaca (bpm)	71,2	74,1	73
Máximo da frequência cardíaca (bpm)	139,0	130,6	133,9
Mínimo da frequência cardíaca (bpm)	49,6	54,3	52,4
Moda da frequência cardíaca (bpm)	60,7	66,6	64,3
Desvio padrão da frequência cardíaca (bpm)	15,6	12,7	13,8
Soma da distância diária (km)	8,7	5,3	6,7
Média do nível de atividade física (0 a 4)	0,33	0,22	0,26
Soma do dispêndio calórico diário (kcal)	2.866,7	2.418,6	2.596,8
Peso no <i>Baseline</i> (kg)	83,47	85,61	84,80
Peso ao fim de 12 meses (kg)	82,37	85,92	84,47
Variação do peso (%)	-1,32	0,36	-0,39

Na análise dos valores médios amostrais dos indivíduos classificados em cada um dos 2 *clusters*, foi verificado que os indivíduos do *cluster* Ativ fizeram mais passos diariamente e percorreram uma maior distância, tinham frequências cardíacas mais baixas, mas atingiram um valor de pico superior, com um maior dispêndio calórico diário, e finalmente, uma variação de peso negativa (esta significa perda de peso ao fim dos 12 meses). De notar que os indivíduos classificados no *cluster* I-Ativ apresentam uma variação de peso positiva (que significa, aumento de peso ao fim dos 12 meses).

Em seguida foi utilizado o teste *t-Student* para avaliar se as diferenças verificadas na tabela 4, entre os valores médios das variáveis: soma de passos, média da frequência cardíaca, soma da distância, média do nível de atividade física e soma do dispêndio calórico nos 2 *clusters* eram estatisticamente significativas. De notar que não foram incluídas na análise inferencial as variáveis referentes ao peso, pois estas seriam analisadas por outros investigadores no âmbito do projeto *NoHoW*. Foi, então, avaliada a seguinte hipótese nula, H_0 : Não há diferença dos valores médios entre os dois grupos (Ativ e I-Ativ), considerando-se que os 2 *clusters* são 2 amostras independentes, 2 grupos totalmente distintos. A aplicação do teste *t-Student* foi precedida do teste de *Levene*, para testar a igualdade de variâncias, e verificou-se que estava comprovado pelo teste a não existência de igualdade das variâncias para os 2 grupos para todas as variáveis testadas. Adicionalmente, foi também efetuado o teste à magnitude do efeito.

Assim, da aplicação do teste *t-Student* verificaram-se os seguintes resultados:

Tabela 5 – Resultados obtidos da aplicação do teste *t-Student* e magnitude do efeito (*Cohen's d*)

Variável	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>Cohen's d</i>
Soma de passos	1139	9.423,61	2.960,65	34,20	<0,001	2,19
Ativ	453	12.086,82	2.344,11			
I-Ativ	686	7.664,96	1.774,40			
Média da FC	1139	72,95	7,06	-7,28	<0,001	0,43
Ativ	453	71,17	6,25			
I-Ativ	686	74,13	7,31			
Soma da distância percorrida	1139	6,66	2,26	33,75	<0,001	2,20
Ativ	453	8,70	1,86			
I-Ativ	686	5,32	1,27			
Média do nível de AF	1139	0,26	0,07	33,33	<0,001	2,15
Ativ	453	0,33	0,06			
I-Ativ	686	0,22	0,04			
Soma do dispêndio calórico	1139	2.596,79	537,14	14,46	<0,001	0,91
Ativ	453	2.866,66	551,03			
I-Ativ	686	2.418,58	446,10			

M = média; *SD* = desvio-padrão

Da análise da tabela 5, podemos apurar os resultados do teste efetuado à igualdade das médias, os valores de *p* (*p-value*) eram todos inferiores a 0,001, pelo que se rejeitou a hipótese nula, ou seja, foram verificadas diferenças nos valores médios dos 2 grupos, existiu efetivamente evidência para afirmar que os diversos indicadores físicos de prática de atividade física diferiam quando comparados os Ativ dos I-Ativ (resultados em anexo). Por exemplo, verificou-se que o número médio de passos era diferente nos 2 grupos, sendo superior para os Ativ; a mesma leitura pôde ser feita para as restantes variáveis testadas. Este resultado permitiu reforçar a premissa inicial de que os Ativ são diferentes dos I-Ativ, no que respeita a prática de atividade física. No que respeita a análise da magnitude do efeito, através da análise do valor *d* de *Cohen*, verificou-se que as diferenças entre os 2 *clusters* eram todas grandes exceto no caso da frequência cardíaca, que era pequena/média.

Em seguida, foi efetuada a associação das variáveis motivacionais a cada um dos 2 grupos obtidos, cumprindo-se assim um outro objetivo deste trabalho. Após a identificação dos 2 grupos, foi necessário encontrar a caracterização dos mesmos quanto à componente motivacional.

Assim, foram calculadas as variações referentes às diversas variáveis motivacionais desde o *Baseline*, após 6 e 12 meses, nos 2 grupos, de modo a caracterizar os indivíduos classificados em cada um deles. Foi feita uma comparação dos valores médios motivacionais, entre grupos, nos 3 momentos temporais medidos. O perfil motivacional dos participantes foi identificado utilizando dois questionários: BREQ 3 – Amotivação, Formas de regulação (externa, introjetada, identificada e integrada) e motivação intrínseca no comportamento de exercício e BPN - Satisfação das necessidades psicológicas básicas da teoria da autodeterminação: autonomia, competência e relacionamento positivo (*Basic psychological needs*) no âmbito da gestão do peso.

III.3.b) Comparação dos indicadores de motivação dos Ativ face aos I-Ativ

Foram calculadas as variações dos indicadores de motivação dos Ativ face aos I-Ativ nos 3 períodos em análise (*Baseline*, 6 e 12 meses) para verificar em que grupo e em que momento esses indicadores eram superiores/inferiores. Foram utilizados testes de hipóteses para validar se as diferenças verificadas eram estatisticamente significativas (identificado na tabela 6) através do teste *t-Student* e considerando amostras independentes (leia-se os 2 grupos).

Tabela 6 –Variação dos indicadores de motivação dos Ativ face aos I-Ativ, no *Baseline*, 6 e 12 meses

Variáveis	<i>Baseline</i>	6 meses	12 meses
RAI (BREQ3)	39,49* (1)	42,35*	39,65*
Motivação Controlada (BREQ3)	0,56	-0,94	0,53
Motivação Autónoma (BREQ3)	17,64*	19,25*	18,19*
Regulação Intrínseca (BREQ3)	16,80*	18,96*	18,00*
Regulação Integrada (BREQ3)	22,38*	24,87*	23,52*
Regulação Identificada (BREQ3)	14,53*	14,88*	14,01*
Regulação Introjetada (BREQ3)	7,12*	3,52	6,65
Regulação Externa (BREQ3)	-12,67*	-10,61*	-11,98*
Amotivação (BREQ3)	-14,62*	-14,78* (2)	-14,78*
Global (BPN)	2,36	4,51*	5,38*
Relações de Pertença (BPN)	2,64	3,89	5,11
Necessidades Psicológicas Básicas (BPN)	-0,01	1,11	0,87
Competência (BPN)	2,57	6,25*	7,36*
Autonomia (BPN)	2,01	3,49*	3,99*

Os valores identificados como **(1)** e **(2)** foram utilizados para explicar a leitura dos valores.

* Variação estatisticamente significativa (obtido do teste *t-Student*)

Nota: foram colocadas a *itálico* as variáveis com variações negativas (ou seja, valores superiores nos I-Ativ face aos Ativ). As restantes variáveis apresentam variações positivas (ou seja, valores superiores nos Ativ face aos I-Ativ).

O que se apresentou na tabela 6 tem a seguinte leitura: **(1)** O valor do RAI é 39,49% mais elevado nos Ativ do que nos I-Ativ, esta variação foi marcada com * por ser estatisticamente significativa, já **(2)** o valor da amotivação aos 6 meses é 14,78% mais baixo nos Ativ do que nos I-Ativ, esta variação foi marcada com * por ser estatisticamente significativa.

Assim, fazendo uma análise global comparativa entre ambos os grupos (apenas diferenças estatisticamente significativas), verificou-se a separação dos seguintes indicadores motivacionais por grupo e nos 3 períodos temporais em análise:

Tabela 7 – Indicadores motivacionais - valores mais elevados nos Ativ

Baseline	6 meses	12 meses
RAI (BREQ3)	RAI (BREQ3)	RAI (BREQ3)
Motivação Autónoma (BREQ3)	Motivação Autónoma (BREQ3)	Motivação Autónoma (BREQ3)
Regulação Intrínseca (BREQ3)	Regulação Intrínseca (BREQ3)	Regulação Intrínseca (BREQ3)
Regulação Integrada (BREQ3)	Regulação Integrada (BREQ3)	Regulação Integrada (BREQ3)
Regulação Identificada (BREQ3)	Regulação Identificada (BREQ3)	Regulação Identificada (BREQ3)
Regulação Introjetada (BREQ3)		
	Global (BPN)	Global (BPN)
	Competência (BPN)	Competência (BPN)
	Autonomia (BPN)	Autonomia (BPN)

Tabela 8 – Indicadores motivacionais - valores mais elevados nos I-Ativ

Baseline	6 meses	12 meses
Regulação Externa (BREQ3)	Regulação Externa (BREQ3)	Regulação Externa (BREQ3)
Amotivação (BREQ3)	Amotivação (BREQ3)	Amotivação (BREQ3)

Como era expectável através do que é preconizado pela teoria da autodeterminação foram os comportamentos mais autónomos que foram encontrados nos indivíduos Ativ. Os Ativ têm tipos mais autónomos de regulação extrínseca, que se aproximam da Regulação Intrínseca. No caso dos I-Ativ,

verificou-se exatamente o oposto, estes apresentam níveis superiores de amotivação e regulação externa.

Analisando os resultados da tabela 6 e respectivos resultados dos testes de hipóteses efetuados, podemos concluir que:

- O RAI, a Motivação Autônoma, a Regulação Intrínseca, Regulação Integrada, Regulação Identificada eram superiores nos Ativ nos 3 períodos temporais avaliados;
- A Regulação Introjetada era superior nos Ativ no *Baseline*;
- A Regulação Externa e a Amotivação eram superiores nos I-Ativ nos 3 períodos temporais avaliados;
- O valor Global, Competência e Autonomia (BPN), no âmbito da manutenção de perda de peso, eram superiores nos Ativ nos 6 meses e nos 12 meses.

III.3.c) Evolução dos indicadores de motivação dos indivíduos Ativ

Na próxima tabela foram utilizados apenas os valores médios motivacionais dos indivíduos classificados como Ativ, e foi verificada a sua variação ao longo do tempo, ou seja, foi medida a variação da motivação ao longo do tempo.

Tabela 9 – Variação dos valores médios dos indicadores motivacionais dos Ativ – variações estatisticamente significativas

Variáveis	Baseline vs 6 meses	6 meses vs 12 meses	Baseline vs 12 meses
RAI (BREQ3)	6,71 (1)	---	6,22
Motivação Autônoma (BREQ3)	3,67	---	3,47
Regulação Intrínseca (BREQ3)	2,94	---	---
Regulação Integrada (BREQ3)	6,02	---	6,96
Regulação Identificada (BREQ3)	2,30	---	---
Autonomia (BPN)	---	-2,20	-2,99

O valor identificado como **(1)** foi utilizado para explicar a leitura dos valores

Nota: considerou-se que a variação era estatisticamente significativa para um p-value<0,005.

Os valores em falta não eram estatisticamente significativos (---).

O que se apresenta na tabela tem a seguinte leitura: **(1)** Para os Ativ, o valor do RAI aumentou 6,71% entre o momento do *Baseline* e os 6 meses seguintes.

No que respeita a análise da tabela 9 que apresenta a variação dos valores médios dos indicadores motivacionais dos Ativ, e para verificar que as diferenças eram estatisticamente significativas, foi efetuado um teste às médias das variáveis considerando amostras emparelhadas (apenas o grupo de Ativ em momentos temporais distintos), ou seja, foi utilizado o teste *t-Student* para amostras emparelhadas.

Avaliando os resultados da tabela 9 podemos concluir que, nos Ativ:

- O RAI aumentou entre o *Baseline* e os 6 meses, e entre o *Baseline* e os 12 meses;
- A Motivação Autónoma aumentou entre o *Baseline* e os 6 meses, e entre o *Baseline* e os 12 meses;
- A Regulação Intrínseca aumentou entre o *Baseline* e os 6 meses;
- A Regulação Integrada aumentou entre o *Baseline* e os 6 meses, e entre o *Baseline* e os 12 meses;
- A Regulação Identificada aumentou entre o *Baseline* e os 6 meses;
- A Autonomia (BPN), no âmbito da manutenção de perda de peso, diminuiu entre os 6 e os 12 meses e entre o *Baseline* e os 12 meses.

Resumindo, ao traçar um perfil motivacional de um Ativ, através da informação analisada, podemos concluir que é um indivíduo com maior RAI, Motivação Autónoma, Regulação Intrínseca, Regulação Integrada e Regulação Identificada ao longo dos 12 meses de análise, com aumento do valor dessas variáveis sobretudo nos primeiros 6 meses. É também verificado que a sua Autonomia diminui, no âmbito da manutenção de perda de peso, aos 6 e 12 meses.

III.3.d) Evolução dos indicadores de motivação dos indivíduos I-Ativ

Na próxima tabela foram utilizados apenas os valores médios motivacionais dos indivíduos classificados como I-Ativ, e foi verificada a sua variação ao longo do tempo, ou seja, foi medida a variação da motivação ao longo do tempo.

Tabela 10 – Variação dos valores médios dos indicadores motivacionais dos I-Ativ – variações estatisticamente significativas

Variáveis	<i>Baseline</i> vs 6 meses	6 meses vs 12 meses	<i>Baseline</i> vs 12 meses
RAI (BREQ3)	4,57 (1)	---	6,09
Motivação Autónoma (BREQ3)	---	---	2,98

Regulação Integrada (BREQ3)	3,91	---	5,97
Regulação Identificada (BREQ3)	1,99	---	2,30
Global (BPN)	---	---	-4,35
Competência (BPN)	---	---	-5,13
Autonomia (BPN)	-2.22	---	-4,83

O valor identificado como **(1)** foi utilizado para explicar a leitura dos valores

Nota: considerou-se que a variação era estatisticamente significativa para um $p\text{-value} < 0,005$.

Os valores em falta não eram estatisticamente significativos (---).

O que se apresenta na tabela tem a seguinte leitura: **(1)** Para os I-Ativ, o valor do RAI aumentou 4,57% entre o momento do *Baseline* e os 6 meses seguintes.

No que respeita a análise da tabela 10 que apresenta variação dos valores médios dos indicadores motivacionais dos I-Ativ, e para verificar que as diferenças eram estatisticamente significativas, foi efetuado um teste às médias das variáveis considerando amostras emparelhadas (apenas o grupo de I-Ativ em momentos temporais distintos), ou seja, foi utilizado o teste *t-Student* para amostras emparelhadas.

Avaliando os resultados da tabela 10 podemos concluir que, nos I-Ativ:

- O RAI aumentou entre o *Baseline* e os 6 meses, e entre o *Baseline* e os 12 meses;
- A Motivação Autónoma aumentou entre o *Baseline* e os 12 meses;
- A Regulação Integrada aumentou entre o *Baseline* e os 6 meses, e entre o *Baseline* e os 12 meses;
- A Regulação Identificada aumentou entre o *Baseline* e os 6 meses, e entre o *Baseline* e os 12 meses;
- O valor Global (BPN) diminuiu entre o *Baseline* e os 12 meses;
- A Competência (BPN) diminuiu entre o *Baseline* e os 12 meses;
- A Autonomia (BPN) diminuiu entre o *Baseline* e os 6 meses e entre o *Baseline* e os 12 meses.

Resumindo, ao traçar um perfil motivacional de um I-Ativ, através da informação analisada, podemos concluir que é um indivíduo com maior RAI, Motivação Autónoma, Regulação Integrada e Regulação Identificada ao longo dos 12 meses de análise, com aumento do valor de algumas dessas variáveis sobretudo nos primeiros 6 meses. É também verificado que a nível de BPN, no âmbito da manutenção de perda de peso, o valor Global, a sua Competência e a sua Autonomia diminuem ao longo dos 12 meses. Sendo essa a principal diferença, relativamente ao grupo dos Ativ.

III.3.e) Comparação dos indicadores de motivação em ambos os grupos

Em seguida, apresentam-se 3 gráficos nos quais estão representadas, em simultâneo, as variações dos valores médios dos indicadores motivacionais para ambos os grupos.

Gráfico 4 – Comparação das variações dos indicadores motivacionais para os Ativ e os I-Ativ (entre o *Baseline* e os 6 meses)

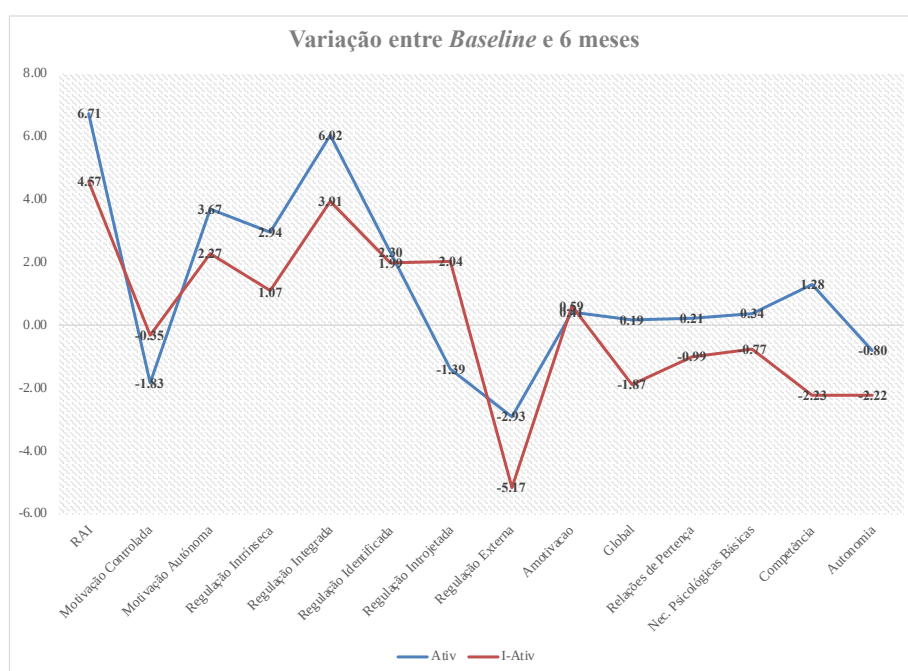


Gráfico 5 – Comparação das variações dos indicadores motivacionais para os Ativ e os I-Ativ (entre os 6 e os 12 meses)

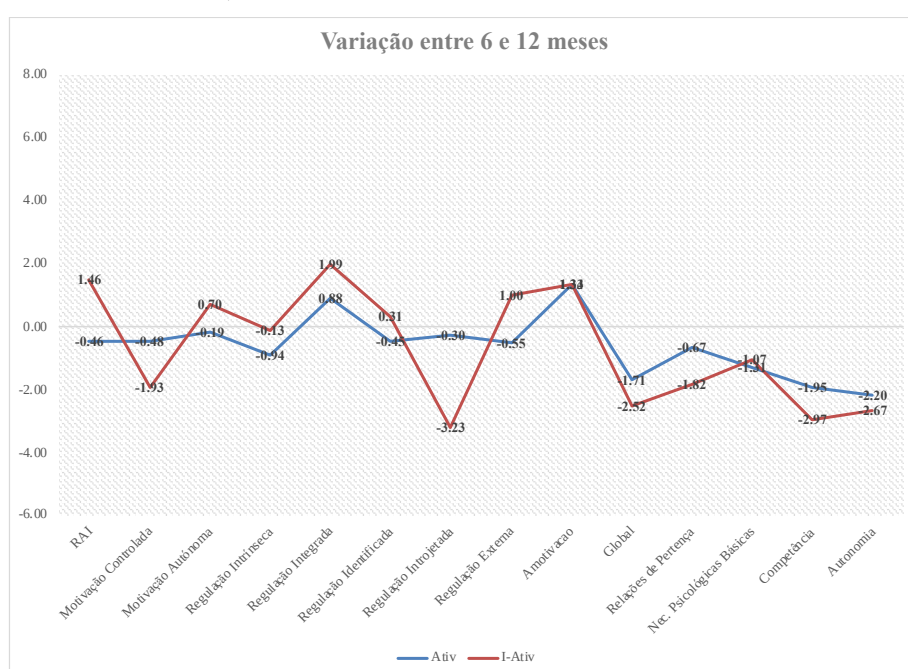
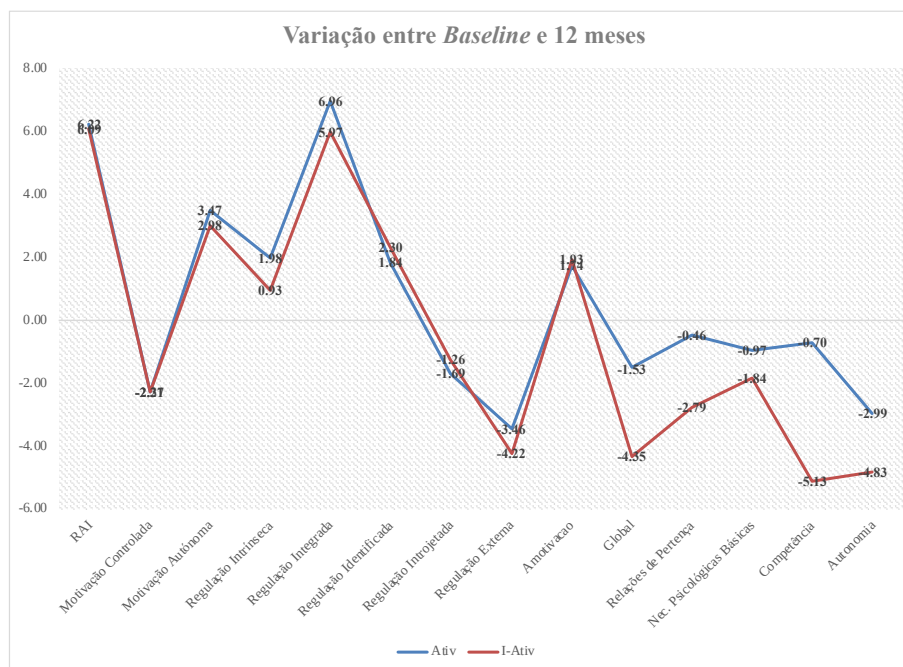


Gráfico 6 – Comparação das variações dos indicadores motivacionais para os Ativ e os I-Ativ (entre o *Baseline* e os 12 meses)



Da análise visual dos 3 gráficos foi possível verificar que os indicadores de motivação variaram de forma distinta nos 2 grupos identificados, se avaliarmos o período temporal *Baseline* e 6 meses, conseguimos visualizar que quase todas as variáveis motivacionais tiveram variações superiores nos Ativ, excetuando-se a Motivação Controlada, Regulação Introjetada e Regulação Externa.

No período temporal 6 e 12 meses, notou-se que as variações foram menos acentuadas, verificando-se maiores decréscimos na Motivação Autônoma, Motivação Introjetada, e, no âmbito da manutenção de perda de peso, nas variáveis Global, Competência e Autonomia do BPN, no grupo dos I-Ativ.

Finalmente, ao compararmos a variação dos indicadores motivacionais nos 2 grupos, entre o *Baseline* e os 12 meses, verificámos que as maiores diferenças estavam em todos os indicadores do BPN, estes diminuíram nos 2 grupos, verificando-se a maior diminuição no grupo dos I-Ativ. Nas restantes variáveis a variação era bastante semelhante.

Resumindo, verificou-se que é na fase *Baseline*-6 meses que os 2 grupos se distinguiram mais a nível do perfil motivacional, a partir dos 6 meses e até aos 12 meses, os 2 grupos apresentaram perfis motivacionais semelhantes, exceto no que concerne as variáveis motivacionais referentes à gestão do peso, que eram claramente menores nos I-Ativ que nos Ativ..

III.4 Discussão dos Resultados

Este estudo teve como propósito ajudar na explicação do fenómeno da manutenção do peso perdido, através da caracterização do perfil motivacional de indivíduos que apresentam níveis diferentes de prática de atividade física ao longo de 12 meses após a perda de peso. Neste trabalho foram aplicados algoritmos de *Machine Learning* à informação recolhida durante 12 meses pelas *FitBit* utilizadas no âmbito do projeto *NoHoW*, usadas por indivíduos que se encontram em fase de manutenção do peso perdido. Estes algoritmos permitiram identificar quais os indivíduos com maiores níveis de prática de atividade física, denominados “Ativos” (Ativ), e os com menores níveis, apelidados de “Insuficientemente Ativos” (I-Ativ). Foram, portanto, identificadas as variáveis motivacionais presentes nos Ativ e I-Ativ, através da análise da informação discreta obtida em questionários motivacionais recolhidos no *Baseline*, após 6 e 12 meses.

Foi possível com este trabalho, identificar 2 grupos de indivíduos, os Ativ e distingui-los dos I-Ativ, e validar que nos grupos os indivíduos tinham comportamentos diferentes, face à prática de AF diferente.

A nível da definição do perfil dos indivíduos Ativ, os resultados mostraram que eram indivíduos com uma média diária de mais de 12 mil passos, uma frequência cardíaca a rondar os 70bpm, uma distância percorrida de mais de 8 km diariamente e um dispêndio calórico próximo das 3000 kcal diárias. Eram indivíduos que apresentaram valores motivacionais mais elevados nas variáveis RAI, Motivação Autónoma, Regulações Intrínseca, Integrada e Identificada nos 3 períodos em análise, e que aos 6 e 12 meses, tinham valores mais elevados nas variáveis motivacionais no contexto da gestão do peso (BPN): Global, Competência e Autonomia, quando comparados com os indivíduos I-Ativ. Ao longo do tempo, verifica-se que os indicadores RAI, Motivação Autónoma e Regulação Integrada aumentaram, e que a Autonomia, no âmbito da manutenção de perda de peso, diminuiu no final dos 12 meses.

Considerando novamente a revisão sistemática de Teixeira, Carraça, Markland, Silva e Ryan (2012), a TAD permite o entendimento do comportamento no exercício, e na revisão efetuada foi possível clarificar a importância e os benefícios de desenvolver regulação autónoma através de formas de regulação extrínseca (regulações identificada e integrada) ou aumentando a motivação intrínseca. Através deste estudo percebeu-se que todas as formas de regulação autónoma eram preditores para a prática de atividade física e que eram os indivíduos com perfis motivacionais com elevada motivação autónoma que exibiam comportamentos de prática ativa de atividade física ao longo do tempo. Para que a prática se mantenha ao longo do tempo, a motivação intrínseca revelou-se de extrema importância. De notar que estas conclusões são concordantes com os resultados obtidos na definição do perfil

motivacional do grupo Ativ que identificamos, já que os tipos mais elevados de motivação que apresentam são os referentes à motivação intrínseca.

A nível da definição do perfil dos indivíduos I-Ativ, os resultados mostraram que eram indivíduos com uma média diária de mais de 7 mil passos, uma frequência cardíaca a rondar os 74 bpm, uma distância percorrida de mais de 5 km diariamente e um dispêndio calórico próximo das 2500 kcal diárias. Isto mostra que eram indivíduos com uma atividade física significativa, não sendo sedentários, apresentaram níveis inferiores de prática de atividade física que os Ativ. Eram indivíduos que apresentaram valores motivacionais mais elevados nas variáveis Regulação Externa e Amotivação nos 3 períodos em análise, quando comparados com os indivíduos Ativ. Ao longo do tempo, e sobretudo nos primeiros 6 meses, verificou-se que os indicadores RAI, e Regulação Integrada e Identificada aumentaram, e que os valores motivacionais no âmbito da manutenção de perda de peso: Global, Competência e Autonomia diminuíram no final dos 12 meses.

Importa referir que no âmbito do projeto *NoHoW*, houve uma preocupação inicial em apoiar e motivar os participantes neste projeto, através de intervenções motivacionais, o que teve impacto no perfil motivacional dos indivíduos de ambos os grupos. Essas intervenções permitiram melhorar os níveis de motivação, pelo que os nossos resultados refletiram também essa componente.

É também verificado que a nível das variáveis motivacionais, no âmbito da gestão do peso (BPN), o valor Global, a sua Competência e a sua Autonomia diminuem ao longo dos 12 meses no grupo dos I-Ativ. Sendo essa a principal diferença encontrada entre os 2 grupos. Verificou-se que é na fase Baseline-6 meses que os 2 grupos se distinguem mais a nível do perfil motivacional, a partir dos 6 meses e até aos 12 meses, os 2 grupos apresentaram perfis motivacionais semelhantes, exceto no que concerne as variáveis motivacionais referentes à gestão do peso, que foram claramente menores nos I-Ativ que nos Ativ..

No trabalho de Teixeira, Carraça, Markland, Silva e Ryan (2012), houve também indicação que quando os indivíduos perceberam que a sua autonomia poderia estar comprometida, houve um maior risco de adotarem comportamentos de desistência. Fazendo a ligação com o perfil motivacional identificado do grupo dos I-Ativ, apesar de apresentarem níveis de atividade física assinalável, podem ter tendência, por apresentarem menores níveis de autonomia, para adotar comportamentos de desistência e uma menor adesão à prática sustentada de atividade física. No estudo de Teixeira, Silva, Mata, Palmeira e Markland (2012) foi também descrita a importância da autonomia e regulações autónomas na mudança comportamental a longo prazo. A evidência existente aponta para o papel da motivação autónoma e intrínseca na adesão à atividade física, sendo também comportamentos essenciais quando se trata da gestão do peso. Estas conclusões podem ser enquadradas no nosso estudo, no qual a avaliação do comportamento é feita ao longo de 12 meses, e onde se verificou que a manutenção de um comportamento de prática de exercício estava associada a esses tipos de motivação. Verificou-se que as diferenças entre os indivíduos Ativ e I-Ativ são visíveis sobretudo no seu perfil motivacional, mais do que nos indicadores recolhidos pela *FitBit*, que apesar de serem diferentes, não

apresentam diferenças muito discrepantes. Ficou evidente a extrema relevância da definição do perfil motivacional do indivíduo para determinar o seu comportamento face à prática de atividade física e à gestão do peso.

Avaliando a evolução dos indicadores motivacionais ao longo do tempo, foi possível verificar que no início do projeto e nos primeiros 6 meses ocorreu um aumento dos níveis de motivação, o que é expectável pois é um processo natural numa situação nova, em que os indivíduos tendem a estar mais e melhor motivados, no entanto verificou-se uma estagnação aos 6 meses, e globalmente, no final dos 12 meses. Ainda que a motivação tivesse aumentado no início do projeto, as variáveis motivacionais, no contexto da gestão do peso, reduziram-se significativamente. Estes resultados podem indicar que estes períodos (6 e 12 meses) são momentos importantes de intervenção motivacional, são fases em que há necessidade de reforçar a motivação individual independentemente do nível de atividade física do indivíduo sob pena que poder ocorrer um decréscimo na prática de atividade física.

Importa referir que existem diferenças conhecidas nas intervenções que têm como objetivo promover estilos de vida saudável e a prática de atividade física quando comparadas com intervenções cujo objetivo é a manutenção desses mesmos comportamentos (Molanorouzi, Khoo, & Morris, 2015). Este estudo mostrou através de uma RSL que as intervenções de modificação comportamental são modestamente efetivas no curto e longo prazo e que existe uma grande heterogeneidade nas mesmas. No entanto, foi mostrado que intervenções de modificação comportamental que enfatizam um estilo de comunicação centrado no indivíduo e que promovem a autonomia, como a Entrevista Motivacional, a Teoria da Autodeterminação entre outras, estão associadas a efeitos positivos a longo prazo (Molanorouzi et al., 2015).

Adicionalmente, e de acordo com Rebar, Elavsky, Maher, Doerksen, e Conroy (2014), a atividade física é regulada por processos controlados como intenções, e por processos automáticos como hábitos. Quando os hábitos são inexistentes ou fracos, as intenções devem prevalecer (situação menos presente em pessoas com hábitos fortes), no entanto essas intenções são mutáveis ao longo do tempo, o que dificulta a prática diária de atividade física. Verificou-se, neste estudo, que nos dias em que os indivíduos tinham intenções mais fracas do que o normal, a força do hábito estava positivamente relacionada à manutenção da prática de atividade física, mas nos dias em que as pessoas tinham intenções típicas ou mais fortes do que era normal para elas, a força do hábito não estava relacionada com a prática de atividade física diária. Assim, os esforços para promover a atividade física devem ter em consideração não só o perfil motivacional dos indivíduos e os seus hábitos, algo de carácter mais estático, mas também a dinâmica das suas intenções diárias e isso só é possível com análises que recorrem a inteligência artificial.

Concluindo, neste trabalho através da aplicação de algoritmos de *Machine Learning*, foi possível efetuar o *clustering* de um conjunto de indivíduos com base na respetiva atividade diária recolhida durante 12 meses de forma ininterrupta, e de forma relativamente simples. Este tipo de algoritmos permitem abrir um caminho no sentido de definição de intervenções mais acertadas e eficazes, do que

tradicionalmente se fariam com a análise de informação recolhida em apenas um dia ou uma semana. A facilidade de utilização destas técnicas de inteligência artificial, aliada à capacidade computacional existente, permite nos dias de hoje analisar muito mais informação, que é recolhida de forma simples com pequenos dispositivos ao alcance de todos e assim obter muito melhores resultados do que os alcançados com análises tradicionais. Atualmente já se pratica medicina individualizada, com tratamentos adequados a cada um dos pacientes, porque não aplicar o mesmo tipo de procedimentos nas intervenções relacionadas com a atividade física e com a motivação individual? Parece ser possível torná-las cada vez mais aperfeiçoadas e adequadas a cada indivíduo ou a um conjunto de indivíduos com características semelhantes. Essa é a grande vantagem que estas técnicas nos aportam. Importa referir que, este tipo de abordagens implica a necessidade de envolver equipas multidisciplinares na análise da informação, mais onerosas, com inclusão de cientistas de dados, com fortes conhecimentos matemáticos, de estatística e de programação, nestes projetos da área da saúde e bem-estar.

Neste trabalho, ainda que tenha sido possível tirar conclusões acerca da caracterização motivacional de 2 grupos distintos de indivíduos no que respeita a prática de atividade física e de ter sido possível encontrar algumas pistas no que respeitam possíveis momentos chave de intervenções motivacionais, seria interessante complementar esta análise no futuro. Seria interessante por um lado revalidar qual a tendência dos indicadores recolhidos pela *FitBit*, ou seja, se com o decréscimo da motivação também se verifica uma redução da atividade física no final dos 12 meses ou num momento posterior, por exemplo, 18 meses. Com outras palavras, olhar para a evolução dos indicadores de prática de atividade física em conjunto com a evolução das variáveis motivacionais. Seria também interessante utilizar outras variáveis motivacionais de outros questionários e avaliar os seus resultados para completar o perfil motivacional dos indivíduos.

Finalmente, seria também muito aliciante e valorizável, utilizar esta primeira análise para construir um modelo genérico de identificação de padrões de atividade física, com base na recolha de informação de dispositivos como a *FitBit* (entre outros), e de respostas a inquéritos motivacionais. Este modelo genérico poderia servir nos contextos médicos, nos quais é importante que pacientes pratiquem atividade física no âmbito de recuperação física e mental, mas também em ambientes mais recreativos como ginásios ou *health clubs*, nos quais se pretende diminuir a taxa de abandono dos clientes.

III.5 Conclusão e Limitações

Os resultados deste estudo permitiram por um lado caracterizar os indivíduos no que respeita padrões de atividade física, como associar o respetivo perfil motivacional. São a Motivação Autónoma, Regulação Intrínseca, Integrada e Identificada que permitem que os indivíduos façam e permaneçam empenhados na prática diária de atividade física. Este estudo permitiu validar a informação já existente que suporta que ao promovermos nos indivíduos a sua Autonomia e Motivação Intrínseca, potenciamos igualmente a sua capacidade de adotar um comportamento em prol da prática de atividade física.

No que respeita as limitações deste estudo, consideramos que poderiam ser aplicados outros algoritmos a este conjunto de dados, de modo a avaliar que metodologia de *clustering* seria adequada e produziria melhores resultados, já que neste caso apenas foi utilizado apenas o método *K-Means*.

Por outro lado, poderiam ser avaliadas outras variáveis motivacionais disponíveis no âmbito do projeto *NoHoW* com o objetivo de definir de forma mais completa o perfil motivacional dos indivíduos que praticam ativamente atividade física.

Segundo a literatura disponível, é possível verificar que os motivos que explicam a prática de atividade física diferem consoante o tipo de atividade, a idade e o sexo nos adultos e que o entendimento desses motivos é crítico no desenvolvimento de intervenções que promovam elevados níveis de comprometimento (Molanorouzi et al., 2015). Nesse sentido, poderia ser interessante em estudos futuros, diferenciar os tipos de atividade física praticados, as idades e o sexo dos indivíduos, e reavaliar as respetivas motivações associadas.

Outras análises futuras poderiam também incluir variáveis socioeconómicas e a identificação das barreiras para a prática de atividade física, de modo a definir de forma mais exata o perfil motivacional dos indivíduos no que respeita a prática de atividade física. Segundo Herazo-Beltrán (2017) existem barreiras diferentes no que concerne a prática de atividade física, em indivíduos que estão em estratos sociais diferentes. Indivíduos de estratos sociais mais baixos têm um risco acrescido de identificar como barreiras à prática de atividade física, a falta de motivação e de recursos financeiros. Estas barreiras diferem em indivíduos com menor nível de educação, estes identificam suporte social, a falta de recursos e o medo de lesões como as principais barreiras à prática de atividade física.

III.6 Referências Bibliográficas

- Alley, S., Jennings, C., Plotnikoff, R. C., & Vandelanotte, C. (2014). My Activity Coach - using video-coaching to assist a web-based computer-tailored physical activity intervention: a randomised controlled trial protocol. *BMC Public Health*, 14, 738. doi:10.1186/1471-2458-14-738
- Alley, S., Jennings, C., Plotnikoff, R. C., & Vandelanotte, C. (2016). Web-Based Video-Coaching to Assist an Automated Computer-Tailored Physical Activity Intervention for Inactive Adults: A Randomized Controlled Trial. *J Med Internet Res*, 18(8), e223. doi:10.2196/jmir.5664
- Aswani, A., Kaminsky, P., Mintz, Y., Flowers, E., & Fukuoka, Y. (2019). Behavioral Modeling in Weight Loss Interventions. *Eur J Oper Res*, 272(3), 1058-1072. doi:10.1016/j.ejor.2018.07.011
- Batterham, M., Tapsell, L., Charlton, K., O'Shea, J., & Thorne, R. (2017). Using data mining to predict success in a weight loss trial. *J Hum Nutr Diet*, 30(4), 471-478. doi:10.1111/jhn.12448
- Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *Int J Environ Res Public Health*, 15(12). doi:10.3390/ijerph15122796
- Chin, S. O., Keum, C., Woo, J., Park, J., Choi, H. J., Woo, J. T., & Rhee, S. Y. (2016). Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity. *Sci Rep*, 6, 34563. doi:10.1038/srep34563
- Cid, L., Monteiro, D., Teixeira, D., Teques, P., Alves, S., Moutão, J., . . . Palmeira, A. (2018). The Behavioral Regulation in Exercise Questionnaire (BREQ-3) Portuguese-Version: Evidence of Reliability, Validity and Invariance Across Gender. *Front Psychol*, 9, 1940. doi:10.3389/fpsyg.2018.01940
- Deo, R. C. (2015). Machine Learning in Medicine. *Circulation*, 132(20), 1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593
- Dijkhuis, T. B., Blaauw, F. J., van Ittersum, M. W., Velthuisen, H., & Aiello, M. (2018). Personalized Physical Activity Coaching: A Machine Learning Approach. *Sensors (Basel)*, 18(2). doi:10.3390/s18020623
- Dugan, T. M., Mukhopadhyay, S., Carroll, A., & Downs, S. (2015). Machine Learning Techniques for Prediction of Early Childhood Obesity. *Appl Clin Inform*, 6(3), 506-520. doi:10.4338/ACI-2015-03-RA-0036
- Flannery, M. (2017). Self-Determination Theory: Intrinsic Motivation and Behavioral Change. *Oncol Nurs Forum*, 44(2), 155-156. doi:10.1188/17.ONF.155-156
- Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media. ISBN: 978-1491962299

- Goldstein, S. P., Zhang, F., Thomas, J. G., Butryn, M. L., Herbert, J. D., & Forman, E. M. (2018). Application of Machine Learning to Predict Dietary Lapses During Weight Loss. *J Diabetes Sci Technol*, 12(5), 1045-1052. doi:10.1177/1932296818775757
- Gortler, J., Spinner, T., Streeb, D., Weiskopf, D., & Deussen, O. (2019). Uncertainty-Aware Principal Component Analysis. *IEEE Trans Vis Comput Graph*. doi:10.1109/TVCG.2019.2934812
- Gray, V. (2017). *Principal component analysis : methods, applications and technology*. Novinka. ISBN: 978-1-53610-889-7
- Hall, K. D., & Jordan, P. N. (2008). Modeling weight-loss maintenance to help prevent body weight regain. *Am J Clin Nutr*, 88(6), 1495-1503. doi:10.3945/ajcn.2008.26333
- Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism*, 69S, S36-S40. doi:10.1016/j.metabol.2017.01.011
- Haynes, A., Kersbergen, I., Sutin, A., Daly, M., & Robinson, E. (2018). A systematic review of the relationship between weight status perceptions and weight loss attempts, strategies, behaviours and outcomes. *Obes Rev*, 19(3), 347-363. doi:10.1111/obr.12634
- Herazo-Beltrán, Y., Pinillos, Y., Vidarte, J., Crissien, E., Suarez, D., & García, R. (2017). Predictors of perceived barriers to physical activity in the general adult population: a cross-sectional study. *Braz J Phys Ther*, 21(1), 44-50. doi:10.1016/j.bjpt.2016.04.003
- Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed. ed.). Berlin ; London: Springer.
- Kerr, J., Marinac, C. R., Ellis, K., Godbole, S., Hipp, A., Glanz, K., . . . Berrigan, D. (2017). Comparison of Accelerometry Methods for Estimating Physical Activity. *Med Sci Sports Exerc*, 49(3), 617-624. doi:10.1249/MSS.0000000000001124
- Manutenção do peso perdido com recurso às novas tecnologias da informação. (2015). NoHoW. Retrieved from: <http://nohow.fmh.ulisboa.pt/>
- MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms*. Cambridge: Cambridge University Press.
- Metrics and Methods: Questionnaires. Center for Self-Determination Theory. Retrieved from: <http://selfdeterminationtheory.org/questionnaires/>
- Molanorouzi, K., Khoo, S., & Morris, T. (2015). Motives for adult participation in physical activity: type of activity, age, and gender. *BMC Public Health*, 15, 66. doi:10.1186/s12889-015-1429-7
- Moutão, J. M., Serra, L. F., Alves, J. A., Leitão, J. C., & Vlachopoulos, S. P. (2012). Validation of the Basic Psychological Needs in Exercise Scale in a Portuguese sample. *Span J Psychol*, 15(1), 399-409. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/22379729>
- Ng, J. Y., Ntoumanis, N., Thøgersen-Ntoumani, C., Stott, K., & Hindle, L. (2013). Predicting psychological needs and well-being of individuals engaging in weight management: the role of important others. *Appl Psychol Health Well Being*, 5(3), 291-310. doi:10.1111/aphw.12011

- Novatchkov, H., & Baca, A. (2013). Artificial intelligence in sports on the example of weight training. *J Sports Sci Med*, 12(1), 27-37. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/24149722>
- Painter, S. L., Ahmed, R., Hill, J. O., Kushner, R. F., Lindquist, R., Brunning, S., & Margulies, A. (2017). What Matters in Weight Loss? An In-Depth Analysis of Self-Monitoring. *J Med Internet Res*, 19(5), e160. doi:10.2196/jmir.7457
- Poulimeneas, D., Yannakoulia, M., Anastasiou, C. A., & Scarmeas, N. (2018). Weight Loss Maintenance: Have We Missed the Brain? *Brain Sci*, 8(9). doi:10.3390/brainsci8090174
- Ramage, S., Farmer, A., Eccles, K. A., & McCargar, L. (2014). Healthy strategies for successful weight loss and weight maintenance: a systematic review. *Appl Physiol Nutr Metab*, 39(1), 1-20. doi:10.1139/apnm-2013-0026
- Rebar, A. L., Elavsky, S., Maher, J. P., Doerksen, S. E., & Conroy, D. E. (2014). Habits predict physical activity on days when intentions are weak. *J Sport Exerc Psychol*, 36(2), 157-165. doi:10.1123/jsep.2013-0173
- RIAP. Repositório de Instrumentos de Avaliação Psicossocial e Autoregulação na Atividade Física Alimentação e Gestão do Peso. FMH - Faculdade de Motricidade Humana. Retrieved from: <http://repositorio.fmh.ulisboa.pt/catalogo-de-instrumentos/lista-completa-por-temas>
- Rodrigues, F., Bento, T., Cid, L., Pereira Neiva, H., Teixeira, D., Moutão, J., . . . Monteiro, D. (2018). Can Interpersonal Behavior Influence the Persistence and Adherence to Physical Exercise Practice in Adults? A Systematic Review. *Front Psychol*, 9, 2141. doi:10.3389/fpsyg.2018.02141
- Rosa, J. P., de Souza, A. A., de Lima, G. H., Rodrigues, D. F., de Aquino Lemos, V., da Silva Alves, E., . . . de Mello, M. T. (2015). Motivational and evolutionary aspects of a physical exercise training program: a longitudinal study. *Front Psychol*, 6, 648. doi:10.3389/fpsyg.2015.00648
- Rossi, A., & Calogiuri, G. (2018). Patterns and correlates of physical activity in adult Norwegians: a forecasted evolution up to 2025 based on machine learning approach. *BMC Public Health*, 18(1), 913. doi:10.1186/s12889-018-5854-2
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78. <https://dx.doi.org/10.1037/0003-066X.55.1.68>
- Sainsbury, K., Evans, E. H., Pedersen, S., Marques, M. M., Teixeira, P. J., Lähteenmäki, L., . . . Sniehotta, F. F. (2018). Attribution of weight regain to emotional reasons amongst European adults with overweight and obesity who regained weight following a weight loss attempt. *Eat Weight Disord*. doi:10.1007/s40519-018-0487-0
- Scott, S. E., Duarte, C., Encantado, J., Evans, E. H., Harjumaa, M., Heitmann, B. L., . . . Stubbs, R. J. (2019). The NoHoW protocol: a multicentre 2×2 factorial randomised controlled trial

- investigating an evidence-based digital toolkit for weight loss maintenance in European adults. *BMJ Open*, 9(9), e029425. doi:10.1136/bmjopen-2019-029425
- Serrano, K. J., Yu, M., Coa, K. I., Collins, L. M., & Atienza, A. A. (2016). Mining Health App Data to Find More and Less Successful Weight Loss Subgroups. *J Med Internet Res*, 18(6), e154. doi:10.2196/jmir.5473
- Silva, M. N., Markland, D., Carraça, E. V., Vieira, P. N., Coutinho, S. R., Minderico, C. S., . . . Teixeira, P. J. (2011). Exercise autonomous motivation predicts 3-yr weight loss in women. *Med Sci Sports Exerc*, 43(4), 728-737. doi:10.1249/MSS.0b013e3181f3818f
- Silva, M. N., Markland, D., Minderico, C. S., Vieira, P. N., Castro, M. M., Coutinho, S. R., . . . Teixeira, P. J. (2008). A randomized controlled trial to evaluate self-determination theory for exercise adherence and weight control: rationale and intervention description. *BMC Public Health*, 8, 234. doi:10.1186/1471-2458-8-234
- Soeliman, F. A., & Azadbakht, L. (2014). Weight loss maintenance: A review on dietary related strategies. *J Res Med Sci*, 19(3), 268-275. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/24949037>
- Sorgente, A., Pietrabissa, G., Manzoni, G. M., Re, F., Simpson, S., Perona, S., . . . Castelnovo, G. (2017). Web-Based Interventions for Weight Loss or Weight Loss Maintenance in Overweight and Obese People: A Systematic Review of Systematic Reviews. *J Med Internet Res*, 19(6), e229. doi:10.2196/jmir.6972
- Stein, N., & Brooks, K. (2017). A Fully Automated Conversational Artificial Intelligence for Weight Loss: Longitudinal Observational Study Among Overweight and Obese Adults. *JMIR Diabetes*, 2(2), e28. doi:10.2196/diabetes.8590
- Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: a systematic review. *Int J Behav Nutr Phys Act*, 9, 78. doi:10.1186/1479-5868-9-78
- Teixeira, P. J., Silva, M. N., Mata, J., Palmeira, A. L., & Markland, D. (2012). Motivation, self-determination, and long-term weight control. *Int J Behav Nutr Phys Act*, 9, 22. doi:10.1186/1479-5868-9-22
- Thomas, D. M., Martin, C. K., Heymsfield, S., Redman, L. M., Schoeller, D. A., & Levine, J. A. (2011). A Simple Model Predicting Individual Weight Change in Humans. *J Biol Dyn*, 5(6), 579-599. doi:10.1080/17513758.2010.508541
- Trost, S. G., Wong, W. K., Pfeiffer, K. A., & Zheng, Y. (2012). Artificial neural networks to predict activity type and energy expenditure in youth. *Med Sci Sports Exerc*, 44(9), 1801-1809. doi:10.1249/MSS.0b013e318258ac11
- Tudor-Locke, C., Craig, C.L., Brown, W.J. et al. How many steps/day are enough? for adults. *Int J Behav Nutr Phys Act* 8, 79 (2011) doi:10.1186/1479-5868-8-79

- Willetts, M., Hollowell, S., Aslett, L., Holmes, C., & Doherty, A. (2018). Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Sci Rep*, 8(1), 7961. doi:10.1038/s41598-018-26174-1
- Wing, R. R., & Phelan, S. (2005). Long-term weight loss maintenance. *Am J Clin Nutr*, 82(1 Suppl), 222S-225S. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/16002825>
- World Health Organization. Health topics. Obesity. Retrieved from: <http://www.who.int/topics/obesity/en/>
- World Health Organization. (2016). Obesity and overweight, Key facts. *Fact sheet* - Reviewed February 2018.
- Yoon, S., Co, M. C., Suero-Tejeda, N., & Bakken, S. (2016). A Data Mining Approach for Exploring Correlates of Self-Reported Comparative Physical Activity Levels of Urban Latinos. *Stud Health Technol Inform*, 225, 553-557. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/27332262>
- Yoon, S., Suero-Tejeda, N., & Bakken, S. (2015). A Data Mining Approach for Examining Predictors of Physical Activity Among Urban Older Adults. *J Gerontol Nurs*, 41(7), 14-20. doi:10.3928/00989134-20150420-01
- Zhou, M., Fukuoka, Y., Mintz, Y., Goldberg, K., Kaminsky, P., Flowers, E., & Aswani, A. (2018). Evaluating Machine Learning-Based Automated Personalized Daily Step Goals Delivered Through a Mobile Phone App: Randomized Controlled Trial. *JMIR Mhealth Uhealth*, 6(1), e28. doi:10.2196/mhealth.9117

CAPÍTULO IV

CONCLUSÃO GERAL

IV.1 Discussão geral

Em primeiro lugar, neste trabalho foi efetuada uma RSL rápida, na qual foi avaliada a aplicação de *Machine Learning* no contexto da gestão do peso e da adesão á atividade física, os principais resultados obtidos dos estudos avaliados apontaram para a necessidade de intervenções personalizadas, com inclusão de atividade física regular, para o sucesso individual da gestão do peso, e também para o fenómeno de que a prática de atividade física deve ser ajustada aos padrões de vida atuais, ou seja, no futuro deverá ocorrer uma redução da duração da prática, mas aumentará a sua intensidade. Os preditores para a prática de atividade física obtidos com esta RSL rápida parecem estar mais associados a fatores psicológicos e comportamentais do que a fatores sociodemográficos.

Em seguida, foi efetuado um estudo que teve como objetivo ajudar a explicar o fenómeno da manutenção do peso perdido, através da identificação de diferentes níveis de atividade física, em indivíduos adultos ao longo de 12 meses, e da sua associação a variáveis motivacionais. Este projeto iniciou-se com a aplicação de algoritmos de *Machine Learning* à informação recolhida durante 12 meses pelas *FitBit* utilizadas no âmbito do projeto *NoHoW* e permitiu identificar quais os indivíduos que praticam ativamente atividade física (Ativ), e os que não conseguiram fazê-lo de forma tão ativa, ou fizeram-no de forma insuficiente (I-Ativ). Em seguida, foram associadas aos Ativ e aos I-Ativ as respetivas variáveis motivacionais, cruzando a informação obtida do *clustering* dos participantes com a informação discreta obtida em questionários comportamentais e motivacionais recolhidos no *Baseline*, após 6 e 12 meses.

Verificou-se que os Ativ são indivíduos com níveis altos de Motivação Autónoma, Regulação Intrínseca, Integrada e Identificada sobretudo nos primeiros 6 meses de avaliação, fase em que estes índices aumentam. Apesar de terem também estes índices motivacionais elevados, os I-Ativ não atingem os mesmos valores que os Ativ. Os Ativ apresentaram também uma característica interessante, o seu nível de Autonomia, no contexto da gestão do peso, diminuiu entre os 6 e 12 meses, apresentando valor mais baixo sobretudo no final dos 12 meses. Este parece ser, portanto, um momento importante para intervir a nível motivacional. Os I-Ativ exibiram igualmente descidas significativas no que respeita as variáveis motivacionais no contexto da gestão do peso, já que o valor Global, a Autonomia e a Competência decresceram no final do período dos 12 meses.

IV.2 Conclusões finais

Concluindo, é possível aferir que a inteligência artificial e as metodologias estatísticas mais avançadas podem trazer valor acrescentado quando se analisa informação, designadamente quando pensamos na aplicação de métodos de *Machine Learning* e quando pretendemos compreender os fenómenos associados à gestão do peso e à prática de atividade física. É essencial reter a importância do reforço da Motivação Autónoma e a Regulação Intrínseca para que os indivíduos façam e permaneçam empenhados na prática diária de atividade física. É também muito relevante avaliar as variáveis motivacionais no contexto da gestão do peso, pois estas diferenciam indivíduos mais ativos dos outros, menos ativos.

Considerando que o sedentarismo é uma realidade inevitável nos tempos modernos e que os impactos negativos associados a esse comportamento são conhecidos, sobretudo por não ajudarem na manutenção de um peso adequado e saudável, considera-se que mais análises deverão ser feitas no futuro no sentido de obter uma caracterização mais exata do perfil dos praticantes ativos de atividade física, não só na componente motivacional, que deve ser mais aprofundada, mas noutros fatores como a caracterização individual socioeconómica, demográfica, grau académico, entre outros.

O próximo passo desta análise poderá ser a criação de uma rede neuronal que permita através da conjugação dos dados da *FitBit* e das respostas aos inquéritos motivacionais prever quando é que um indivíduo classificado como Ativ começa a apresentar características que indiciam a mudança para um perfil de I-Ativ. Só assim será possível atuar para antecipar e intervir nessas alterações comportamentais, de modo ajudar os indivíduos a manterem uma prática de atividade física consistente, permanente e duradoura ao longo da vida e assim, terem um estilo de vida mais saudável.

IV.3 Referências Bibliográficas

- Alley, S., Jennings, C., Plotnikoff, R. C., & Vandelanotte, C. (2014). My Activity Coach - using video-coaching to assist a web-based computer-tailored physical activity intervention: a randomised controlled trial protocol. *BMC Public Health*, 14, 738. doi:10.1186/1471-2458-14-738
- Alley, S., Jennings, C., Plotnikoff, R. C., & Vandelanotte, C. (2016). Web-Based Video-Coaching to Assist an Automated Computer-Tailored Physical Activity Intervention for Inactive Adults: A Randomized Controlled Trial. *J Med Internet Res*, 18(8), e223. doi:10.2196/jmir.5664
- APA (2009). *Publication Manual of the American Psychological Association - 6th Edition*. Washington - DC: APA
- Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *Int J Environ Res Public Health*, 15(12). doi:10.3390/ijerph15122796
- Cone, J. D. , & Foster, S. L. (1993). *Dissertations and Theses from Start to Finish: Psychology and Related Fields*. Washington, DC: APA
- Deo, R. C. (2015). Machine Learning in Medicine. *Circulation*, 132(20), 1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593
- Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism*, 69S, S36-S40. doi:10.1016/j.metabol.2017.01.011
- Kerr, J., Marinac, C. R., Ellis, K., Godbole, S., Hipp, A., Glanz, K., . . . Berrigan, D. (2017). Comparison of Accelerometry Methods for Estimating Physical Activity. *Med Sci Sports Exerc*, 49(3), 617-624. doi:10.1249/MSS.0000000000001124
- Painter, S. L., Ahmed, R., Hill, J. O., Kushner, R. F., Lindquist, R., Brunning, S., & Margulies, A. (2017). What Matters in Weight Loss? An In-Depth Analysis of Self-Monitoring. *J Med Internet Res*, 19(5), e160. doi:10.2196/jmir.7457
- Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. *Journal of global health*, 8(2), 020303. doi:10.7189/jogh.08.020303
- Ramage, S., Farmer, A., Eccles, K. A., & McCargar, L. (2014). Healthy strategies for successful weight loss and weight maintenance: a systematic review. *Appl Physiol Nutr Metab*, 39(1), 1-20. doi:10.1139/apnm-2013-0026
- Sorgente, A., Pietrabissa, G., Manzoni, G. M., Re, F., Simpson, S., Perona, S., . . . Castelnovo, G. (2017). Web-Based Interventions for Weight Loss or Weight Loss Maintenance in Overweight and Obese People: A Systematic Review of Systematic Reviews. *J Med Internet Res*, 19(6), e229. doi:10.2196/jmir.6972

Willetts, M., Hollowell, S., Aslett, L., Holmes, C., & Doherty, A. (2018). Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Sci Rep*, 8(1), 7961. doi:10.1038/s41598-018-26174-1

CAPÍTULO V

ANEXOS

V.1 Website do projeto *NoHoW*



V.2 Questionários motivacionais – Questionário de Regulação Comportamental para o Exercício (BREQ-3)

14.5.5. Questionário de Regulação Comportamental para o Exercício

Porque se exercita?

Estamos interessados nas razões fundamentais das pessoas na decisão de se envolverem ou não envolverem no exercício físico. Usando a escala abaixo, por favor indique qual o nível mais verdadeiro para si. Relembramos que não há respostas certas ou erradas nem perguntas traiçoeiras. Queremos apenas saber como é que se sente em relação ao exercício.

Escala: De 1 a 7; 1=Não é verdade para mim; 4= Algumas vezes é verdade para mim; 7=Muitas vezes é verdade para mim.

Eu faço exercício (ou eu faria exercício), porque...

1. Eu faço exercício porque é divertido
2. Eu faço exercício porque faz parte da minha identidade
3. É importante para mim fazer exercício regularmente
4. Sinto-me culpado/a quando não faço exercício
5. Faço exercício porque outras pessoas dizem que devo fazer
6. Não percebo porque é que tenho de fazer exercício
7. Gosto das minhas sessões de exercício
8. Eu faço exercício porque é uma parte fundamental da pessoa que sou
9. Dou valor aos benefícios/vantagens do exercício
10. Sinto-me fracassado/a quando não faço exercício durante algum tempo
11. Participo no exercício porque os meus amigos/família dizem que devo fazer
12. Não percebo o objectivo de fazer exercício
13. Acho o exercício uma atividade agradável
14. Eu faço exercício porque é consistente com os meus valores pessoais
15. Penso que vale a pena o esforço de me exercitar
16. Eu sinto que tenho, obrigatoriamente, de fazer exercício
17. Sinto-me pressionado/a pela minha família e amigos para fazer exercício
18. Penso que o exercício é uma perda de tempo

V.3 Questionários motivacionais – Escala de Necessidades Psicológicas Básicas (BPN)

14.5.3. Escala de Necessidades Psicológicas Básicas

Usando a escala abaixo, por favor indique até que ponto concorda com as seguintes afirmações sobre a forma como se sente em relação à sua gestão do peso.

Relativamente à gestão do meu peso:

Escala: 1 (Discordo totalmente) a 7 (Concordo totalmente)

1. Sinto que fazer esforços para manter o meu peso é uma decisão minha
2. Sinto-me capaz de controlar o meu peso
3. Existem pessoas que se preocupam comigo
4. Sinto-me apoiado/a.
5. Faço esforços, de forma intencional, para manter o meu peso
6. Sinto-me eficaz quando estou a tentar controlar o meu peso.
7. Sinto-me próximo/a de outras pessoas
8. Sinto-me valorizado/a (por terceiros).
9. É de livre vontade que faço esforços para manter o meu peso
10. Consigo superar desafios.
11. Tenho uma relação próxima com outras pessoas
12. Sinto-me escutado/a.
13. Sinto que estou a tentar atingir os meus próprios objetivos
14. Sou capaz de atingir objetivos de manutenção do peso que são desafiantes para mim
15. Existem pessoas em quem posso confiar
16. Sinto-me compreendido/a

V.4 Machine Learning – Código Fonte

```
import os as os
import numpy as np
import pandas as pd
from datetime import datetime
from scipy import stats
import json
import time
import pdb
import plotly as py
from plotly import tools
import plotly.graph_objects as go
import sklearn as sklearn
from sklearn.preprocessing import StandardScaler
from sklearn.cluster import KMeans
from sklearn import datasets

INPUT_DIRECTORY = 'meta_v2'
#INPUT_DIRECTORY = 'meta_v2_19_07_2019'
INPUT_FILE = INPUT_DIRECTORY + '/fbext-global-data.csv'
INPUT_DICTIONARY = INPUT_DIRECTORY + '/fbext-global-dict.json'
OUTPUT_FILE = INPUT_DIRECTORY + '/fbext-global-kmeans.csv'

def loadMatrix():

    matrix_dict = dict({})
    with open(INPUT_DICTIONARY) as ifile:
        matrix_dict = json.load(ifile)

    np_types = np.dtype(**matrix_dict)

    return np.genfromtxt(INPUT_FILE, delimiter=',', skip_header=0, dtype=np_types)

def cleanNaN(matrix):
    isNaN = np.isnan(matrix).any()
    print("\nExists NaN in array: %s" % isNaN)
    if isNaN:
```

```

matrix[np.isnan(matrix)] = 0.00000001

print('Exists NaN in array: %s\n' % np.isnan(matrix).any())
return matrix

def main():

    start = datetime.now()

    os.chdir("../dev")
    #os.chdir("../dev")
    #os.chdir("../sample")
    #os.chdir("../ver15may")

    print('-----')
    print('-----')
    print('-----')
    print('Running on directory: %s' % os.getcwd())
    print('Started at: %s \n\n' % start.strftime("%m/%d/%Y, %H:%M:%S"))

    # for debug
    np.seterr(invalid="raise")

    work_matrix = loadMatrix()

    # create pandas data frame from numpy array
    data_frame = pd.DataFrame(work_matrix)
    # participant_id column
    participant_ids = data_frame['participant_id']
    # add column names
    data_frame.columns = list(work_matrix.dtype.names)
    # remove std columns
    work_data_frame = data_frame.drop(['steps_sum_std', 'hr_average_std', 'hr_max_std', 'hr_min_std',
'hr_mode_std', 'hr_std_std', 'distance_sum_std', 'level_average_std', 'calories_sum_std'], 1)
    # remove participant_id column
    work_data_frame = work_data_frame.iloc[:,1:].values
    # standardizing values
    work_data_frame_std = StandardScaler().fit_transform(work_data_frame)
    # covariance matrix
    covariance_matrix = np.cov(work_data_frame_std.T)

```



```

# calculate eigenvectors and eigenvalues
eig_vals, eig_vecs = np.linalg.eig(covariance_matrix)

print('Eigenvectors \n%s' % eig_vecs)
print("\nEigenvalues \n%s' % eig_vals)

# calculate correlation matrix
correlation_matrix = np.corrcoef(work_data_frame_std.T)

eig_vals, eig_vecs = np.linalg.eig(correlation_matrix)

print('Eigenvectors \n%s' % eig_vecs)
print("\nEigenvalues \n%s' % eig_vals)

# singular vector decomposition
u,s,v = np.linalg.svd(work_data_frame_std.T)

print("\nU: \n%s' % u)
print('S: \n%s' % s)
print("V: \n%s' % v)

for ev in eig_vecs:
    np.testing.assert_array_almost_equal(1.0, np.linalg.norm(ev))

print("\nTest passed!\n')
# make a list of (eigenvalue, eigenvector) tuples
eig_pairs = [(np.abs(eig_vals[i]), eig_vecs[:,i]) for i in range(len(eig_vals))]

# sort the (eigenvalue, eigenvector) tuples from high to low
eig_pairs.sort()
eig_pairs.reverse()

# visually confirm that the list is correctly sorted by decreasing eigenvalues
print('Eigenvalues in descending order:')
for i in eig_pairs:
    print(i[0])

print("\n')

```

```

tot = sum(eig_vals)
var_exp = [(i / tot)*100 for i in sorted(eig_vals, reverse=True)]
cum_var_exp = np.cumsum(var_exp)

trace1 = dict(
    type='bar',
    x=['PC %s' %i for i in range(1,8)],
    y=var_exp,
    name='Individual'
)

trace2 = dict(
    type='scatter',
    x=['PC %s' %i for i in range(1,8)],
    y=cum_var_exp,
    name='Cumulative'
)

data = [trace1, trace2]

layout=dict(
    title='Explained variance by different principal components',
    yaxis=dict(
        title='Explained variance in percent'
    ),
    annotations=list([
        dict(
            x=1.16,
            y=1.05,
            xref='paper',
            yref='paper',
            text='Explained Variance',
            showarrow=False,
        )
    ])
)

fig = dict(data=data, layout=layout)
py.offline.ipplot(fig, filename='selecting-principal-components')

```

```

# 2 PC
#matrix_w = np.hstack((eig_pairs[0][1].reshape(9,1), eig_pairs[1][1].reshape(9,1)))

# 3 PC
matrix_w = np.hstack((eig_pairs[0][1].reshape(9,1), eig_pairs[1][1].reshape(9,1),
eig_pairs[2][1].reshape(9,1)))

print('Matrix W:\n', matrix_w)

# projection onto the new feature space
Y = work_data_frame_std.dot(matrix_w)

print('Matrix Y:\n', Y)

N = 100000
fig = go.Figure(data=go.Scattergl(
    x = Y[:,0],
    y = Y[:,1],
    mode='markers',
    marker=dict(
        color=np.random.randn(N),
        colorscale='Viridis',
        line_width=1
    )
))

fig.update_layout(title='Projection based on PC1 and PC2 values (no clustering)',
    yaxaxis_zeroline=False, xaxis_zeroline=False)

fig.show()

# calculate and plot k-means
fig = tools.make_subplots(rows=1, cols=2,
    print_grid=True,
    specs=[[{'is_3d': True}, {'is_3d': True}]]
scene = dict(
    camera = dict(
        up=dict(x=0, y=0, z=1),
        center=dict(x=0, y=0, z=0),
        eye=dict(x=2.5, y=1.1, z=2.1)
    ),

```

```

xaxis=dict(
    range=[-10, 10],
    title='PC 1',
    gridcolor='rgb(255, 255, 255)',
    zerolinecolor='rgb(255, 255, 255)',
    showbackground=True,
    backgroundcolor='rgb(227, 236, 245)',
    showticklabels=False, ticks=""
),
yaxis=dict(
    range=[-10, 10],
    title='PC 2',
    gridcolor='rgb(255, 255, 255)',
    zerolinecolor='rgb(255, 255, 255)',
    showbackground=True,
    backgroundcolor='rgb(227, 236, 245)',
    showticklabels=False, ticks=""
),
zaxis=dict(
    range=[-10, 10],
    title='PC 3',
    gridcolor='rgb(255, 255, 255)',
    zerolinecolor='rgb(255, 255, 255)',
    showbackground=True,
    backgroundcolor='rgb(227, 236, 245)',
    showticklabels=False, ticks=""
)
)

estimators = {'k_means_pc_2': KMeans(n_clusters=2),
              'k_means_pc_3': KMeans(n_clusters=3)}

k_labels = [0, 0]
k_fits = [0, 0]
fignum = 1
for name, est in estimators.items():
    est.fit(Y)
    labels = est.labels_
    k_labels[fignum - 1] = est.labels_
    k_fits[fignum - 1] = Y

```

```

trace = go.Scatter3d(x=Y[:, 0], y=Y[:, 1], z=Y[:, 2],
                    showlegend=False,
                    mode='markers',
                    marker=dict(
                        color=labels.astype(np.float),
                        line=dict(color='black', width=1)
                    ),
                    text=participant_ids.values,
                    textposition="top center"
                )
fig.append_trace(trace, 1, fignum)

fignum += 1

fig['layout'].update(height=1100, width=1500, margin=dict(l=1,r=1))

fig['layout']['scene1'].update(scene)
fig['layout']['scene2'].update(scene)

fig.update_layout(title='k-means based on 2 and 3 clusters')

py.offline.iplot(fig)

# create two additional columns to append to final output
data_frame['km_2'] = k_labels[0]
data_frame['km_3'] = k_labels[1]

if os.path.exists(OUTPUT_FILE):
    os.remove(OUTPUT_FILE)

data_frame.to_csv(OUTPUT_FILE, index=False)

# k-means for PC1 and PC2 2D
fig = go.Figure(data=go.Scattergl(
    x = k_fits[0][:,0],
    y = k_fits[1][:,1],
    mode='markers',
    marker=dict(
        color=k_labels[0].astype(np.float),

```

```

        colorscale='Viridis',
        line_width=1
    ),
    text=participant_ids.values
))

fig.update_layout(title='Projection based on PC1 and PC2 values with k=2',
                  yaxis_zeroline=False, xaxis_zeroline=False)

```

```
fig.show()
```

```

finish = datetime.now()
elapsed = finish - start
print("\n\nElapsed: %s" % str(elapsed))
print("Finished at: %s" % finish.strftime("%m/%d/%Y, %H:%M:%S"))
print('-----')
print('-----')
print('-----')

```

```

if __name__ == '__main__':
    main()

```

```

import os as os
import numpy as np
from datetime import datetime
import multiprocessing
from scipy import stats
import json
import time
import pdb

```

```
multiprocessing.set_start_method('spawn', True) #workrround for debugging
```

```

NUMBER_OF_PROCESS = 2
OUTPUT_VALUES_PER_DAY = 8
OUTPUT_DIRECTORY = 'meta_v2'
OUTPUT_FILE_FORMAT = OUTPUT_DIRECTORY + '/fbext-aggr-{:04d}.csv'
OUTPUT_DICTIONARY_FILE = OUTPUT_DIRECTORY + '/fbext-aggr-dict.json'
OUTPUT_GLOBAL_DICTIONARY_FILE = OUTPUT_DIRECTORY + '/fbext-global-dict.json'
OUTPUT_GLOBAL_MATRIX_FILE = OUTPUT_DIRECTORY + '/fbext-global-data.csv'

```

```

UTILIZATION_RATIO_TARGET = 0.5
UTILIZATION_DAYS_TARGET = 30
REPLACEMENT_MASK_VALUE = 0.0000000001
MULTIPROCESSING = True

```

```

counter = None
total_files = None

```

```

def init(ctr, tfiles):

```

```

    """ store the counter for later use """

```

```

    global counter

```

```

    counter = ctr

```

```

    global total_files

```

```

    total_files = tfiles

```

```

# Print iterations progress

```

```

def printProgressBar (iteration, total, prefix = "", suffix = "", decimals = 1, length = 100, fill = '█'):

```

```

    """

```

```

    Call in a loop to create terminal progress bar

```

```

    @params:

```

```

        iteration - Required : current iteration (Int)

```

```

        total     - Required : total iterations (Int)

```

```

        prefix    - Optional : prefix string (Str)

```

```

        suffix    - Optional : suffix string (Str)

```

```

        decimals  - Optional : positive number of decimals in percent complete (Int)

```

```

        length    - Optional : character length of bar (Int)

```

```

        fill      - Optional : bar fill character (Str)

```

```

    """

```

```

    percent = ("{0:." + str(decimals) + "f").format(100 * (iteration / float(total)))

```

```

    filledLength = int(length * iteration // total)

```

```

    bar = fill * filledLength + '-' * (length - filledLength)

```

```

    print('\r%s |%s| %s%% %s' % (prefix, bar, percent, suffix), end = '\r')

```

```

# Print New Line on Complete

```

```

    if iteration == total:

```

```

        print()

```

```

def generateOutputMatrix(line_count=0):

```

```

    matrix_dict = dict(

```

```

        dtype={

```

```

        'names': {},
        'formats': {}
    })

matrix_dict['dtype']['names'] = []
matrix_dict['dtype']['formats'] = []

# first column: set the id column type
matrix_dict['dtype']['names'].append('participant_id') #0
matrix_dict['dtype']['formats'].append('i4')

matrix_dict['dtype']['names'].append('steps_sum')      #1
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_average')      #2
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_max')          #3
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_min')          #4
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_mode')         #5
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_std')          #6
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('distance_sum')    #7
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('level_average')   #8
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('calories_sum')    #9
matrix_dict['dtype']['formats'].append('f16')

np_types = np.dtype(**matrix_dict)
output_matrix = np.zeros((line_count), dtype=np_types)

```



```

return output_matrix, matrix_dict

def calculateOutputMatrix(input_matrix, output_matrix):

    # set participant_id
    output_matrix['participant_id'] = int(input_matrix['participant_id'][0])

    fbdays_column = input_matrix['fbdays']
    fbdays_unique_values = np.unique(fbdays_column)

    row_idx = 0
    for day in fbdays_unique_values:

        # if average HR is 0 then ignore record

        #print('day: ', day)
        #print('-----')

        day_indexes = np.isin(fbdays_column, day)
        day_data = input_matrix[day_indexes]
        # steps
        day_steps_column = day_data['steps']
        day_steps_sum = np.sum(day_steps_column)
        # steps to matrix
        output_matrix[row_idx]['steps_sum'] = day_steps_sum

        # heart rate
        day_hr_column = day_data['HR']
        day_hr_filtered_indexes = np.where(day_hr_column > 0) # because when NA is transformed to -999 it
only uses above 0 values
        day_hr_sum = np.sum(day_hr_filtered_indexes)

        if day_hr_sum > 0:
            # hr filtered average
            day_hr_filtered_column = day_hr_column[day_hr_filtered_indexes]

            # hr filtered average to matrix
            output_matrix[row_idx]['hr_average'] = np.average(day_hr_filtered_column)

```

```

# hr filtered max to matrix
output_matrix[row_idx]['hr_max'] = np.max(day_hr_filtered_column)

# hr filtered min to matrix
output_matrix[row_idx]['hr_min'] = np.min(day_hr_filtered_column)

# hr filtered mode
day_hr_filtered_mode = stats.mode(day_hr_filtered_column)
# hr filtered mode to matrix
output_matrix[row_idx]['hr_mode'] = day_hr_filtered_mode[0][0]

# hr filtered std to matrix
output_matrix[row_idx]['hr_std'] = np.std(day_hr_filtered_column)
else:
    output_matrix[row_idx]['hr_average'] = REPLACEMENT_MASK_VALUE
    output_matrix[row_idx]['hr_max'] = REPLACEMENT_MASK_VALUE
    output_matrix[row_idx]['hr_min'] = REPLACEMENT_MASK_VALUE
    output_matrix[row_idx]['hr_mode'] = REPLACEMENT_MASK_VALUE
    output_matrix[row_idx]['hr_std'] = REPLACEMENT_MASK_VALUE

# distance sum
day_distance_column = day_data['distance']
# day_distance_sum to matrix
output_matrix[row_idx]['distance_sum'] = np.sum(day_distance_column)

# level average
day_level_column = day_data['level']
# level average to matrix
output_matrix[row_idx]['level_average'] = np.average(day_level_column)

# calories sum
day_calories_column = day_data['calories_overall']
# calories sum to matrix
output_matrix[row_idx]['calories_sum'] = np.sum(day_calories_column)

row_idx += 1

def loadInputMatrix(ifile):

    #str2date = lambda x: datetime.strptime(x.decode("utf-8"), '%Y-%m-%d').timestamp()

```

```

kwargs = dict(delimiter=","
              skip_header=1,
              usecols=(0, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13),
              dtype={
                  'names':
                      ('participant_id', 'date_time', 'steps', 'distance', 'level', 'calories_activity', 'calories_overall',
'sedentary', 'lightly_active', 'fairly_active', 'very_active', 'HR', 'fbdays', 'weekday'),
                  'formats':
                      (int, '|S20', int, float, int, float, float, int, int, int, int, '<U20', '|S5')},
                  names=('participant_id', 'date_time', 'steps', 'distance', 'level', 'calories_activity', 'calories_overall',
'sedentary', 'lightly_active', 'fairly_active', 'very_active', 'HR', 'fbdays', 'weekday'),
                  missing_values={5:"NA", 11:"NA"},
                  filling_values={5:0, 11:-999})
              #converters = {12: str2date})

return np.genfromtxt(ifile, **kwargs)

def saveOutputMatrix(file_name, output_matrix):
    separator = ','
    matrix_header = separator.join(np.array(output_matrix.dtype.descr)[:][:,0])
    np.savetxt(file_name, output_matrix, delimiter=",", header=matrix_header, fmt="%6.10f")

def generateGlobalMatrix(line_count=0):

    matrix_dict = dict(
        dtype={
            'names':{},
            'formats':{}
        })

    matrix_dict['dtype']['names'] = []
    matrix_dict['dtype']['formats'] = []

    # first column: set the id column type
    matrix_dict['dtype']['names'].append('participant_id')    #0
    matrix_dict['dtype']['formats'].append('i4')

    matrix_dict['dtype']['names'].append('steps_sum_avg')    #1
    matrix_dict['dtype']['formats'].append('f16')

```

```

matrix_dict['dtype']['names'].append('steps_sum_std')    #2
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_average_avg')    #3
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_average_std')    #4
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_max_avg')        #5
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_max_std')        #6
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_min_avg')        #7
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_min_std')        #8
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_mode_avg')       #9
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_mode_std')       #10
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_std_avg')        #11
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('hr_std_std')        #12
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('distance_sum_avg')  #13
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('distance_sum_std')  #14
matrix_dict['dtype']['formats'].append('f16')

```

```

matrix_dict['dtype']['names'].append('level_average_avg') #15
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('level_average_std') #16
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('calories_sum_avg') #17
matrix_dict['dtype']['formats'].append('f16')

matrix_dict['dtype']['names'].append('calories_sum_std') #18
matrix_dict['dtype']['formats'].append('f16')

np_types = np.dtype(**matrix_dict)
output_matrix = np.zeros((line_count), dtype=np_types)

return output_matrix, matrix_dict

def appendToGlobalMatrix(o_matrix, idx, queue):

    # set participant_id
    participant_id = o_matrix['participant_id'][0]

    # steps sum avg and std
    steps_sum_column = o_matrix['steps_sum']
    steps_sum_avg = np.average(steps_sum_column)
    steps_sum_std = np.std(steps_sum_column)

    # hr average avg and std
    hr_average_column = o_matrix['hr_average']
    hr_average_avg = np.average(hr_average_column)
    hr_average_std = np.std(hr_average_column)

    # hr max avg and std
    hr_max_column = o_matrix['hr_max']
    hr_max_avg = np.average(hr_max_column)
    hr_max_std = np.std(hr_max_column)

    # hr min avg and std
    hr_min_column = o_matrix['hr_min']
    hr_min_avg = np.average(hr_min_column)
    hr_min_std = np.std(hr_min_column)

    # hr mode avg and std
    hr_mode_column = o_matrix['hr_mode']

```

```

hr_mode_avg = np.average(hr_mode_column)
hr_mode_std = np.std(hr_mode_column)
# hr std avg and std
hr_std_column = o_matrix['hr_std']
hr_std_avg = np.average(hr_std_column)
hr_std_std = np.std(hr_std_column)
# distance sum avg and std
distance_sum_column = o_matrix['distance_sum']
distance_sum_avg = np.average(distance_sum_column)
distance_sum_std = np.std(distance_sum_column)
# level average avg and std
level_average_column = o_matrix['level_average']
level_average_avg = np.average(level_average_column)
level_average_std = np.std(level_average_column)
# calories sum avg and std
calories_sum_column = o_matrix['calories_sum']
calories_sum_avg = np.average(calories_sum_column)
calories_sum_std = np.std(calories_sum_column)

g_matrix = queue.get(block=True)
# set participant_id
g_matrix['participant_id'][idx] = participant_id

# steps sum avg and std
steps_sum_column = o_matrix['steps_sum']
g_matrix['steps_sum_avg'][idx] = steps_sum_avg
g_matrix['steps_sum_std'][idx] = steps_sum_std
# hr average avg and std
hr_average_column = o_matrix['hr_average']
g_matrix['hr_average_avg'][idx] = hr_average_avg
g_matrix['hr_average_std'][idx] = hr_average_std
# hr max avg and std
hr_max_column = o_matrix['hr_max']
g_matrix['hr_max_avg'][idx] = hr_max_avg
g_matrix['hr_max_std'][idx] = hr_max_std
# hr min avg and std
hr_min_column = o_matrix['hr_min']
g_matrix['hr_min_avg'][idx] = hr_min_avg
g_matrix['hr_min_std'][idx] = hr_min_std
# hr mode avg and std

```

```

hr_mode_column = o_matrix['hr_mode']
g_matrix['hr_mode_avg'][idx] = hr_mode_avg
g_matrix['hr_mode_std'][idx] = hr_mode_std
# hr std avg and std
hr_std_column = o_matrix['hr_std']
g_matrix['hr_std_avg'][idx] = hr_std_avg
g_matrix['hr_std_std'][idx] = hr_std_std
# distance sum avg and std
distance_sum_column = o_matrix['distance_sum']
g_matrix['distance_sum_avg'][idx] = distance_sum_avg
g_matrix['distance_sum_std'][idx] = distance_sum_std
# level average avg and std
level_average_column = o_matrix['level_average']
g_matrix['level_average_avg'][idx] = level_average_avg
g_matrix['level_average_std'][idx] = level_average_std
# calories sum avg and std
calories_sum_column = o_matrix['calories_sum']
g_matrix['calories_sum_avg'][idx] = calories_sum_avg
g_matrix['calories_sum_std'][idx] = calories_sum_std

queue.put(g_matrix)

```

```
def worker(argz):
```

```

    my_idx = argz[0]
    ifile = argz[1]
    queue = argz[2]

    input_matrix = loadInputMatrix(ifile)

    # only process if row count > 0
    if np.size(input_matrix, 0) > 0:

        fbdays_column = input_matrix['fbdays']
        fbdays_unique_values = np.unique(fbdays_column)
        # create the structure of the output matrix
        output_matrix, matrix_dict = generateOutputMatrix(len(fbdays_unique_values))
        # process the input matrix
        calculateOutputMatrix(input_matrix, output_matrix)
        # total number of rows in the output matrix (one row per day)

```

```
output_matrix_row_count = output_matrix.size
```

```
hr_column = output_matrix['hr_average']
```

```
hr_zero_values = hr_column[np.where(hr_column == REPLACEMENT_MASK_VALUE)]
```

```
non_utilization_ratio = (len(hr_zero_values)/output_matrix_row_count)
```

```
utilization_ratio = 1 - non_utilization_ratio
```

```
utilization_days = output_matrix_row_count - len(hr_zero_values)
```

```
if utilization_ratio > UTILIZATION_RATIO_TARGET and utilization_days >
```

```
UTILIZATION_DAYS_TARGET:
```

```
    # remove all HR zero value rows from the output matrix
```

```
    non_zero_day_indexes = np.where(hr_column > REPLACEMENT_MASK_VALUE)
```

```
    output_matrix = output_matrix[non_zero_day_indexes]
```

```
    # save output matrix
```

```
    participant_id = int(input_matrix['participant_id'][0])
```

```
    saveOutputMatrix(OUTPUT_FILE_FORMAT.format(participant_id), output_matrix)
```

```
    # add values to global matrix
```

```
    appendToGlobalMatrix(output_matrix, my_idx, queue)
```

```
global total_files
```

```
global counter
```

```
try:
```

```
    MULTIPROCESSING
```

```
    # += operation is not atomic, so we need to get a lock:
```

```
    with counter.get_lock():
```

```
        counter.value += 1
```

```
except NameError:
```

```
    counter.value += 1
```

```
# update Progress Bar
```

```
printProgressBar(counter.value, total_files.value, prefix = 'Progress:', suffix = 'Complete', length = 50)
```

```
#print('Finished processing %s file...' % ifile)
```

```
def main():
```

```
    start = datetime.now()
```



```

os.chdir("./dev")
#os.chdir("./dev")
#os.chdir("./sample")
#os.chdir("./ver15may")

print('-----')
print('-----')
print('-----')
print("There are %d CPUs on this machine." % multiprocessing.cpu_count())
print('Running on directory: %s' % os.getcwd())
print('Started at: %s \n\n' % start.strftime("%m/%d/%Y, %H:%M:%S"))

if not os.path.exists(OUTPUT_DIRECTORY):
    os.mkdir(OUTPUT_DIRECTORY)
else: # clean-up directory
    flz = [f for f in os.listdir(OUTPUT_DIRECTORY)]
    for fl in flz:
        os.remove(OUTPUT_DIRECTORY + '/' + fl)

included_extensions = ['csv']
filez = [f for f in os.listdir('.') if any(f.endswith(ext) for ext in included_extensions)]

files_count = len(filez)

# initial call to print 0% progress
printProgressBar(0, files_count, prefix = 'Progress:', suffix = 'Complete', length = 50)

# output matrix structure
o_matrix, o_matrix_dict = generateOutputMatrix()

json_dict = json.dumps(o_matrix_dict)
with open(OUTPUT_DICTIONARY_FILE, 'w') as ofile:
    ofile.write(json_dict)

# global matrix structure
g_matrix, g_matrix_dict = generateGlobalMatrix(files_count)

json_dict = json.dumps(g_matrix_dict)
with open(OUTPUT_GLOBAL_DICTIONARY_FILE, 'w') as ofile:

```

```

ofile.write(json_dict)

counter = multiprocessing.Value('i', 0)
total_files = multiprocessing.Value('i', files_count)

# multi-threaded process

manager = multiprocessing.Manager()
work_queue = manager.Queue()
work_queue.put(g_matrix)
queuez = [work_queue] * len(filez)
indexes = range(len(filez))

try:
    MULTIPROCESSING

    pool = multiprocessing.Pool(NUMBER_OF_PROCESS, initializer=init, initargs=(counter, total_files))
    argz = list(zip(indexes, filez, queuez))
    pool.map(worker, argz)
    pool.close()
    pool.join()
except NameError:
    init(counter, total_files)
    for ifile in filez:
        worker([indexes, ifile, work_queue])

g_matrix = work_queue.get(block=True)
# remove all participant_id zero value rows from the output matrix
zero_id_indexes = np.where(g_matrix['participant_id'] > 0)
g_matrix = g_matrix[zero_id_indexes]
saveOutputMatrix(OUTPUT_GLOBAL_MATRIX_FILE, g_matrix)

finish = datetime.now()
elapsed = finish - start
print("\n\nElapsed: %s" % str(elapsed))
print('Finished at: %s' % finish.strftime("%m/%d/%Y, %H:%M:%S"))
print('-----')
print('-----')
print('-----')

```

```
if __name__ == '__main__':  
    main()
```

V.5 Análises Estatísticas – Testes de Hipóteses (output do SPSS)

CAPÍTULO III – III.3 Resultados (Tabela 5)

Group Statistics

	km_final	N	Mean	Std. Deviation	Std. Error Mean
steps_sum_avg	1.00	453	12086.823298880290	2344.1125897732227	110.1360169852733
	2.00	686	7664.960051507309	1774.3958939068527	67.7467500356931
hr_average_avg	1.00	453	71.174467698683	6.2493331888361	.2936192865629
	2.00	686	74.125327535632	7.3118418997776	.2791674209717
distance_sum_avg	1.00	453	8.695455036599	1.8622317696294	.0874952810310
	2.00	686	5.317118334333	1.2731208755061	.0486079808989
level_average_avg	1.00	453	.325092253016115	.058964618890030	.002770399466278
	2.00	686	.217670585548397	.043117012166612	.001646215174172
calories_sum_avg	1.00	453	2866.657546232527	551.0282908905862	25.8895675359872
	2.00	686	2418.580758429903	446.1017280677006	17.0322431232360

		Levene's Test for Equality of Variances		t-test for Equality of Means	
		F	Sig.	t	df
steps_sum_avg	Equal variances assumed	25.492	.000	36.154	1137
	Equal variances not assumed			34.197	784.635
hr_average_avg	Equal variances assumed	10.293	.001	-7.055	1137
	Equal variances not assumed			-7.283	1064.543
distance_sum_avg	Equal variances assumed	38.261	.000	36.362	1137
	Equal variances not assumed			33.753	728.286
level_average_avg	Equal variances assumed	34.930	.000	35.472	1137
	Equal variances not assumed			33.334	764.644
calories_sum_avg	Equal variances assumed	27.741	.000	15.089	1137
	Equal variances not assumed			14.459	825.848

CAPÍTULO III – III.3 Resultados (Tabela 6, 9 e 10)

T-Test

Group Statistics					
	km_final	N	Mean	Std. Deviation	Std. Error Mean
var_RAI_0006	1.00	433	-4.595539794290255	186.242410101809440	8.950239048832630
	2.00	661	-24.644886776606278	391.246618946059130	15.217729768238321
var_RAI_0612	1.00	403	-8.188003093385852	89.116993996663350	4.439233670426515
	2.00	611	-16.595337344184780	158.858512629337240	6.426727516464044
var_RAI_0012	1.00	433	-33.310645901474310	484.013066297464830	23.260183562662660
	2.00	661	-34.454023068055356	360.824518105313700	14.034447185953853
var_CM_0006	1.00	436	-8.088060551248612	42.527843887179270	2.036714336446099
	2.00	663	-2.497123197708330	51.144157247500010	1.986274478896572
var_CM_0612	1.00	405	-3.437934242460017	42.679205065020334	2.120746750023084
	2.00	620	-12.070007506509590	46.852935828197210	1.881659183651821
var_CM_0012	1.00	436	-9.105352429332761	45.123156471870686	2.161007267045233
	2.00	663	-15.513784704053226	54.074188147128120	2.100067449814949
var_AM_0006	1.00	436	-3.696400727588327	40.068735671815446	1.918944412102161
	2.00	663	-2.239634212122359	44.402087820568750	1.724434199216981
var_AM_0612	1.00	405	-6.534314895422499	36.875207387620580	1.832343786736046
	2.00	618	-11.631067221175678	41.293806643468700	1.661080509217976
var_AM_0012	1.00	436	-5.194883900371313	43.779230316908010	2.096644877214261
	2.00	663	-11.540316392780143	53.016444336941820	2.058988084185426
var_ITI_0006	1.00	436	-3.937523871793443	45.187340397530310	2.164081119599506
	2.00	663	-.503026170239946	52.849428866516970	2.052501740791783
var_ITI_0612	1.00	405	-6.520677090564865	40.667634865288505	2.020791023176125
	2.00	619	-11.856133799762402	44.307390494717225	1.780864383456502
var_ITI_0012	1.00	436	-5.989876825184558	46.697411845098095	2.236400425851625
	2.00	663	-11.145459791530579	60.194755761060770	2.337770599151906
var_ITE_0006	1.00	436	2.740596285074456	52.489536671048320	2.513792896130481
	2.00	663	4.438324631981095	58.504763652190830	2.272136744259342
var_ITE_0612	1.00	405	-3.858452301084278	42.505625784094040	2.112121526208935
	2.00	619	-7.634497022920351	50.465917587188780	2.028396486590451
var_ITE_0012	1.00	436	2.580835571633476	59.853942787883900	2.866483983817908
	2.00	663	-3.256505571901374	71.199358401664090	2.765153951460378
var_IDE_0006	1.00	436	-4.636182510760393	40.533591984536950	1.941206991860943
	2.00	663	-2.777442176250289	44.981939169219630	1.746953759560998
var_IDE_0612	1.00	405	-6.956668914263860	34.470560164854070	1.712855905580224
	2.00	620	-10.821164721042923	42.313332011543880	1.699344307097045
var_IDE_0012	1.00	436	-6.522912429960385	43.034832683092760	2.060994696198027
	2.00	663	-12.755747355663518	50.120062636029550	1.946501939857644
var_ITO_0006	1.00	436	-5.952300552843960	48.228857690652280	2.309743380114852
	2.00	663	2.479884421890669	64.006423524851870	2.485803508651478

var_ITO_0612	1.00	405	-.794908996489389	53.363253262484040	2.651641373232261
	2.00	620	-11.816430484560506	48.560284385659770	1.950227951778207
var_ITO_0012	1.00	436	-6.119294056948041	52.224193699711670	2.501085272503235
	2.00	663	-11.775909269163469	61.854387996853400	2.402225374276197
var_EXT_0006	1.00	436	-2.519310485041731	60.319356040671520	2.888773236168422
	2.00	663	-1.211945307967288	65.216068747284870	2.532782236294024
var_EXT_0612	1.00	405	-1.327761092768596	52.829453476741314	2.625116648625693
	2.00	620	-3.284409934362495	66.506324580064460	2.670958269848557
var_EXT_0012	1.00	436	-5.241104341482108	60.633805201531030	2.903832619750270
	2.00	663	-11.866498254028512	69.917359026903810	2.715365221386580
var_AMO_0006	1.00	435	-.779913025936299	60.400978024772640	2.896005233670363
	2.00	663	.044217690318424	65.030664495010730	2.525581731790863
var_AMO_0612	1.00	405	-1.471602129009536	53.130944706708280	2.640097868288496
	2.00	619	-4.820385521006219	61.555076569520930	2.474107417738082
var_AMO_0012	1.00	435	-3.726769828449639	56.146545742497990	2.692020818869823
	2.00	663	-9.209072443402840	72.868290579888240	2.829969906420277
var_GLO_0006	1.00	437	-8.185192109595965	37.151277813265540	1.777186630165880
	2.00	659	-10.076725972717915	37.522420330361020	1.461665972943800
var_GLO_0612	1.00	404	-9.525672841552728	42.483152222601640	2.113615820942040
	2.00	613	-10.515258737715037	55.110101971521000	2.225876053980966
var_GLO_0012	1.00	437	-12.606102807338258	40.066182661619670	1.916625438993364
	2.00	659	-20.653698908716430	44.049110680593220	1.715909732191813
var_REL_0006	1.00	437	-6.422210972237493	43.133322715996480	2.063346645323941
	2.00	661	-7.439870998382808	44.804313187873190	1.742685808713845
var_REL_0612	1.00	404	-5.443002839491745	48.794617526609560	2.427622956052434
	2.00	615	-8.063613051157633	62.419949303994240	2.517015289840286
var_REL_0012	1.00	437	-9.093850253678776	44.611026649312400	2.134034810797885
	2.00	661	-17.936854104517376	50.672616745376870	1.970936363253193
var_SOS_0006	1.00	437	-6.060432292057198	48.397752911715640	2.315178493646880
	2.00	663	-4.685131189952244	59.936575340626580	2.327743702480649
var_SOS_0612	1.00	404	-5.345070941663153	54.405252930354350	2.706762500424179
	2.00	619	-6.580109323037367	72.205733551099440	2.902193465395684
var_SOS_0012	1.00	437	-9.600938187714513	49.594267966083756	2.372415570877530
	2.00	663	-16.306380268968640	54.952798440749580	2.134189846136211
var_COM_0006	1.00	437	-6.282911654072989	40.159316260635820	1.921080623222091
	2.00	661	-8.583271745996424	38.975018896604810	1.515952538780370
var_COM_0612	1.00	404	-8.694780355677858	46.935282621806905	2.335117587085724
	2.00	620	-10.458633982132627	58.242348179003620	2.339068990902249
var_COM_0012	1.00	437	-10.811745879512044	42.557220042286240	2.035787917075970
	2.00	661	-20.553273759743590	45.693480898134750	1.777270424345423
var_AUT_0006	1.00	437	-7.449724462587655	48.056543551109925	2.298856236393607
	2.00	661	-8.637849979521691	42.079856771987720	1.636716735769590
var_AUT_0612	1.00	404	-8.496677436864008	48.554614459830270	2.415682357191941
	2.00	619	-9.003941912869145	69.838913239054100	2.807062925124632
var_AUT_0012	1.00	437	-11.576904236316777	51.912373957845430	2.483305618766953

2.00	661	-20.110927371633510	47.408274337712680	1.843968158992025
------	-----	---------------------	--------------------	-------------------

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
var_RAI_000 6	Equal variances assumed	6.534	.011	.995	1092	.320	20.049346 982316024	20.15145641 3323945	- 19.490606 777206395	59.5893007 41838440
	Equal variances not assumed			1.136	1010.798	.256	20.049346 982316024	17.65463333 8882320	- 14.594581 507818702	54.6932754 72450750
var_RAI_061 2	Equal variances assumed	21.822	.000	.967	1012	.334	8.4073342 50798929	8.696720609 216438	- 8.6583352 60100468	25.4730037 61698330
	Equal variances not assumed			1.076	989.240	.282	8.4073342 50798929	7.810865646 746497	- 6.9204346 90788095	23.7351031 92385953
var_RAI_001 2	Equal variances assumed	.582	.446	.045	1092	.964	1.1433771 66581047	25.59334552 3745775	- 49.074318 160456960	51.3610724 93619055
	Equal variances not assumed			.042	739.633	.966	1.1433771 66581047	27.16618941 2283984	- 52.188647 783536176	54.4754021 16698270
var_CM_000 6	Equal variances assumed	2.826	.093	- 1.892	1097	.059	- 5.5909373 53540282	2.954297969 289975	- 11.387650 600607271	.205775893 526708
	Equal variances not assumed			- 1.965	1038.598	.050	- 5.5909373 53540282	2.844906253 956520	- 11.173356 678658926	- .008518028 421636
var_CM_061 2	Equal variances assumed	3.989	.046	2.986	1023	.003	8.6320732 64049573	2.891103669 626531	2.9589021 08233795	14.3052444 19865353
	Equal variances not assumed			3.045	918.816	.002	8.6320732 64049573	2.835173374 091029	3.0679059 97050824	14.1962405 31048325
var_CM_001 2	Equal variances assumed	15.350	.000	2.049	1097	.041	6.4084322 74720465	3.127001455 657241	.27285253 6308871	12.5440120 13132058
	Equal variances not assumed			2.127	1036.907	.034	6.4084322 74720465	3.013342944 637164	.49548668 6917482	12.3213778 62523448
var_AM_000 6	Equal variances assumed	2.290	.130	-.553	1097	.580	- 1.4567665 15465968	2.635095668 041113	- 6.6271637 21504690	3.71363069 0572753
	Equal variances not assumed			-.565	994.910	.572	- 1.4567665 15465968	2.579926542 397520	- 6.5194885 84685324	3.60595555 3753387

var_AM_061 2	Equal variances assumed	22.178	.000	2.013	1021	.044	5.0967523 25753179	2.531975087 824014	.12828249 3353531	10.0652221 58152828
	Equal variances not assumed			2.061	929.725	.040	5.0967523 25753179	2.473190694 405517	.24306900 3171513	9.95043564 8334846
var_AM_001 2	Equal variances assumed	31.284	.000	2.076	1097	.038	6.3454324 92408830	3.055835440 728000	.34948964 4307404	12.3413753 40510256
	Equal variances not assumed			2.159	1041.871	.031	6.3454324 92408830	2.938596888 306795	.57918979 0671481	12.1116751 94146180
var_ITI_000 6	Equal variances assumed	3.649	.056	- 1.115	1097	.265	- 3.4344977 01553497	3.080004208 880586	- 9.4778627 86670865	2.60886738 3563870
	Equal variances not assumed			- 1.152	1024.732	.250	- 3.4344977 01553497	2.982618059 383460	- 9.2872345 09681115	2.41823910 6574118
var_ITI_061 2	Equal variances assumed	23.476	.000	1.946	1022	.052	5.3354567 09197537	2.742143228 596601	- .04541775 1759203	10.7163311 70154277
	Equal variances not assumed			1.981	914.581	.048	5.3354567 09197537	2.693524514 760004	.04925001 5260828	10.6216634 03134246
var_ITI_001 2	Equal variances assumed	26.692	.000	1.514	1097	.130	5.1555829 66346021	3.405974937 838664	- 1.5273786 87008337	11.8385446 19700379
	Equal variances not assumed			1.594	1067.504	.111	5.1555829 66346021	3.235221513 128335	- 1.1925322 02682687	11.5036981 35374730
var_ITE_000 6	Equal variances assumed	4.940	.026	-.490	1097	.624	- 1.6977283 46906639	3.465044942 491357	- 8.4965929 59960374	5.10113626 6147098
	Equal variances not assumed			-.501	998.283	.616	- 1.6977283 46906639	3.388474599 174312	- 8.3470783 39330249	4.95162164 5516971
var_ITE_061 2	Equal variances assumed	22.351	.000	1.244	1022	.214	3.7760447 21836073	3.034439080 812526	- 2.1783983 51376272	9.73048779 5048417
	Equal variances not assumed			1.289	959.380	.198	3.7760447 21836073	2.928386867 933889	- 1.9707381 19657342	9.52282756 3329487
var_ITE_001 2	Equal variances assumed	17.051	.000	1.414	1097	.158	5.8373411 43534850	4.126920341 765301	- 2.2602082 78824402	13.9348905 65894102
	Equal variances not assumed			1.466	1033.306	.143	5.8373411 43534850	3.982813930 471963	- 1.9779850 27203202	13.6526673 14272902
var_IDE_000 6	Equal variances assumed	1.754	.186	-.697	1097	.486	- 1.8587403 34510104	2.668169394 541172	- 7.0940324 53156046	3.37655178 4135840

	Equal variances not assumed			-.712	995.750	.477	-	2.611538248	-	3.26600973
							1.8587403 34510104	483855	6.9834904 08633672	9613465
var_IDE_061 2	Equal variances assumed	24.390	.000	1.535	1023	.125	3.8644958 06779063	2.517493291 145678	-	8.80453668 9551124
	Equal variances not assumed			1.602	974.509	.110	3.8644958 06779063	2.412808825 278991	-	8.59939494 2310685
									8752559	
var_IDE_001 2	Equal variances assumed	27.795	.000	2.131	1097	.033	6.2328349 25703132	2.924952393 114110	.49370147 9972951	11.9719683 71433311
	Equal variances not assumed			2.199	1022.534	.028	6.2328349 25703132	2.834884290 341665	.66997924 4817998	11.7956906 06588266
var_ITO_000 6	Equal variances assumed	4.515	.034	- 2.347	1097	.019	- 8.4321849 74734628	3.592489222 970155	- 15.481111 685990300	- 1.38325826 3478957
	Equal variances not assumed			- 2.485	1076.918	.013	- 8.4321849 74734628	3.393248232 241252	- 15.090312 326050960	- 1.77405762 3418298
var_ITO_061 2	Equal variances assumed	.758	.384	3.415	1023	.001	11.021521 488071116	3.227231784 885535	4.6887709 65323664	17.3542720 10818570
	Equal variances not assumed			3.348	805.462	.001	11.021521 488071116	3.291593996 247729	4.5604069 66985812	17.4826360 09156420
var_ITO_001 2	Equal variances assumed	14.314	.000	1.576	1097	.115	5.6566152 12215426	3.590210446 986180	-	12.7010706 71393373
	Equal variances not assumed			1.631	1031.151	.103	5.6566152 12215426	3.467868839 669256	-	12.4615006 51632235
									27201382	
var_EXT_00 06	Equal variances assumed	2.716	.100	-.335	1097	.738	- 1.3073651 77074443	3.904248754 329090	-	6.35327390 1760069
	Equal variances not assumed			-.340	980.222	.734	- 1.3073651 77074443	3.841874108 620626	-	6.23187887 0160980
									24309866	
var_EXT_06 12	Equal variances assumed	12.177	.001	.498	1023	.618	1.9566488 41593900	3.927354459 981087	-	9.66324002 1669890
	Equal variances not assumed			.522	984.684	.601	1.9566488 41593900	3.745030773 994279	-	9.30580760 8998910
									25811110	
var_EXT_00 12	Equal variances assumed	8.959	.003	1.618	1097	.106	6.6253939 12546404	4.093661599 863013	-	14.6576853 98241904
	Equal variances not assumed			1.667	1017.252	.096	6.6253939 12546404	3.975607144 706517	-	14.4267228 59380417
									34287610	

var_AMO_006	Equal variances assumed	1.552	.213	-.211	1096	.833	-	3.901911170	-	6.83192943
							.82413071 6254724	570312	8.4801908 70134416	7624970
	Equal variances not assumed			-.214	975.333	.830	-	3.842578482	-	6.71654231
							.82413071 6254724	920273	8.3648037 44088002	1578555
var_AMO_0612	Equal variances assumed	13.040	.000	.898	1022	.370	3.3487833 91996683	3.730528909 610723	-	10.6691551 09362642
									25369276	
	Equal variances not assumed			.926	947.476	.355	3.3487833 91996683	3.618193509 010825	-	10.4493829 06728587
									22735222	
var_AMO_012	Equal variances assumed	23.482	.000	1.331	1096	.183	5.4823026 14953202	4.118588120 536894	-	13.5635112 89746266
									59839862	
	Equal variances not assumed			1.404	1068.096	.161	5.4823026 14953202	3.905855317 401419	-	13.1463230 53674980
									23768576	
var_GLO_006	Equal variances assumed	.292	.589	.820	1094	.412	1.8915338 63121950	2.305698126 439961	-	6.41562435 5753361
									29509461	
	Equal variances not assumed			.822	940.273	.411	1.8915338 63121950	2.301056221 586536	-	6.40733400 8146921
									81903021	
var_GLO_0612	Equal variances assumed	6.331	.012	.306	1015	.760	.98958589 6162309	3.234639882 546843	-	7.33693248 1231476
									88906859	
	Equal variances not assumed			.322	990.406	.747	.98958589 6162309	3.069510717 723978	-	7.01307744 6298569
									53973951	
var_GLO_0012	Equal variances assumed	14.383	.000	3.069	1094	.002	8.0475961 01378170	2.622269848 800170	2.9023492	13.1928429 81718112
									21038229	
	Equal variances not assumed			3.128	992.523	.002	8.0475961 01378170	2.572508363 917809	2.9994163	13.0957758 80047187
									22709156	
var_REL_006	Equal variances assumed	.031	.860	.374	1096	.709	1.0176600 26145316	2.721838911 094070	-	6.35826403 3145502
									80854871	
	Equal variances not assumed			.377	957.893	.706	1.0176600 26145316	2.700806029 070246	-	6.31783957 5923748
									23633117	
var_REL_0612	Equal variances assumed	7.089	.008	.713	1017	.476	2.6206102 11665888	3.676527535 456142	-	9.83505774 0934670
									17602893	
	Equal variances not assumed			.749	986.733	.454	2.6206102 11665888	3.496958562 242700	-	9.48294046 5145287
									41813511	

var_REL_00 12	Equal variances assumed	10.318	.001	2.966	1096	.003	8.8430038 50838600	2.981105837 309747	2.9936842 18021493	14.6923234 83655710
	Equal variances not assumed			3.044	1011.065	.002	8.8430038 50838600	2.904943152 918983	3.1425959 74027825	14.5434117 27649378
var_SOS_00 06	Equal variances assumed	.253	.615	-.401	1098	.688	- 1.3753011 02104953	3.428470806 778842	- 8.1023957 81661990	5.35179357 7452082
	Equal variances not assumed			-.419	1053.800	.675	- 1.3753011 02104953	3.283053792 109315	- 7.8173673 22421650	5.06676511 8211744
var_SOS_06 12	Equal variances assumed	4.626	.032	.294	1021	.769	1.2350383 81374214	4.205806960 769199	- 7.0179752 86085155	9.48805204 8833582
	Equal variances not assumed			.311	1000.199	.756	1.2350383 81374214	3.968537532 175797	- 6.5525760 32806702	9.02265279 5555128
var_SOS_00 12	Equal variances assumed	6.954	.008	2.058	1098	.040	6.7054420 81254125	3.258912311 671278	.31104268 0319997	13.0998414 82188253
	Equal variances not assumed			2.101	997.124	.036	6.7054420 81254125	3.191100427 798075	.44339911 4226668	12.9674850 48281581
var_COM_0 006	Equal variances assumed	.039	.843	.946	1096	.344	2.3003600 91923435	2.432266551 928769	- 2.4720650 65339801	7.07278524 9186671
	Equal variances not assumed			.940	913.946	.347	2.3003600 91923435	2.447174464 715180	- 2.5023739 72540419	7.10309415 6387288
var_COM_0 612	Equal variances assumed	8.122	.004	.510	1022	.610	1.7638536 26454768	3.456958870 219235	- 5.0196949 14465384	8.54740216 7374920
	Equal variances not assumed			.534	977.044	.594	1.7638536 26454768	3.305150206 831381	- 4.7221564 58454363	8.24986371 1363900
var_COM_0 012	Equal variances assumed	12.164	.001	3.553	1096	.000	9.7415278 80231546	2.741888204 622352	4.3615845 36535777	15.1214712 23927315
	Equal variances not assumed			3.605	978.412	.000	9.7415278 80231546	2.702429019 338987	4.4383040 12829186	15.0447517 47633907
var_AUT_00 06	Equal variances assumed	.066	.797	.433	1096	.665	1.1881255 16934036	2.746896701 524323	- 4.2016451 52881103	6.57789618 6749176
	Equal variances not assumed			.421	846.381	.674	1.1881255 16934036	2.821981868 962669	- 4.3507780 05852595	6.72702903 9720668
var_AUT_06 12	Equal variances assumed	5.376	.021	.127	1021	.899	.50726447 6005137	3.985438594 404468	- 7.3133225 11759166	8.32785146 3769440

	Equal variances not assumed			.137	1016.970	.891	.50726447 6005137	3.703393513 584219	- 6.7599024 00236931	7.77443135 2247205
var_AUT_00 12	Equal variances assumed	7.498	.006	2.811	1096	.005	8.5340231 35316735	3.036412714 528233	2.5761841 74304342	14.4918620 96329127
	Equal variances not assumed			2.759	873.850	.006	8.5340231 35316735	3.093060841 234126	2.4633270 00682436	14.6047192 69951033

T-Test

km_final = 2 (FILTER)

		Paired Samples Statistics			
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BRE00RAI	14.5939	602	10.92644	.44533
	BRE06RAI	15.5104	602	10.46771	.42663
Pair 2	BRE06RAI	15.4605	531	10.45811	.45384
	BRE12RAI	15.5461	531	10.95092	.47523
Pair 3	BRE00RAI	14.4082	534	11.13488	.48185
	BRE12RAI	15.7308	534	10.89105	.47130
Pair 4	BRE00CM	2.87169	605	.969975	.039435
	BRE06CM	2.87169	605	.967224	.039323
Pair 5	BRE06CM	2.84902	534	.956027	.041371
	BRE12CM	2.80595	534	1.021922	.044223
Pair 6	BRE00CM	2.84725	536	.980156	.042336
	BRE12CM	2.81157	536	1.028422	.044421
Pair 7	BRE00AM	4.55845	603	1.371083	.055834
	BRE06AM	4.69596	603	1.333460	.054302
Pair 8	BRE06AM	4.68289	534	1.329780	.057545
	BRE12AM	4.70177	534	1.394892	.060362
Pair 9	BRE00AM	4.53258	537	1.391735	.060057
	BRE12AM	4.72454	537	1.396373	.060257
Pair 10	BRE00ITI	4.6714	604	1.54440	.06284
	BRE06ITI	4.7632	604	1.45390	.05916
Pair 11	BRE06ITI	4.7463	535	1.45516	.06291
	BRE12ITI	4.7248	535	1.58891	.06869
Pair 12	BRE00ITI	4.6315	538	1.55344	.06697
	BRE12ITI	4.7495	538	1.58595	.06838
Pair 13	BRE00ITE	4.0571	604	1.66151	.06761
	BRE06ITE	4.2541	604	1.64343	.06687
Pair 14	BRE06ITE	4.2416	536	1.63223	.07050
	BRE12ITE	4.3102	536	1.66632	.07197
Pair 15	BRE00ITE	4.0367	538	1.68150	.07249
	BRE12ITE	4.3374	538	1.67529	.07223
Pair 16	BRE00IDE	4.9409	605	1.27902	.05200
	BRE06IDE	5.0669	605	1.26656	.05149

Pair 17	BRE06IDE	5.0497	538	1.26346	.05447
	BRE12IDE	5.0548	538	1.27567	.05500
Pair 18	BRE00IDE	4.9296	540	1.28712	.05539
	BRE12IDE	5.0782	540	1.27020	.05466
Pair 19	BRE00ITO	3.8269	605	1.42750	.05804
	BRE06ITO	3.9281	605	1.48126	.06022
Pair 20	BRE06ITO	3.9065	535	1.48971	.06441
	BRE12ITO	3.7907	535	1.49785	.06476
Pair 21	BRE00ITO	3.7854	537	1.43887	.06209
	BRE12ITO	3.8026	537	1.50633	.06500
Pair 22	BRE00EXT	1.9165	605	1.14522	.04656
	BRE06EXT	1.8153	605	1.06479	.04329
Pair 23	BRE06EXT	1.7934	536	1.04551	.04516
	BRE12EXT	1.8232	536	1.11033	.04796
Pair 24	BRE00EXT	1.9154	538	1.14501	.04936
	BRE12EXT	1.8225	538	1.10646	.04770
Pair 25	BRE00AMO	1.6175	604	.93044	.03786
	BRE06AMO	1.6035	604	.95246	.03876
Pair 26	BRE06AMO	1.6156	534	.95837	.04147
	BRE12AMO	1.6320	534	.93749	.04057
Pair 27	BRE00AMO	1.6341	537	.95415	.04117
	BRE12AMO	1.6243	537	.91603	.03953
Pair 28	BPN00GLO	5.5400	595	.90616	.03714
	BPN06GLO	5.4229	595	1.10828	.04543
Pair 29	BPN06GLO	5.4213	534	1.12898	.04885
	BPN12GLO	5.3018	534	1.18995	.05149
Pair 30	BPN00GLO	5.5338	537	.90725	.03915
	BPN12GLO	5.2985	537	1.19321	.05149
Pair 31	BPN00REL	5.1999	599	1.27360	.05204
	BPN06REL	5.1227	599	1.43365	.05858
Pair 32	BPN06REL	5.1164	535	1.44666	.06254
	BPN12REL	5.0430	535	1.49977	.06484
Pair 33	BPN00REL	5.1870	540	1.27738	.05497
	BPN12REL	5.0296	540	1.50565	.06479
Pair 34	BPN00SOS	5.8700	604	1.13637	.04624
	BPN06SOS	5.7968	604	1.26790	.05159
Pair 35	BPN06SOS	5.8072	538	1.26875	.05470
	BPN12SOS	5.7500	538	1.30302	.05618
Pair 36	BPN00SOS	5.8427	542	1.13752	.04886
	BPN12SOS	5.7426	542	1.30527	.05607
Pair 37	BPN00COM	5.3416	603	1.02866	.04189
	BPN06COM	5.2156	603	1.20705	.04915
Pair 38	BPN06COM	5.2161	538	1.22489	.05281
	BPN12COM	5.0878	538	1.31133	.05654
Pair 39	BPN00COM	5.3349	539	1.02812	.04428

	BPN12COM	5.0724	539	1.32036	.05687
Pair 40	BPN00AUT	6.0681	602	.93270	.03801
	BPN06AUT	5.9174	602	1.10706	.04512
Pair 41	BPN06AUT	5.9336	538	1.12556	.04853
	BPN12AUT	5.7742	538	1.24477	.05367
Pair 42	BPN00AUT	6.0699	540	.92737	.03991
	BPN12AUT	5.7755	540	1.24145	.05342

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BRE00RAI & BRE06RAI	602	.740	.000
Pair 2	BRE06RAI & BRE12RAI	531	.803	.000
Pair 3	BRE00RAI & BRE12RAI	534	.717	.000
Pair 4	BRE00CM & BRE06CM	605	.549	.000
Pair 5	BRE06CM & BRE12CM	534	.679	.000
Pair 6	BRE00CM & BRE12CM	536	.568	.000
Pair 7	BRE00AM & BRE06AM	603	.770	.000
Pair 8	BRE06AM & BRE12AM	534	.814	.000
Pair 9	BRE00AM & BRE12AM	537	.736	.000
Pair 10	BRE00ITI & BRE06ITI	604	.737	.000
Pair 11	BRE06ITI & BRE12ITI	535	.764	.000
Pair 12	BRE00ITI & BRE12ITI	538	.705	.000
Pair 13	BRE00ITE & BRE06ITE	604	.748	.000
Pair 14	BRE06ITE & BRE12ITE	536	.800	.000
Pair 15	BRE00ITE & BRE12ITE	538	.719	.000
Pair 16	BRE00IDE & BRE06IDE	605	.687	.000
Pair 17	BRE06IDE & BRE12IDE	538	.724	.000
Pair 18	BRE00IDE & BRE12IDE	540	.665	.000
Pair 19	BRE00ITO & BRE06ITO	605	.594	.000
Pair 20	BRE06ITO & BRE12ITO	535	.706	.000
Pair 21	BRE00ITO & BRE12ITO	537	.612	.000
Pair 22	BRE00EXT & BRE06EXT	605	.532	.000
Pair 23	BRE06EXT & BRE12EXT	536	.596	.000
Pair 24	BRE00EXT & BRE12EXT	538	.509	.000
Pair 25	BRE00AMO & BRE06AMO	604	.524	.000
Pair 26	BRE06AMO & BRE12AMO	534	.575	.000
Pair 27	BRE00AMO & BRE12AMO	537	.456	.000
Pair 28	BPN00GLO & BPN06GLO	595	.477	.000
Pair 29	BPN06GLO & BPN12GLO	534	.456	.000
Pair 30	BPN00GLO & BPN12GLO	537	.420	.000
Pair 31	BPN00REL & BPN06REL	599	.552	.000
Pair 32	BPN06REL & BPN12REL	535	.562	.000
Pair 33	BPN00REL & BPN12REL	540	.525	.000
Pair 34	BPN00SOS & BPN06SOS	604	.374	.000

Pair 35	BPN06SOS & BPN12SOS	538	.381	.000
Pair 36	BPN00SOS & BPN12SOS	542	.475	.000
Pair 37	BPN00COM & BPN06COM	603	.450	.000
Pair 38	BPN06COM & BPN12COM	538	.469	.000
Pair 39	BPN00COM & BPN12COM	539	.403	.000
Pair 40	BPN00AUT & BPN06AUT	602	.346	.000
Pair 41	BPN06AUT & BPN12AUT	538	.311	.000
Pair 42	BPN00AUT & BPN12AUT	540	.280	.000

		Paired Differences					t	df
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
					Lower	Upper		
Pair 1	BRE00RAI - BRE06RAI	-.91653	7.72311	.31477	-1.53471	-.29834	-2.912	601
Pair 2	BRE06RAI - BRE12RAI	-.08569	6.74084	.29253	-.66034	.48897	-.293	530
Pair 3	BRE00RAI - BRE12RAI	-1.32257	8.28371	.35847	-2.02676	-.61838	-3.689	533
Pair 4	BRE00CM - BRE06CM	.000000	.920359	.037418	-.073485	.073485	.000	604
Pair 5	BRE06CM - BRE12CM	.043071	.795165	.034410	-.024525	.110667	1.252	533
Pair 6	BRE00CM - BRE12CM	.035681	.934404	.040360	-.043603	.114965	.884	535
Pair 7	BRE00AM - BRE06AM	-.137506	.918069	.037386	-.210931	-.064082	-3.678	602
Pair 8	BRE06AM - BRE12AM	-.018882	.834158	.03609	-.089793	.052028	-.523	533
Pair 9	BRE00AM - BRE12AM	-.191961	1.01269	.04370	-.277807	-.106115	-4.393	536
Pair 10	BRE00ITI - BRE06ITI	-.09189	1.09112	.04440	-.17908	-.00470	-2.070	603
Pair 11	BRE06ITI - BRE12ITI	.02150	1.05291	.04552	-.06793	.11092	.472	534
Pair 12	BRE00ITI - BRE12ITI	-.11803	1.20685	.05203	-.22024	-.01582	-2.268	537
Pair 13	BRE00ITE - BRE06ITE	-.19702	1.17273	.04772	-.29073	-.10331	-4.129	603
Pair 14	BRE06ITE - BRE12ITE	-.06856	1.04315	.04506	-.15707	.01995	-1.522	535
Pair 15	BRE00ITE - BRE12ITE	-.30065	1.25937	.05430	-.40731	-.19399	-5.537	537
Pair 16	BRE00IDE - BRE06IDE	-.12603	1.00790	.04098	-.20651	-.04556	-3.076	604
Pair 17	BRE06IDE - BRE12IDE	-.00511	.94339	.04067	-.08501	.07479	-.126	537
Pair 18	BRE00IDE - BRE12IDE	-.14861	1.04736	.04507	-.23715	-.06007	-3.297	539
Pair 19	BRE00ITO - BRE06ITO	-.10124	1.31069	.05329	-.20589	.00341	-1.900	604
Pair 20	BRE06ITO - BRE12ITO	.11589	1.14467	.04949	.01867	.21310	2.342	534
Pair 21	BRE00ITO - BRE12ITO	-.01723	1.29915	.05606	-.12735	.09290	-.307	536
Pair 22	BRE00EXT - BRE06EXT	.10124	1.07094	.04354	.01573	.18675	2.325	604

Pair 23	BRE06EXT - BRE12EXT	-.02985	.97026	.04191	-.11218	.05248	-.712	535
Pair 24	BRE00EXT - BRE12EXT	.09294	1.11614	.04812	-.00159	.18746	1.931	537
Pair 25	BRE00AMO - BRE06AMO	.01407	.91876	.03738	-.05935	.08749	.376	603
Pair 26	BRE06AMO - BRE12AMO	-.01639	.87456	.03785	-.09073	.05796	-.433	533
Pair 27	BRE00AMO - BRE12AMO	.00978	.97569	.04210	-.07293	.09249	.232	536
Pair 28	BPN00GLO - BPN06GLO	.117086	1.04421	.042808	.033012	.201161	2.735	594
Pair 29	BPN06GLO - BPN12GLO	.119538	1.21058	.052387	.016627	.222448	2.282	533
Pair 30	BPN00GLO - BPN12GLO	.235257	1.15648	.049905	.137222	.333292	4.714	536
Pair 31	BPN00REL - BPN06REL	.07721	1.28853	.05265	-.02619	.18061	1.467	598
Pair 32	BPN06REL - BPN12REL	.07336	1.37973	.05965	-.04382	.19054	1.230	534
Pair 33	BPN00REL - BPN12REL	.15741	1.37023	.05897	.04158	.27324	2.669	539
Pair 34	BPN00SOS - BPN06SOS	.07326	1.34914	.05490	-.03455	.18107	1.335	603
Pair 35	BPN06SOS - BPN12SOS	.05716	1.43110	.06170	-.06405	.17836	.926	537
Pair 36	BPN00SOS - BPN12SOS	.10009	1.25963	.05411	-.00619	.20638	1.850	541
Pair 37	BPN00COM - BPN06COM	.12604	1.18225	.04814	.03148	.22059	2.618	602
Pair 38	BPN06COM - BPN12COM	.12825	1.30883	.05643	.01741	.23910	2.273	537
Pair 39	BPN00COM - BPN12COM	.26252	1.30642	.05627	.15198	.37306	4.665	538
Pair 40	BPN00AUT - BPN06AUT	.15075	1.17500	.04789	.05670	.24480	3.148	601
Pair 41	BPN06AUT - BPN12AUT	.15939	1.39442	.06012	.04129	.27748	2.651	537
Pair 42	BPN00AUT - BPN12AUT	.29444	1.32572	.05705	.18238	.40651	5.161	539

T-Test

km_final = 1 (FILTER)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BRE00RAI	20.4688	392	9.38408	.47397
	BRE06RAI	21.9962	392	8.82694	.44583
Pair 2	BRE06RAI	21.9926	373	8.83705	.45757
	BRE12RAI	21.8351	373	9.21165	.47696
Pair 3	BRE00RAI	20.5706	386	9.46707	.48186
	BRE12RAI	21.8944	386	9.16382	.46643
Pair 4	BRE00CM	2.87117	392	.923410	.046639
	BRE06CM	2.83291	392	.958697	.048422
Pair 5	BRE06CM	2.80965	373	.943130	.048833
	BRE12CM	2.83009	373	.959440	.049678
Pair 6	BRE00CM	2.85881	386	.939646	.047827
	BRE12CM	2.81768	386	.964967	.049116
Pair 7	BRE00AM	5.35607	392	1.224685	.061855
	BRE06AM	5.59056	392	1.155169	.058344
Pair 8	BRE06AM	5.58288	374	1.162221	.060097
	BRE12AM	5.56974	374	1.225968	.063393
Pair 9	BRE00AM	5.36412	387	1.232200	.062636
	BRE12AM	5.57149	387	1.226206	.062331
Pair 10	BRE00ITI	5.4719	392	1.26197	.06374
	BRE06ITI	5.6429	392	1.22514	.06188
Pair 11	BRE06ITI	5.6364	374	1.23501	.06386
	BRE12ITI	5.5889	374	1.30923	.06770
Pair 12	BRE00ITI	5.4612	387	1.29045	.06560
	BRE12ITI	5.5833	387	1.29454	.06581
Pair 13	BRE00ITE	4.9560	392	1.60685	.08116
	BRE06ITE	5.3151	392	1.53421	.07749
Pair 14	BRE06ITE	5.3088	374	1.55006	.08015
	BRE12ITE	5.3469	374	1.56113	.08072
Pair 15	BRE00ITE	4.9851	387	1.62326	.08252
	BRE12ITE	5.3534	387	1.56249	.07943
Pair 16	BRE00IDE	5.6403	392	1.12889	.05702
	BRE06IDE	5.8138	392	.99250	.05013
Pair 17	BRE06IDE	5.8035	374	.99672	.05154
	BRE12IDE	5.7734	374	1.08423	.05606
Pair 18	BRE00IDE	5.6460	387	1.11905	.05688
	BRE12IDE	5.7778	387	1.09463	.05564
Pair 19	BRE00ITO	4.0969	392	1.43392	.07242
	BRE06ITO	4.0574	392	1.46000	.07374
Pair 20	BRE06ITO	4.0294	374	1.45905	.07545
	BRE12ITO	4.0635	374	1.47334	.07618
Pair 21	BRE00ITO	4.0665	387	1.43600	.07300
	BRE12ITO	4.0517	387	1.49792	.07614
Pair 22	BRE00EXT	1.6454	392	.94438	.04770
	BRE06EXT	1.6084	392	.89384	.04515

Pair 23	BRE06EXT	1.5979	373	.86591	.04484
	BRE12EXT	1.6032	373	.90978	.04711
Pair 24	BRE00EXT	1.6509	386	.96095	.04891
	BRE12EXT	1.5900	386	.89795	.04570
Pair 25	BRE00AMO	1.3705	392	.77214	.03900
	BRE06AMO	1.3673	392	.69636	.03517
Pair 26	BRE06AMO	1.3650	374	.67098	.03470
	BRE12AMO	1.3817	374	.78581	.04063
Pair 27	BRE00AMO	1.3630	387	.78377	.03984
	BRE12AMO	1.3747	387	.79187	.04025
Pair 28	BPN00GLO	5.6501	393	.90182	.04549
	BPN06GLO	5.6666	393	.96958	.04890
Pair 29	BPN06GLO	5.7098	367	.95506	.04985
	BPN12GLO	5.5660	367	1.15282	.06017
Pair 30	BPN00GLO	5.6605	382	.89734	.04591
	BPN12GLO	5.5684	382	1.13399	.05802
Pair 31	BPN00REL	5.2958	393	1.29405	.06528
	BPN06REL	5.3219	393	1.34679	.06794
Pair 32	BPN06REL	5.3689	370	1.32647	.06896
	BPN12REL	5.3074	370	1.45387	.07558
Pair 33	BPN00REL	5.2987	385	1.29253	.06587
	BPN12REL	5.2844	385	1.46051	.07443
Pair 34	BPN00SOS	5.8594	393	1.16666	.05885
	BPN06SOS	5.8817	393	1.19603	.06033
Pair 35	BPN06SOS	5.8896	369	1.19435	.06218
	BPN12SOS	5.8069	369	1.29070	.06719
Pair 36	BPN00SOS	5.8522	384	1.17113	.05976
	BPN12SOS	5.7852	384	1.29195	.06593
Pair 37	BPN00COM	5.4695	393	1.02846	.05188
	BPN06COM	5.5541	393	1.09284	.05513
Pair 38	BPN06COM	5.5894	369	1.09383	.05694
	BPN12COM	5.4255	369	1.27097	.06616
Pair 39	BPN00COM	5.4954	384	1.02553	.05233
	BPN12COM	5.4414	384	1.24790	.06368
Pair 40	BPN00AUT	6.1851	393	.91695	.04625
	BPN06AUT	6.1240	393	.95216	.04803
Pair 41	BPN06AUT	6.1692	368	.92539	.04824
	BPN12AUT	5.9776	368	1.15740	.06033
Pair 42	BPN00AUT	6.1880	383	.90947	.04647
	BPN12AUT	5.9915	383	1.13348	.05792

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BRE00RAI & BRE06RAI	392	.704	.000

Pair 2	BRE06RAI & BRE12RAI	373	.809	.000
Pair 3	BRE00RAI & BRE12RAI	386	.713	.000
Pair 4	BRE00CM & BRE06CM	392	.652	.000
Pair 5	BRE06CM & BRE12CM	373	.698	.000
Pair 6	BRE00CM & BRE12CM	386	.625	.000
Pair 7	BRE00AM & BRE06AM	392	.770	.000
Pair 8	BRE06AM & BRE12AM	374	.804	.000
Pair 9	BRE00AM & BRE12AM	387	.741	.000
Pair 10	BRE00ITI & BRE06ITI	392	.707	.000
Pair 11	BRE06ITI & BRE12ITI	374	.752	.000
Pair 12	BRE00ITI & BRE12ITI	387	.702	.000
Pair 13	BRE00ITE & BRE06ITE	392	.765	.000
Pair 14	BRE06ITE & BRE12ITE	374	.810	.000
Pair 15	BRE00ITE & BRE12ITE	387	.746	.000
Pair 16	BRE00IDE & BRE06IDE	392	.680	.000
Pair 17	BRE06IDE & BRE12IDE	374	.714	.000
Pair 18	BRE00IDE & BRE12IDE	387	.633	.000
Pair 19	BRE00ITO & BRE06ITO	392	.681	.000
Pair 20	BRE06ITO & BRE12ITO	374	.698	.000
Pair 21	BRE00ITO & BRE12ITO	387	.641	.000
Pair 22	BRE00EXT & BRE06EXT	392	.475	.000
Pair 23	BRE06EXT & BRE12EXT	373	.618	.000
Pair 24	BRE00EXT & BRE12EXT	386	.447	.000
Pair 25	BRE00AMO & BRE06AMO	392	.358	.000
Pair 26	BRE06AMO & BRE12AMO	374	.556	.000
Pair 27	BRE00AMO & BRE12AMO	387	.553	.000
Pair 28	BPN00GLO & BPN06GLO	393	.523	.000
Pair 29	BPN06GLO & BPN12GLO	367	.545	.000
Pair 30	BPN00GLO & BPN12GLO	382	.461	.000
Pair 31	BPN00REL & BPN06REL	393	.607	.000
Pair 32	BPN06REL & BPN12REL	370	.633	.000
Pair 33	BPN00REL & BPN12REL	385	.547	.000
Pair 34	BPN00SOS & BPN06SOS	393	.564	.000
Pair 35	BPN06SOS & BPN12SOS	369	.521	.000
Pair 36	BPN00SOS & BPN12SOS	384	.516	.000
Pair 37	BPN00COM & BPN06COM	393	.495	.000
Pair 38	BPN06COM & BPN12COM	369	.553	.000
Pair 39	BPN00COM & BPN12COM	384	.430	.000
Pair 40	BPN00AUT & BPN06AUT	393	.324	.000
Pair 41	BPN06AUT & BPN12AUT	368	.313	.000
Pair 42	BPN00AUT & BPN12AUT	383	.229	.000

Paired Differences

t

df

		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
					Lower	Upper		
Pair 1	BRE00RAI - BRE06RAI	-1.52742	7.02785	.35496	-2.22529	-.82955	-4.303	391
Pair 2	BRE06RAI - BRE12RAI	.15751	5.58646	.28926	-.41127	.72629	.545	372
Pair 3	BRE00RAI - BRE12RAI	-1.32383	7.06827	.35977	-2.03119	-.61648	-3.680	385
Pair 4	BRE00CM - BRE06CM	.038265	.785734	.039686	-.039758	.116289	.964	391
Pair 5	BRE06CM - BRE12CM	-.020442	.739255	.038277	-.095709	.054824	-.534	372
Pair 6	BRE00CM - BRE12CM	.041127	.824523	.041967	-.041386	.123640	.980	385
Pair 7	BRE00AM - BRE06AM	-.234481	.809060	.040863	-.314821	-.154141	-5.738	391
Pair 8	BRE06AM - BRE12AM	.013146	.749685	.038765	-.063079	.089372	.339	373
Pair 9	BRE00AM - BRE12AM	-.207364	.884795	.044976	-.295794	-.118934	-4.610	386
Pair 10	BRE00ITI - BRE06ITI	-.17092	.95241	.04810	-.26549	-.07634	-3.553	391
Pair 11	BRE06ITI - BRE12ITI	.04746	.89790	.04643	-.04384	.13876	1.022	373
Pair 12	BRE00ITI - BRE12ITI	-.12209	.99795	.05073	-.22183	-.02235	-2.407	386
Pair 13	BRE00ITE - BRE06ITE	-.35906	1.07773	.05443	-.46608	-.25204	-6.596	391
Pair 14	BRE06ITE - BRE12ITE	-.03810	.95862	.04957	-.13557	.05937	-.769	373
Pair 15	BRE00ITE - BRE12ITE	-.36822	1.13660	.05778	-.48181	-.25462	-6.373	386
Pair 16	BRE00IDE - BRE06IDE	-.17347	.85817	.04334	-.25869	-.08825	-4.002	391
Pair 17	BRE06IDE - BRE12IDE	.03008	.79137	.04092	-.05038	.11054	.735	373
Pair 18	BRE00IDE - BRE12IDE	-.13178	.94818	.04820	-.22655	-.03702	-2.734	386
Pair 19	BRE00ITO - BRE06ITO	.03954	1.15642	.05841	-.07529	.15437	.677	391
Pair 20	BRE06ITO - BRE12ITO	-.03409	1.14030	.05896	-.15003	.08185	-.578	373
Pair 21	BRE00ITO - BRE12ITO	.01486	1.24491	.06328	-.10956	.13928	.235	386
Pair 22	BRE00EXT - BRE06EXT	.03699	.94285	.04762	-.05664	.13062	.777	391
Pair 23	BRE06EXT - BRE12EXT	-.00536	.77705	.04023	-.08448	.07375	-.133	372
Pair 24	BRE00EXT - BRE12EXT	.06088	.97893	.04983	-.03708	.15885	1.222	385
Pair 25	BRE00AMO - BRE06AMO	.00319	.83453	.04215	-.07968	.08606	.076	391
Pair 26	BRE06AMO - BRE12AMO	-.01671	.69411	.03589	-.08729	.05386	-.466	373
Pair 27	BRE00AMO - BRE12AMO	-.01163	.74493	.03787	-.08608	.06282	-.307	386

Pair 28	BPN00GLO - BPN06GLO	-.01653	.91549	.04618	-.107331	.07425	-.358	392
Pair 29	BPN06GLO - BPN12GLO	.14373	1.01985	.05323	.039046	.24841	2.700	366
Pair 30	BPN00GLO - BPN12GLO	.092059	1.07365	.05493	-.015950	.200069	1.676	381
Pair 31	BPN00REL - BPN06REL	-.02608	1.17088	.05906	-.14220	.09004	-.442	392
Pair 32	BPN06REL - BPN12REL	.06149	1.19582	.06217	-.06076	.18373	.989	369
Pair 33	BPN00REL - BPN12REL	.01429	1.31923	.06723	-.11791	.14648	.212	384
Pair 34	BPN00SOS - BPN06SOS	-.02226	1.10324	.05565	-.13168	.08715	-.400	392
Pair 35	BPN06SOS - BPN12SOS	.08266	1.21902	.06346	-.04213	.20745	1.302	368
Pair 36	BPN00SOS - BPN12SOS	.06706	1.21655	.06208	-.05501	.18912	1.080	383
Pair 37	BPN00COM - BPN06COM	-.08461	1.06755	.05385	-.19048	.02127	-1.571	392
Pair 38	BPN06COM - BPN12COM	.16396	1.12871	.05876	.04841	.27950	2.790	368
Pair 39	BPN00COM - BPN12COM	.05404	1.22867	.06270	-.06924	.17732	.862	383
Pair 40	BPN00AUT - BPN06AUT	.06107	1.08713	.05484	-.04675	.16888	1.114	392
Pair 41	BPN06AUT - BPN12AUT	.19158	1.23512	.06439	.06497	.31819	2.975	367
Pair 42	BPN00AUT - BPN12AUT	.19648	1.28029	.06542	.06785	.32510	3.003	382