



UNIVERSIDADE DO MINHO

MESTRADO INTEGRADO DE ENGENHARIA INFORMÁTICA

## **Análise de Dados**

### ***Autores:***

Adriana Guedes	A74545
Marco Barbosa	A75278
Ricardo Certo	A75315

11 de Janeiro de 2018

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>5</b>
<b>2</b>	<b>Modelação Lógica</b>	<b>6</b>
2.1	Dataset Inicial . . . . .	6
2.2	Modelo Lógico . . . . .	6
2.3	Entidades . . . . .	7
2.3.1	Landmass . . . . .	7
2.3.2	Zone . . . . .	7
2.3.3	Language . . . . .	7
2.3.4	Religion . . . . .	7
2.3.5	Country . . . . .	7
2.3.6	Flag . . . . .	7
2.3.7	Color . . . . .	7
2.4	Relacionamentos . . . . .	8
2.4.1	Relação 1:1 . . . . .	8
2.4.2	Relação 1:N . . . . .	8
2.5	Decisões tomadas . . . . .	8
2.6	Normalização . . . . .	9
<b>3</b>	<b>Data Warehouse</b>	<b>10</b>
3.1	Processos para a criação de um Data Warehouse . . . . .	10
3.2	Perguntas às quais se pretende que o modelo criado responda no futuro . . . . .	11
3.3	Decisões . . . . .	11
3.4	Modelo de DW . . . . .	12
<b>4</b>	<b>Talend</b>	<b>13</b>
4.1	Metadata . . . . .	13
4.1.1	Flags_dw . . . . .	15
4.1.2	Flags . . . . .	15
4.1.3	Povo . . . . .	15
4.1.4	Paisesemfalta . . . . .	15
4.2	Jobs . . . . .	16
4.2.1	Dim_country . . . . .	16
4.2.2	Dim_flags . . . . .	17
4.2.3	Dim_landmass . . . . .	17
4.2.4	Dim_language . . . . .	17
4.2.5	Dim_Religion . . . . .	18

4.2.6	Flags_dim . . . . .	18
4.3	Parte Lógica - tMaps . . . . .	19
4.3.1	Dim_country . . . . .	19
4.3.2	Dim_flags . . . . .	20
4.3.3	Dim_landmass . . . . .	20
4.3.4	Dim_language . . . . .	20
4.3.5	Dim_Religion . . . . .	21
4.3.6	Flags_dim . . . . .	21
<b>5</b>	<b>Power BI</b>	<b>23</b>
5.1	Perguntas . . . . .	23
5.1.1	Países cuja população é menor que 25M . . . . .	23
5.1.2	Países de África que tem uma área entre 300 e 500 . . . . .	24
5.1.3	Países da Europa que tem vermelho na bandeira . . . . .	24
5.1.4	Países que falam Alemão . . . . .	25
5.1.5	Mapa mundo com países em estudo . . . . .	25
5.1.6	Número de listas e barras em função de um continente . . . . .	26
5.1.7	População Total . . . . .	26
5.1.8	Área Total . . . . .	26
5.1.9	Cores Totais . . . . .	27
5.1.10	Total Bars (barras) . . . . .	27
5.1.11	Total Stripes (listas) . . . . .	27
5.1.12	Listar todos os Países . . . . .	28
5.1.13	Relação entre uma língua, os países, o total da população, da área, das listas, das colunas e das cores da bandeira . . . . .	29
5.1.14	Área por Continente . . . . .	29
5.1.15	Análise da população por continente . . . . .	30
5.1.16	População por continente . . . . .	30
5.1.17	Lista de países de um continente que tem na bandeira verde, laranja, vermelho e não tem branco . . . . .	31
5.1.18	Relação com as stripes, bars entre continentes . . . . .	31
5.1.19	Nome do país, área e população dos países que tem a religião 'Muslism' . . . . .	32
5.1.20	População de uma dada religião . . . . .	33
<b>6</b>	<b>Conclusão</b>	<b>34</b>

## Lista de Figuras

1	Modelo Lógico do Dataset . . . . .	6
2	Modelo de DW . . . . .	12
3	Metadata . . . . .	13
4	Import CSV . . . . .	14
5	Import Tables CSV . . . . .	14
6	Conexão à Base de dados . . . . .	15
7	Jobs . . . . .	16
8	Dim Country . . . . .	16
9	Dim Flags . . . . .	17
10	Dim Landmass . . . . .	17
11	Dim Language . . . . .	17
12	Dim Language . . . . .	18
13	Job Final . . . . .	18
14	tMap Country - Parte 1 . . . . .	19
15	tMap Country - Parte 2 . . . . .	19
16	tMap Flags . . . . .	20
17	tMap landmass . . . . .	20
18	tMap language . . . . .	20
19	tMap religion . . . . .	21
20	tMap flags final - Parte 1 . . . . .	21
21	tMap flags final - Parte 2 . . . . .	22
22	Países cuja população é menor que 25M . . . . .	23
23	Países de África que tem uma área entre 300 e 500 . . . . .	24
24	Países da Europa que tem vermelho na bandeira . . . . .	24
25	Países que falam Alemão . . . . .	25
26	Mapa mundo com países em estudo . . . . .	25
27	Número de listas e barras em função de um continente . . . . .	26
28	População Total . . . . .	26
29	Área Total . . . . .	26
30	Cores Totais . . . . .	27
31	Total bars . . . . .	27
32	Total Stripes . . . . .	27
33	Países Totais . . . . .	28
34	Relação 1 . . . . .	29
35	Área por continente . . . . .	29
36	População por continente 1 . . . . .	30
37	População por continente 2 . . . . .	30

38	Lista países por uma característica na bandeira . . . . .	31
39	Relação stripes,bars com continentes . . . . .	31
40	Informação país em função da Religião . . . . .	32
41	População de uma dada religião . . . . .	33

## 1 Introdução

Com este trabalho prático, pretende-se o desenvolvimento de um sistema que permita o tratamento e análise dos dados presentes num dataset à escolha dos grupos. No nosso caso em particular, foi escolhido um dataset contendo informação relacionada com países e bandeiras, proveniente do livro "Collins Gem Guide to Flags": Collins Publishers (1986).

Ainda antes de se obter informação útil do dataset, é necessário que um processo BI (Business Intelligence) extraia, transforme e carregue os dados (ETL) em Data Warehouses.

Após este ponto, os dados tornam-se muito mais credíveis e passíveis de serem usados para conhecimentos de negócio.

## 2 Modelação Lógica

### 2.1 Dataset Inicial

O Dataset inicial era composto por 194 linhas, onde cada uma fazia referência a 30 características sobre um dado país. As características fornecidas eram bastante diversas o que nos levou, na passagem para uma base de dados relacional, a criar várias tabelas de modo a separar, melhor organizar e reduzir redundâncias na informação.

### 2.2 Modelo Lógico

Após discussão/análise dos dados do Dataset e estruturação de como seria guardada a informação proveniente do Dataset, foi obtido o seguinte modelo lógico, sendo que as várias decisões tomadas serão detalhadas numa fase posterior deste relatório.

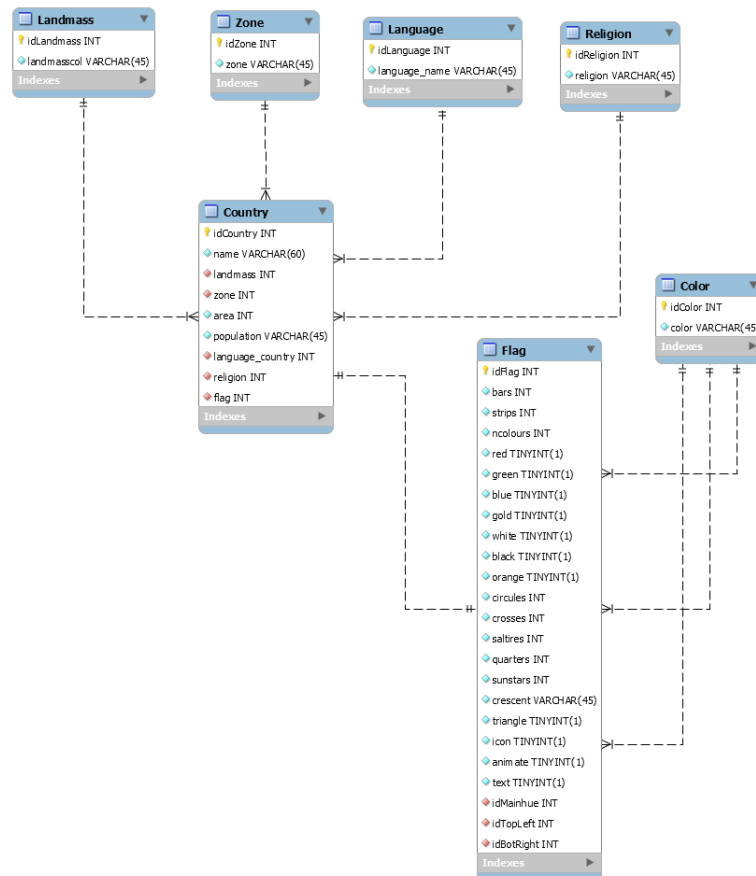


Figura 1: Modelo Lógico do Dataset

## **2.3 Entidades**

Houve necessidade de se criarem várias entidades de forma a melhor organizar as informações retiradas do Dataset e também para não tornamos essa mesma redundante. Para isso, foram criadas 7 entidades para guardar essa informação e a manter organizada, de modo a facilitar a procura de uma dada característica.

### **2.3.1 Landmass**

Esta tabela contém informação relativa ao continente a que um dado país pertence.

### **2.3.2 Zone**

Esta tabela contém informação relativa ao posicionamento de um dado país mediante os pontos cardeais.

### **2.3.3 Language**

Esta tabela contém informação sobre a língua falada num determinado país.

### **2.3.4 Religion**

Esta tabela contém informação sobre a religião predominante num determinado país.

### **2.3.5 Country**

Esta tabela contém informação sobre o nome de um país, a sua área e também sua população em milhões de habitantes.

### **2.3.6 Flag**

Esta tabela contém informação sobre a bandeira de um país, desde as cores que contem, número de barras e listas, se possui triângulos ou não, etc.

### **2.3.7 Color**

Esta tabela contém informação sobre uma dada cor que permite identificar qual é a cor predominante na bandeira.



## 2.4 Relacionamentos

### 2.4.1 Relação 1:1

A única relação que temos deste tipo é entre a tabela país e a tabela flag, pois um país só pode ter a si associado uma bandeira.

### 2.4.2 Relação 1:N

Todas as relações são de 1:N exceto a referida em cima, sendo que estas relações ocorrem, por exemplo, quando se sabe que cada país tem uma linguagem e uma linguagem pode pertencer a vários países, obtendo por isso: Country 1 <-> N Language.

## 2.5 Decisões tomadas

No tópico seguinte, serão apresentadas e explicadas todas as decisões tomadas pelo grupo que foram preponderantes para a obtenção do modelo lógico referido cima.

- **Uso da relação PK/FK** : o uso desta relação funciona como uma prevenção de erros de associação entre tabelas.
- **Criação de várias tabelas no modelo lógico**: foi decidido serem criadas várias tabelas de forma a evitar a repetição de dados e no sentido de existir uma maior organização deles.
- **Relação Country - Flag 1:1** : esta relação é uma relação 1:1 de forma a garantir que cada país tem a sua própria bandeira, excluindo assim a hipótese de 2 países terem a mesma bandeira.
- **Inclusão da lista das cores na tabela Flag**: A lista de cores inclui-se na tabela Flag pois se fosse criada uma nova tabela com todas as permutações de cores, iriam existir 7! registos na tabela (provenientes das diferentes combinações). Para além disso, como esses atributos são binários, podemos concluir de uma forma mais eficaz se uma bandeira tem uma dada cor.
- **Atributos são todos não nulos**: Os atributos são todos não nulos pois, dessa forma, é certo que a informação para os diferentes países existe.
- **Separação da Bandeira do País**: Foi feita a separação na base de dados da Entidade Flag da entidade Country pelo facto de simplificar o acesso às tabelas, pois se forem necessárias informações sobre um dado

país é feito um acesso à tabela Country, enquanto que se forem necessárias informações sobre a sua bandeira é feito um acesso à tabela Flag.

- **Atributos *Cor de Mainhue*, *top\_left* e *botright* numa outra tabela:** na tabela flag estes atributos estão na forma de ID por serem um FK, e assim evitamos a redundância com a criação da tabela Color, que possui um ID e o respetivo nome da cor.

## 2.6 Normalização

- **Primeira Forma Normal(1FN)** : visto que o nosso modelo não contem valores multi valor nem grupos repetidos podemos afirmar que este está de acordo com a 1ª forma normal.
- **Segunda Forma Normal(2FN)**: uma relação está na 2ª forma normal quando essa relação tem uma chave primária composta e não existem atributos na relação apenas dependentes de uma das chaves primárias. Como não existem chaves primárias compostas no nosso modelo não há relações que quebrem a 2ª forma normal, assim sendo o nosso modelo respeita a 2ª forma normal.
- **Terceira Forma Normal(3FN)** : uma relação não está na 3ª forma normal, quando um atributo não-chave influencia outro atributo não-chave. No entanto para ser considerado da 3ª forma norma, a relação tem de estar na 2ª forma. Como no nosso sistema não existem relações de 2ª forma não existem também relações da 3ª forma. Após verificação de todas as relações do nosso modelo, concluímos que este se encontra de acordo com a 3ª forma normal.

## 3 Data Warehouse

### 3.1 Processos para a criação de um Data Warehouse

De forma a ser possível a criação de um data Warehouse, há que ter em conta os seguintes passos:

- **Levantamento das necessidades**

Antes de tudo deve-se fazer o levantamento de todas as informações desejadas pelo utilizador. Assim, faz-se o cruzamento das dimensões, trabalhando naquilo que o DW irá ter e não como irá ter, havendo por isso uma preocupação com os desejos requisitados e não com a existência efetiva dos dados.

- **Mapeamento dos dados**

Nesta etapa é feito o mapeamento dos dados, identificando assim a fonte e como chegar até eles (no nosso caso é base de dados inicial). Neste momento do desenvolvimento é também avaliada a viabilidade dos requisitos apontados na primeira etapa, verificando a existência ou não dos dados para os cumprir.

- **Construção da Staging Area**

Após o mapeamento, é altura de dedicar esforços à construção da estrutura chamada Staging Area, que se trata da área de transição dos dados do DW. Para que tal processo possa ser feito, será usada a ferramenta Talend.

- **Construção das Dimensões**

Neste momento é construída a estrutura das Dimensões que farão parte do DW.

- **Construção do(s) Fato(s)**

Após a construção das Dimensões, nesta fase passamos para a estrutura do fatos. A tabela dos fatos é a responsável por receber todas as ligações das tabelas de dimensão(sendo essencialmente uma tabela com as FK todas dessas tabelas).

- **Povoamento do DW** Nesta fase será utilizada a ferramenta talend para povoamento do DW e também para que seja feito o relacionamento dos atributos da base de dados inicial e do dataset com o DW a ser criado. Num momento mais à frente neste relatório este processo será concretamente detalhado.

### 3.2 Perguntas às quais se pretende que o modelo criado responda no futuro

Para a criação do DW foi determinado que ele respondesse a uma série de perguntas para que fosse possível, através dele, ser representada a informação pretendida. Essas perguntas são as seguintes:

- Quais os países cuja população seja menor que 25M ?
- Quais os países europeus que têm vermelho na sua bandeira ?
- Qual a população de uma dada religião ?
- Quais os países africanos que têm uma área entre 300 e 500 ?
- Quais os países que constam no estudo ?
- Qual a população de um dado continente?
- Quais os países que têm como língua-mãe o alemão?

### 3.3 Decisões

No tópico seguinte serão apresentadas e explicadas todas as decisões tomadas pelo grupo para obter o modelo de DW que poderá ser visto no próximo tópico.

- **Escolha das perguntas para orientar a criação do DW:** Tiveram que ser definidas quais as perguntas que queríamos que o nosso modelo de DW responde-se e, em função disso começamos a idealizar o modelo com base nos passos referidos a cima.

- **Quais tabelas de dimensão a representar:** As tabelas de dimensão que foram decididas de representar são as seguintes: Landmass, Language, Country, Flag, Religion.

A tabela Landmass contém informação relativa aos continentes, a tabela Language possui informação relativa á linguagem falada em cada país, a tabela Country tem informação sobre a área e a população de um país, a Flag apresenta as características de uma dada bandeira e por fim a tabela Religion possui informação sobre uma dada religião.

- **Seleção dos atributos para as tabelas de dimensão:** Para todas as tabelas referidas no tópico anterior, tivemos de fazer uma seleção dos atributos que achamos relevantes para colocar em cada tabela

### 3.4 Modelo de DW

Na figura seguinte poderá ser observado o modelo de data Warehouse elaborado de forma a que seja possível ser dada resposta às perguntas inicialmente propostas.

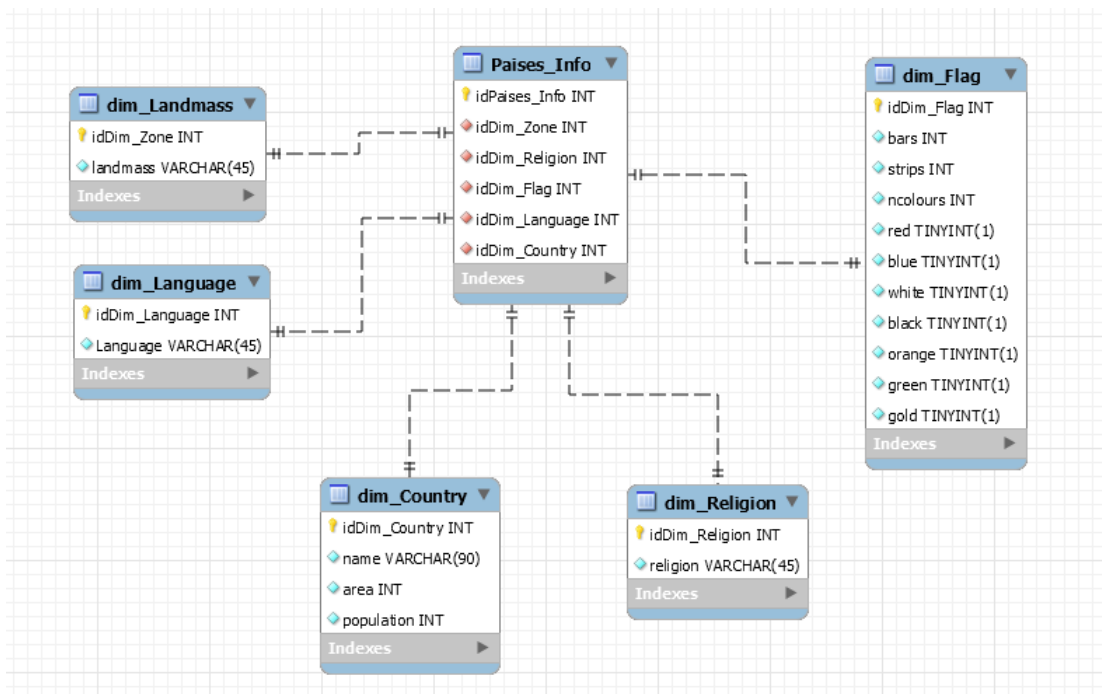


Figura 2: Modelo de DW

## 4 Talend

Nesta fase do relatório será explicado detalhadamente todo o processo de povoamento do data Warehouse através da base de dados inicial e também através de um ficheiro .csv. Para efetuar este processo são necessárias estabelecer duas conexões, uma à base de dados inicial e outra à base de dados do data Warehouse, sendo também necessário importar os ficheiros .csv pois a base de dados inicial não está totalmente povoada.

### 4.1 Metadata

O Metadata no talend é onde estão guardadas as conexões e os ficheiros importados. Neste projeto em particular foram usados quatro como se pode ver na imagem seguinte.

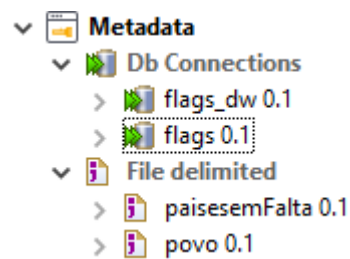


Figura 3: Metadata

Nas imagens seguintes podemos visualizar o estabelecimento da conexão a uma base de dados e a importação de um ficheiro.

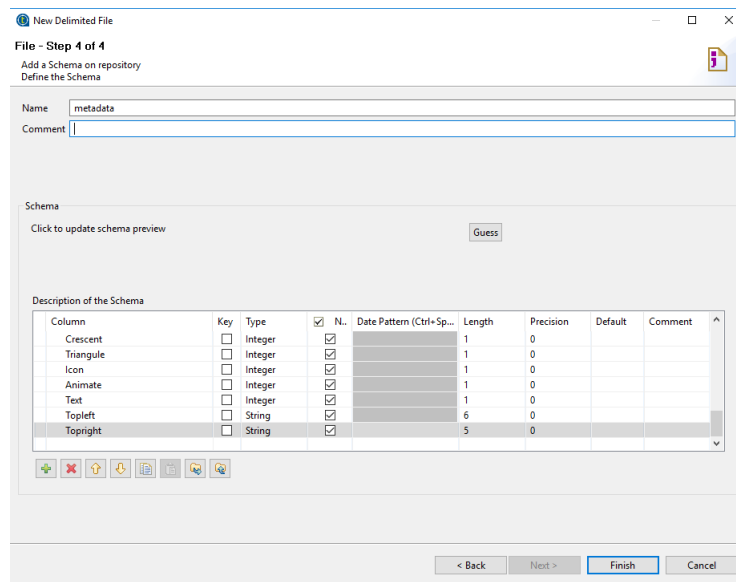


Figura 4: Import CSV

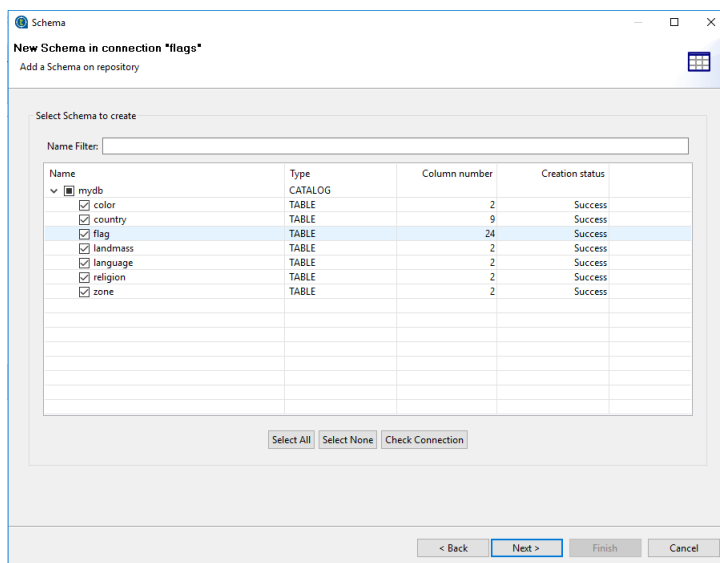


Figura 5: Import Tables CSV

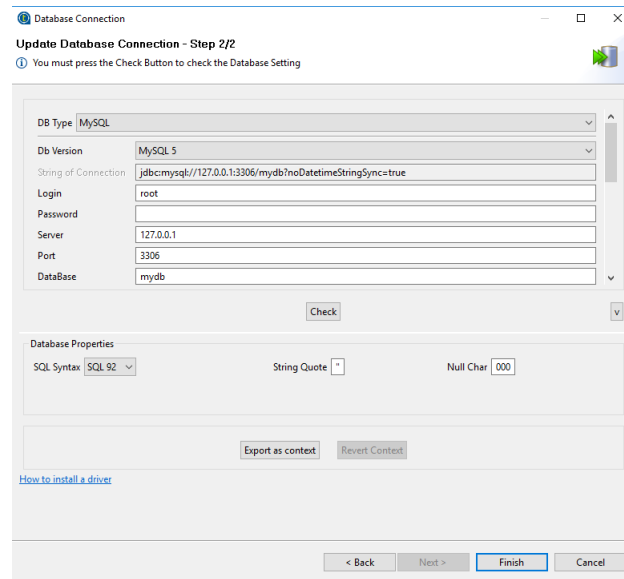


Figura 6: Conexão à Base de dados

#### 4.1.1 Flags\_dw

O Flags\_dw representa a conexão que estabelecemos com a base de dados do data warehouse.

#### 4.1.2 Flags

O Flags representa a conexão estabelecida com a base de dados inicial.

#### 4.1.3 Povo

O Povo representa o ficheiro .csv inicial (dataset).

#### 4.1.4 Paísesemfalta

O Paísesemfalta representa a lista de todos os países e atributos da tabela country.



## 4.2 Jobs

Os jobs servem para efetuar o povoamento da nossa data Warehouse através da base de dados inicial.

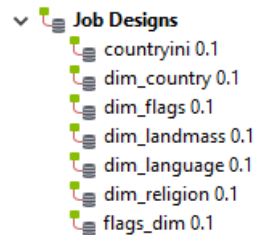


Figura 7: Jobs

### 4.2.1 Dim\_country

Como a tabela Country não está totalmente povoada será necessário também de utilizar o ficheiro .csv, de forma a que o povoamento final da nossa dim\_country fique completo.

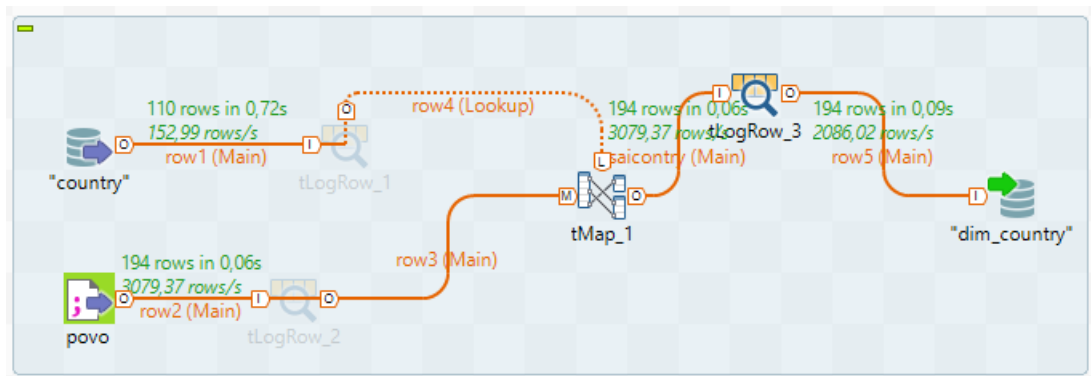


Figura 8: Dim Country

#### 4.2.2 Dim\_flags

Esta tabela vai conter toda a informação sobre as bandeiras definidas como relevantes numa fase anterior do relatório.

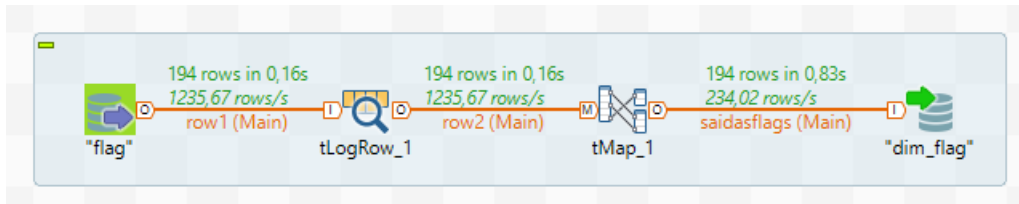


Figura 9: Dim Flags

#### 4.2.3 Dim\_landmass

Esta tabela vai conter toda a informação sobre os continentes.

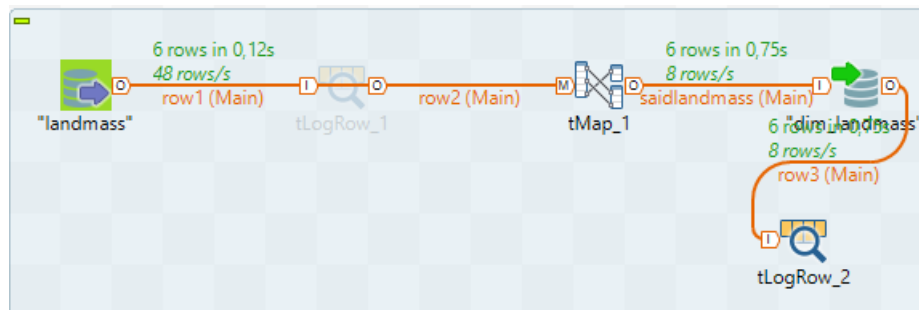


Figura 10: Dim Landmass

#### 4.2.4 Dim\_language

Esta tabela vai conter toda a informação sobre as diferentes línguas.

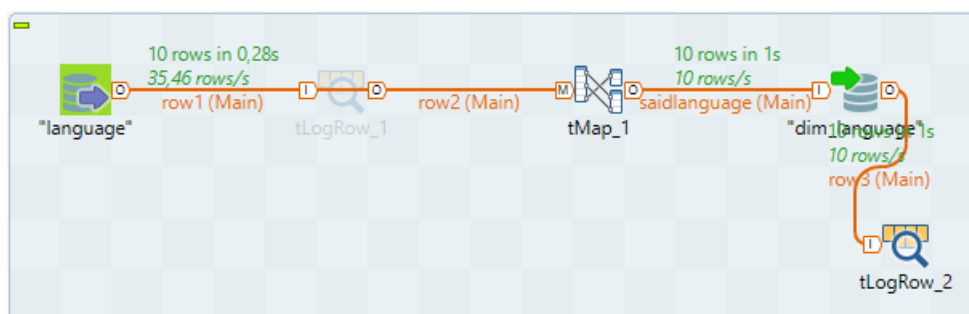


Figura 11: Dim Language

#### 4.2.5 Dim\_Religion

Esta tabela vai conter toda a informação sobre as religiões.

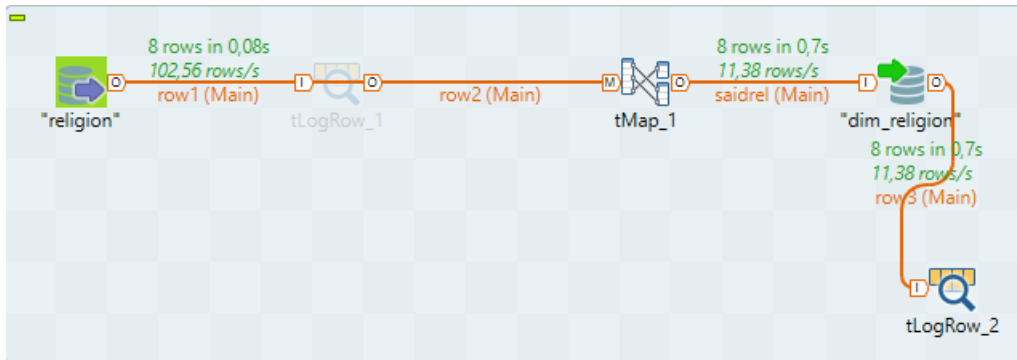


Figura 12: Dim Language

#### 4.2.6 Flags\_dim

Para obter o resultado pretendido neste dim final teve que se acrescentar um input extra ao tMap para além de todas as outras tabelas de dim. Este input que contém a lista de todos os atributos da tabela Country num ficheiro .csv serviu para criar relações entre chave primária e estrangeira com todas as tabelas de dimensão e assim garantir que o resultado era o pretendido.

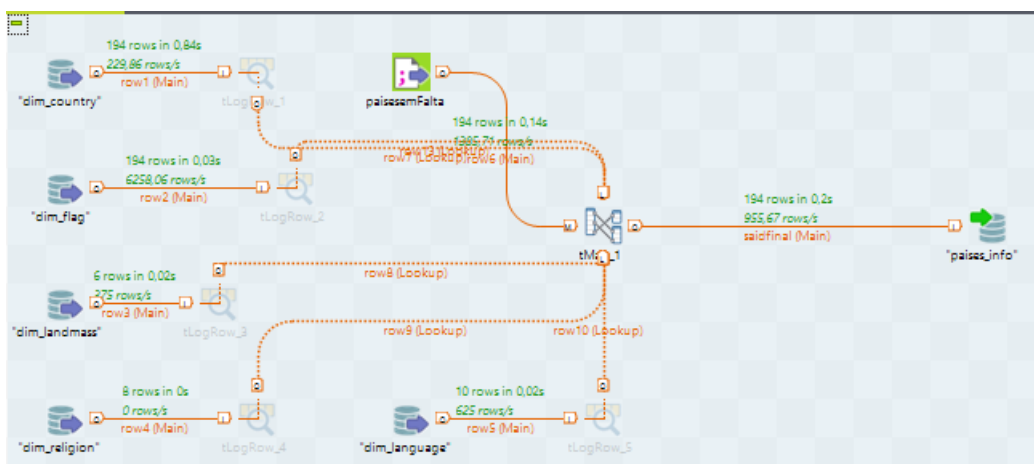


Figura 13: Job Final

### 4.3 Parte Lógica - tMaps

Os tMaps vão servir para fazer a associação lógica entre as tabelas e ficheiros de origem e as tabelas de dimensão criadas por nós na fase anterior a esta. Ou seja, servem para que se indiquem quais os atributos da tabela inicial que vão pertencer à tabela final.

#### 4.3.1 Dim\_country

Este tMap representa a associação dos atributos da tabela Country e do ficheiro .csv com a tabela final dim\_country. De notar ainda que a tabela Country e o ficheiro .csv têm ligações entre eles, de forma a garantir que o resultado final é o pretendido.

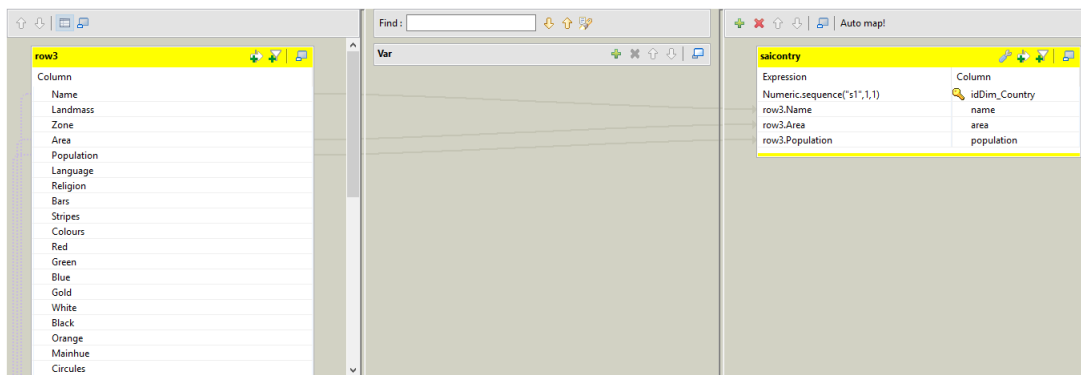


Figura 14: tMap Country - Parte 1

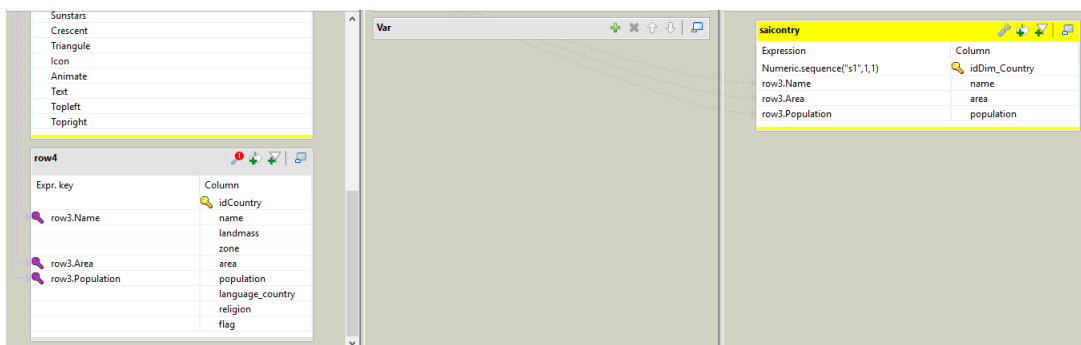


Figura 15: tMap Country - Parte 2

### 4.3.2 Dim\_flags

Este tMap representa a associação dos atributos da tabela Flags com a tabela final dim\_flags.

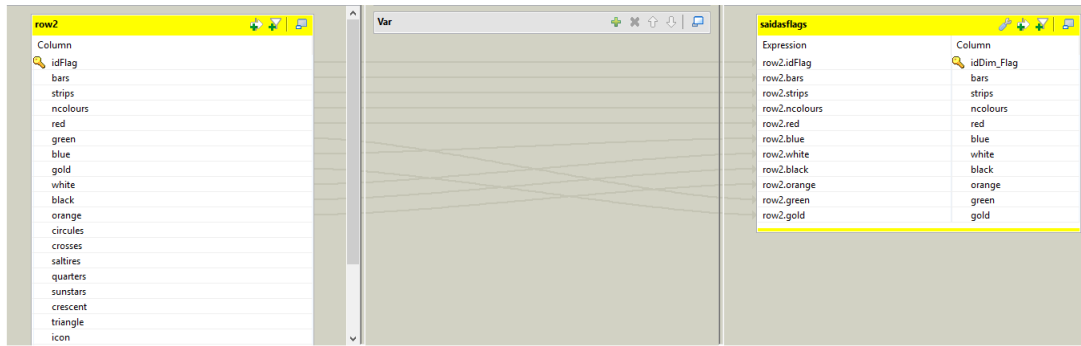


Figura 16: tMap Flags

### 4.3.3 Dim\_landmass

Este tMap representa a associação dos atributos da tabela Landmass com a tabela final dim\_landmass.

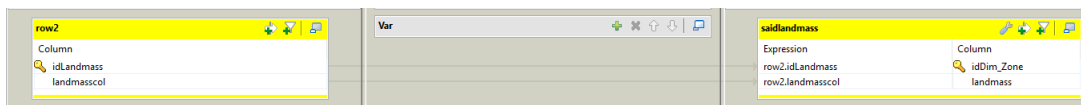


Figura 17: tMap landmass

### 4.3.4 Dim\_language

Este tMap representa a associação dos atributos da tabela Language com a tabela final dim\_language.

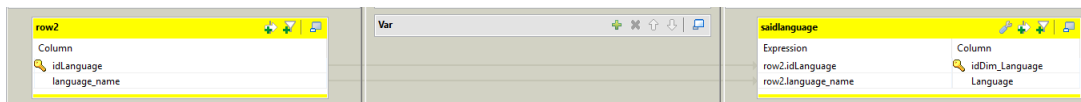


Figura 18: tMap language

#### 4.3.5 Dim\_Religion

Este tMap representa a associação dos atributos da tabela Religion com a tabela final dim\_religion.

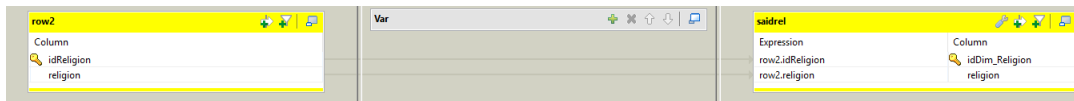


Figura 19: tMap religion

#### 4.3.6 Flags\_dim

Este tMap representa a associação dos atributos de todas as tabelas de dimensão e o ficheiro .csv (contendo a informação toda sobre os países) com a tabela final dim\_religion.

De notar que existem associações entre todas as tabelas de input e o ficheiro .csv, de modo a garantir que o resultado obtido é o pretendido.

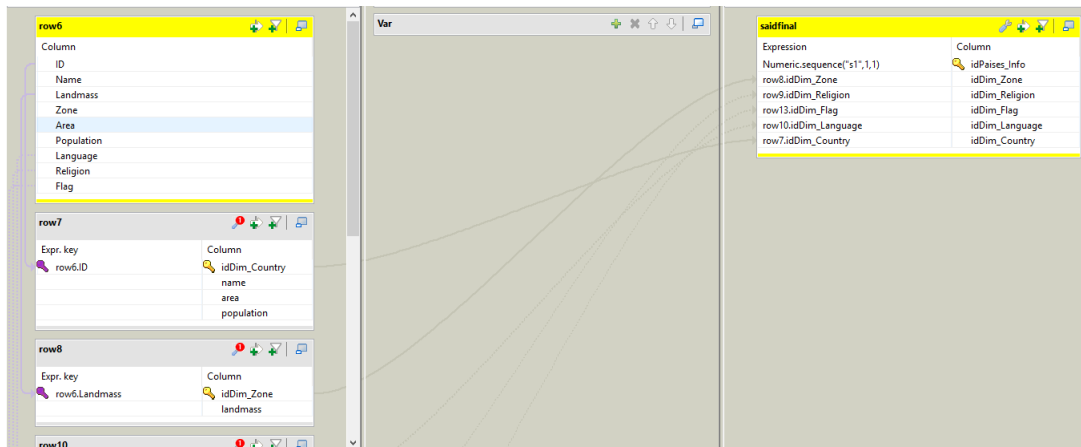


Figura 20: tMap flags final - Parte 1

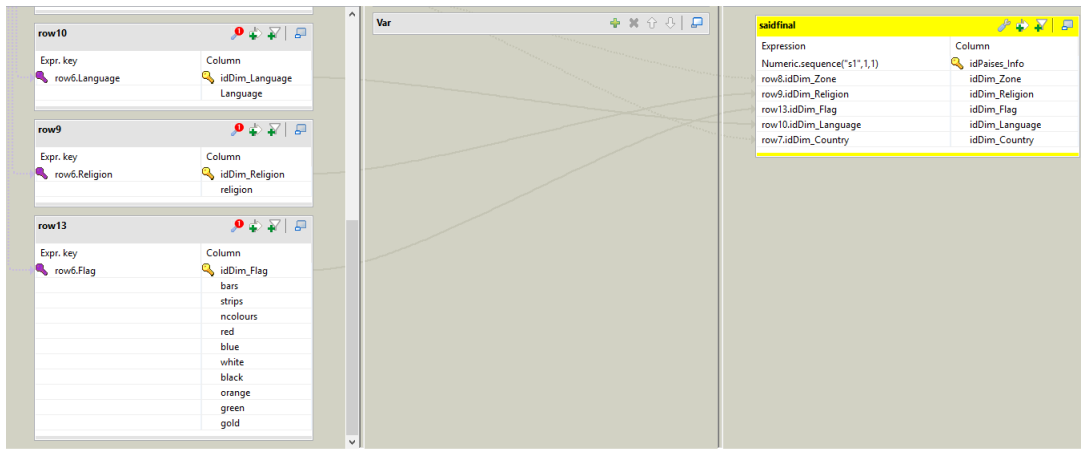


Figura 21: tMap flags final - Parte 2

## 5 Power BI

O Power BI foi a ferramenta escolhida para representar e visualizar os dados obtidos através do DW quando desejado. Pois esta informação pode ser publicada de modo a ser partilhada com todos os intervenientes num futuro negócio onde os intervenientes não estejam por dentro dos conceitos de elaboração e funcionamento de uma base de dados.

### 5.1 Perguntas

Estas "perguntas" apresentadas de seguida surgem em resposta às perguntas que serviram de base para criar o modelo de data Warehouse, sendo que ainda serão apresentadas mais algumas que foram consideradas oportunas de ver respondidas.

#### 5.1.1 Países cuja população é menor que 25M

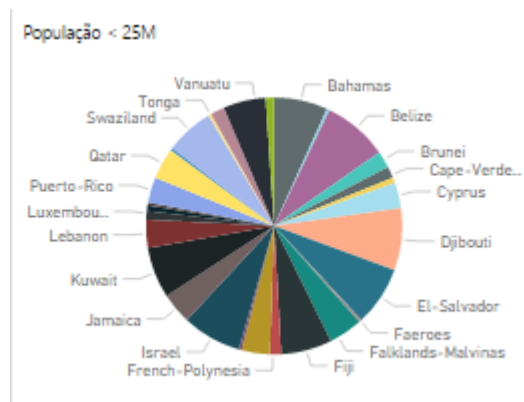


Figura 22: Países cuja população é menor que 25M



### 5.1.2 Países de África que tem uma área entre 300 e 500

Países Africanos com 300<area<500

194	Zimbabwe
idDim_Country	name
391	Africa
area	landmass
118	Morocco
idDim_Country	name
447	Africa
area	landmass
89	Ivory-Coast
idDim_Country	name
323	Africa
area	landmass
41	Congo
idDim_Country	name
342	Africa

Figura 23: Países de África que tem uma área entre 300 e 500

### 5.1.3 Países da Europa que tem vermelho na bandeira

Países Europeus com Vermelho na Bandeira

red	name	landmass
True	Albania	Europe
True	Andorra	Europe
True	Austria	Europe
True	Belgium	Europe
True	Bulgaria	Europe
True	Czechoslovakia	Europe
True	Denmark	Europe
True	Faeroes	Europe
True	France	Europe
True	Germany-DDR	Europe
True	Germany-FRG	Europe
True	Gibraltar	Europe
True	Hungary	Europe
True	Iceland	Europe
True	Italy	Europe
True	Liechtenstein	Europe
True	Luxembourg	Europe
True	Malta	Europe
True	Monaco	Europe
True	Netherlands	Europe
True	Norway	Europe
True	Poland	Europe
True	Portugal	Europe

Figura 24: Países da Europa que tem vermelho na bandeira

#### 5.1.4 Países que falam Alemão

Países que falam Alemão

Language	name
German	Austria
German	Germany-DDR
German	Germany-FRG
German	Liechtenstein
German	Luxembourg
German	Switzerland

Figura 25: Países que falam Alemão

#### 5.1.5 Mapa mundo com países em estudo



Figura 26: Mapa mundo com países em estudo

### 5.1.6 Número de listas e barras em função de um continente



Figura 27: Número de listas e barras em função de um continente

### 5.1.7 População Total



Figura 28: População Total

### 5.1.8 Área Total



Figura 29: Área Total

### 5.1.9 Cores Totais



Figura 30: Cores Totais

### 5.1.10 Total Bars (barras)



Figura 31: Total bars

### 5.1.11 Total Stripes (listas)



Figura 32: Total Stripes

### 5.1.12 Listar todos os Países



name
Afghanistan
Albania
Algeria
American-Samoa
Andorra
Angola
Anguilla
Antigua-Barbuda
Argentina
Argentine
Australia
Austria
Bahamas
Bahrain
Bangladesh
Barbados
Belgium
Belize
Benin
Bermuda
Bhutan
Bolivia
Botswana
Brazil
British-Virgin-Isles
Brunei
Bulgaria
Burkina
Burma
Burundi
Cameroon
Canada
Cape-Verde-Islands
Cayman-Islands
Central-African-Republic
Chad
Chile
China

Figura 33: Países Totais

### 5.1.13 Relação entre uma língua, os países, o total da população, da área, das listas, das colunas e das cores da bandeira

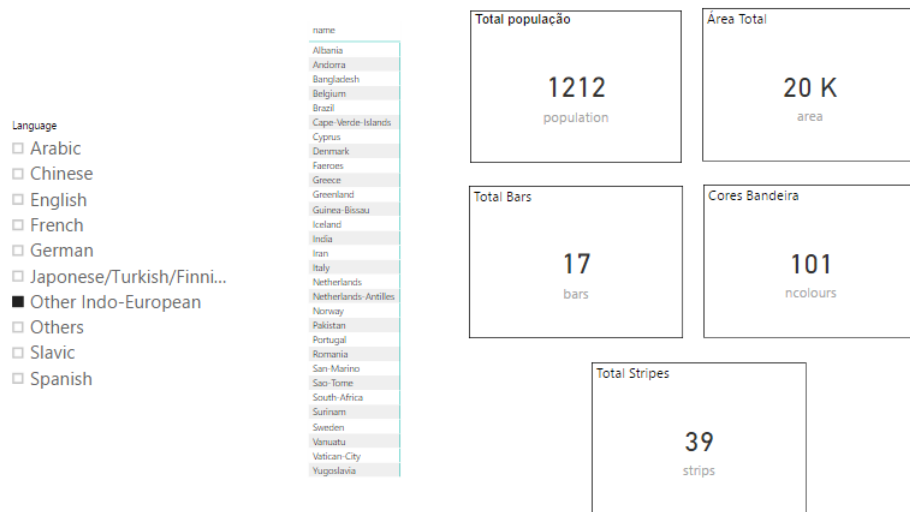


Figura 34: Relação 1

### 5.1.14 Área por Continente



Figura 35: Área por continente

### 5.1.15 Análise da população por continente

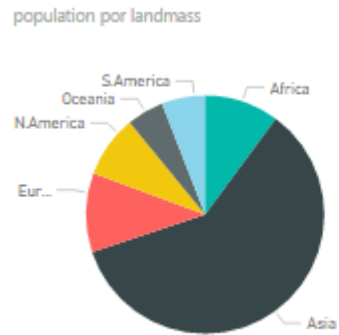


Figura 36: População por continente 1

### 5.1.16 População por continente

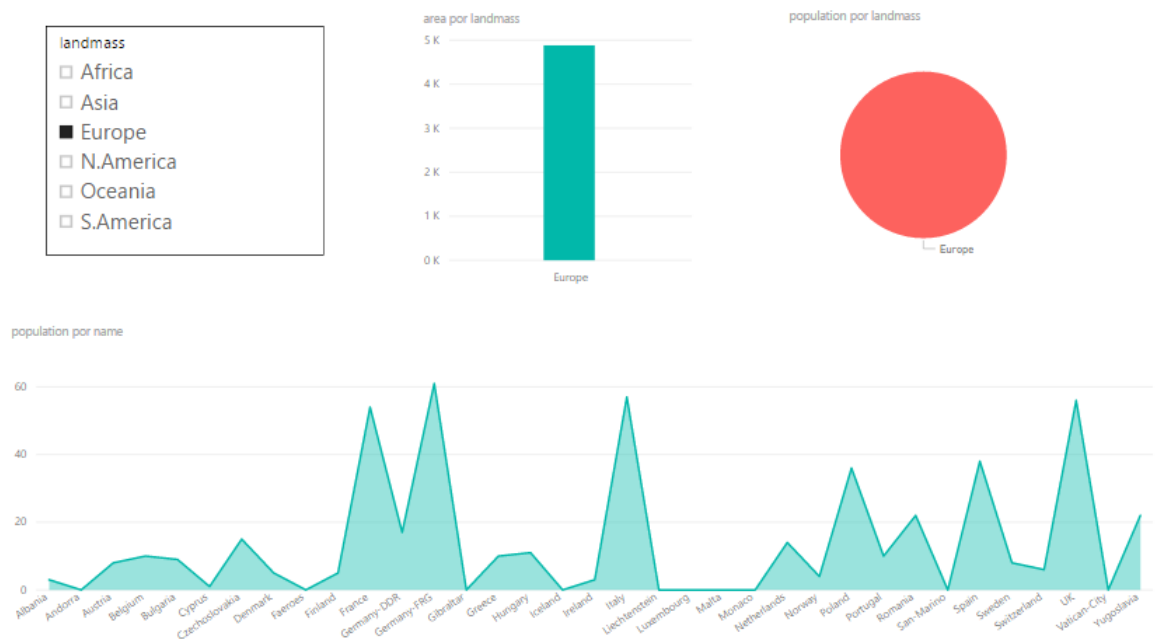


Figura 37: População por continente 2

### 5.1.17 Lista de países de um continente que tem na bandeira verde, laranja, vermelho e não tem branco

Cape-Verde-Isla...	True
name	green
True	True
red	orange
Falso	Africa
white	landmass

Zaire	True
name	green
True	True
red	orange
Falso	Africa
white	landmass

Zambia	True
name	green
True	True
red	orange
Falso	Africa
white	landmass

Figura 38: Lista países por uma característica na bandeira

### 5.1.18 Relação com as stripes, bars entre continentes

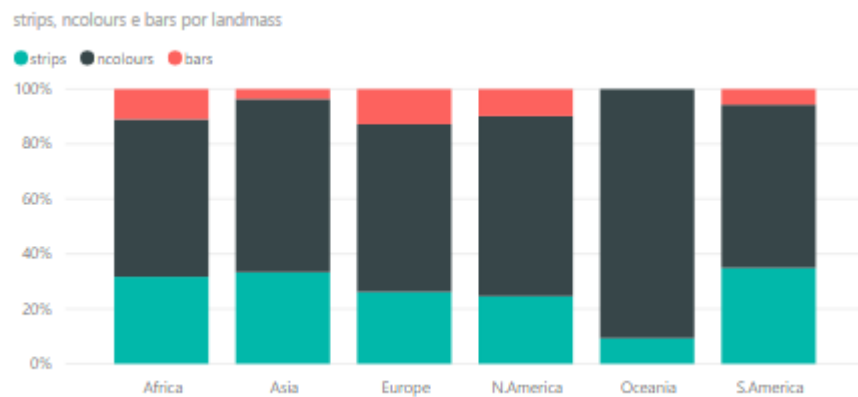


Figura 39: Relação stripes,bars com continentes



### 5.1.19 Nome do país, área e população dos países que tem a religião 'Muslism'

name	population	area	religion
Afghanistan	16	648	Muslim
Algeria	20	2388	Muslim
Bahrain	0	1	Muslim
Bangladesh	90	143	Muslim
Brunei	0	6	Muslim
Comorro-Islands	0	2	Muslim
Djibouti	0	22	Muslim
Egypt	47	1001	Muslim
Guinea	6	246	Muslim
Indonesia	157	1904	Muslim
Iran	39	1648	Muslim
Iraq	14	435	Muslim
Jordan	2	98	Muslim
Kuwait	2	18	Muslim
Lebanon	3	10	Muslim
Libya	3	1760	Muslim
Malaysia	13	333	Muslim
Maldivo-Islands	0	0	Muslim
Mali	7	1240	Muslim
Mauritania	2	1031	Muslim
Morocco	20	447	Muslim
Niger	5	1267	Muslim
Nigeria	56	925	Muslim
North-Yemen	9	195	Muslim
Oman	1	212	Muslim
Pakistan	84	804	Muslim
Qatar	0	11	Muslim
Saudi-Arabia	9	2150	Muslim
Senegal	6	196	Muslim
Somalia	5	637	Muslim
South-Yemen	2	288	Muslim
Sudan	20	2506	Muslim
Syria	10	185	Muslim
Tunisia	7	164	Muslim
Turkey	45	781	Muslim
UAE	1	84	Muslim
<b>Total</b>	<b>701</b>	<b>23786</b>	

Figura 40: Informação país em função da Religião

### 5.1.20 População de uma dada religião

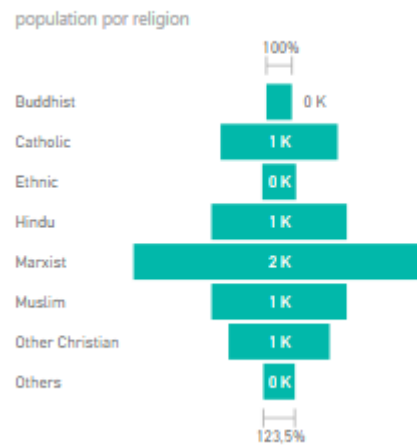


Figura 41: População de uma dada religião

## 6 Conclusão

Com a realização deste trabalho prático foram perceptíveis e compreendidas as grandes vantagens da existência de um sistema analítico (BI), mais concretamente na forma de um Data Warehouse, onde por um lado há necessidade de respostas em tempo útil face ao grande volume de dados tratado, e por outro lado se lidam com fontes de dados heterogêneas (os dados são provenientes de uma base de dados relacional e de um ficheiro .csv).

Em jeito de conclusão, pode-se dizer que este trabalho foi extremamente importante, na medida em que nos permitiu conhecer um "novo tipo" de forma de armazenamento de dados que não as típicas bases de dados relacionais e não relacionais, e um "novo tipo" de forma de apresentação de dados a um futuro cliente.