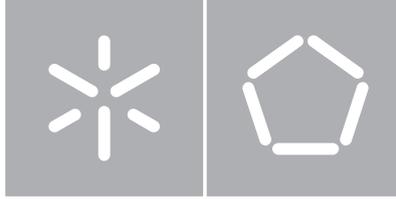


Universidade do Minho
Escola de Engenharia

Adriana Pereira

**Monitorização de saúde física e
psicológica utilizando dispositivos móvel**



Universidade do Minho

Escola de Engenharia

Departamento de Informática

Adriana Pereira

**Monitorização de saúde física e
psicológica utilizando dispositivos móvel**

Dissertação de Mestrado

Mestrado Integrado em Engenharia Informática

Trabalho realizado sob orientação de

Professor Cesar Analide Rodrigues

Supervisor Bruno Fernandes

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



Atribuição

CC BY

<https://creativecommons.org/licenses/by/4.0/>

Agradecimentos

Antes de mais, quero expressar os meus agradecimentos a todos aqueles que, de alguma forma, permitiram que esta dissertação se concretizasse.

Quero agradecer ao meu coordenador, o Professor Cesar Analide Rodrigues e ao meu supervisor Bruno Fernandes, por todo apoio durante esta jornada. Por estarem sempre disponíveis e me guiarem neste projeto.

Quero agradecer aos meus amigos de Engenharia Informática, pelo continuo apoio e inspiração. E aos meus amigos que me acompanham ao longo dos anos, pela ajuda, pela motivação e por demonstrarem estar continuamente disponíveis para fornecer as suas opiniões e companhia.

À minha família, sem eles nada seria possível, em especial à minha mãe pela paciência e por sempre acreditar no meu sucesso. E aos meus irmãos, principalmente ao João Carlos, por me acompanhar e pelo cuidado. E o meu avô pelo constante carinho.

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

Resumo

Monitorização de saúde física e psicológica utilizando dispositivos móveis.

Na última década, tem sido cada vez mais frequente a utilização de testes de personalidade, tais como o *OCEAN/BIG FIVE*, desenvolvido por *Goldenberg* em 1992, para avaliar aquilo que os psicólogos consideram ser as cinco principais dimensões da personalidade (Extroversão, Agradabilidade, Abertura, Consciencioso e Estabilidade). Assim sendo, o propósito fundamental desta dissertação é o de conceber e desenvolver uma aplicação móvel focada no tratamento de vários dados biométricos do utilizador assim como na sua personalidade. Para isso, será implementado na aplicação o teste de personalidade, baseado no modelo de *Saucier*, que permite ao utilizador escolher de entre quarenta adjetivos, os que melhor se adequam a si. Aquando da realização do teste, que se repetirá ao longo do tempo, será tida em consideração a informação biométrica do utilizador e o seu histórico. O objetivo é o de oferecer, ao utilizador, a possibilidade de conhecer os traços que compõe a sua personalidade ao mesmo tempo que lhe é fornecido informação biométrica, sendo possível analisar relações e correlações entre as cinco dimensões fundamentais da personalidade e os dados biométricos recolhidos, assim como a evolução de cada um destes atributos ao longo do tempo.

Palavras-chave: Aplicações de Saúde, *Big Five Mini-Markers*, Dados Biométricos, *Machine Learning*, Teste Psicológico.

Abstract

Physical and psychological health monitoring using mobile devices.

In the last decade, there has been an increased use of personality tests, such as the *OCEAN/BIG FIVE*, developed by *Goldenberg* in 1992, to evaluate what psychologists consider to be the five traits, or dimensions, of personality (Extroversion, Agreeableness, Openness, Conscientiousness and Neuroticism). Therefore, the main purpose of this dissertation is to design and develop a mobile application focused on the collection and treatment of several biometric data of the user as well as on his personality. For this, a personality test, based on the *Saucier* model, will be implemented, which allows the user to choose from forty adjectives the ones that best suit him. When performing the test, which will be repeated over time, the user's biometric information and history will be taken into consideration. The goal is to offer, the user, the possibility of knowing the dimensions that make his personality while providing biometric data, being than possible to understand relations and correlations between the five fundamental personality dimensions and the collected data, as well as the evolution of these features over time.

Keywords: Big Five Mini-Markers, Biometric Data, Heath Apps, Machine Learning, Psychological Tests.

Índice

1 Introdução	1
1.1 Contextualização	1
1.2 Motivação	2
1.3 O Problema	4
1.4 Objetivos.....	5
1.5 Estrutura do Documento	5
2 Estado da Arte	7
2.1 As Cinco Dimensões da Personalidade.....	7
2.2 Dados Biométricos.....	10
2.2.1 Batimento Cardíaco	11
2.2.2 Monitorização do Sono.....	15
2.2.3 Monitorização de Atividade.....	15
2.2.4 Outras Informações	16
2.3 Dispositivos de Monitorização	18
2.3.1 <i>Smartwatch</i> e <i>Smartband</i>	19
2.3.2 Sensores.....	22

2.4	<i>Machine Learning</i>	24
2.4.1	Aprendizagem Supervisionada	24
2.4.2	Aprendizagem Não-Supervisionada	25
2.4.3	Aprendizagem por Reforço	27
3	Ferramentas e Tecnologias	29
3.1	Aplicações <i>Mobile</i> de Saúde e <i>Fitness</i>	29
3.2	<i>APIs</i> de Saúde e <i>Fitness</i>	32
3.3	Ferramentas e Tecnologias Utilizadas	34
3.3.1	Plataforma <i>Web Big Five</i>	34
3.3.2	<i>Android</i>	36
3.3.3	<i>Android Studio</i>	36
3.3.4	<i>Firebase</i>	37
3.3.5	<i>Google Fit</i>	37
3.3.6	<i>KNIME</i>	39
3.3.7	<i>Python</i>	39
3.3.8	<i>Jupyter Notebook</i>	40
4	Design e Implementação	42
4.1	Proposta e Design	42
4.2	Modelo de Dados	46

4.3	Aplicação Móvel (Path).....	52
5	Caso de Estudo.....	61
5.1	Tratamento dos Dados	65
5.2	Análise dos Resultados	66
6	Conclusões e Trabalho Futuro	81
6.1	Conclusões.....	81
6.2	Trabalho Futuro	83
	Bibliografia	84

Lista de Siglas e Acrónimos

API	<i>Application Programming Interface</i>
APK	<i>Android Package</i>
ASR	<i>Arritmia Sinusal Respiratória</i>
BFI	<i>Big Five Inventory</i>
BFI-2-S	<i>Big-Five Inventory-2-S</i>
BFI-2-XS	<i>Big Five Inventory -2-XS</i>
BFI-S	<i>Big-Five Inventory Short</i>
CMOS	<i>Complementary Metal Oxide Semiconductor</i>
CSV	<i>Comma - separated values</i>
DCV	<i>Doenças Cardiovasculares</i>
FC	<i>Frequência cardíaca</i>
GPS	<i>Global Positioning System</i>
IMC	<i>Índice de Massa Corporal</i>
JSON	<i>JavaScript Object Notation</i>
LED	<i>Luz Emissora de Díodo</i>
ML	<i>Machine Learning</i>
Neo FFI	<i>Neo Five Factor Inventory</i>

NEO-PI-R	<i>Revised NEO Personality Inventory</i>
NFC	<i>Near Field Communication</i>
OMS	<i>Organização Mundial de Saúde</i>
PaTH	<i>Personality and Track Health</i>
SB	<i>Smartband</i>
SDK	<i>Software development kit</i>
SO	<i>Sistema Operativo</i>
SW	<i>Smartwatch</i>
VFC	<i>Variabilidade da Frequência Cardíaca</i>

Índice de Figuras

Figura 1 - Os cinco traços de personalidade <i>Big Five Inventory</i>	8
Figura 2 - Fatores que alteram o batimento cardíaco.....	12
Figura 3 - Tabela do Índice de Massa Corporal	18
Figura 4 - Crescimento da venda de dispositivos.	20
Figura 5 - Menu inicial das aplicações <i>Samsung Health, Mi Fit, Huawei Health, Google Fit, Fitbit, Apple Health</i> , respectivamente.	32
Figura 6 - Vista inicial da plataforma <i>web</i>	34
Figura 7 - Resultado da realização do teste de personalidade na plataforma <i>web</i>	35
Figura 8 - Logótipo da aplicação móvel.	42
Figura 9 - <i>Mockup</i> do <i>login</i>	43
Figura 10 - <i>Mockup</i> da vista de menu e da monitorização de atividade.	44
Figura 11 - <i>Mockup</i> da vista da monitorização do batimento cardíaco e sono.	45
Figura 12 - <i>Mockup</i> da vista do teste de personalidade e o resultado.	46
Figura 13 - Árvore da base de dados.	48
Figura 14 - <i>UserBiometrics</i> com as suas subcategorias.....	49
Figura 15 - <i>UserInfo</i> com as respetivas subcategorias.	50
Figura 16 - <i>UserPerso</i> com algumas das suas subcategorias.	51
Figura 17 - Adicionar conta <i>Google Fit</i> a aplicação <i>Mi Fit</i>	53

Figura 18 - <i>Login Path</i> e página inicial.	54
Figura 19 - Vista para inserção das horas de sono.	54
Figura 20 - Vista das opções de visualização no menu principal.....	55
Figura 21 - Realização do teste psicológico.....	56
Figura 22 - Página inicial após a realização do teste de personalidade.....	57
Figura 23 - Vista dos gráficos relativos ao número de passos do diário e semanal.....	58
Figura 24 - Gráfico das variações do batimento cardíaco nas últimas 24h.....	59
Figura 25 - Vista do perfil do utilizador antes de preencher os dados e adicionar a data de nascimento.....	60
Figura 26 - <i>Workflow</i> do <i>KNIME</i>	66
Figura 27 - <i>Metanode</i> expandido referente ao <i>DataInfo</i>	68
Figura 28 - Gráfico relativo a faixa etária e género.	69
Figura 29 – <i>Metanode</i> referente ao <i>dataset</i> <i>DataPerso</i>	69
Figura 30 - <i>Boxplot</i> referente aos cinco traços de personalidade.....	70
Figura 31 - Gráfico de barras proveniente do nodo <i>Data Explorer</i> relativos aos 5 traços de personalidade estabilidade, abertura, extroversão, consciencioso e agradabilidade, por esta ordem.....	71
Figura 32 - Tabela proveniente do nodo <i>Data Explorer</i> com informação estatista dos dados nominais.....	72

Figura 33 - Matriz de Correlação entre os 5 traços de personalidade e os 40 adjetivos clicáveis.	73
Figura 34 - Matriz de Correlação dos 32 adjetivos com os 5 traços de personalidade.	74
Figura 35 -Gráficos com as características com mais correlação ao traço de personalidade.	74
Figura 36 - <i>Line Plot</i> que representa a variação dos traços de personalidade de entre os utilizadores.....	75
Figura 37 - Gráfico do utilizador Think j.c. referente aos resultados ao longo do tempo dos 5 traços de personalidade.	76
Figura 38 - Gráfico do utilizador Adriana referente aos resultados ao longo do tempo dos 5 traços de personalidade.	77
Figura 39 - <i>Metanode</i> referente a análise de dados biométricos.	77
Figura 40 - Gráfico de barras proveniente do nodo <i>Data Explorer</i> relativos aos atributos sleepMinutes, imc, heartbeat, steps, por esta ordem.	78
Figura 41 - Tabela proveniente do nodo <i>Data Explorer</i> , relativos aos dados nominais do dataBiometrics.	79
Figura 42 - Matriz de correlação com os dados biométricos, de personalidade e informação.	80

Índice de Tabelas

Tabela 1 - Tabela interpretativa do resultado do IMC.	17
Tabela 2 - Tabela comparativa entre diferentes <i>smartbands</i> e <i>smartwatches</i>	22
Tabela 3 - Tabela comparativa de diferentes aplicações de saúde.....	30
Tabela 4 - Tabela com os principais atributos de cada <i>dataset</i>	67

Capítulo 1

Introdução

Neste capítulo pretende-se apresentar o contexto do tema da dissertação, a motivação por trás do propósito deste projeto, os diferentes objetivos e, por último, a estrutura do documento.

1.1 Contextualização

Os testes de personalidade são uma ferramenta muito utilizada em processos de seleção de indivíduos, uma vez que podem revelar traços de personalidade do candidato que são menos óbvios ou que não são relatados numa entrevista. Também se torna possível identificar, por exemplo, se um indivíduo está apto para operar máquinas, conduzir veículos e/ou possuir licença de porte de armas de fogo, tendo ainda a finalidade de autoconhecimento ou de diagnosticar algum tipo de transtorno psicológico, sendo que, em último caso, o psicólogo indicará o tratamento mais adequado para trabalhar os traços de personalidade (Oliveira & Filho, 2015).

O modelo *Big Five Inventory* (BFI) (Digman, 1990) foi sofrendo variações ao longo dos anos, gerando novas versões, como por exemplo o *Big Five-Extra -Short* (Soto & John, 2017) ou o *Big Five Mini-Markers* (Thompson, 2008), versão mais curta e simplificada do original. No *Big Five Short* (Hahn, Gottschling, & Spinath, 2012) ocorreu uma diminuição de 44 para 15 dos itens em avaliação. Apesar das versões citadas acima mostrarem níveis aceitáveis

de consistência interna, estabilidade e nível suficiente de utilidade, o BFI permanece mais confiável. Perante a informação dos vários modelos de *Big Five*, será testado um modelo relacionado com o *Big Five Mini-Markers* (Saucier, 1994) de forma a avaliar os adjetivos de personalidade de 1 a 9 em conformidade com a seleção dos termos que mais definem o utilizador e ainda a idade, sexo, língua e a informação biometria correspondente ao batimento cardíaco, peso e qualidade do sono, entre outros.

É de salientar que a crescente utilização de dispositivos móveis, como *smartwatches* e *smartbands* induziram um aumento do interesse na monitorização da saúde pessoal. Esses dispositivos estão equipados com o *hardware* do sensor necessário, como acelerómetro e sensor de frequência cardíaca, para monitorização de atividades físicas e batimentos cardíacos, sendo a partir destes mecanismos que é obtida a variação dos batimentos (Pessemier & Martens, 2017). É também notório um interesse acrescido de cada pessoa em saber como se classifica a sua personalidade, i.e., o que é que a caracteriza e como.

1.2 Motivação

Em diversas cenários, testes e avaliações psicológicas são semelhantes a exames médicos. Se um paciente tiver sintomas físicos, um prestador de cuidados primários pode solicitar radiografias ou exames de sangue para entender sintomas. Os resultados dos testes ajudarão a informar o desenvolvimento de um plano de tratamento. Avaliações psicológicas servem o mesmo propósito, i.e., os psicólogos usam testes e outras ferramentas de avaliação para medir e observar o comportamento de um utente para chegar a um diagnóstico e orientar o tratamento.

Psicólogos administram testes e avaliações por uma ampla variedade de razões. As crianças que estão com dificuldades na escola, por exemplo, podem passar por testes de

aptidão ou testes que identificam dificuldades de aprendizagem. Testes para habilidades como destreza, tempo de reação e memória podem ajudar um neuropsicólogo a diagnosticar condições como lesões cerebrais ou demência (Eabon & Abrahamson, n.d.). Esses testes variam desde instrumentos científicos sérios e amplamente usados à psicologia popular.

É do consenso geral na psicologia académica que existem cinco traços fundamentais de personalidade (Open-Source Psychometrics Project, 2019). O modelo BFI é assumido na maioria das pesquisas sobre personalidade e é a base de muitos dos testes mais utilizados por psicólogos que mantêm uma estreita aceitação com a comunidade academia (Open-Source Psychometrics Project, 2019). No entanto este tipo de teste apesar ser bem aceite tem tendência a ser extenso e demorado.

Cada vez mais as pessoas se preocupam com a sua saúde e existe uma crescente tendência a monitorizá-la. Segundo uma análise de 2018 (“The Rise of mHealth Apps: A Market Snapshot,” 2018) o mercado de saúde móvel (mHealth) tem crescido constantemente nos últimos anos com a adoção de novas tecnologias, novos modelos de negócios e novos fluxos de trabalho que transformaram os serviços de saúde. Em consonância, tem se sentido um crescimento no número e qualidade de dispositivos de monitorização de *fitness*. Segundo artigos e análises existentes (“Wearables based on APIs, an excellent business opportunity,” 2018) está previsto um crescimento do número de dispositivos de *fitness* no mercado. As vendas aumentarão e a taxa de crescimento aumentará nos próximos anos.

Pretende-se, com este trabalho, tirar partido deste interesse para conceber e desenvolver uma aplicação que faça uma recolha exaustiva de vários dados biométricos, permita ao utilizador analisar o seu histórico e que, de forma diferenciada das aplicações atuais, permita ao utilizador saber como se caracteriza a sua personalidade e como esta

tem evoluído. Pessoalmente, uma forte motivação foca-se na possibilidade de ligar os domínios da psicologia à parte física de um utilizador. Na realidade atual, as pessoas, de um modo geral, têm mais tendência a ter problemas psicológicos, existe mais suscetibilidade para depressões, *stress*, ou seja, uma saúde mental mais débil. Esperemos que num futuro próximo as tecnologias, além de conseguirem examinar a saúde física, permitam também que uma pessoa controle a sua saúde mental de forma constante e assertiva.

1.3 O Problema

Nos dias de hoje, as cinco principais doenças mortais não transmissíveis são doenças cardíacas, cancro, doenças crônicas, doenças respiratórias e diabetes (Brundtland, 2000). Há fortes evidências científicas de que uma dieta saudável e atividade física adequada (ou seja, ≥ 30 minutos de atividade física de intensidade moderada, ≥ 5 dias por semana) desempenham papel importante na prevenção dessas doenças (*Preventing Noncommunicable Diseases in the Workplace through Diet and Physical Activity*, 2008). Além desta informação, cinco das dez principais causas de incapacidade no mundo (depressão, esquizofrenia, distúrbios bipolares, uso de álcool e transtornos obsessivos compulsivos) são problemas mentais (Brundtland, 2000).

Posto isto, e visto que cada vez mais as pessoas usam tecnologias móveis e dispositivos para controlar a atividade física, sentiu-se a necessidade de juntar dois temas que estão cada vez mais em voga, i.e., o controlo da saúde mental e física. Isto vai de encontro a uma lacuna que se verifica no mercado das aplicações móveis de saúde e *fitness* como a *Fitbit*, *Samsung Health*, *Health Mate*, *Apple Health*, que não abordam a temática das doenças mentais nem permitem realizar testes psicológicos.

1.4 Objetivos

O propósito fundamental desta dissertação é o de conceber e desenvolver uma aplicação móvel focada no tratamento de vários dados biométricos do utilizador assim como na sua personalidade. Para isso, será implementado um teste de personalidade, baseado no modelo de Saucier, que permite ao utilizador escolher de entre quarenta adjetivos, os que melhor se adequam a si. Aquando da realização do teste, que se repetirá ao longo do tempo, será tida em consideração a informação biométrica do utilizador e o seu histórico. Posto isto, os pontos a focar são os seguintes:

- Conceção e criação da aplicação móvel capaz de recolher um vasto conjunto de dados biométricos;
- Incorporação do teste de personalidade *big five* na aplicação móvel;
- Criação de vistas que permitam perceber a evolução dos dados biométricos e de respostas ao teste de personalidade que são submetidas ao longo do tempo;
- Tratar, limpar e extrair conhecimento dos dados colecionados;
- Avaliar relações entre os dados psicológicos e os dados biométricos.

1.5 Estrutura do Documento

A estrutura deste documento consiste em seis capítulos, cada um composto por diversas seções. O primeiro capítulo introduz o tema deste projeto, revelando as principais motivações para o seu desenvolvimento, bem como os objetivos a alcançar para o produto final. Termina com uma breve descrição da estrutura desta dissertação. No segundo

capítulo, encontra-se explorado o conceito de teste psicológico, dados biométricos e a sua relação com o estado psicológico de um indivíduo. No terceiro capítulo é feita uma análise do estado da arte de várias aplicações de *fitness* e saúde, sendo realizada uma comparação entre as mesmas. No quarto capítulo está descrito o processo de construção da aplicação, desde as *mockups* até ao protótipo final. No quinto capítulo aplica-se um caso de estudo, tendo a aplicação sido distribuída por um conjunto de indivíduos. Os dados são tratados e analisados. No sexto capítulo apresentam-se conclusões sobre o trabalho desenvolvido, analisando aspetos positivos e negativos que o tenham envolvido. Conclui-se o capítulo com uma breve análise relativamente ao futuro do método desenvolvido e da sua aplicação em ambientes reais.

Capítulo 2

Estado da Arte

Este capítulo apresenta uma análise às áreas do domínio do projeto. A metodologia utilizada para elaborar a pesquisa dos conceitos relevantes discutidos neste capítulo, consistiu principalmente na análise de artigos científicos em conferências relevantes, revistas científicas, livros científicos e aplicações práticas. Os termos pesquisados foram *Big Five Personality Test*, dados biométricos, monitorização de saúde e atividade física, monitorização do sono e *Machine Learning*, tanto na língua portuguesa como inglesa. O objetivo desta análise não é apenas reunir informações relevantes, mas também refletir a fim de entender a área e os temas do projeto de dissertação.

2.1 As Cinco Dimensões da Personalidade

Os testes de personalidade sofreram uma grande evolução ao longo dos anos. Inicialmente vários grupos independentes de cientista descobriram e definiram os cinco grandes traços baseados em pesquisas empíricas (Thiel, 2018). Ernest Tupes e Raymond Christal avançaram o modelo inicial, com base no trabalho realizado no final dos anos 50 (Thiel, 2018). Digman propôs o seu modelo de cinco fatores de personalidade em 1990, sendo que Goldberg o estendeu ao mais alto nível organizacional em 1992 (Goldberg, 1992).

Na Figura 1 é possível encontrar os cinco traços de personalidade, nomeadamente extroversão, agradabilidade, abertura, consciencioso e estabilidade.

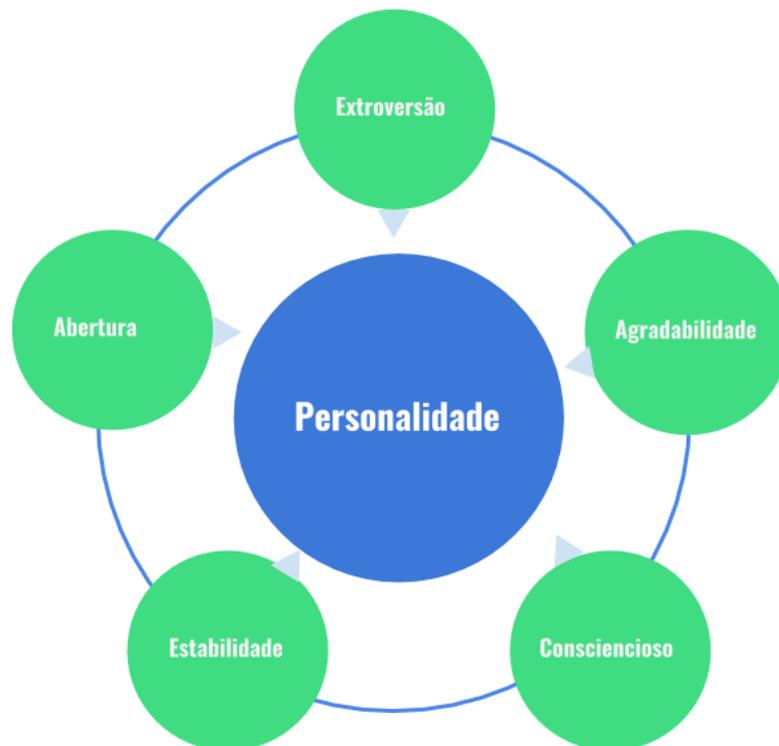


Figura 1 - Os cinco traços de personalidade *Big Five Inventory*.

- **Extroversão** - Este fator representa dois fins: extroversão e introversão. Diz respeito à forma como um indivíduo atrai energia e como interage com os outros. De uma forma geral, extrovertidos consomem energia ou “recarregam” ao interagir com os outros, enquanto os introvertidos se cansam de interagir com os outros e reabastecem a sua energia da solidão (Lebowitz, 2016).
- **Agradabilidade** - Este fator diz respeito à forma como as pessoas se relacionam positivamente com os outros. É uma construção que assenta sobre como o indivíduo geralmente interage com os outros (Lebowitz, 2016) .

- **Conscienzioso** - É uma característica que pode ser descrita como a tendência de controlar impulsos e agir de formas socialmente aceitáveis, i.e., comportamentos que facilitam o comportamento direcionado por objetivos (John & Srivastava, 1999). As pessoas conscientes são excelentes na sua capacidade de retardar a gratificação, trabalhar dentro das regras e planejar, e organizar de forma eficaz.
- **Estabilidade** – Este traço é o único grande fator em que uma baixa pontuação indica um traço negativo superior. Não é um fator de maldade ou incompetência, mas de confiança e de estar confortável dentro da própria pele. Engloba a estabilidade emocional e o temperamento geral (Lebowitz, 2016).
- **Abertura** - A abertura à experiência tem sido descrita como a profundidade e complexidade da vida mental e das experiências de um indivíduo (John & Srivastava, 1999). É também chamado de intelecto ou imaginação. A abertura à experiência diz respeito à disposição de um indivíduo em tentar coisas novas, ser vulnerável e ter capacidade de pensar fora da caixa (Ackerman, 2017).

A designação dos traços de personalidade foi obtida segundo tradução para português de uma especialista, nomeadamente uma psicóloga da Escola de Psicologia da Universidade do Minho.

Uma vez que os teste de personalidade são por norma extensos, houve várias tentativas de os abreviar como é o caso do *Big-Five Inventory Short* (BFI-2-S) (Soto & John, 2017) que é uma abreviação dos 60 item do *Neo Five Factor Inventory* (Neo FFI; Costa and McCrae, 1992) e dos 44 itens do *Big Five Inventory* (BFI; Jonh, Donahue, and Kentle, 1991). O BFI-S com 15 itens, 3 item para cada dimensão, revelou-se com baixa correlação em comparação com o NEO-PI-R em relação à amabilidade. Nessa mesma análise foi ainda concluído que existia uma média de estabilidade de coeficientes de 70% ao refazer o teste passados 18 meses. Outras

tentativas de redução foram o *Big-Five Inventory-2-S* (BFI-2-S) e o *Big Five Inventory -2-XS* (BFI-2-XS) de 30 e 15 itens respetivamente, desenvolvidas em 2017 (Soto & John, 2017) e que pretendia evidenciar a validade e a confiabilidade ao classificar os 5 domínios do *Big Five Personality* e os 15 traços mais específicos. No entanto, apenas o BFI-2-S se mostrou confiável e válido para examinar os traços das facetas para um conjunto razoavelmente grande de amostras, enquanto que o segundo abreviamento não se revelou viável. Uma outra abordagem concretizada (Goldberg, 1992), consistia num conjunto robusto de 100 adjetivos para a estrutura do *Big-Five* encontrada na descrição da personalidade fenotípica. Saucier (Saucier, 1994) reduziu o conjunto dos 100 marcadores para um subconjunto de apenas 40 adjetivos avaliados individualmente de 1 a 9. Esse subconjunto *Mini-Marker* demonstrou ser consistente nas cinco escalas, com favorável correlação e confiabilidade.

Noutro sentido, Thompson (Thompson, 2008) desenvolveu o *International English Mini-Markers* um modelo que pretendia compreender apenas 40 adjetivos descritores de personalidade originalmente selecionados pelas suas qualidades psicométricas (Saucier, 1994). Em que por hipótese os Mini-marcadores se usam sem tradução. No entanto, no teste ele adquiriu a perceção que os não-nativos na língua inglesa tiveram dificuldade em entender certos adjetivos que tinham significado ambíguos. Assim, de forma a obter resultados mais exatos o ideal é realizar o teste na língua nativa do utilizador.

2.2 Dados Biométricos

Autenticação biométrica está se a tornar cada vez mais comum. Um exemplo disso é o desbloquear do *smartphone* que poderá ser feito recorrendo ao *scanner* de impressão digital. De facto, a biometria é a medição e a análise das características fisiológicas e comportamentais das pessoas. É usada principalmente para a identificação e concessão de

acesso a edifícios ou serviços, ou para assinar transações digitalmente. Baseia-se na ideia de que todos temos atributos fisiológicos únicos e podemos ser identificados por eles (Gemalto, n.d.).

Os dispositivos de monitorização de *fitness* são uma forma de tecnologia biométrica, uma vez que pode ler e sensorizar elementos humanos. Com estes dispositivos, é possível acompanhar estatísticas como frequência cardíaca, as calorias consumidas ou as etapas compridas. Tudo isto é possível com um pequeno dispositivo usado no pulso, na cintura ou até no pescoço (Kenney, 2015). Os principais dados biométricos que é possível adquirir através dos dispositivos de monitorização *fitness* são a frequência cardíaca, dados de atividade física, altura e peso e a monitorização do sono. Neste capítulo será dada maior relevância aos dados biométricos referentes à variação do batimento cardíaco, à monitorização do sono e de atividade física e de que forma é que estão relacionados com a personalidade de um indivíduo.

2.2.1 Batimento Cardíaco

O conhecimento sobre a frequência cardíaca (FC) pode ajudar a monitorizar o nível de condição física, e pode até mesmo ajudar a identificar problemas de saúde. A variabilidade da frequência cardíaca (VFC) corresponde à adaptação do coração a qualquer estímulo. De fato, entre as patologias que mais afetam a VFC, estão as doenças cardiovasculares e os transtornos depressivos. Consequentemente, a VFC é hoje amplamente usada como um índice de saúde. Uma alta VFC é o sinal de que o coração está saudável, pois tem mais flexibilidade para reagir a qualquer stress (fisiológico ou ambiental) (Fatisson, Oswald, & Lalonde, 2016). Alternativamente, a redução da VFC é um sintoma de problemas de saúde e pode afetar as funções imunológicas, a autorregulação e as habilidades psicossociais. Dentre as patologias que mais afetam a VFC, estão as doenças cardiovasculares (DCV) e as depressivas (American Heart Association, n.d.). Na Figura 2 é possível verificar os principais fatores que afetam a VFC segundo Fatisson.



Figura 2 - Fatores que alteram o batimento cardíaco.

Fatores fisiológicos e patológicos

As glândulas endócrinas são responsáveis pela produção de hormonas que permitem regular várias funções do corpo humano. Segundo Fatisson (Fatisson et al., 2016) é possível verificar que as relações entre tiroide, hormonas sexuais ou cortisol com a VFC são estabelecidas e significativas em alguns casos. Outro fator que é um grande influenciador da VFC é a respiração. Um fenómeno fisiológico acontece a cada respiração: a arritmia sinusal respiratória (ASR), onde o batimento cardíaco aumenta à medida que o ar é inspirado e diminui a cada expiração (van Ravenswaaij-Arts, Kollée, Hopman, Stoelinga, & van Geijn, 1993). Ainda no mesmo ponto, outro fator importante é o neuronal, uma vez que existem numerosas conexões neuronais entre o coração e o cérebro, e se uma dessas estruturas for

afetada, a VFC também é alterada (Fatisson et al., 2016). De fato, distúrbios neurológicos, associados a danos cerebrais, foram relacionados à redução da VFC. A VFC é menor em indivíduos com pressão arterial elevada do que em indivíduos com pressão normal. Existe também uma relação semelhante com um alto nível de colesterol no sangue ou glicose (diabetes) e uma diminuição da VFC (Thayer, Yamamoto, & Brosschot, 2010).

Fatores de estilo de vida

Estudos revelam que VFC é maior (e o ritmo cardíaco mais baixo) em indivíduos com atividade física de moderada a vigorosa. O consumo de álcool e tabaco também afeta a VFC. Os fumadores crônicos e os consumidores de álcool têm baixa VFC, mas esse efeito é reversível quando eles cessam o consumo (Thayer et al., 2010)(Valentini & Parati, 2009).

Fatores não modificáveis

Com a evolução da idade a VFC tem tendência diminuir na população em geral (Thayer et al., 2010)).

Fatores ambientais

Fatores relacionados com o ambiente no local de trabalho podem influenciar a VFC. Entre esses fatores, está o contacto com vários componentes químicos, campos eletromagnéticos, ferramentas vibratórias, carga psicossocial, tempo de trabalho, fadiga, entre outros. A exposição prolongada a esses fatores leva a uma diminuição da VFC (Togo & Takahashi, 2009)

Fatores Neuropsicológicos

Consequentemente, e de um ponto de vista geral, a VFC está ligada à personalidade, i.e., uma lista de fatores neuropsicológicos que influenciam a VFC inclui stress, depressão e emoções negativas. De facto, fontes de stress como ansiedade, hostilidade e depressão induzem baixa VFC. A depressão e a utilização de antidepressivos têm uma correlação significativa com baixa VFC (van Zyl, Hasegawa, & Nagata, 2008). Na psicologia, a baixa estabilidade está associado à experiência de emoções negativas e diz respeito a pacientes que sofrem de raiva, depressão ou ansiedade o que se traduz numa baixa VFC. Embora as emoções negativas, como raiva, ansiedade, frustração e preocupação, produzam eletrocardiograma irregular e uma VFC baixa. (McCraty, Atkinson, Tomasino, & Bradley, 2009).

Outros fatores

A temperatura corporal é outro fator que afeta a VFC. Uma vez que, se a temperatura diminuir o suficiente, induz uma bradicardia, os batimentos tornam-se mais lentos. Alternativamente, quando as temperaturas sobem, como com febre, os batimentos cardíacos aceleram (Fatisson et al., 2016).

A personalidade e a variabilidade da frequência cardíaca (VFC) são fortes preditores de bem-estar, particularmente saúde cardíaca e longevidade. Os autores (Zohar & McCraty, 2013) exploram as correlações dos traços de personalidade sobre a VFC para esclarecer como a regulação autonómica pode mediar o desenvolvimento e a manutenção da saúde e da doença. O estudo concluiu que existem relações entre a personalidade e a VFC apoiam a validade das medidas de forma a esclarecer as fortes relações entre personalidade, VFC e saúde.

2.2.2 Monitorização do Sono

Uma boa noite de sono é importante para uma mente e um corpo são. Existem vários dispositivos e aplicações que permitem seguir os movimentos durante o sono, e desta forma medir a qualidade do sono. A ideia é que a maneira como nos movemos durante a noite corresponde às diferentes fases do sono (Timo-laria et al., 1970). Algumas aplicações permitem ainda gravar barulhos, o que aliado a variação dos batimentos cardíacos, permite em teoria estabelecer as fases do sono ultrapassadas.

Embora pesquisas anteriores tenham mostrado que a personalidade e o sono são preditivos substanciais de saúde ao longo da vida, pouco se sabe sobre as ligações entre a personalidade e os padrões saudáveis de sono. O estudo realizado por (Duggan, Friedman, McDevitt, & Mednick, 2014) examinou o *Big Five* e uma série de fatores relacionados com a saúde do sono. O resultado mostrou que baixo consciencioso e baixa estabilidade foram os melhores preditivos de sono insatisfatório (má higiene do sono, baixa qualidade do sono e aumento da sonolência), consistente com outras pesquisas sobre saúde precária e risco de mortalidade.

Outro estudo realizado por (Hintsanen et al., 2014) examina associações entre os cinco traços de personalidade e duração média do sono, deficiência de sono e problemas de sono. Os resultados mostraram que maior extroversão, agradabilidade e consciencioso foram, em geral, associados com melhor sono, enquanto menor estabilidade foi associado com uma pior noite de sono. A abertura não estava associada ao sono.

2.2.3 Monitorização de Atividade

Foi realizada uma pesquisa por artigos científicos, com o intuito de relacionar a personalidade e a atividade física. O estudo realizado por (Rhodes & Smith, 2006), em que este realizou uma revisão de literatura sobre os principais traços de personalidade e atividade

física, além de fornecer alguns resumos meta-analíticos dos resultados, revelaram diferenças por características como a extroversão, sendo necessárias mais pesquisas para se poder tirar conclusões válidas. Um estudo mais recente que pretendia estimar as correlações populacionais entre os fatores de personalidade *Big Five* e a atividade física, examinou se elas variavam de acordo com as características da amostra e as características do estudo. Os resultados obtidos por Wilson & Dishman, permitiram concluir que variaram de acordo com as características da amostra (por exemplo, idade, sexo ou estado clínico) e/ou com as características do estudo (por exemplo, medir qualidade ou formato do item). Esta análise fornece um novo suporte para uma relação entre atividade física e a dimensão Abertura (Wilson & Dishman, 2015).

2.2.4 Outras Informações

Segundo a Organização Mundial de Saúde as "Recomendações Globais sobre Atividade Física para a Saúde" abordam duas faixas etárias: 5 a 17 anos e 18 ou mais. Essas faixas etárias foram selecionadas levando em consideração a natureza e a disponibilidade de evidências científicas relevantes para a prevenção de doenças não transmissíveis por meio da atividade física (*Global recommendations on physical activity for health, 2010*)).

De forma a melhorar a aptidão cardiorrespiratória e muscular, a saúde óssea e os biomarcadores de saúde cardiovascular e metabólica, crianças e jovens de 5 a 17 anos devem acumular diariamente pelo menos 60 minutos de atividade física de intensidade moderada a vigorosa. Já os adultos devem fazer pelo menos 150 minutos de atividade física aeróbica de intensidade moderada ao longo da semana ou pelo menos 75 minutos de atividade física aeróbica de intensidade vigorosa no mesmo período ou uma combinação equivalente de atividade de intensidade moderada e vigorosa. Desta forma permite melhorar a capacidade cardiorrespiratória e muscular, a saúde óssea reduz o risco de doenças crônicas não

transmissíveis, depressão e declínio cognitivo (*Global recommendations on physical activity for health*, 2010).

Outro indicador de saúde, o índice de massa corporal, ou IMC, é uma medida útil para a maioria das pessoas com mais de 18 anos de idade, sendo utilizada para determinar se o indivíduo está numa faixa de peso saudável para a sua altura. O IMC compara o peso com a altura e é calculado dividindo o peso pela altura. Desta forma, permite uma ideia de se a pessoa está 'abaixo do peso', 'saudável', 'acima do peso' ou 'obeso' para a sua altura. O IMC é um tipo de ferramenta para ajudar os profissionais de saúde a avaliar o risco de doença crônica (“BMI Calculator | The Heart Foundation,” n.d.). Os resultados do IMC, segundo o ministério da saúde do Brasil (Pinheiro, 2019) deverão interpretados a partir da Tabela 1 seguinte forma:

Estado	IMC por kg/m ²
Baixo peso muito grave	< de 16
Baixo peso grave	16 e 16,99
Baixo peso	entre 17 e 18,49
Peso normal	entre 18,50 e 24,99
Sobrepeso	entre 25 e 29,99
Obesidade grau I	entre 30 e 34,99
Obesidade grau II	Entre 35 e 39,99
Obesidade grau III (obesidade mórbida)	> 40

Tabela 1 - Tabela interpretativa do resultado do IMC.

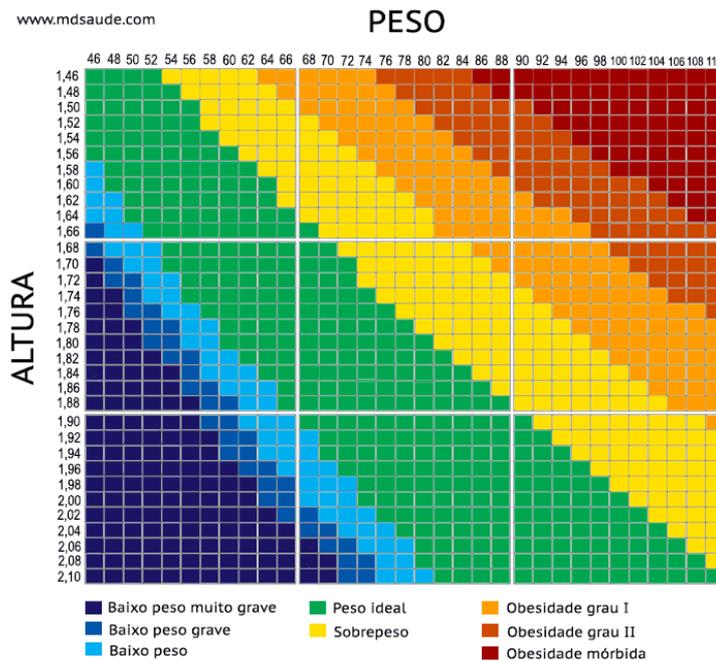


Figura 3 - Tabela do Índice de Massa Corporal

2.3 Dispositivos de Monitorização

Atualmente a facilidade com que o ser humano tem acesso a dispositivos móveis, como *smartphones*, *smartwatches* e *smartbands* induzem um aumento do interesse na monitorização da saúde pessoal. Esses dispositivos estão equipados com os sensores necessários, como acelerómetro e sensor de frequência cardíaca. Pessemier & Martens, em 2017 avaliaram a precisão de vários dispositivos como a *chest strap*, *smartband*, *smartwatch* e um *smartphone* na medição da frequência cardíaca.

Todos estes dispositivos, com a exceção do *smartphone*, fazem a leitura do batimento cardíaco através de um flash embutido com luz emissora de diodo (LED) como fonte de luz e a câmara CMOS (complementary metal oxide semiconductor) como sensor de luz que permite

a leitura das ondas cardíacas. O *flash* do *smartphone* pode ser configurado como fotopletismógrafo, e deste modo, pode então ser usado como um gravador de ondas de pulso (Matsumura & Yamakoshi, 2013) que desta forma, permite medir a variação do ritmo cardíaco. Ao comparar valores detetados pelos vários monitores cardíacos referidos, os resultados mostram medições semelhantes para os quatro dispositivos no estado de repouso.

Em contraste, durante atividades, como atividade física, a *smartband*, o *smartwatch* e o *smartphone* podem registar variações súbitas na frequência cardíaca com um atraso, devido a movimentos físicos do pulso ou da mão. O dispositivo especializado com cinta de peito (*chest strap*) mostra a frequência cardíaca mais precisa, altamente correlacionados com as medidas obtidas com o monitor de pressão arterial aprovado para fins médicos (Pessemier & Martens, 2017). Segundo um artigo da Digital Trends (Hill, 2019) que afirma que a medição de passos é bastante precisa nos dispositivos já referenciados acima. Salvo raros casos de pessoas que se deslocam muito lentamente ou de forma incomum ou então deteta movimentos involuntário, nestes casos os passos são subconta-dos.

2.3.1 *Smartwatch e Smartband*

Os dispositivos de *fitness* continuam em expansão em termos de vendas, um exemplo disso é que em 2016 foi lançada uma notícia (SAPO, 2016), que defere que já há mais vendas de *smartwatches* do que relógios analógicos. No gráfico da Figura 4 é possível ver os dispositivos de monitorização mais vendidos até ao ano de 2017, sendo que a tendência é para aumentar:

Forecast for Wearables Devices Worldwide (millions of units)

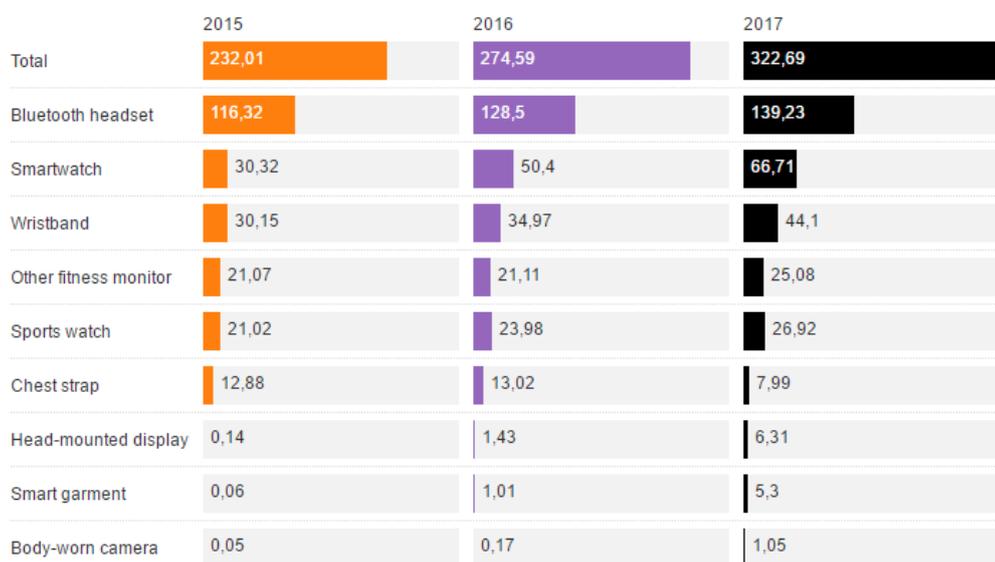
Source: [Gartner](#) [Get the data](#)Created with [Datawrapper](#)

Figura 4 - Crescimento da venda de dispositivos.

Além do crescimento comercial é possível verificar que em termos de dispositivos de monitorização as pulseiras e os relógios inteligentes são os mais populares. Desta forma, nesta secção vai ser dada mais importância aos dispositivos do tipo *smartbands* e *smartwatches*, uma vez que são os mais práticos e mais usados. Tal como o nome indica, *smartwatches* são relógios inteligentes, também conhecidos, como uma espécie de *smartphone* de pulso. O dispositivo é capaz de se conectar aos telemóveis de última geração e usar aplicações como *WhatsApp*, *Facebook*, determinados jogos e ainda receber chamadas e tirar fotos. As funções de contagem de passos, acompanhamento cardíaco e do sono são apenas algumas opções oferecidas (Tecnologia - ig, 2018).

De forma a obter mais informação sobre estes aparelhos, foi realizada uma tabela comparativa sobre os dispositivos mais conhecidos evidenciando as características técnicas,

os tipos de monitorização que permitem, entre outras características, que é possível verificar na Tabela 2.

	 Amazfit Bip	 Apple Watch Series 4	 Fitbit Charge 3	 Fitbit. Versa	 Huawei Band 3 Pro	 Huawei Watch 2	 Samsung Gear Fit 2	 Samsung Gear Sport	 Xiaomi Mi Band 3
Especificações Técnicas									
Vida da Bateria	30-45 dias	18+ horas	+7 dias	+4 dias	20 dias	3-25 dias	3-4 dias	4 dias	20 dias
Tipo	SW	SW	SB	SW	SB	SW	SW	SW	SB
Resistência a água	IP68	5 ATM	5 ATM	5 ATM	5 ATM	IP68	IP68	5 ATM	5 ATM
Sensores									
Acelerómetro	✓	✓	✓	✓	✓	✓	✓	✓	✓
Luminosidade	✓	✓	X	X	X	✓	X	✓	X
Barómetro	✓	✓	X	X	X	✓	✓	✓	X
Oxímetro	✓	✓	✓	✓	✓	✓	✓	✓	✓
GPS	✓	✓	X	X	✓	✓	✓	✓	X
GLONASS	✓	✓	X	X	X	✓	✓	X	X
Altímetro	X	✓	✓	✓	X	X	X	X	X
Bluetooth	✓	✓	✓	✓	✓	✓	✓	✓	✓
Giroscópio	X	✓	X	X	X	✓	✓	✓	X
Monitorização de atividade									
Passos	✓	✓	✓	✓	✓	✓	✓	✓	✓
Distancia	✓	✓	✓	✓	✓	✓	✓	✓	✓
Calorias	✓	✓	✓	✓	✓	✓	✓	✓	✓
Atividade	✓	✓	✓	✓	✓	✓	✓	✓	✓
Piso	X	✓	✓	✓	✓	✓	✓	✓	X
Batimento Cardíaco	✓	✓	24/7	24/7	✓	24/7	✓	✓	✓

Vo2Max	✓	X	✓	✓	✓	✓	X	X	X
GPS	✓	✓	X	X	✓	✓	✓	✓	X
Outras Funções									
Notificações	✓	✓	✓	✓	✓	✓	✓	✓	✓
Música em Memoria	X	X	✓	✓	X	X	X	✓	X
Pagamento NFC	X	✓	✓	✓	X	X	X	X	✓
SO	<i>Android Wear</i>	<i>Watch OS</i>	<i>Fitbit OS</i>	<i>Fitbit OS</i>	<i>Android Wear</i>	<i>SO Produtor</i>	<i>Tizen</i>	<i>Tizen</i>	<i>SO Produtor</i>
Apps	<i>Mi Fit</i>	<i>Health</i>	<i>Fitbit</i>	<i>Fitbit</i>	<i>Huawei Health</i>	<i>Huawei Health</i>	<i>Samsung Health</i>	<i>Samsung Health</i>	<i>Mi Fit</i>

Tabela 2 - Tabela comparativa entre diferentes *smartbands* e *smartwatches*.

A partir Tabela 2 foi possível concluir que os dispositivos mais usados, mesmo os de mais baixa gama, contêm sensores que permitem monitorizar a atividade, como a contagem das calorias, número de passos e distância, a qualidade e a duração do ciclo de sono, a monitorização do batimento cardíaco, sendo que neste último existem dispositivos como a *Fitbit* que monitorizam 24h por dia e outros que só verificam a pulsação quando são invocados como é o caso da *Xiaomi Amazfit Bip*. A qualidade da monitorização depende sempre da qualidade dos sensores e do tipo. Por exemplo, o *Apple Watch* tem sensor ótico e elétrico, sendo que existem ainda outros dispositivos que permitem a medição da pressão arterial como é o caso da *Fitbit Charge 3*.

2.3.2 Sensores

Da grande variedade de sensores usados por *smartbands*, *smartwatches* e *smartphones* pode-se definir os mais importantes como sendo:

- **Acelerómetro** - serve para medir a aceleração de um corpo em relação à gravidade. Este sensor controla a deteção de movimento baseada em eixos e podem ser

encontrados tanto em monitores *fitness* como no *smartphone* sendo por causa deles que é possível registar os passos e a velocidade do carro quando usamos aplicações como *Google Maps* ou *SnapChat*.

- **Altímetro** - aparelho responsável pela medição de altitude em relação nível do mar, é especialmente útil a utilizadores que pratiquem exercício físico ao ar livre. Com ele é possível, por exemplo, medir com exatidão a altura alcançada durante o percurso de uma corrida, ou então saber precisamente o momento mais alto durante um passeio de bicicleta.
- **Giroscópio** – este dispositivo complementa o acelerómetro, ajudando a entender de que forma o *smartphone* ou um monitor de *fitness* está orientado, desta forma adiciona uma camada extra de precisão.
- **Sensor de luminosidade** - este sensor, permite medir a luz do ambiente e deste modo ajustar o brilho do ecrã em conformidade com a luz atual.
- **Barómetro** - Responsável por indicar a pressão atmosférica na tela do relógio, ele transmite informações importantes, como variações abruptas na pressão atmosférica que levam a mudanças climáticas.
- **Oxímetro** - Sensor responsável por medir a pulsação, determina a quantidade de oxigénio no sangue, e permite a classificação da condição arterial.
- **GPS** - sistema de posicionamento global é um sistema de satélites e outros dispositivos que tem como função básica prestar informações precisas sobre o posicionamento individual no globo terrestre.
- **GLONASS** - Sistema de Navegação Global via Satélite é o equivalente russo do GPS. Praticamente todos os dispositivos que são compatíveis com o GPS também o são com GLONASS.
- **Bluetooth** - tecnologia de comunicação sem fios que permite que computadores, *smartphones*, *tablets*, *smartwatches* e *smartbands* troquem dados entre si e se

conectem a ratos, teclados, fones, impressoras, colunas e outros acessórios a partir de ondas de rádio. O conceito permite que os dispositivos se interliguem de forma rápida sem a necessidade de usar cabos, sendo suficiente que estejam próximos um do outro.

2.4 Machine Learning

É um ramo da inteligência artificial baseado na ideia de que os sistemas podem aprender com dados, identificar padrões e tomar decisões com mínima intervenção humana. É um ingrediente crítico para aplicações inteligentes e oferece a oportunidade de acelerar ainda mais os processos de descoberta, bem como aprimorar os processos de tomada de decisão (Bari, 2018). Os três métodos mais adotados de *Machine Learning* são a aprendizagem supervisionada, a aprendizagem não-supervisionada e aprendizagem por reforço.

2.4.1 Aprendizagem Supervisionada

Na aprendizagem supervisionada um conjunto de dados de treino contém dados de entrada e o valor que se pretende prever (que pode ser numérico ou não). O modelo a desenvolver usa os dados de treino para aprender ligações entre as entradas e as saídas. A ideia subjacente é que os dados de treino podem ser generalizados e que o modelo pode ser usado em novos dados com alguma precisão. Alguns exemplos de algoritmos de aprendizagem supervisionada incluem:

- **Support Vector Machine (SVM)** é um algoritmo de *machine learning* supervisionado que pode ser usado para desafios de classificação ou regressão. No entanto, é usado principalmente em problemas de classificação. Neste algoritmo, alocamos cada item de dados como um ponto no espaço n-dimensional, sendo que o valor de cada recurso corresponde ao valor de uma determinada coordenada. Os Vetores de Suporte são

simplesmente as coordenadas da observação individual. SVM é a fronteira que melhor separa as duas classes (hiperplano/linha).

- **Classificador de Naive Bayes** é baseado no teorema de *Bayes*. A abordagem probabilística seguida pelo Classificador *Naive Bayes* determina a probabilidade da classificação e ajuda a encontrar a incerteza sobre o modelo. É um algoritmo eficiente que usa o conhecimento prévio dos dados observados (Sahu, Nautiyal, & Prasad, 2017).
- **Árvores de decisão** é um algoritmo que gera árvores de decisão a partir dos dados de treino para resolver problemas de classificação e regressão. O modelo de árvore de decisão é bom para manipular dados tabulares com recursos numéricos ou recursos categóricos com poucas centenas de categorias. Ao contrário dos modelos lineares, as árvores de decisão são capazes de capturar interação não-linear.

2.4.2 Aprendizagem Não-Supervisionada

Ao contrário do que acontece na aprendizagem supervisionada, estes paradigmas não possuem dados de entrada correspondentes com as variáveis de saída. A aprendizagem não-supervisionada é um tipo de *Machine Learning* que visa descobrir padrões em grandes conjuntos de dados ou classificar os dados em várias categorias sem ser explicitamente treinado (Wang, 2016). Este tipo de aprendizagem está dividido em dois grandes grupos, nomeadamente *clustering* e *dimensionally reduction*.

A análise de *clustering* é utilizada para análise exploratória de dados, com o objetivo de encontrar padrões ou agrupamentos ocultos nos dados, de que é exemplo:

- **K-Means** é um método em que *clustering* requer pouco esforço.

Para calcular o algoritmo *K-means* são necessárias três etapas, que são iteradas até existir convergência, com os seguintes passos (Johor Bahru, Darul Ta, Bin Mohamad, Usman, & Bahru, 2013):

- Determinar a coordenada do centroide;
 - Determinar a distância de cada objeto ao centroides;
 - Agrupar o objeto com base na distância mínima;

O objetivo do *clustering* é descobrir pontos comuns e os aglomerados dos grandes conjuntos de dados, dividindo-o os em grupos. Como se supõe que os conjuntos de dados não são rotulados, o *clustering* é frequentemente considerado como o mais valioso problema de aprendizagem não supervisionada (Cios et al., 2007).

A *dimensionality reduction* é uma forma de reduzir a complexidade de um modelo e evitar o ajuste excessivo. Existem duas categorias principais de redução de dimensionalidade: seleção e extração de recursos. Por meio da seleção de recursos, seleciona-se um subconjunto dos recursos originais, enquanto que, na extração de recursos, derivam-se informações do conjunto de recursos para construir um novo subespaço de recursos.

Uma das técnicas de *dimensionality reduction* encontra-se descrita a seguir:

- **Principal Component Analysis (PCA)** é provavelmente a técnica de redução de dimensão mais conhecida e mais usada para reduzir consideravelmente o número de variáveis, mantendo muitas das informações do conjunto de dados original (Jolliffe, 2011). Esta técnica é caracterizada por identificar as dimensões ao longo das quais os dados se encontram mais dispersos. Desta forma, é possível identificar as dimensões que melhor diferenciam o conjunto de dados em análise, ou seja, os seus componentes principais. Através desta técnica é possível realçar as semelhanças e diferenças neles existentes através da identificação de padrões. A sua identificação em dados

caracterizados por grandes dimensões é complexa, uma vez que a sua representação gráfica não é viável, o que inviabiliza a análise visual dos dados. Quando identificados os padrões no conjunto, o número de dimensões a analisar pode ser reduzido, sem que haja uma perda significativa de informação, pois o foco recai sobre a análise das dimensões principais que caracterizam o conjunto de dados.

2.4.3 Aprendizagem por Reforço

A aprendizagem por reforço é uma forma de sistemas inteligentes em que um agente tenta aprender uma política que maximiza um sinal de recompensa numérico (Sutton & Barto, 1998). Este algoritmo descobre através de testes do tipo 'tentativa e erro' quais as ações que rendem as maiores recompensas. Este tipo de sistema possui três componentes principais: o agente (o aprendiz ou tomador de decisão), o ambiente (tudo com que o agente interage) e ações (o que o agente pode fazer). O objetivo é que o agente escolha ações que maximizem a recompensa esperada num período de tempo determinado. O agente atingirá o objetivo muito mais rápido se seguir uma boa política. Abaixo estão dois exemplos de algoritmos deste tipo:

- ***Q-learning*** é um dos mais populares algoritmos de diferença temporal. Este algoritmo é livre de modelo, o que significa que ele aprende um controlador sem aprender um modelo. Para aprender este controlador, o *Q-learning* treina uma função de valor de ação que retorna o valor esperado para tomar uma ação num estado s . O agente usará essa função para formar uma política que maximizará os retornos. Uma vez que todos os pares estado-ação tenham sido visitados um número finito de vezes, garante-se que o método gerará estimativas Q_t que convergem para o valor de Q (Watkins, C. J. C. H. e Dayan, P. (1992)). Na prática, a política de ações converge para a política ótima em tempo finito, embora de forma lenta (Monteiro & Ribeiro, 2004).

- **Sarsa**, o algoritmo que é uma modificação do algoritmo *Q-learning* que utiliza um mecanismo de iteração de política (Sutton, R. S. e Barto, A. G. (1998)). A forma de procedimento do algoritmo *Sarsa* é similar à do algoritmo *Q-learning*. Idealmente, o algoritmo *Sarsa* converge para uma política e valor de função de ação ótimos assim que todos os pares estado-ação tenham sido visitados um número finito de vezes e a política de escolha da próxima ação converge, no limite, para uma política que utilize a melhor ação (ou seja, aquela que maximize a recompensa futura esperada) (Monteiro & Ribeiro, 2004).

Capítulo 3

Ferramentas e Tecnologias

Este capítulo aborda várias aplicações móveis, presentes tanto em *Android* como em *IOS*, que são grandes aliadas na monitorização da saúde física aliadas a pulseiras ou relógios inteligentes.

3.1 Aplicações *Mobile* de Saúde e *Fitness*

Existem várias aplicações que permitem monitorizar a saúde física, que incentivam ao desporto e o controlo da alimentação. Nesta secção irá ser feita a comparação das aplicações nativas usadas por cada marca de *smartwatch* ou *smartband* referenciadas na Tabela 2. Além disso foi também avaliada a *Google Fit*, por ser bastante usada e associar-se a outras aplicações que usam o sistema operativo *Wear OS*.

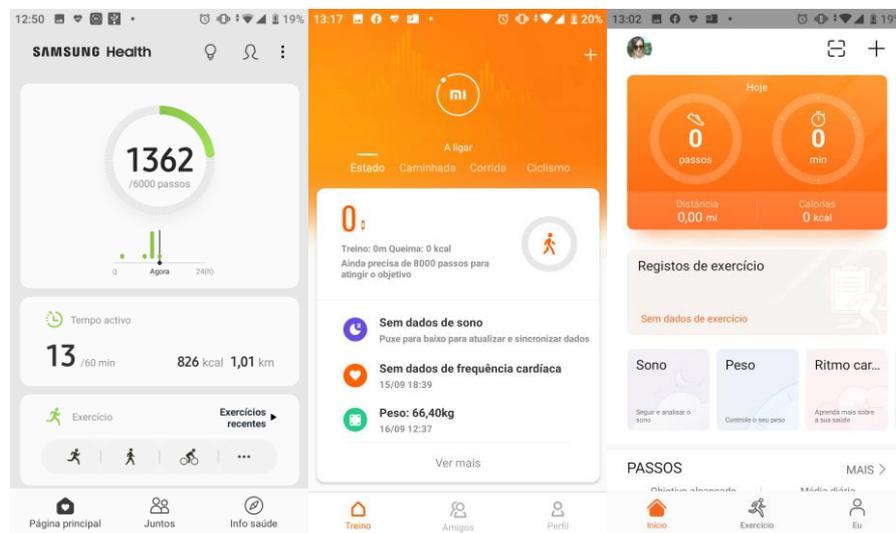
Aplicações Móveis	Monitorização de						Integração com outras aplicações	SO
	Saúde	Atividade	Alimentação	Sono	Objetivos	Personalidade		
<i>Fitbit</i>	√	√	√	√	√	X	Integrado com outras aplicações.	<i>IOS</i> e <i>Android</i>
<i>Huawei Health</i>	√	√	X	√	√	X	Interage com <i>Apple Health</i>	<i>IOS</i> e <i>Android</i>

							e ainda outras.
<i>Samsung Health</i>	✓	✓	✓	✓	✓	X	Integrado com outras aplicações <i>IOS e Android</i>
<i>Apple Health</i>	✓	✓	✓	✓	✓	X	<i>Mi Fit, Strava, MyFitnessPal, Sleep Cycle</i> <i>IOS</i>
<i>Mi Fit</i>	✓	✓	X	✓	✓	X	<i>Apple Health, Google fit</i> <i>IOS e Android</i>
<i>Google Fit</i>	✓	✓	X	✓	✓	X	Integrado com outras aplicações <i>Android</i>

Tabela 3 - Tabela comparativa de diferentes aplicações de saúde.

Através da Tabela 3 foi possível perceber que todas as aplicações possuem monitorização de atividade, saúde e sono embora umas sejam mais completas que outras. As principais diferenças entre as aplicações sentem-se na monitorização da alimentação onde o *scanner* de código de barras da *Fitbit* permite realizar tal monitorização. Ainda é possível retirar que a *Google Fit* apesar de não ser uma aplicação nativa de nenhum *Wearable* possui agregação com várias aplicações, o que é bastante significativo para abranger a informação de mais utilizadores e vários tipos de dispositivos *fitness*. Outra diferenciação está presente na monitorização de objetivos e recomendações, uma vez que todas as aplicações analisadas possuem metas diárias, desafios globais para motivar os utilizadores, algumas permitem partilhar dados com os amigos e ainda os desafias com é o caso da *Fitbit*, *Apple Health* e *Samsung Health*. A *Fitbit* permite ainda realizar treinos pela aplicação, uma vez que têm vídeo de orientação de exercício físico e *mindfulness* e comunidades de exercício e alimentação para os utilizadores se juntarem e competirem. *Apple Health* também tem tutoriais *mindfulness*.

As aplicações que controlam a alimentação (*Fitbit*, *Apple Health* e *Samsung Health*) também controlam a hidratação, nomeadamente o número de copos de água. A *Samsung Health* controla até o nível de cafeína. Todas as aplicações referidas, com exceção da *Mi Fit*, conseguem controlar os passos e minutos de atividade através dos próprios sensores do telemóvel. A *Apple Health* e a *Fitbit* permitem inserir dados relativos à saúde do aparelho reprodutor feminino, permitindo, entre outras funções, controlar a toma de contraceptivos orais. Na Figura 5 encontra-se representado o menu inicial das aplicações referenciadas acima.



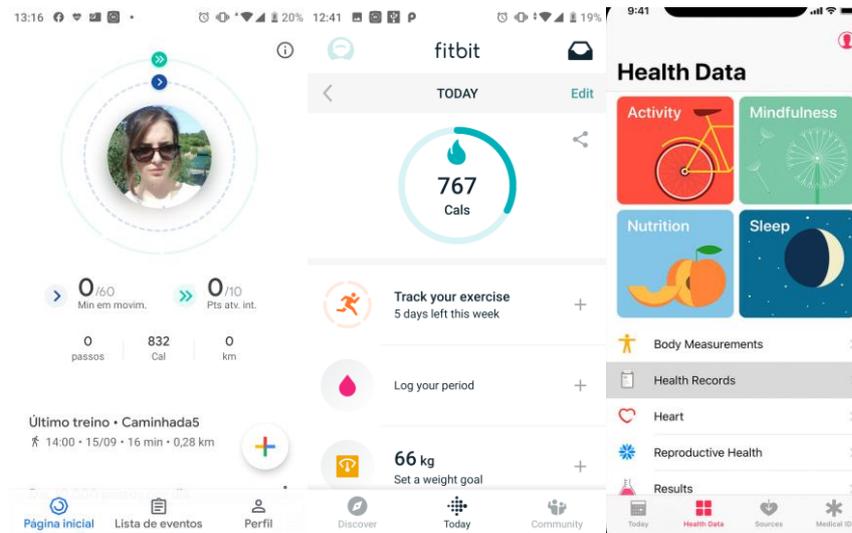


Figura 5 - Menu inicial das aplicações *Samsung Health*, *Mi Fit*, *Huawei Health*, *Google Fit*, *Fitbit*, *Apple Health*, respectivamente.

3.2 APIs de Saúde e Fitness

As *APIs* (*Application Programming Interface*) são a maneira mais eficaz de enviar dados de um sensor para uma equipa de desenvolvimento que precisam das informações para criar uma aplicação. A *API* atua como o veículo para os fabricantes de dispositivos gerarem negócios em torno de *smartwatches*, *headsets* de realidade virtual ou *smartbands* que medem frequência cardíaca. Sem eles, os *wearables* não fariam muito sentido uma vez que é incoerente colecionar dados sem lhes dar uma forma. Para este estudo foi feita uma pesquisa de forma a conhecer as diversas *APIs* usadas pelas aplicações móveis pesquisadas no capítulo anterior. A *Google* disponibiliza várias *APIs* conectadas a *wearables*. A primeira é a *API REST* da *Google Fit*, a interface de programação de aplicações do *Google Fit*, a plataforma para a atividade física dos utilizadores, o segundo é o antigo *Android Wear APIs* agora designado por *Wear OS* para o sistema operacional da empresa para *wearables*:

- **API REST da Google Fit:** API REST, que permite armazenar e ter acesso a dados da *Google Fitness Store*. Esta API oferece a possibilidade de realizar operações (criação, agregação ou exclusão) usando as fontes/conjuntos de dados colecionados pelo sensor. A API da *Google* funciona com três tipos de dados: completamente públicos (dados padrão fornecidos pela plataforma para qualquer aplicação), dados privados e informações que podem ser compartilhadas com outras pessoas. Esta API opera sob o protocolo *OAuth2*. Por este motivo, os *developers* precisam ter uma Conta *Google*, ter acesso a consola do *Google Developers* e solicitar as credenciais por meio de um *token* de acesso.
- **Wear OS API:** o *Android* oferece uma série de APIs para o desenvolvimento de aplicações e jogos para os dispositivos da *Google*. Desenvolvido sempre num ambiente de desenvolvimento na linguagem de programação *Java*.
- **Fitbit API:** assim como todas as outras APIs, a API da *Fitbit* permite aos desenvolvedores conectar e desenvolver aplicações e serviços a partir dos dados capturados pelos *wearables* da marca (pulseiras ou *smartwatches*). Os protocolos de autorização usados são *OAuth2* e *OAuth1*. É capaz de interpretar vários idiomas: espanhol, francês, alemão, japonês e inglês (“Wearables based on APIs, an excellent business opportunity,” 2018).
- **Tizen APIs:** O sistema operacional baseado em Linux para dispositivos móveis possui todos os tipos de APIs - funções de autenticação, ligadas ao *framework*, conteúdo, desenvolvimento gráfico, experiência do utilizador, interface, entre outros.
- **Health API:** A API da Apple permite que aplicações de saúde e *fitness* compartilhem dados, tanto com o novo *iOS 8 Health* como entre si. O *Health* pode servir como um *hub* para todas as aplicações de saúde e *fitness* num dispositivo

IOS. Isto permite que os consumidores tenham mais supervisão de sua saúde e condicionamento físico.

3.3 Ferramentas e Tecnologias Utilizadas

No desenvolvimento deste projeto de dissertação, foram várias as ferramentas utilizadas, de entre as quais se podem destacar a plataforma *Web Big Five*, o *Android Studio* para a criação da aplicação móvel e o *Firebase* para modelação da base de dados.

3.3.1 Plataforma *Web Big Five*

Plataforma desenvolvida para ambiente *Web*, que se encontra no seguinte endereço *web* <http://crowdsensing.di.uminho.pt/main>.



Figura 6 - Vista inicial da plataforma *web*.

Nesta plataforma é possível realizar o teste de personalidade BFI com *mini-markers* desenvolvidos por (Saucier, 1994) em três línguas diferentes, designadamente português, inglês e espanhol. Além do utilizador avaliar de 1 a 9 os adjetivos, ainda possível seleccionar os marcadores que mais caracterizam o indivíduo que está a realizar o teste. São ainda pedidas informações relativas a idade e ao sexo do mesmo. No final, o indivíduo obtém o resultado do teste que efetuou.



Figura 7 - Resultado da realização do teste de personalidade na plataforma web.

Na Figura 7 é possível ver o resultado de um teste realizado na plataforma, onde é possível observar a expressão de cada traço de personalidade e de que forma essa informação se distribui graficamente no pentágono.

3.3.2 Android

Android é o sistema operativo (SO) móvel do *Google* baseado em *Linux* desenvolvido pela *Open Handset Alliance* (OHA). Este SO está presente em diversos aparelhos de diversas fabricantes, como *Samsung*, *Motorola*, *LG*, e *Sony*, *Huawei*, *Xiaomi*, sendo atualmente a plataforma mobile mais popular do mundo. As razões que levaram à escolha da aplicação para este software foram, para além da sua popularidade por todo o mundo, com cerca de 80% da quota de mercado, a ampla gama de marcas que o utilizam. A *Play Store*, do *Android*, permite que qualquer programador possa publicar os seus conteúdos, tendo um custo reduzido, o que possibilita que milhares de apps de programadores independentes, estejam presentes não só na *PlayStore*, como também em plataformas dedicadas (Bacelar, 2019).

3.3.3 Android Studio

O *Android Studio* é o ambiente de desenvolvimento integrado (IDE) oficial para o desenvolvimento de aplicações *Android* e é baseado no *IntelliJ IDEA*. O *Android Studio* oferece vários recursos para criação de aplicações *Android*, tais como:

- Editor de código e das ferramentas de desenvolvimento avançados do *IntelliJ*;
- Sistema de compilação flexível baseado no *Gradle*;
- Emulador rápido, um ambiente consolidado para todos os dispositivos *Android*;
- *Instant Run* para aplicar alterações a aplicações em execução;
- Integração com *GitHub* para ajudar a criar recursos comuns das aplicações e importar exemplos de código;
- Ferramentas tanto de teste, como de verificação de código suspeito;
- Compatibilidade embutida com o *Google Cloud Platform* (*Android Developers*, n.d.).

3.3.4 **Firestore**

O *Firestore* é um *BaaS (Backend as a Service)* para aplicações Web e Mobile do Google, que foi lançado em 2004 e tem sofrido um grande crescimento ao longo dos anos. Esta ferramenta não é apenas uma base de dados *NoSQL*, uma vez que oferece um conjunto de produtos interligados entre si, como é o caso *Realtime Database, Authentication, Cloud Firestore, Storage, Cloud Functions*, entre outros.

A *Realtime Database*, base de dados utilizada nesta dissertação, permite o armazenamento e sincronismo dos dados entre utilizadores e dispositivos em tempo real, com uma base de dados *NoSQL* hospedada na nuvem. Os dados atualizados são sincronizados em todos os dispositivos conectados em segundos. Também possibilita o acesso aos dados estando a aplicação off-line, o que é vantajoso uma vez que um utilizador pode ter acesso, independentemente da conectividade de rede.

O *Firestore* é um sistema completo, robusto, adequado para prototipagens e de fácil uso. As principais vantagens desta ferramenta seguem-se por ser facilmente escalável, rápida implementação, segura, com uma estrutura pronta. No entanto, tal como qualquer outra ferramenta tem desvantagens, sendo que a principal é a limitação da plataforma, uma vez que o plano gratuito é limitado a 100 utilizadores em simultâneo e 1GB de armazenamento. Apesar desta desvantagem, esta ferramenta adequa-se a este projeto, uma vez que não se prevê um volume muito grande de dados, nem de utilizadores.

3.3.5 **Google Fit**

A *Google Fit* é a aplicação de *fitness* que se centraliza na simplicidade e em executar algumas peças essenciais. Esta tem um *design* intuitivo, o que faz com que tenha uma baixa curva de aprendizagem. Existem três seções principais da aplicação *Google Fit*, designadamente Página inicial, Diário e Perfil.

A Página inicial mostra uma visão geral das atividades e métricas de saúde do próprio dia, incluindo minutos de movimento e pontos cardíacos, calorias queimadas, distância percorrida, dados de frequência cardíaca e peso. Na secção Diário é apresentado uma lista de todas as atividades gravadas, caminhadas, voltas de bicicletas, treino de mergulho. Por último na página Perfil, o utilizador pode ajustar os minutos de movimento e pontos cardíacos que pretende alcançar, bem como as informações pessoais, como peso e a altura.

O que torna o *Google Fit* diferente das outras aplicações são os Minutos ativos e os Pontos cardio, que resultaram do trabalho que a Google realizou com a *American Heart Association* para criar dois objetivos, com base nas recomendações de atividades da *Heart Association* (Google Commerce Ltd, 2019). A partir da versão 2.16.22, o *Google Fit* passou a ter um modo escuro que ajuda a reduzir o gasto da bateria. Esta aplicação é ainda compatível diversas aplicações, destacamos as seguintes:

- *Strava;*
- *Instant;*
- *Google Calendar;*
- *7 Minute Workout;*
- *Glow;*
- *Instant Heart Rate;*
- *Withings Health Mate;*
- *Mi Fit*
- *Sleep as Android;*
- *Calm;*

- *Daily Yoga*;
- *Lose it!*.

Além da compatibilidade com estas aplicações, todos os *smartwatches* com *Wear OS* funcionam com o *Google Fit*. Já com as *smartbands*, este tem baixa conexão, uma vez que apenas as pulseiras dos fabricantes *Huawei* e *Xiaomi* têm melhor compatibilidade de recursos, devido à partilha de dados com as suas aplicações de gestão de origem (Sumra, 2018).

3.3.6 KNIME

O *Konstanz Information Miner (KNIME)* é um ambiente modular, que permite fácil montagem visual e execução interativa de um *pipeline* de dados. Este é projetado como uma plataforma de ensino, pesquisa e colaboração, que permite a integração simples de novos algoritmos e ferramentas, bem como manipulação de dados ou métodos de visualização na forma de novos módulos ou nós (Lartillot, Toiviainen, & Eerola, 2008).

Algumas de suas principais características incluem a combinação de dados, uma poderosa análise, extensões de *big data* e vinculação de meta-nó, funções estatísticas, algoritmos de *Machine Learning* e algoritmos preditivos avançados.

O *KNIME* é uma plataforma *open source* que auxilia o utilizador a descobrir o potencial de dados ocultos, extrair novos *insights* e prever com precisão novas regras (FinancesOnline, n.d.).

3.3.7 Python

O *Python* é uma linguagem de programação interpretada de alto nível e *open source* que é adequada para código científico e de engenharia, uma vez que é rápida o suficiente para ser útil imediatamente, mas também flexível o suficiente para ser acelerada com extensões adicionais (Oliphant, 2007). Esta linguagem, possui muitas bibliotecas distintas e bem

documentadas que facilitam as análises de computação científica de alto nível, como é o caso da análise de dados. Existem várias bibliotecas que permitem visualizar, processar e analisar de dados, alguns exemplos:

- **numpy**, uma biblioteca de processamento numérico que suporta matrizes multidimensionais;
- **scipy**, uma biblioteca de processamento científico fornecendo ferramentas avançadas para análise de dados, incluindo regressão, integradores e solucionadores de equações diferenciais ordinárias, álgebra linear e funções estatísticas;
- **pandas**, dados e uma biblioteca de manipulação de dados e tabelas que oferece funcionalidade semelhante a *spreadsheets* como o *Excel*TM;
- **matplotlib**, uma biblioteca de plotagem com ferramentas para exibir dados de várias maneiras (Badenhorst et al., 2019).

3.3.8 Jupyter Notebook

O *Jupyter* suporta computação científica e ciência de dados numa vasta diversidade de linguagens de programação (*Python*, *Julia*, *R*, etc.) e fornece ainda recursos básicos de reprodutibilidade e colaboração (Fernando Perez, 2015).

O *Jupyter Notebook* é uma aplicação *web open-source* para computação exploratória, que permite criar e partilhar documentos que contém código ativo, equações, permite a visualização gráfica e descrição textual. As principais utilidades desta ferramenta são as seguintes:

- Limpeza de dados e transformação;
- Simulação numérica;
- Modelos estatísticos;

- Visualização de dados;
- *Machine Learning* (Budkewicz, 2018).

Capítulo 4

Design e Implementação

4.1 Proposta e Design

Após estudo e análise de múltiplas aplicações e APIs, determinou-se usar a API da *Google Fit*, uma vez que é mais abrangente e intuitiva. Tal como tem vindo a ser partilhado, o foco máximo deste projeto permanece nos testes psicológicos *Big Five* e a sua possível correlação entre eles e alguns dados biométricos. Estes dados são possíveis obter a partir da plataforma móvel que será desenvolvida. A aplicação móvel estará disponível para dispositivos *Android*, sendo designada de *Personality and Track Health (Path)*. Na Figura 8 é possível ver logótipo criado que vai ser o rosto desta aplicação.



Figura 8 - Logótipo da aplicação móvel.

Seguidamente, iniciou-se a criação dos *mockups* de forma a criar a imagem geral da aplicação e a adquirir a perceção do conteúdo que será necessário. Os *mockups* são representações de objetos e produtos, bem elaborada, com um design muito próximo ao design final do produto em questão. A simulação torna-se palpável e as pessoas são estimuladas a rever a parte visual do projeto. Na sequência de imagens abaixo é possível ter a perceção da imagem geral da aplicação, e das suas principais vistas e funções.



Figura 9 - *Mockup do login.*

A Figura 9 é referente à vista de *login*, onde será realizada a autenticação na aplicação por meio de uma conta *Google*, de forma a permitir a conexão entre a *Google Fit* e a *Path*. Ao ser validada a autenticação, é possível observar as características da página inicial na Figura 10, que se pretende que o utilizador possa ver a informação do batimento cardíaco, das horas de sono, do número de passos, e o resultado do teste psicológico caso já tenha sido realizado. Caso contrário a imagem de um pentágono perfeito é a que irá aparecer. Dentro da vista

principal existem vários componentes que vão ser apresentadas em subvistas do menu principal, com mais informação.



Figura 10 - Mockup da vista de menu e da monitorização de atividade.

Começando pela monitorização de atividade que além do gráfico, da informação da quantidade de passos dada, dos quilómetros percorridos, das calorias consumidas e do tempo de atividade, ainda é possível recuar uma vista de forma a ver os dados do dia anterior. O mesmo acontece na Figura 11, além do gráfico da variação do batimento cardíaco também é possível visualizar o último batimento e a média diária. Ainda em termos de gráficos, além dos diários, existe também a vista semanal, mensal e anual.

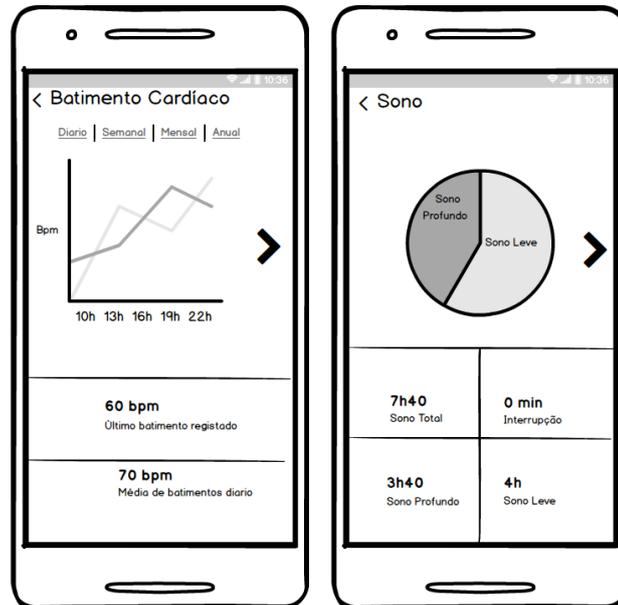


Figura 11 - *Mockup* da vista da monitorização do batimento cardíaco e sono.

Na vista relativa à monitorização do sono (Figura 11), é possível visualizar um gráfico com a representação do sono leve e sono profundo, e ainda com a representação do tempo destes dois tipos de sono, tempo total de sono e interrupções nos mesmos.

Através da Figura 12, é possível verificar o *mockup* que representa o teste de personalidade com os quarenta marcadores, em que após serem seleccionas a sua cor fica realçada.

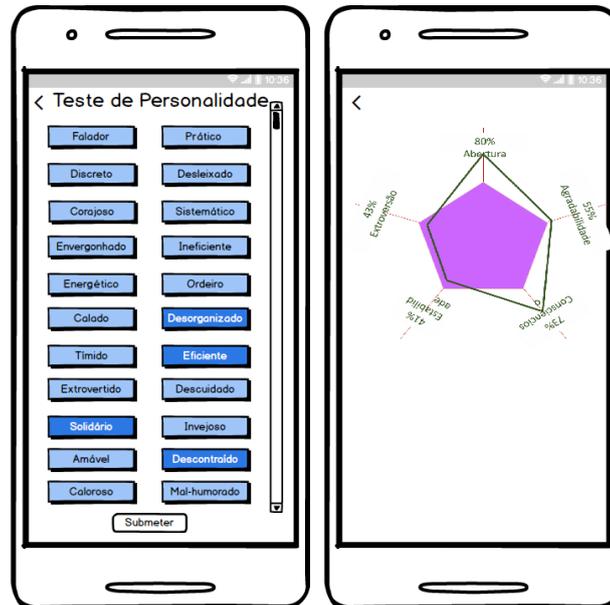


Figura 12 - *Mockup* da vista do teste de personalidade e o resultado.

Após submeter os testes, a vista que será apresentada será a visível na Figura 12, com uma representação mais significativa em alguns traços de personalidade caso a aresta do pentágono seja mais acentuada.

4.2 Modelo de Dados

Tendo os *mockups* definidos foi altura de definir o modelo de dados. Para isso utilizou-se uma base de dados *NoSQL* criada no *Firebase* com auxílio da ferramenta *Realtime Database*. Esta ferramenta possui é uma base de dados hospedada na *Cloud*, os dados são armazenados como *JSON* e sincronizados em tempo real com todos os clientes conectados. A *Realtime Database* possui diversas funcionalidades incorporadas na *API* do *Firebase*, como a possibilidade de funcionar *offline* e sincronizar automaticamente ao retornar a conexão, atualizar e sincronizar em tempo real a base de dados para os utilizadores conectados.

A base de dados está no formato de *NoSQL*, ou seja, produz uma Base de Dados que não possui como padrão o sistema de tabelas e relacionamentos entre dados, tratando cada informação como um nó de um gigantesco tronco que, basicamente, é a *Root* da Base de Dados. Este formato tem a vantagem de lidar bem com um grande número de dados, uma vez que o facto de não existir um relacionamento entre os nós, diminuiu o tempo de *delay* na procura/inserção. Para a integração do *Firebase* na aplicação foi necessário adicionar e configurar o *SDK* do *Firebase Realtime Database*. De seguida foram criadas referências do *Realtime Database*. Estas referências são relativas aos dados *JSON*, como por exemplo:

“https://diesel-amulet-235316.firebaseio.com/UserPerso/2019-07-27_15:44_joaocdomingues31@gmail.com”, para definir ou aceder aos dados relativos ao resultado dos testes psicológico do utilizador *joaocdomingues31@gmail.com* na data 2019-07-27.

As referências são geradas nas atividades correspondentes, dependendo da categoria em que estão inseridas, por exemplo a referência que usa *UserPerso* só é criada na atividade *PersonalityActivity*, uma vez que é nela que está implementado o código necessário para gerar o teste psicológico. A estrutura de dados está organizada como se vê na Figura 13 abaixo:

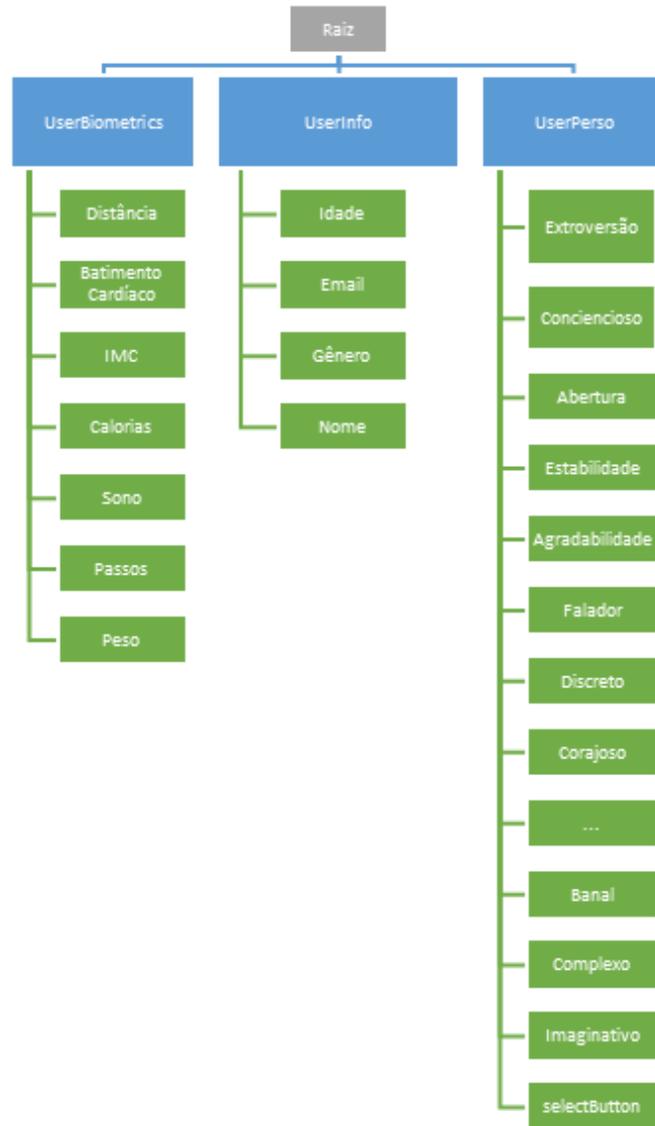


Figura 13 - Árvore da base de dados.

Como é possível ver através da figura anterior (Figura 13), os dados estão organizados três categorias, denominadas por:

UserBiometrics - Nesta secção são guardados os dados biométricos dos utilizadores, sendo que a informação é guardada sempre que o utilizador acede ao menu principal da

aplicação. As informações que se encontram nesta categoria são: a distância percorrida, o batimento cardíaco, o valor do IMC, calorias consumidas, horas de sono, passos percorridos e peso atual. Tal como descrito, na Figura 14, é possível ver um exemplo no *Firebase* essas subcategorias referidas acima.

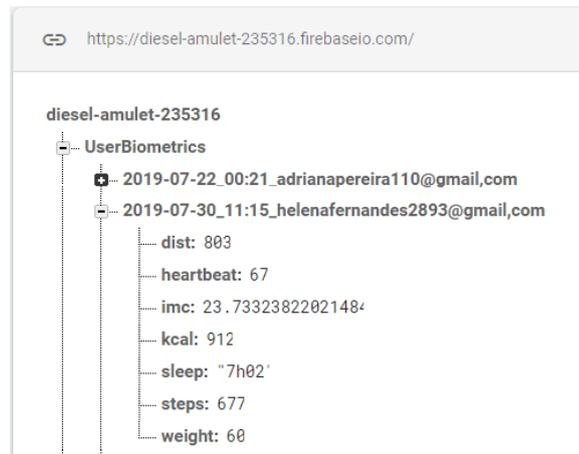


Figura 14 - UserBiometrics com as suas subcategorias.

UserInfo – Nesta secção são guardados os dados do perfil dos utilizadores. Esta informação é guardada quando o utilizador efetua o *login* ou quando atualiza as informações na atividade perfil, tal como é possível verificar na Figura 15, está categoria é composta pela idade, email, género, e nome do utilizador.

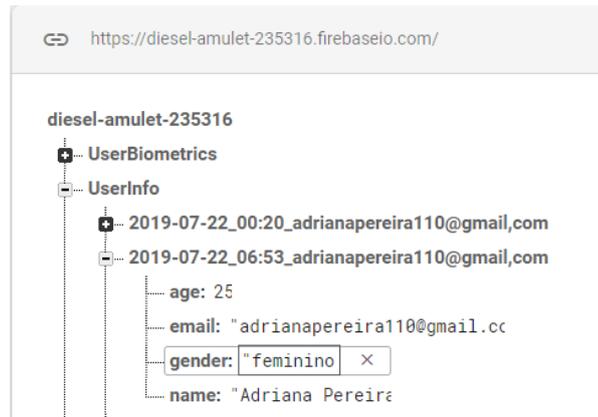


Figura 15 - UserInfo com as respetivas subcategorias.

UserPerso – Nesta categoria são guardados os resultados dos testes psicológicos, cada vez que este é efetuado e guardado. Verifica-se na Figura 16, que nesta categoria são guardados os valores 0 a 9 dos 40 adjetivos do teste psicológico, bem como nome dos botões selecionados e resultado de cada um dos 5 traços de personalidade.

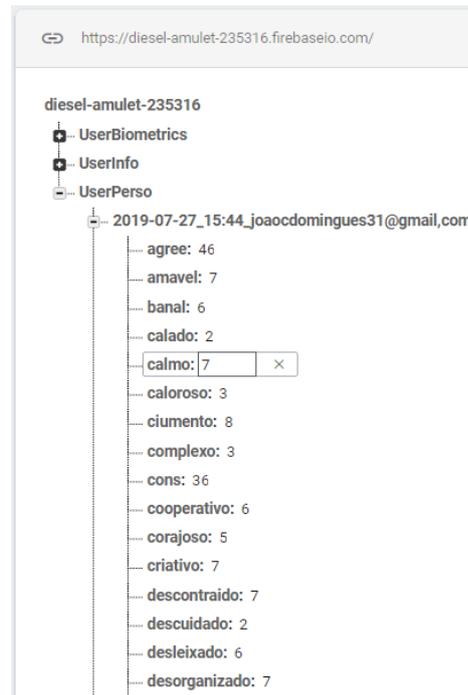


Figura 16 - UserPerso com algumas das suas subcategorias.

Cada inserção na BD é alocada à categoria correspondente, com o identificador, estruturado da forma como se pode observar no exemplo:

"2019-07-28_23:37_helenafernandes2893@gmail,com".

Este identificador é composto por data (YYYY-MM-DD) separado por "_", seguido de hora (00:00) novamente separado por "_", e é finalizado com o email em que todos os pontos finais tiveram de ser substituídos por vírgulas, uma vez que o ponto é um carácter especial no *Firebase*.

4.3 Aplicação Móvel (Path)

A aplicação móvel criada para este projeto destina-se a dispositivos mobile *Android*, versão mínima especificada pelo identificador do nível da *API* da plataforma executado é a 23, e está projetada para o nível 28 da *API*, o que corresponde a versão do *Android 6.0* até ao *Android 9.0*.

A aplicação para recolha de dados biométricos, em conjunto com os dados dos testes psicológicos, foi desenvolvida em *Java* através da ferramenta *Android Studio*. Para o desenvolvimento foi usada a *API* do *Google Fit*, de forma a conseguir ler os dados relativos aos passos, batimento cardíaco, calorias, distância e os dados corporais relativos ao peso e altura. Uma vez que *API* do *Google Fit* não está suficientemente desenvolvida e com falta de documentação relativamente ao registo de sono, estes são inseridos manualmente pelo utilizador.

Relativamente aos requisitos mínimos para a utilização desta aplicação, é necessário ter um *smartphone Android* com versão mínima 6 e ter já instalada a aplicação *Google Fit*, uma vez que a *PaTH* lê os dados registados a partir dessa aplicação. É ainda necessário ter uma *smartwatch* ou *smartband* que utilize o software *Wear OS*, ou então um dispositivo que use uma das aplicações associadas com o *Google Fit*, como é o caso da *Mi Fit*. A aplicação foi testada com vários *smartphones* e com dispositivos de monitorização da marca *Xiaomi*, como a *Mi Band 2*, *Mi Band 4* e *AMAZFIT* e, como tal, requer a instalação da *Mi Fit*.

De seguida, é necessário adicionar a conta *Google* à aplicação *Mi Fit* sendo que, para isso, é necessário ir à área de perfil do utilizador, tal como se pode ver na primeira imagem da Figura 17, e adicionar a conta *Google Fit*. Desta forma as aplicações já comunicam entre si.

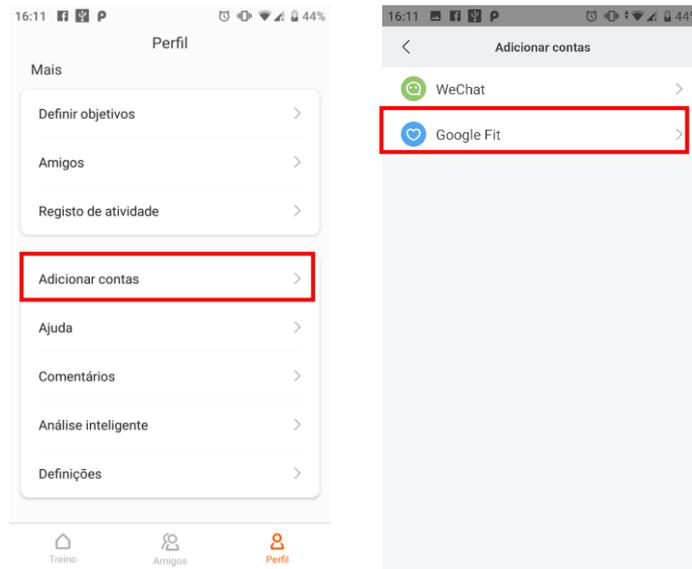


Figura 17 - Adicionar conta *Google Fit* a aplicação *Mi Fit*.

Seguidamente, após todos os requisitos cumpridos, é possível aceder à aplicação. Ao iniciá-la, o utilizador depara-se com a vista da Figura 18, onde é efetuado o *login* através de uma conta *Google*, após autenticado e depois de ter aceite as permissões. Temos a acesso a página inicial (Figura 18).

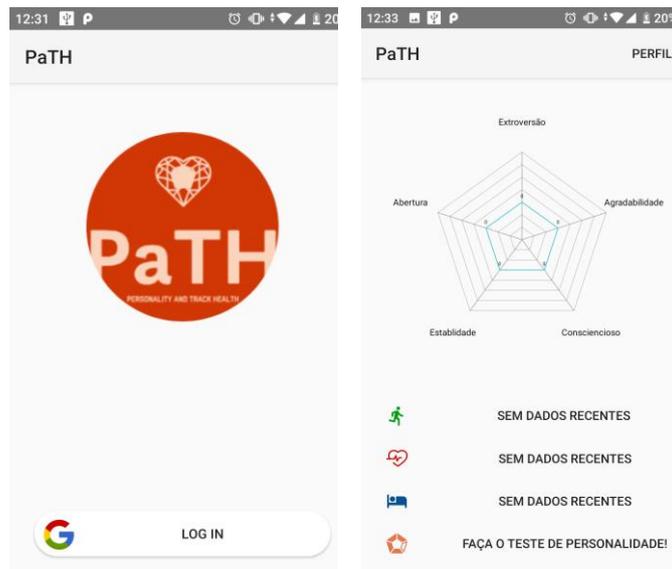


Figura 18 - Login Path e página inicial.

Os dados relativos ao sono são inseridos pelo utilizador, tal como aparece na Figura 19.



Figura 19 - Vista para inserção das horas de sono.

A partir dos dados recolhidos da aplicação *Google Fit* relativos ao peso e à altura, é calculado o IMC do utilizador, este pode escolher quais as opções que quer ver no menu principal, além das 4 opções de base: os passos, o último registo do batimento cardíaco, o último registo de sono inserido e o teste de personalidade. As informações opcionais são o IMC, as calorias consumidas, a distância percorrida e o peso.

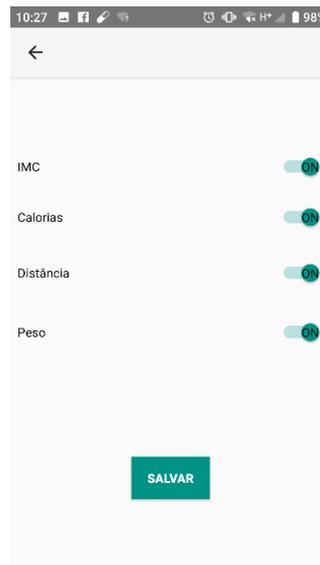


Figura 20 - Vista das opções de visualização no menu principal.

Na secção da realização do teste psicológico, o utilizador tem que escolher os botões com as características que mais o caracterizam e ainda avaliá-las de 1 a 9, arrastando a *seekbar* consoante o valor mais adequado. Após ter clicado no botão salvar, é calculado e guardado em *Shared Preference* o valor dos 5 traços de personalidade: Agradabilidade, Extroversão, Estabilidade, Consciencioso e Abertura, consoante os valores escolhidos pelo utilizador. É também guardado numa *String* o nome da variável dos botões selecionados. Posteriormente, o utilizador vai ser redirecionado para o menu principal onde já será visível o resultado do teste realizado.



Figura 21 - Realização do teste psicológico.

Após ter sido realizado pelo menos um teste psicológico, o resultado vai aparecer no pentágono, expressando o ou os traços mais relevantes de entre 5 possíveis (Figura 22). O sistema vai efetuar o cálculo de um dos traços de personalidade, sendo que, para isso, é usada a formulação de Saucier. Abaixo encontra-se desmistificada a formulação para cada traço de personalidade. Para tal, é necessário ter em conta que cada adjetivo que aparece na fórmula é substituído pelo valor que o utilizador colocou para o mesmo ao realizar o teste, sendo que o valor pode ir de 0, caso não tenha avaliado esse adjetivo, a 9.

- **Cálculo da Extroversão**

Extroversão = Falador + Extrovertido+ Corajoso + Enérgico + (10 - Tímido) + (10- Calado) + (10- Envergonhado) + (10 - Discreto);

- **Cálculo da Agradabilidade**

Agradabilidade = Solidário + Caloroso + Amável+ Cooperativo+ (10- Frio) + (10- Distante) + (10- Grosseiro) + (10- Ríspido);

- **Cálculo da Consciencioso**

Consciencioso = Ordeiro + Eficiente + Sistemático + Prático + (10- Desorganizado) + (10- Desleixado) + (10 - Ineficiente) + (10 - Descuidado);

- **Cálculo da Estabilidade**

Estabilidade = Calmo + Descontraído + (10 - Mal-humorado) + (10 - Ciumento) + (10- Temperamental) + (10 - Invejoso) + (10 - Melindroso) + (10 - Inquieto);

- **Cálculo da Abertura**

Abertura = Criativo + Imaginativo + Filosófico+ Intelectual+ Complexo + Profundo + (10 - Mediano) + (10 - Banal);



Figura 22 - Página inicial após a realização do teste de personalidade.

Esta aplicação, além de permitir ao utilizador ver as alterações da sua saúde mental, permite ainda visualizar gráficos da sua saúde física. Para isso, basta clicar no botão referente aos passos e somos direcionados para a vista da Figura 23, onde é possível visualizar a evolução diária de um utilizador e ainda a evolução semanal.

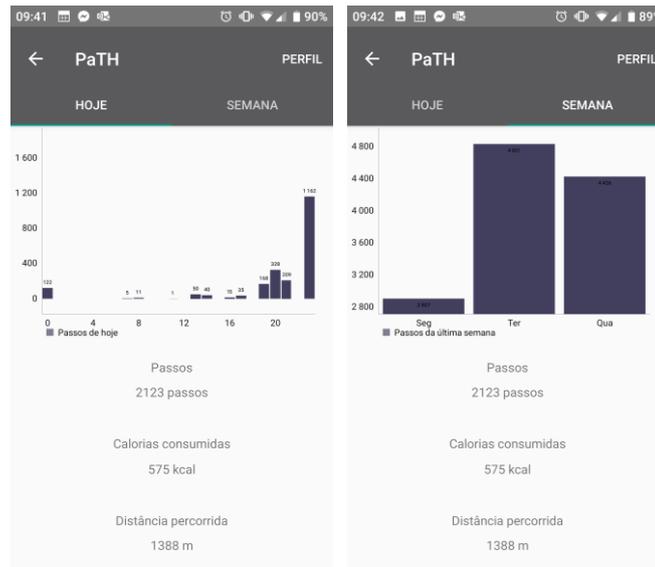


Figura 23 - Vista dos gráficos relativos ao número de passos do diário e semanal.

Entretanto, é ainda possível verificar a evolução do batimento cardíaco ao longo do dia, e o valor do último registo, bastando, para isso, clicar no botão onde aparece o batimento cardíaco e ao mesmo tempo somos redirecionadas para o gráfico da variação dos batimentos (Figura 24).

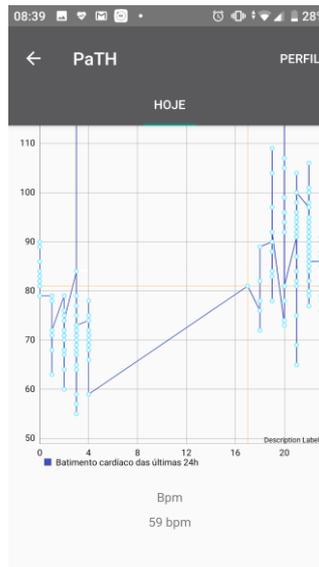


Figura 24 - Gráfico das variações do batimento cardíaco nas últimas 24h.

Por sua vez, o utilizador pode ainda aceder ao seu perfil e inserir ou alterar a data de nascimento e o sexo, como se pode ver na Figura 25. Ao guardar as alterações, o sistema vai atualizar os dados, calcular a idade do utilizador e guardar em *Shared Preference* a informação da data de nascimento e o género. Esta vista dá ainda a opção ao utilizador de terminar a sessão.

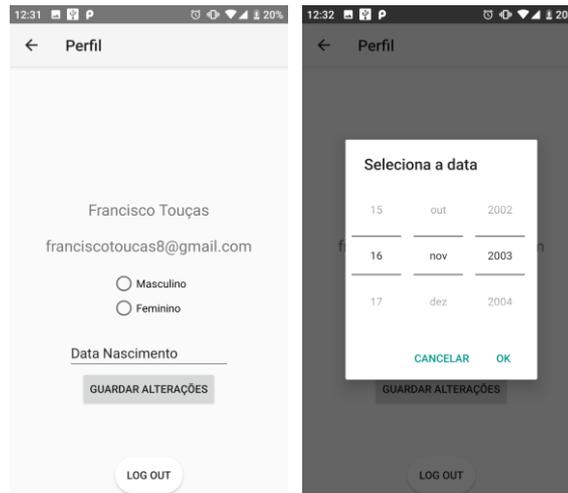


Figura 25 - Vista do perfil do utilizador antes de preencher os dados e adicionar a data de nascimento.

Capítulo 5

Caso de Estudo

Após a aplicação ter sido desenvolvida, esta foi libertada e distribuído um protótipo *APK* por uma população de, aproximadamente, 15 pessoas que usam o português como sua língua nativa, sendo que cada pessoa poderia, ou não, instalar e utilizar a aplicação. O período de recolha de dados detém-se entre 22 de julho de 2019 até 16 de setembro de 2019. Das cerca de 15 pessoas que receberam o *APK* apenas 11 realmente entraram na aplicação, dessas 11 apenas 9 efetuaram o teste de personalidade.

Os dados recolhidos através do *Firebase* estão dividido em três *datasets* *DataPerso*, *DataInfo*, *DataBiometrics*. Destes existem 31 registos no *DataPerso*, 90 registos no *DataInfo*, uma vez que cada vez que se abre a aplicação e inserido informação neste *dataset* independentemente de ser ou não um novo utilizador, e por último temos 241 registos biométricos.

O *DataInfo* referente a informação do utilizador é constituída pelas seguintes colunas:

- *id* – o identificador é uma *string* composta por data, hora e email;
- *age* – inteiro correspondente à idade do utilizador;
- *email* – *string* composta email do utilizador;
- *gender* – *string* com o género do utilizador;
- *name* – *string* com nome do utilizador.

O DataPerso contém a informação acerca do resultado do teste psicológico, que é constituído pelas seguintes colunas:

- id – o identificador é uma *string* composta por data, hora e email;
- agree – inteiro com o resultado de agradabilidade do teste psicológico;
- amavel – inteiro de 0 a 9;
- banal – inteiro de 0 a 9;
- calado – inteiro de 0 a 9;
- calmo – inteiro de 0 a 9;
- caloroso – inteiro de 0 a 9;
- ciumento – inteiro de 0 a 9;
- complexo – inteiro de 0 a 9;
- cons – inteiro com o resultado de consciencioso do teste psicológico;
- cooperativo – inteiro de 0 a 9;
- corajoso – inteiro de 0 a 9;
- criativo – inteiro de 0 a 9;
- descontraído – inteiro de 0 a 9;
- descuidado – inteiro de 0 a 9;
- desleixado – inteiro de 0 a 9;
- desorganizado – inteiro de 0 a 9;

- discreto – inteiro de 0 a 9;
- distante – inteiro de 0 a 9;
- deficiente – inteiro de 0 a 9;
- enérgico – inteiro de 0 a 9;
- envergonhado – inteiro de 0 a 9;
- extra – inteiro com o resultado de extroversão do teste psicológico;
- extrovertido – inteiro de 0 a 9;
- falador – inteiro de 0 a 9;
- filosofico – inteiro de 0 a 9;
- frio – inteiro de 0 a 9;
- grosseiro – inteiro de 0 a 9;
- imaginativo – inteiro de 0 a 9;
- ineficiente – inteiro de 0 a 9;
- inquieto – inteiro de 0 a 9;
- intelectual – inteiro de 0 a 9;
- invejoso – inteiro de 0 a 9;
- malHumorado – inteiro de 0 a 9;
- mediano – inteiro de 0 a 9;
- melindroso – inteiro de 0 a 9;
- open – inteiro com o resultado de abertura no teste psicológico;

- ordeiro – inteiro de 0 a 9;
- pratico – inteiro de 0 a 9;
- profundo – inteiro de 0 a 9;
- ríspido – inteiro de 0 a 9;
- selectButton – *string* composta pelo nome dos botões seleccionadas na realização do teste psicológico separado por “_”;
- sistemático – inteiro de 0 a 9;
- solidário – inteiro de 0 a 9;
- stable - inteiro com o resultado de estabilidade do teste psicológico;
- temperamental - inteiro de 0 a 9;
- tímido - inteiro de 0 a 9.

DataBiometrics inclui os registos biométricos de cada utilizador composto pelas seguintes colunas.

- id - o identificador é uma *string* composta por data, hora e email;
- dist – *float* com a distância percorrida pelo utilizador;
- heartbeat – registo do último batimento cardíaco;
- imc – *float* com o valor calculado do IMC do utilizador;
- kcal – inteiro com as calorias consumidas;
- sleep – *string* com as horas de sono dormidas;

- `steps` – inteiro com os passos percorridos;
- `weight` – *float* com o peso atual do utilizador.

5.1 Tratamento dos Dados

Uma vez que os dados exportados do *Firebase* se encontram no formato *JSON*, estes tiveram de se convertidos para *CSV*. Seguidamente foi criado um script em *Python* através do *Jupyter Notebook*, para tratamento dos dados, para isso foi necessário exportar a três *datasets*. Após o carregamento dos dados, estes tiveram que ser formatados. Iniciou-se por alterar o nome de todas as colunas para remover os caracteres especiais “`_|`”. Seguidamente no primeiro, *DataInfo*, foram removidos os registos duplicados, mantendo apenas o último registo. Visto que apenas faz sentido manter um registo por utilizador, e como a idade pode alterar com o decorrer do tempo, decidiu-se manter apenas a última observação para cada utilizador.

No seguinte, *DataPerso*, foi dividido o `id`, que é composto por data, hora e email, em mais duas colunas `email` e `dia`, desta forma é mais simples aceder aos dados por utilizador. A coluna `selectButton` também foi repartida em 40 colunas que representam os 40 adjetivos, onde estas foram preenchidas como 0 ou 1, de acordo se estavam presentes ou não na coluna `selectButton`. Os valores dos 5 traços de personalidade foram convertidos em percentagem, tendo em conta o máximo valor que cada um pode ter. Tendo em conta que existiam dois testes psicológicos que não continham informação acerca dos botões clicados, estes dois registos foram eliminados.

No *dataset* *DataBiometrics*, uma vez que existiam registos incompletos e muito próximos, estes foram agrupados por dia, tendo em conta o email do utilizador com batimento cardíaco igual. Desta forma, foi possível diminuir os registos de 241 para 91. Foram preenchidos todos os valores em falta, tendo sido substituída a *string* “Sem Registo Recente” por “0”, separado

as horas dos minutos, convertidas as horas em minutos e juntar na coluna *sleepMinute*, o total em minutos de sono. Seguidamente, preenchidos os valores em falta das colunas *sleepMinute*, *heartbeat*, *steps*, *dist*, *kcal* substituídos os pela média dos registos de cada utilizador.

5.2 Análise dos Resultados

Iniciou-se a análise de dados através da criação de um *workflow* no *KNIME* (Figura 26) de forma a explorar melhor a visualização e análise dos dados. Para isso, foi configurado um *metanode* para cada ficheiro um dos *datasets*.

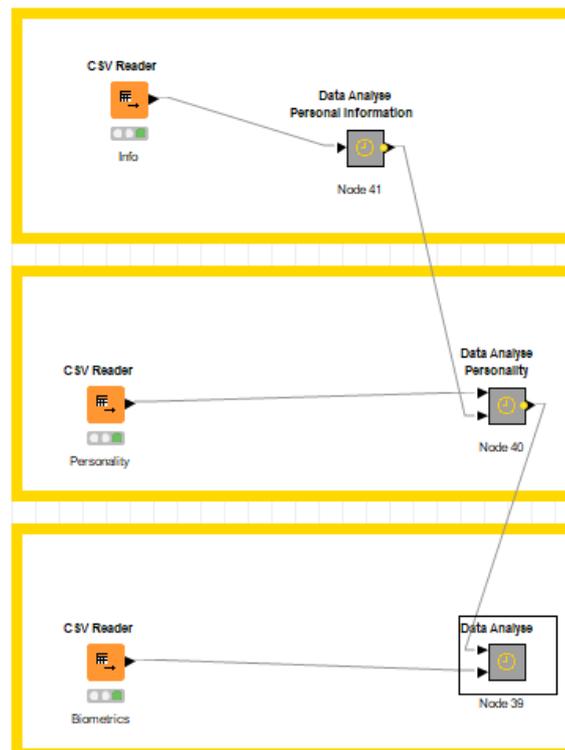


Figura 26 - Workflow do KNIME.

Após o carregamento dos 3 *datasets*, iniciou-se o estudo utilizando o nodo *Statistics* para realizar uma análise estatística dos vários dados, visto que os dados tem muitas colunas, principalmente o *DataPerso* que tem cerca de 90, foi criado a Tabela 4, que concentra os principais dados estatísticos dos três *dataset*.

Column	Min	Max	Mean	Std.Dev.	Variance	Skewness	Kurtosis
agree	32	62	44,69	8,35	69,79	-0,21	-0,83
extra	35	61	48,28	8,40	70,42	-0,17	-1,43
cons	29	59	42,17	7,32	53,58	0,75	-0,23
open	27	47	36,66	7,15	51,16	0,42	-1,47
stable	40	62	48,76	5,2281	27,33	0,52	0,93
age	13	81	30,75	19,59	383,66	1,75	3,264
dist	0	20305	2441,32	3510,30	1,23E7	2,43	7,83
heartbeat	55	219	85,46	22,21	493,36	2,86	14,93
imc	0	28	18,83	8,66	74,98	-1,64	1,07
kcal	5	3174	1306,96	702,34	493278,49	0,43	0,06
steps	0	26079	3509,43	4870,06	2,372E7	2,31	6,25
weight	45	75	66,38	6,43	41,40	-0,99	2,50
sleepMinutes	0	649	389,85	117,81	13879,14	-1,17	2,64

Tabela 4 - Tabela com os principais atributos de cada *dataset*.

Ao explorar a tabela descobriu-se que distância percorrida pelos utilizadores em média é bastante baixa, cerca de 2 km, no entanto existe pelo menos um registo em que o utilizador percorreu 20km num só dia. Outra referência interessante que é possível ver na tabela é o grande intervalo de batimento cardíaco, mínimo 55 e máximo de 219.

De seguida, explorando o *metanode* referente a análise do *dataset* *DataInfo* (Figura 27), foi explorado o *DataExplorer* mas não foi obtida nenhuma informação adicional e foi criado um node *String Manipulation* para substituir no campo email o “.” por “,” para juntar futuramente os dados de *DataInfo* com *DataPerso*. Ainda neste *metanode* foi gerado um

gráfico de barras para verificar a faixa etária e género dos participantes do estudo. (Figura 28).

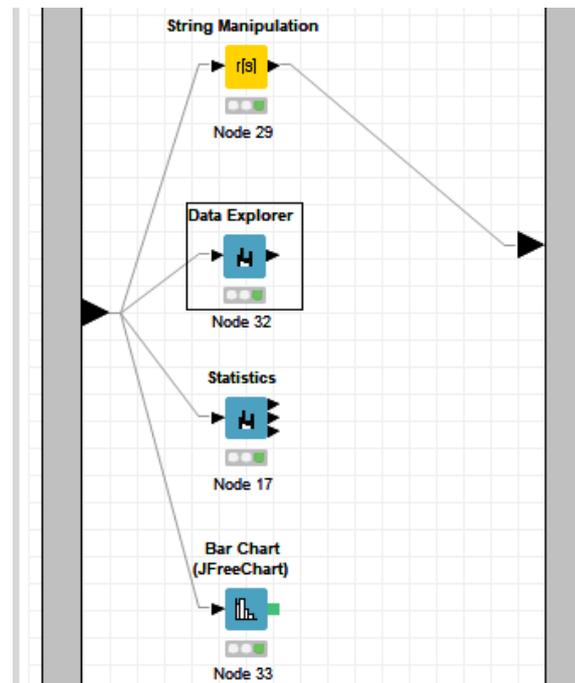


Figura 27 - Metanode expandido referente ao DataInfo.

Como se pode ver na Figura 28, existem variações de idades que vão desde os 13 anos aos 81 e em termos de género estamos equilibrados com 5 pessoas do género feminino e 6 do masculino.

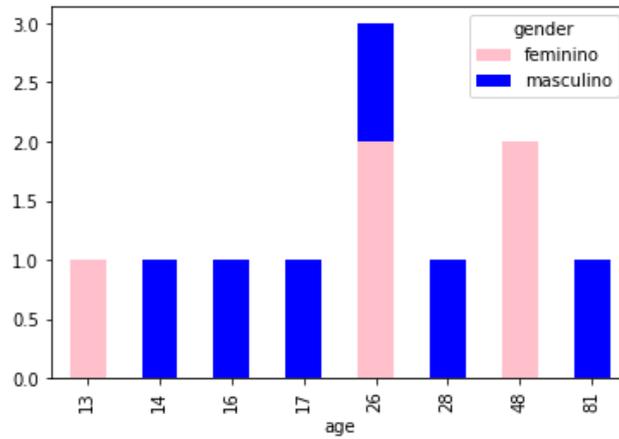


Figura 28 - Gráfico relativo a faixa etária e género.

Redirecionando agora para o *metanodo* que se encontra relacionado com os dados relativos ao teste de personalidade (Figura 29).

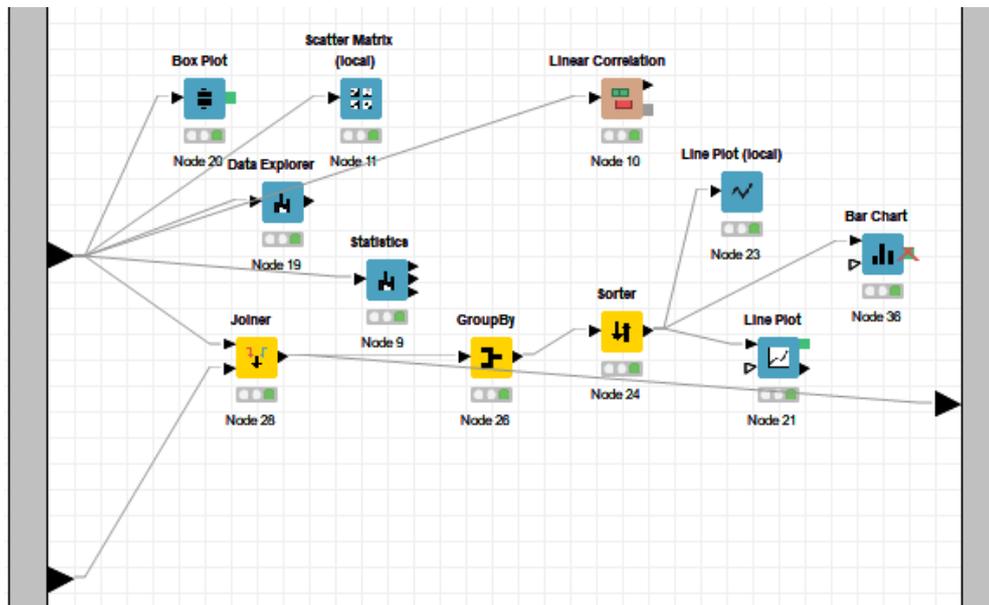


Figura 29 – Metanode referente ao dataset DataPerso.

Posteriormente e seguido a estrutura realizada no *KNIME* analisou-se o Data Perso, criando um *BoxPlot* com as entradas referentes aos cinco traços de personalidade, como se pode ver na Figura 30.

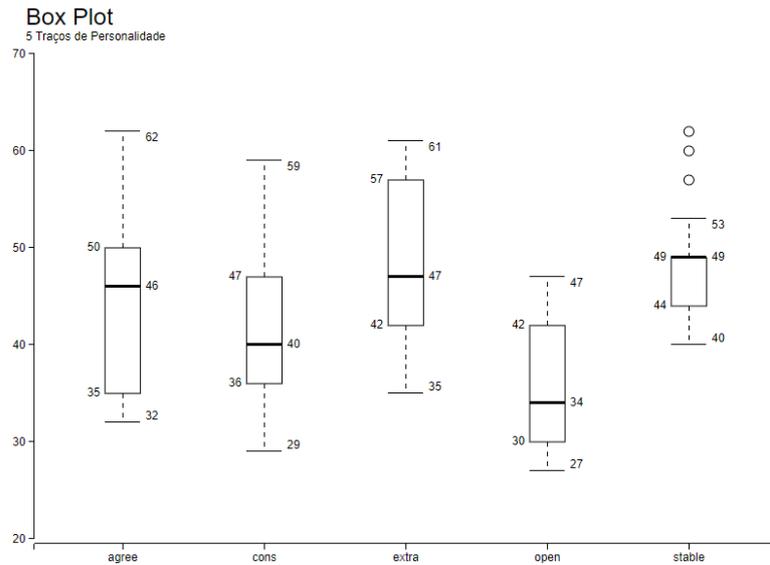


Figura 30 - *Boxplot* referente aos cinco traços de personalidade.

Nela é possível ver a variação entre o valor mínimo e máximo do resultado dos testes de personalidade de todos os participantes para os cinco traços de personalidade. É possível ver que o consciencioso tem bastante variação, mas que a maioria dos registos se concentram entre 36 e 47. A agradabilidade é a que tem mais variedade de valores, 35 e 50. Quanto a extroversão tem os seu valores mais centralizados entre o valor 42 e 57. A estabilidade é o traço que tem o intervalo de valores centralizados mais pequenos, entre 44 e 49. A abertura é a característica que tem valores mais baixo, e centralizados entre 30 e 42. Ainda foi verificado a que utilizadores pertenciam os *outliers* do traço estabilidade, para a averiguar uma possível relação entre eles, no entanto, este eram completamente aleatórios, tanto em idade como género.

Uma vez que maior parte da população era relativamente jovem entre os 13 e os 26, foi possível identificar que apesar da idade a abertura foi o traço com menos expressão entre a amostra.

Seguidamente foi utilizado nodo *Data Explorer* para o *dataset* *DataPerso*, mas mais uma vez como esta tabela é muito extensa, teve que se seleccionar a informação mais relevante. No caso deste *dataset* as entradas mais relevantes são os resultados do teste psicológico presente nos cinco traços no caso de variáveis numéricas. Na Figura 31 é possível observar em que valores cada um dos 5 traços é mais expresso e ainda verificar que os dados estão tratados, uma vez que não há valores em falta, nem nulos.

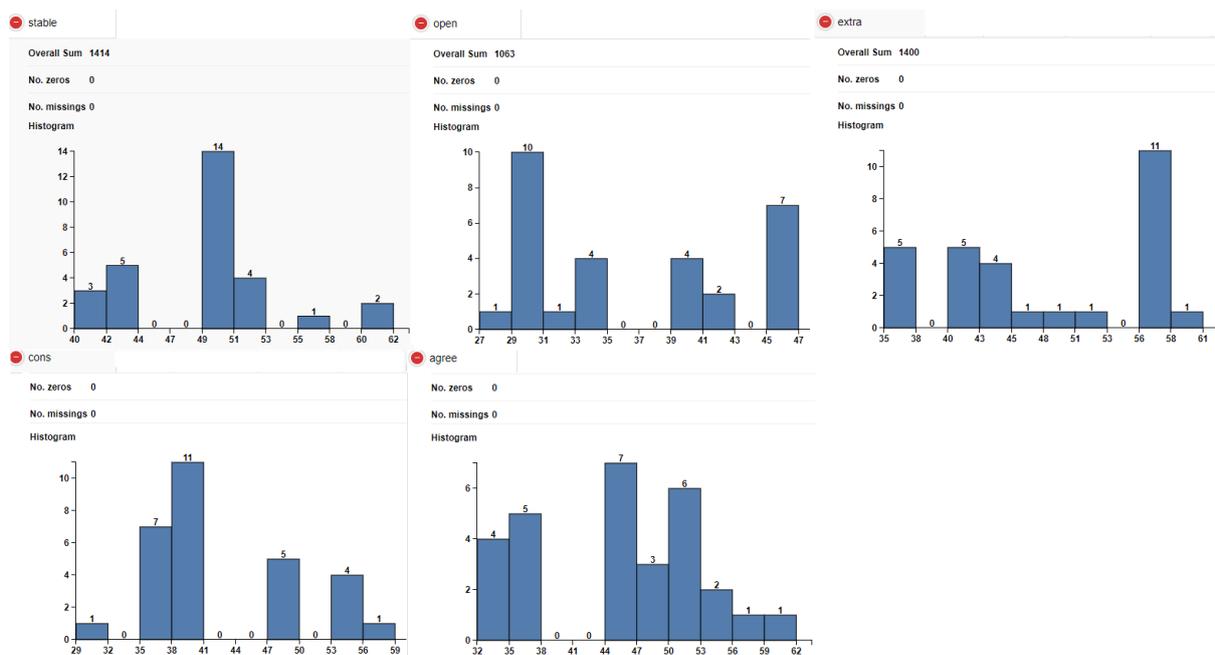


Figura 31 - Gráfico de barras proveniente do nodo *Data Explorer* relativos aos 5 traços de personalidade estabilidade, abertura, extroversão, consciencioso e agradabilidade, por esta ordem.

Seguidamente e ainda no mesmo nodo, foram analisados os valores nominais, e com a tabela da Figura 32, foi possível perceber que dos 29 testes realizados por 9 utilizadores, estes foram realizados em 19 dias distintos, e que e dos 29 teste apenas dois tinham a mesma combinação de botões selecionados.

Column	No. missings	Unique values	All nominal values	Frequency Bar Chart
id	0	29	2019-09-02_22:02_joaocdomingues31@gmail.com, 2019-09-05_21:40_mjpmfermandes24@gmail.com, 2019-07-30_11:15_helenafermandes2893@gmail.com. [...] 2019-08-01_18:47_joaocdomingues31@gmail.com, 2019-09-16_12:35_franciscotoucas8@gmail.com 2019-09-16_12:35_franciscotoucas8@gmail.com	
selectButton	0	27	['Falador', 'Distante', 'Eficiente', 'Ciumento', ''], ['Eficiente', 'Ciumento', 'Falador', ''], ['Falador', 'Calado', 'Cooperativo', 'Sistemático', 'Descontraído', 'Ciumento', 'Filosófico', 'Imaginativo', ''], [...] ['Falador', 'Frio', 'Mal-humorado', 'Temperamental', 'Imaginativo', ''], ['Prático', 'Desorganizado', 'Descontraído', ''], ['Falador', 'Eficiente', 'Ciumento', '']	
data	0	19	2019-07-28, 2019-07-29. [...] 2019-08-08, 2019-07-27.	
email	0	9	joaocdomingues31@gmail.com, adrianapereira110@gmail.com, mjpmfermandes24@gmail.com, helenafermandes2893@gmail.com, nunocovas15@gmail.com, marialvestermendes13@gmail.com, franciscotoucas8@gmail.com, carlosaalves@gmail.com, tmimarcelino1993@gmail.com	

Figura 32 - Tabela proveniente do nodo *Data Explorer* com informação estatista dos dados nominais.

Continuamente foi criado uma matriz de correlação com o nodo *Linear Correlation* entre os traços de personalidade e os 40 adjetivos correspondentes aos botões clicados (1 caso tenha sido clicado e 0 caso contrário), a partir desta foi possível verificar que existiam botões que nunca tinham sido clicados, como é o caso dos botões Caloroso, Complexo, Energético, Invejoso, Mal-humorado, Melindroso, Ordeiro e Solidário, como se pode ver na Figura 33.

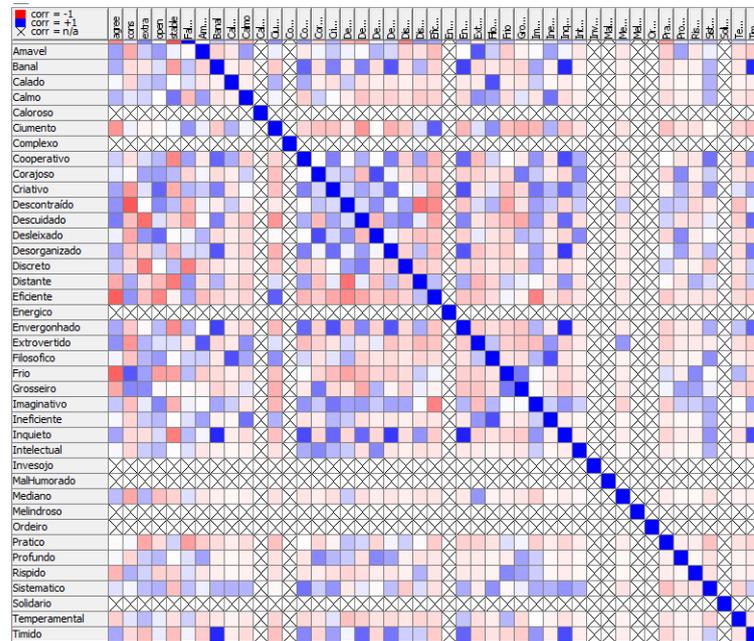


Figura 33 - Matriz de Correlação entre os 5 traços de personalidade e os 40 adjetivos clicáveis.

Refez-se a matriz sem os casos que não tinham qualquer significância (removidos 8 adjetivos), e dando particular atenção apenas a correlação entre os traços e os adjetivos, e não adjetivos com adjetivos como é possível examinar na Figura 34.

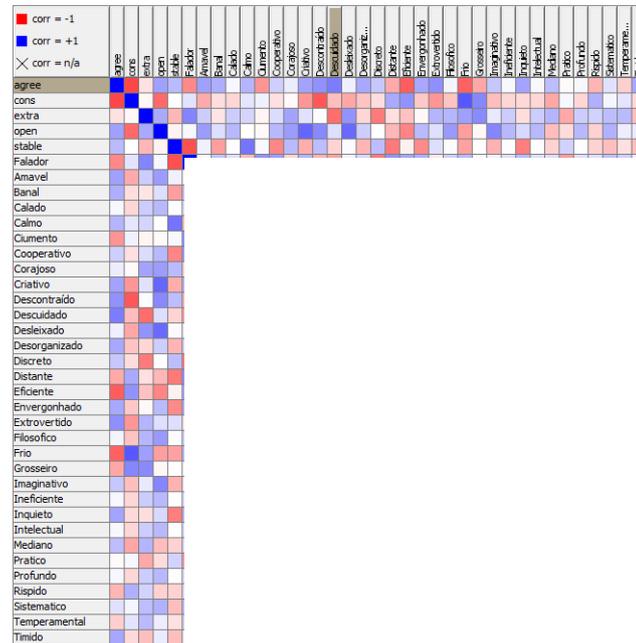


Figura 34 - Matriz de Correlação dos 32 adjetivos com os 5 traços de personalidade.

Ao analisar a matriz acima (Figura 34), foi possível verificar alguns casos em que a correlação é mais elevada, como é o caso de falador e *extra* com 0,48 de correlação, calmo e *stable* com 0,54 de correlação e criativo e *open* com 0,6 de correlação .A partir dessa informação foram criados três gráficos de barras com base nessas correlações (Figura 35).

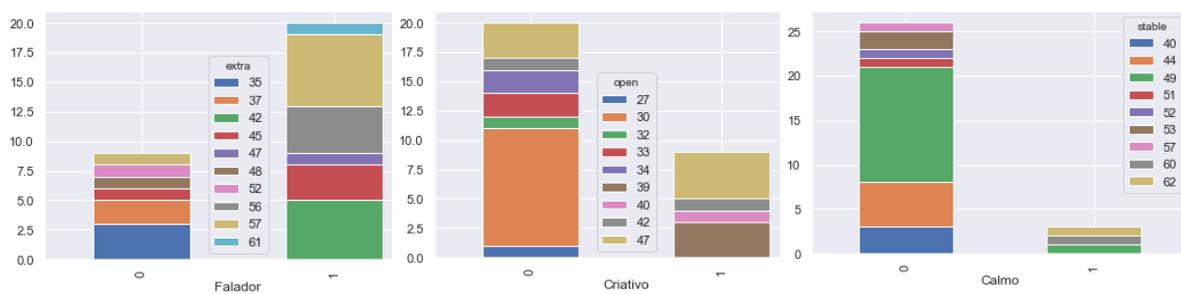


Figura 35 -Gráficos com as características com mais correlação ao traço de personalidade.

É possível identificar que nos 3 gráficos, sempre que o traço tem mais expressão é acionado o botão, isto é, no primeiro gráfico podemos ver o produto do teste psicólogo com mais expressão de estabilidade, 62 e 60 o utilizador clicou no botão calmo como uma das características que mais o identifica. O mesmo aconteceu nos outros dois casos, em que o botão Criativo foi selecionado 9 vezes, que corresponde aos casos em que a abertura foi mais expressa. Da mesma forma, Falador foi o botão mais vezes clicado e que tem bastante relação com a extroversão.

De seguida foram agregadas as entradas do *dataset* DataInfo com o DataPerso, criado de seguida um nodo *GroupBy* para agregar os resultados do teste psicológicos por nome, e exprimido num *LinePlot* como se pode ver na Figura 36.

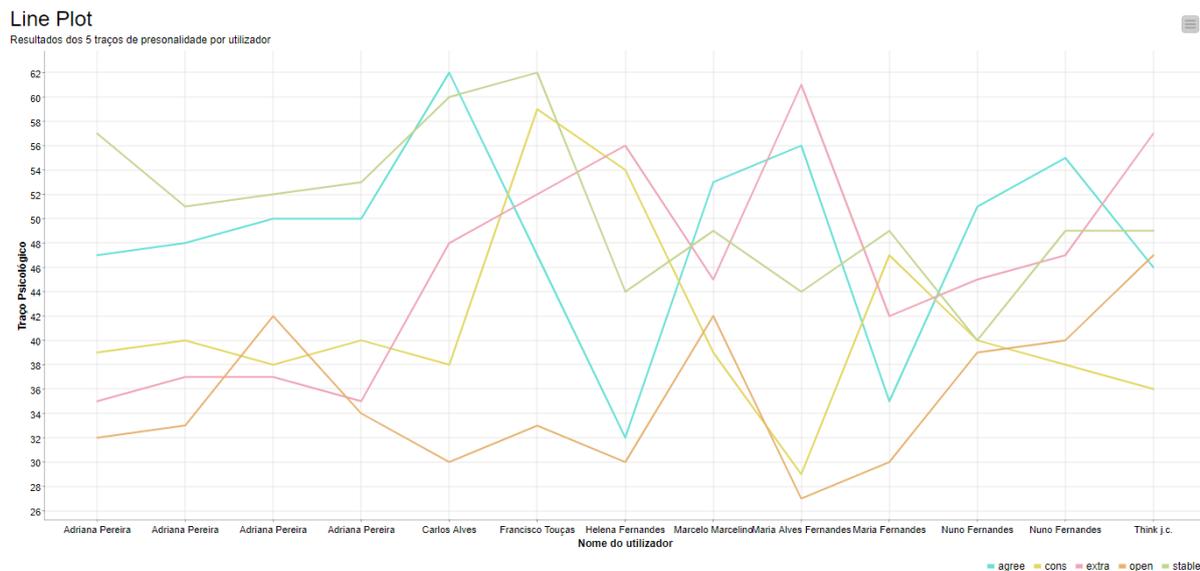


Figura 36 - *Line Plot* que representa a variação dos traços de personalidade de entre os utilizadores.

Ao analisar este *Line Plot*, é possível identificar a variação de personalidade de cada um do utilizador e é possível observar que existem inserções no gráfico com o mesmo nome de

utilizador, o que significa que 2 indivíduos em 9 alteraram o teste psicológico ao longo do tempo.

Foi ainda criado gráficos individuais da evolução da personalidade dos utilizadores, sendo que vamos mostrar apenas dois exemplos. O caso mais comum em que não ocorreram alterações nos cinco traços de personalidade ao longo do período analisado (Figura 37).

Tal como se pode ver na Figura 37 de 27 de julho de 2019 até 25 de setembro do mesmo ano, não se verificou qualquer alteração em nenhum dos cinco traços de personalidade.



Figura 37 - Gráfico do utilizador Think j.c. referente aos resultados ao longo do tempo dos 5 traços de personalidade.

E o caso com menos ocorrências, em que existe evolução da personalidade ao longo do tempo estudados, como é o caso que se verifica no gráfico da Figura 38, em que o utilizador teve uma leve variação em todos os cinco traços de personalidade.

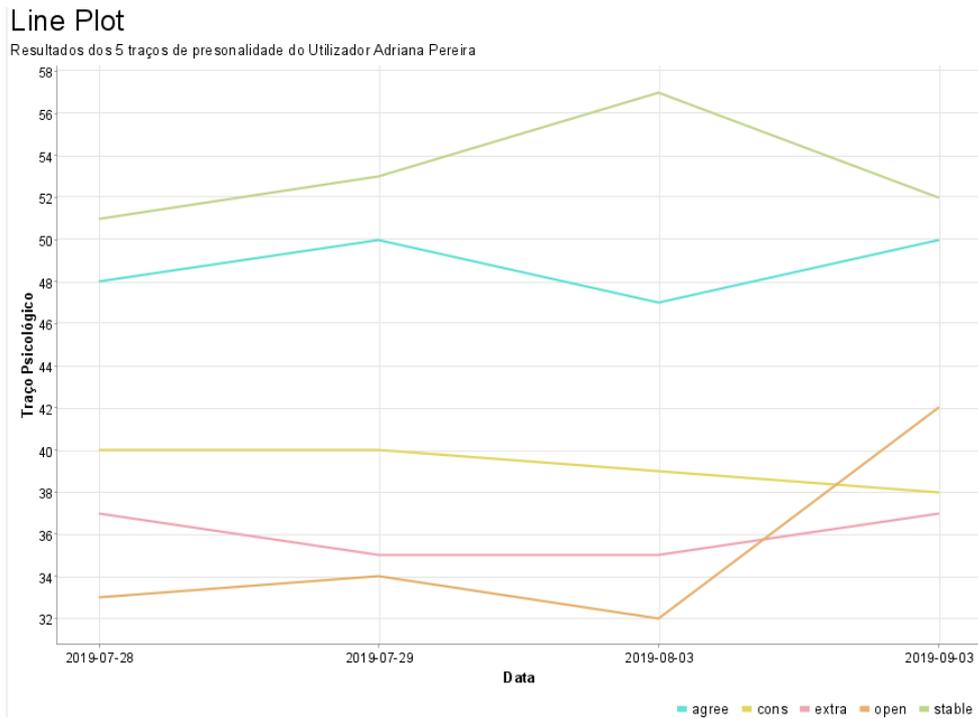


Figura 38 - Gráfico do utilizador Adriana referente aos resultados ao longo do tempo dos 5 traços de personalidade.

Continuamente segue-se para a análise do *dataset* DataBiometrics, como tal, pode ver na Figura 39.

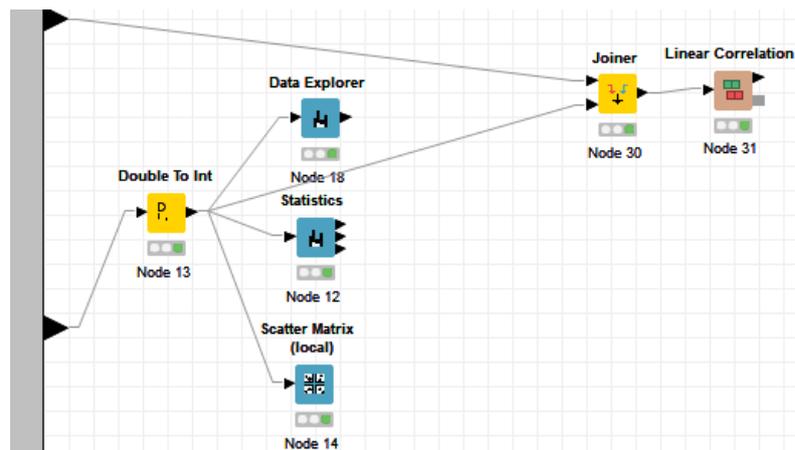


Figura 39 - Metanode referente a análise de dados biométricos.

Foi criado um nodo *Double to Int*, para transformar as colunas *dist*, *heartbeat*, *imc* e *steps* de decimal para inteiro. Seguidamente, com o nodo *Data Explorer* exprimiou-se os gráficos mais significantes perante a finalidade deste projeto.

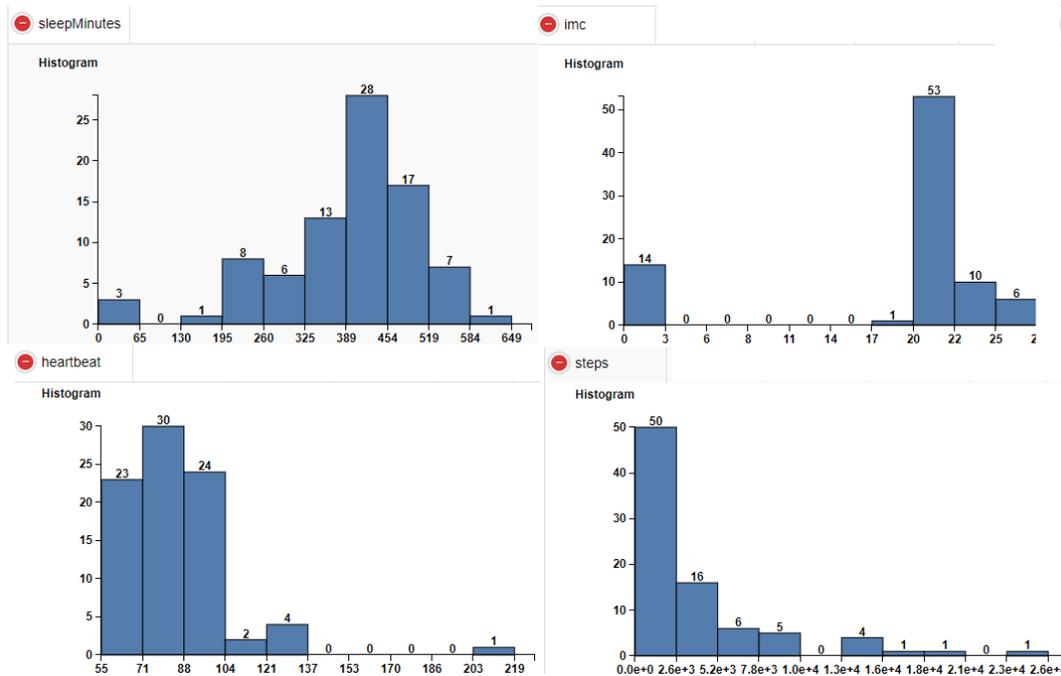


Figura 40 - Gráfico de barras proveniente do nodo *Data Explorer* relativos aos atributos *sleepMinutes*, *imc*, *heartbeat*, *steps*, por esta ordem.

Ao examinar a Figura 40 é possível ver a maior fatia dos registos acerca da duração em minutos de sono são no intervalo 349 e 454 minutos, o que corresponde aproximadamente entre 6h e 7h30 o que é um pouco abaixo do recomendado. A maior parte dos registos de *imc* revelam-se normais, o mesmo acontece com os registos de batimento cardíaco, embora haja um registo extremamente alto, chegando aos 219 batimentos por minuto. Quanto aos passos percorridos que se pode ver no último gráfico da Figura 40, estão muito concentrados numa zona muito baixa.

Seguidamente, analisou-se os valores nominais deste *dataset*, como se pode ver na Figura 41. Pode-se concluir que existem 91 registos, que foram inseridos em 43 dias distintos, de 10 pessoas diferentes e com inseridos 35 registos de horas de sono distintos.

Column	Exclude Column	No. missings	Unique values	All nominal values	Frequency Bar Chart
id	<input type="checkbox"/>	0	91	2019-08-06_23:16_tmimarcelino1993@gmail.com, 2019-08-03_13:55_adrianapereira110@gmail.com, [...]	
sleep	<input type="checkbox"/>	0	35	0h0, 4h15, 6h38, [...]	
email	<input type="checkbox"/>	0	10	adrianapereira110@gmail.com, joaocdomingues31@gmail.com, helenafernandes2893@gmail.com, tmimarcelino1993@gmail.com, mjpmfernandes24@gmail.com, [...]	
data	<input type="checkbox"/>	0	43	2019-07-28, 2019-08-03, 2019-07-29, 2019-08-05, 2019-08-08, [...]	

Figura 41 - Tabela proveniente do nodo *Data Explorer*, relativos aos dados nominais do *dataBiometrics*.

Realizada a análise estatística, e foram agrupados todos os *dataset* através do nodo *Join*, usando como modo *inner join*, e seguidamente foi criada uma matriz de correlação com o *Linear Correlation*, que se pode ver na Figura 42. Como se pode ver na primeira parte da figura, o primeiro *Linear Correlation* tinha atributos sem nenhuma correlação. O género e o nome, foram retirados e foi produzida uma nova matriz.

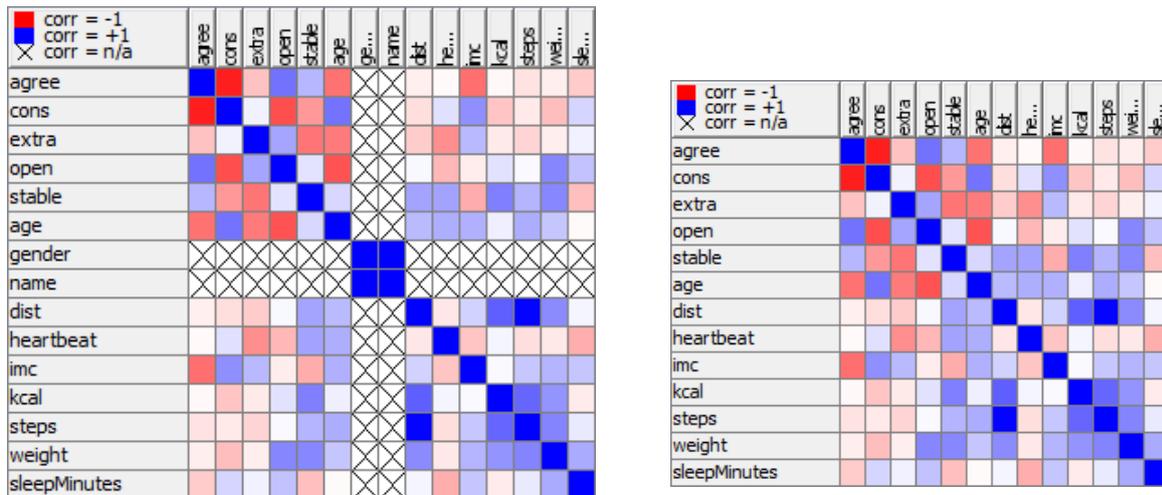


Figura 42 - Matriz de correlação com os dados biométricos, de personalidade e informação.

Ao avaliar a segunda matriz de correlação tendo em conta os traços e os valores biométricos é possível ver que se encontra maior relação entre o imc e o consciencioso, o peso e abertura, e as calorias e a estabilidade. Pensamos, contudo, que essa que essa relação é tendenciosa, uma vez que esse atributo imc e peso são atributos constantes, não tiveram grande variação e visto que o volume de dados é pequeno, o imc acaba por ser único de pessoa para pessoa.

Aparentemente os dados de sono, batimento cardíaco e passos não relevaram nenhuma correlação significativa. Este facto deve-se à falta de um volume significativo de dados. Outro facto, é que quando os utilizadores acedem a aplicação, muito vezes é no início ou meio do dia, daí também o facto de média total dos passos ser abaixo do esperado.

Capítulo 6

Conclusões e Trabalho Futuro

6.1 Conclusões

Atualmente, as pessoas utilizam cada vez mais sistemas inteligentes, consciente ou inconscientemente. Desde obter uma recomendação de um produto em compras *online* para simples pesquisas no *Google*.

Tendo por base o primeiro capítulo, foi possível ter uma noção das principais motivações que se ligam a este projeto, uma vez que existe uma crescente preocupação global e pessoal em analisar as alterações físicas e psicológicas que as pessoas vão sofrendo ao longo do tempo. Adicionalmente, a constante ascensão das tecnologias e a facilidade com que estas se podem alcançar, garante uma constante procura.

Analisando a literatura referente a vários testes psicológicos foi possível examinar vários tipos de teste, aceites pela comunidade de psicologia. De entre os vários avaliados, seguiu-se o teste de *Saucier*, pela simplicidade de realização e pela enorme aceitação. Foram ainda avaliadas as várias perspetivas relativas à possível relação entre os testes psicológicos BFI e os aspetos físicos de um indivíduo. Desta avaliação, obtemos que existe uma possível relação entre a baixa variação da frequência cardíaca e baixa Estabilidade e, ainda, que a alta atividade física pode estar relacionada com o traço de Abertura mais expresso. Visto que para investigar os estados físicos de um indivíduo são necessários dispositivos de monitorização, foram avaliados essencialmente *smartbands* e *smartwatches*. Conclui-se que os dispositivos com

mais aderência estão associados a grandes empresas fabricantes de *smartphones*, tais como, *Samsung, Apple, Huawei e Xiaomi*. Desta forma, a maioria dispõe de *software* próprio, o que tornou difícil a criação de uma aplicação móvel, que fosse compatível com todos os aparelhos.

Seguidamente, foram examinadas as aplicações que por definição comunicam com os dispositivos referidos acima. Desta forma, foi possível verificar que muitas delas possuem a capacidade de transmitir informação entre elas, como é o caso *Apple Health* com *Mi Fit e Huawei Health*, entre outras. Foi ainda possível verificar que a *Google Fit*, apesar de não ser aplicação por definição de nenhum dos dispositivos mencionados, é a que mais comunica com outras aplicações. Assim, e após analisar *APIs* dos dispositivos de *fitness*, a *API* da *Google Fit* apresenta os atributos que se pretendem utilizar na aplicação, de forma mais simplificada, e alcança maior número de utilizadores.

Tendo por base a utilização da *API* do *Google Fit*, foi implementada uma aplicação móvel dirigida para *Android*, uma vez que aloja maior volume de utilizadores e que concentra o teste psicológico com registo de dados de sono, de batimento cardíaco, número de passos, IMC, idade, género. Estes dados são recolhidos para a base de dados do *Firebase*.

Realizou-se uma análise de resultados no *KNIME*, que permitiu concluir que de modo geral os testes psicológicos não se alteram ao longo do tempo, uma vez que 7 pessoas mantiveram os seus resultados inalterados mesmo refazendo o teste várias vezes, apenas 2 tiveram uma evolução psicológica. Foi também descoberto que existe uma correlação entre a seleção de botões realizada paralelamente a realização do teste psicológico. Os adjetivos mais significativos e mais invocados foram a criatividade que se correlaciona com alta abertura, o falador que se correlaciona com a alta extroversão e calmo que se correlaciona com alta estabilidade. No entanto, foi também possível verificar que cada vez que um utilizador realiza o teste psicológico tem tendência a selecionar um conjunto diferente de adjetivos. Quanto à

correlação entre os dados psicológicos e biométricos, não se conseguiu provar que existe relação, mas com volume de dados maior e mais diversidade de população talvez o seja.

Perante as etapas propostas inicialmente, pode dar-se por concluída a análise da literatura acerca de uma diversidade de testes psicológicos, de aplicações móveis no âmbito da saúde física, do planeamento visual da aplicação e a implementação e recolha de dados da mesma. No entanto, o projeto seguiu um curso diferente do estipulado inicialmente, dada a complexidade da aplicação desenvolvida e a insuficiência de dados. Posto isto, não se tornou viável proceder à conceção e implementação de modelos de *machine learning* que permitissem estimar e encontrar padrões na análise de dados biométricos e psicológicos.

6.2 Trabalho Futuro

Como trabalho futuro, pretende-se aprimorar a aplicação de forma a que se torne mais apelativa, que os dados de sono sejam automaticamente lidos e até a possibilidade de recolher outro tipo de informações que se venham a julgar relevantes. Dever-se-á dar continuidade à recolha de dados biométricos e diversificar ainda mais essa recolha, distribuindo para isso a app por uma população maior. Posteriormente aplicar modelos de ML para tentar prever a personalidade de uma pessoa com base na informação obtida a partir da aplicação, especialmente os dados biométricos. Uma vez relacionados os dados biométricos, com os traços psicológicos pretende-se lançar sugestões de forma a por exemplo, incentivar a atividade física para benefício mental, ou a sugerir dicas para uma melhor qualidade do sono.

Bibliografia

- Ackerman, C. (2017). Big Five Personality Traits: The OCEAN Model Explained. Retrieved January 6, 2019, from [positivepsychology.com](https://positivepsychology.com/big-five-personality-theory/) website: <https://positivepsychology.com/big-five-personality-theory/>
- American Heart Association. (n.d.). All About Heart Rate (Pulse). Retrieved January 12, 2019, from <https://www.heart.org> website: <https://www.heart.org/en/health-topics/high-blood-pressure/the-facts-about-high-blood-pressure/all-about-heart-rate-pulse>
- Android Developers. (n.d.). Conheça o Android Studio. Retrieved January 3, 2019, from <https://developer.android.com/studio/intro>
- Bacelar, R. (2019). Android vs iOS: 10 anos depois, melhores apps são do mesmo sistema operativo. Retrieved September 16, 2019, from [pplware](https://pplware.sapo.pt) website: <https://pplware.sapo.pt/smartphones-tablets/ios/android-ios-apps-sistema-operativo/>
- Badenhorst, M., Barry, C. J., Swanepoel, C. J., van Staden, C. T., Wissing, J., & Rohwer, J. M. (2019). Workflow for Data Analysis in Experimental and Computational Systems Biology: Using Python as 'Glue.' *Processes*, 7(7), 460. <https://doi.org/10.3390/pr7070460>
- Bari, A. (2018). *Machine Learning at Work: Speeding Up Discovery*. Retrieved from <https://books.google.pt/books?id=3J5QuAEACAAJ>
- BMI Calculator | The Heart Foundation. (n.d.). Retrieved September 16, 2019, from <https://www.heartfoundation.org.au/your-heart/know-your-risks/healthy-weight/bmi-calculator>
- Brundtland, G. H. (2000). Mental health in the 21st century. *Bulletin of the World Health Organization*, 78, 411–411. <https://doi.org/10.1590/S0042-96862000000400001>
- Budkewicz, M. (2018). Como instalar o Jupyter Notebook? (Windows e Linux). Retrieved October 14, 2019, from Medium website: <https://medium.com/horadecodar/como-instalar-o-jupyter-notebook-windows-e-linux-20701fc583c>
- Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41(1), 417–440. <https://doi.org/10.1146/annurev.ps.41.020190.002221>
- Duggan, K. A., Friedman, H. S., McDevitt, E. A., & Mednick, S. C. (2014). Personality and healthy sleep: The importance of conscientiousness and neuroticism. *PLoS ONE*, 9(3).

<https://doi.org/10.1371/journal.pone.0090628>

- Eabon, M. F., & Abrahamson, D. (n.d.). Understanding psychological testing and assessment. Retrieved December 14, 2018, from American Psychological Association website: <https://www.apa.org/helpcenter/assessment>
- Fatissou, J., Oswald, V., & Lalonde, F. (2016). Influence diagram of physiological and environmental factors affecting heart rate variability: An extended literature overview. *Heart International*, Vol. 11, pp. e32–e40. <https://doi.org/10.5301/heartint.5000232>
- Fernando Perez, B. E. G. (2015). Project Jupyter: Computational narratives as the engine of collaborative data science. Retrieved September, 11(207), 108. Retrieved from <http://archive.ipython.org/JupyterGrantNarrative-2015.pdf>
- FinancesOnline. (n.d.). KNIME Analytics Platform Reviews: Overview, Pricing and Features. Retrieved December 26, 2018, from <https://reviews.financesonline.com/p/knime-analytics-platform/>
- Gemalto. (n.d.). What is biometric data? Retrieved January 13, 2019, from <https://www.justaskgemalto.com/en/what-is-biometric-data/>
- Global recommendations on physical activity for health*. (2010). Geneva.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26–42. <https://doi.org/10.1037/1040-3590.4.1.26>
- Google Commerce Ltd. (2019). Google Fit. Retrieved October 1, 2019, from <https://www.google.com/fit/>
- Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality – Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, 46(3), 355–359. <https://doi.org/10.1016/j.jrp.2012.03.008>
- Hill, S. (2019). How Accurate Are Fitness Trackers? We Asked an Expert | Digital Trends. Retrieved September 26, 2019, from Digital Trends website: <https://www.digitaltrends.com/wearables/how-accurate-are-fitness-trackers/>
- Hintsanen, M., Puttonen, S., Smith, K., Törnroos, M., Jokela, M., Pulkki-Råback, L., ... Keltikangas-Järvinen, L. (2014). Five-factor personality traits and sleep: evidence from two population-based cohort studies. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, 33(10), 1214–1223. <https://doi.org/10.1037/hea0000105>
- John, O. P., & Srivastava, S. (1999). *The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives*.

- Johor Bahru, U., Darul Ta, J., Bin Mohamad, I., Usman, D., & Bahru, J. (2013). Standardization and Its Effects on K-Means Clustering Algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, 6(17), 3299–3303.
- Jolliffe, I. (2011). Principal Component Analysis. In *International Encyclopedia of Statistical Science* (pp. 1094–1096). https://doi.org/10.1007/978-3-642-04898-2_455
- Kenney, B. (2015). Why Biometric Tech Is the Wearable of the Future. Retrieved December 23, 2018, from <https://smartwatches.org/learn/why-biometric-tech-is-the-wearable-of-the-future/%0D>
- Lartillot, O., Toivainen, P., & Eerola, T. (2008). *A Matlab Toolbox for Music Information Retrieval BT - Data Analysis, Machine Learning and Applications* (C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker, Eds.). Springer Berlin Heidelberg.
- Lebowitz, S. (2016). Big 5 personality traits predict who will become a leader - Business Insider. Retrieved September 26, 2019, from Business Insider website: <https://www.businessinsider.com/big-five-personality-traits-predict-leadership-2016-12>
- Matsumura, K., & Yamakoshi, T. (2013). iPhysioMeter: A new approach for measuring heart rate and normalized pulse volume using only a smartphone. *Behavior Research Methods*, 45(4), 1272–1278. <https://doi.org/10.3758/s13428-012-0312-z>
- McCraty, R., Atkinson, M., Tomasino, D., & Bradley, R. T. (2009). The Coherent Heart: Heart–Brain Interactions, Psychophysiological Coherence, and the Emergence of System-Wide Order | HeartMath Institute. Retrieved October 1, 2019, from <https://www.heartmath.org/research/research-library/basic/coherent-heart-heart-brain-interactions-psychophysiological-coherence-emergence-system-wide-order/>
- Monteiro, S. T., & Ribeiro, C. H. C. (2004). Desempenho de algoritmos de aprendizagem por reforço sob condições de ambiguidade sensorial em robótica móvel. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, 15, 320–338. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-17592004000300008&nrm=iso
- Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science and Engineering*, 9(3), 10–20. <https://doi.org/10.1109/MCSE.2007.58>
- Oliveira, V. K., & Filho, A. A. M. (2015). Testes de Personalidade: o que são? Qual a sua finalidade? Quais são os mais usados? - Psicologia e Psiquiatria > AbcMed. Retrieved January 16, 2019, from AbcMed website: <https://www.abc.med.br/p/psicologia-e-psiquiatria/799889/testes+de+personalidade+o+que+sao+qual+a+sua+finalidade+quais+sao+os+mais+usados.htm>

- Open-Source Psychometrics Project. (2019). Recommended test for scientific validity. Retrieved December 28, 2018, from <https://openpsychometrics.org/>
- Pessemier, T. De, & Martens, L. (2017). *Measuring Heart Rate with Mobile Devices for Personal Health Monitoring*. (c), 19–24.
- Pinheiro, P. (2019). CALCULE O SEU PESO IDEAL E IMC » MD.Saúde. Retrieved September 16, 2019, from MD.Saúde website: <https://www.mdsaude.com/obesidade/calcule-o-seu-peso-ideal-e-imc/>
- Preventing Noncommunicable Diseases in the Workplace through Diet and Physical Activity*. (2008). Geneva.
- Rhodes, R. E., & Smith, N. E. I. (2006, December). Personality correlates of physical activity: A review and meta-analysis. *British Journal of Sports Medicine*, Vol. 40, pp. 958–965. <https://doi.org/10.1136/bjbm.2006.028860>
- Sahu, S., Nautiyal, A., & Prasad, M. (2017). Machine Learning Algorithms for Recommender System - a comparative analysis. *International Journal of Computer Applications Technology and Research*, 6, 97–100. <https://doi.org/10.7753/IJCATR0602.1005>
- SAPO. (2016). Pela primeira vez foram vendidos mais relógios inteligentes do que relógios suíços. Retrieved November 30, 2018, from SAPO website: <https://tek.sapo.pt/mobile/equipamentos/artigos/pela-primeira-vez-foram-vendidos-mais-relogios-inteligentes-do-que-relogios-suicos>
- Saucier, G. (1994). Mini-Markers: A Brief Version of Goldberg's Unipolar Big-Five Markers. *Journal of Personality Assessment*, 63(3), 506–516. https://doi.org/10.1207/s15327752jpa6303_8
- Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, 68, 69–81. <https://doi.org/10.1016/j.jrp.2017.02.004>
- Sumra, H. (2018). Essential apps and devices that work with Google Fit. Retrieved October 1, 2019, from <https://www.wareable.com> website: <https://www.wareable.com/sport/the-best-google-fit-compatible-apps-and-devices>
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning An Introduction*. <https://doi.org/10.1145/1068009.1068322>
- Tecnologia - ig. (2018). Smartband ou smartwatch? conheça cada um e saiba qual combina mais com você. Retrieved January 2, 2019, from <https://tecnologia.ig.com.br/dicas/2018-03-27/smartband-smartwatch.html>
- Thayer, J. F., Yamamoto, S. S., & Brosschot, J. F. (2010). The relationship of autonomic

- imbalance, heart rate variability and cardiovascular disease risk factors. *International Journal of Cardiology*, 141(2), 122–131. <https://doi.org/10.1016/j.ijcard.2009.09.543>
- The Rise of mHealth Apps: A Market Snapshot. (2018). Retrieved December 27, 2019, from Digital Health website: <https://liquid-state.com/mhealth-apps-market-snapshot/>
- Thiel, E. van. (2018). Big Five personality test traits. Retrieved January 15, 2019, from <https://www.123test.com/big-five-personality-theory/>
- Thompson, E. R. (2008). *Development and Validation of an International English Big-Five Mini-Markers q. 45*, 542–548. <https://doi.org/10.1016/j.paid.2008.06.013>
- Timo-laria, C., Negrà, N., Schmidek, W. R., Hoshino, K., de Menezes, C. E. L., & da Rocha, T. L. (1970). Phases and states of sleep in the rat. *Physiology and Behavior*, 5(9), 1057–1062. [https://doi.org/10.1016/0031-9384\(70\)90162-9](https://doi.org/10.1016/0031-9384(70)90162-9)
- Togo, F., & Takahashi, M. (2009). Heart rate variability in occupational health --a systematic review. *Industrial Health*, 47(6), 589–602. <https://doi.org/10.2486/indhealth.47.589>
- Valentini, M., & Parati, G. (2009). Variables influencing heart rate. *Progress in Cardiovascular Diseases*, 52(1), 11–19. <https://doi.org/10.1016/j.pcad.2009.05.004>
- van Ravenswaaij-Arts, C. M., Kollée, L. A., Hopman, J. C., Stoeltinga, G. B., & van Geijn, H. P. (1993). Heart rate variability. *Annals of Internal Medicine*, 118(6), 436–447. <https://doi.org/10.7326/0003-4819-118-6-199303150-00008>
- van Zyl, L. T., Hasegawa, T., & Nagata, K. (2008). Effects of antidepressant treatment on heart rate variability in major depression: a quantitative review. *BioPsychoSocial Medicine*, 2, 12. <https://doi.org/10.1186/1751-0759-2-12>
- Wang, L. (2016). Discovering phase transitions with unsupervised learning. *Physical Review B*, 94(19). <https://doi.org/10.1103/PhysRevB.94.195105>
- Wearables based on APIs, an excellent business opportunity. (2018). Retrieved December 20, 2018, from BBVAOpen4u.com website: <https://bbvaopen4u.com/en/actualidad/wearables-based-apis-excellent-business-opportunity>
- Wilson, K. E., & Dishman, R. K. (2015). Personality and physical activity: A systematic review and meta-analysis. *Personality and Individual Differences*, 72, 230–242. <https://doi.org/10.1016/J.PAID.2014.08.023>
- Zohar, A., & McCraty, R. (2013). Personality and Heart Rate Variability: Exploring Pathways from Personality to Cardiac Coherence and Health. *Open Journal of Social Sciences*, 1, 32–39. <https://doi.org/10.4326/jss2013.16007>