

UNIVERSIDADE DO MINHO

APLICAÇÕES INFORMÁTICAS NA BIOMEDICINA

Trabalho Prático - Ficha 05

Autores:

Daniel Vieira
João Palmeira
José Ramos

Números Aluno:

A73974
A73864
A73855

19 de Novembro de 2019

Conteúdo

1	Introdução	3
2	Jobs- Criação e Execução	4
2.1	Job1 - Criação e Execução	4
2.2	Job2 - Criação e Execução	7
2.3	Job3 - Criação e Execução	11
3	Conclusões	14

Lista de Figuras

1	Job1- Design	4
2	Operações efectuadas em tMap_1 com a respectiva seleção de colunas	5
3	Criação de uma tabela denotada "mentalHealthDATA" com as colunas seleccionadas em tMap ₁	5
4	Filtro aplicado na tabela "mentalHealthDATA"	6
5	Filtro aplicado na tabela "mentalHealthDATA"	6
6	Job2- Design	7
7	Job2- DBConnection	7
8	tMap_1	8
9	Criação de uma tabela denotada "mentalHealth_History_and_Stories" com as colunas seleccionadas em tMap_1	8
10	Filter aplicado	9
11	Resultados - Ficheiro de Output	9
12	Refects - Ficheiro de Output	10
13	Job3 - Design	11
14	tMap1 - criação dos outs de sobre os quais trabalhamos	12
15	tFilterRow - filtragem por género Feminino	12
16	tSortRow - organização por idades	12
17	tMap2 - normalização do valor de género para Female	13
18	DBoutput para uma tabela "Table Female"	13
19	Resultados Feminino - Ficheiro de Output	13

1 Introdução

O objetivo deste relatório é apresentar de forma sucinta o problema colocado na ficha5 da cadeira de **Aplicações Informáticas na Biomedicina**.

Neste projeto foi nos facultado um ficheiro *csv*, denominado de **mental_health** que contém informações pessoais e profissionais de imensas pessoas, além de um questionário acerca de perturbações mentais que possam ocorrer nos trabalhadores de **TI**.

O objetivo final deste trabalho é apresentar **Jobs** que consigam incorporar os requisitos expostos no enunciado.

2 Jobs- Criação e Execução

2.1 Job1 - Criação e Execução

Para esta primeira job, do *dataset* dado retiramos as seguintes colunas:

- Age
- Country
- NumberofEmployees
- Tratamento

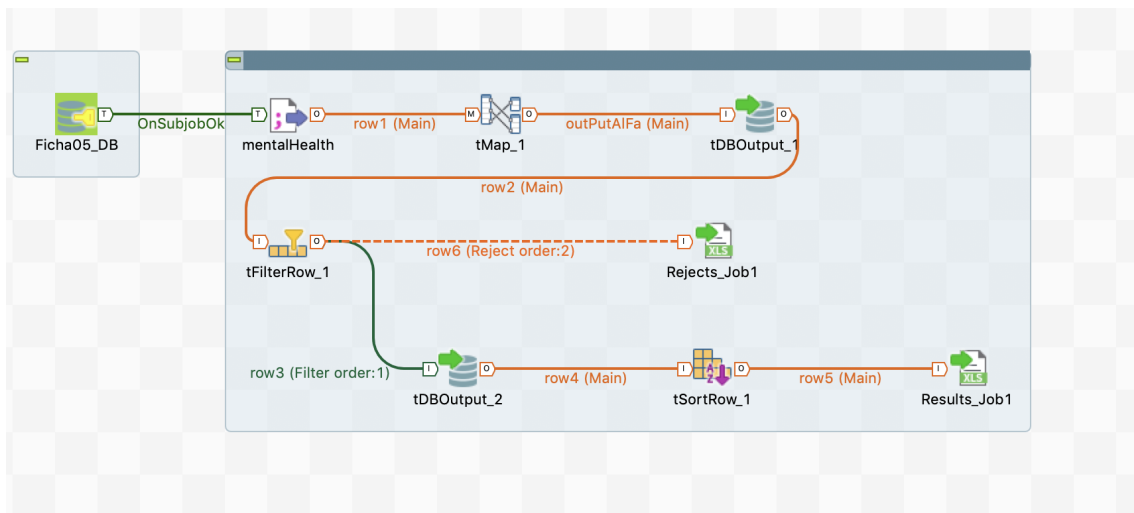


Figura 1: Job1- Design

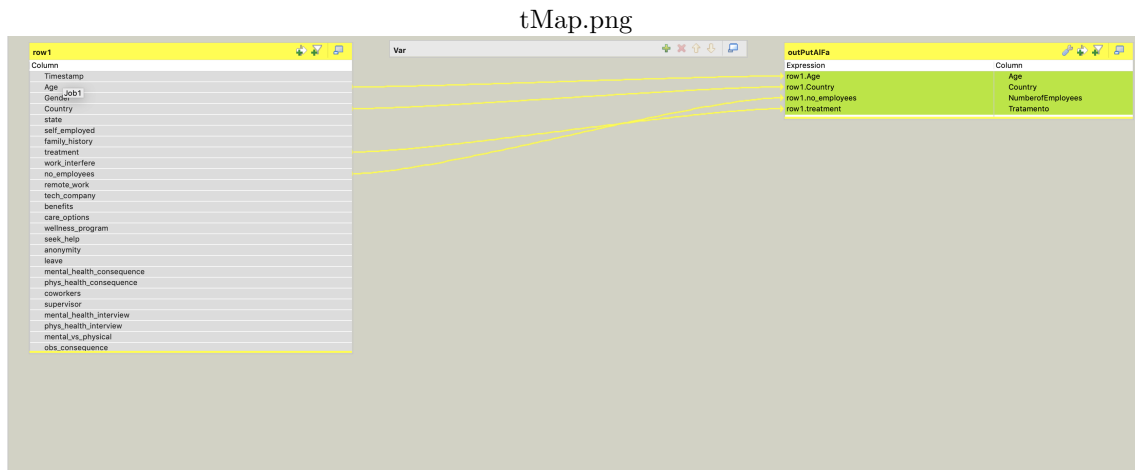


Figura 2: Operações efectuadas em tMap_1 com a respectiva seleção de colunas

Depois de termos feito a conexão com a base de dados criada em **MySQL**, foi a hora de criarmos uma tabela com as colunas escolhidas no **TMAP**.

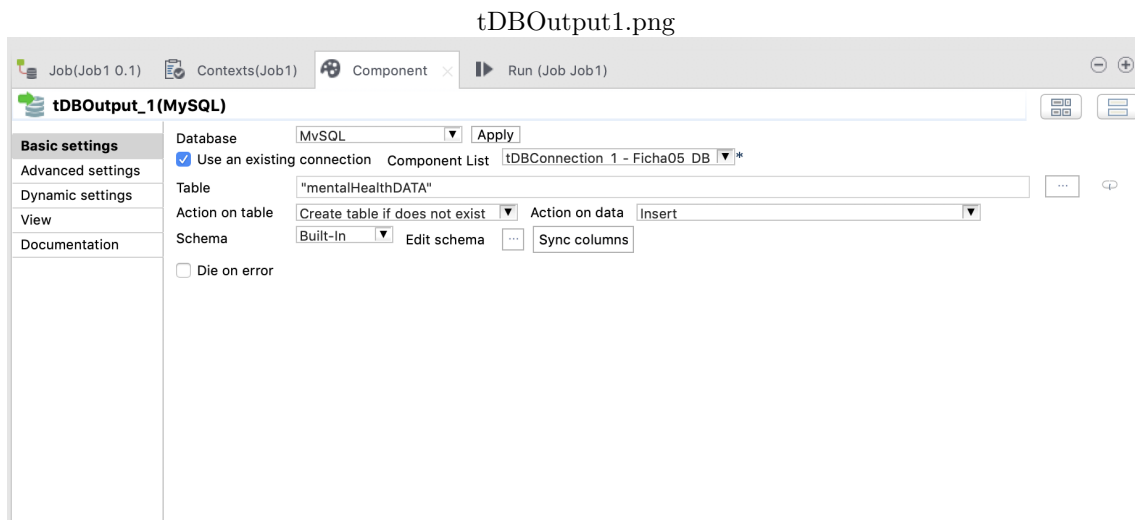


Figura 3: Criação de uma tabela denotada "mentalHealthDATA" com as colunas seleccionadas em tMap₁

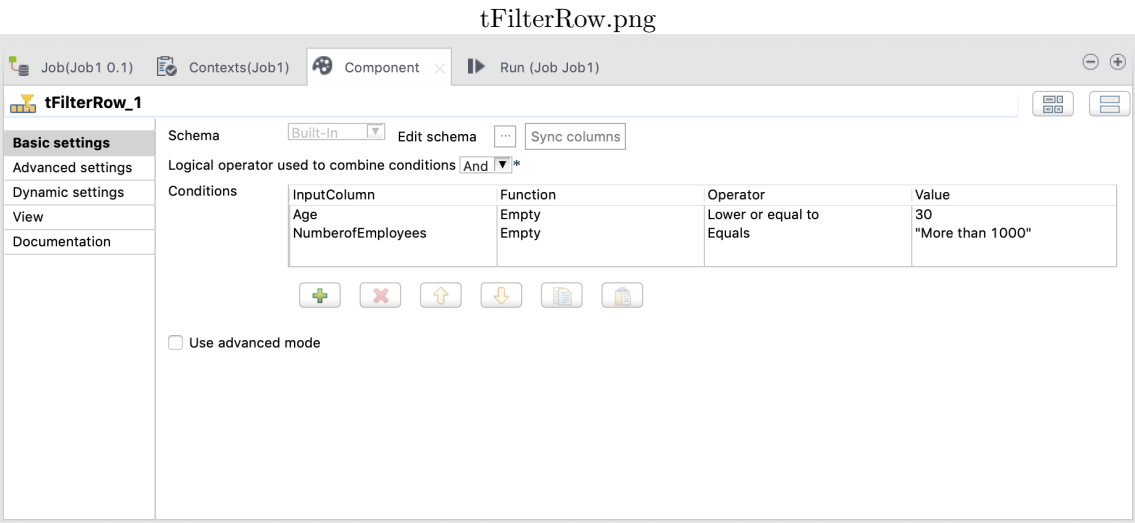


Figura 4: Filtro aplicado na tabela "mentalHealthDATA"

O objetivo final desta **job** é encontrar jovens em empresas com muitos trabalhadores. Portanto de acordo com a nossa premissa anteriormente explicada limitamos as pessoas com idade inferior ou igual a 30 bem como o *numberOfEmployees* ser superior ou igual a 1000. Visto que entendemos que a idade é um fator importante para a saúde mental, achamos que devemos ordenar de acordo com a idade de forma ascendente.

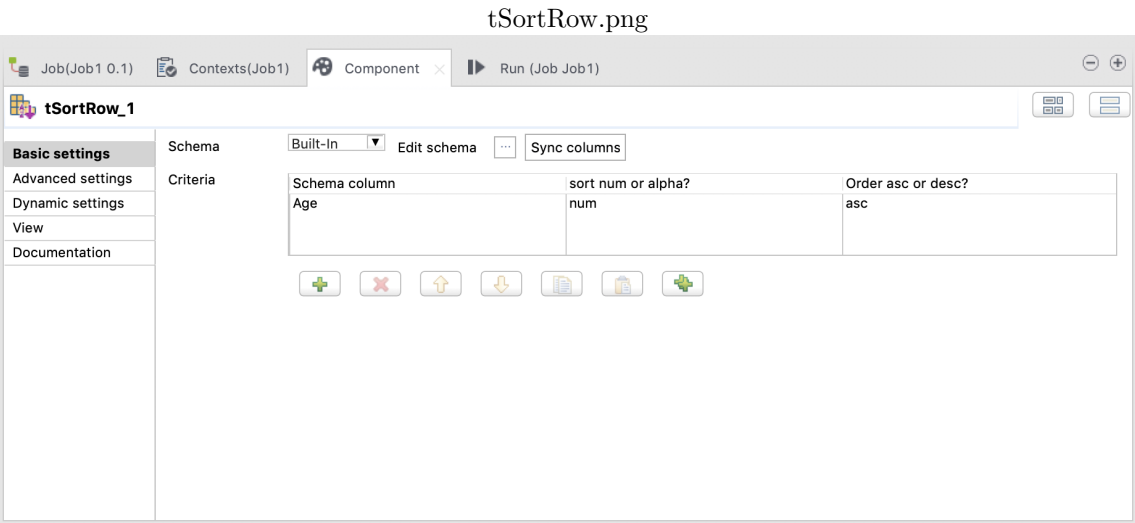


Figura 5: Filtro aplicado na tabela "mentalHealthDATA"

2.2 Job2 - Criação e Execução

A segunda job centra-se no objectivo de criar tabelas permitam investigar se pessoas com family_history e care_options apresentam mental_health_consequence ou não.

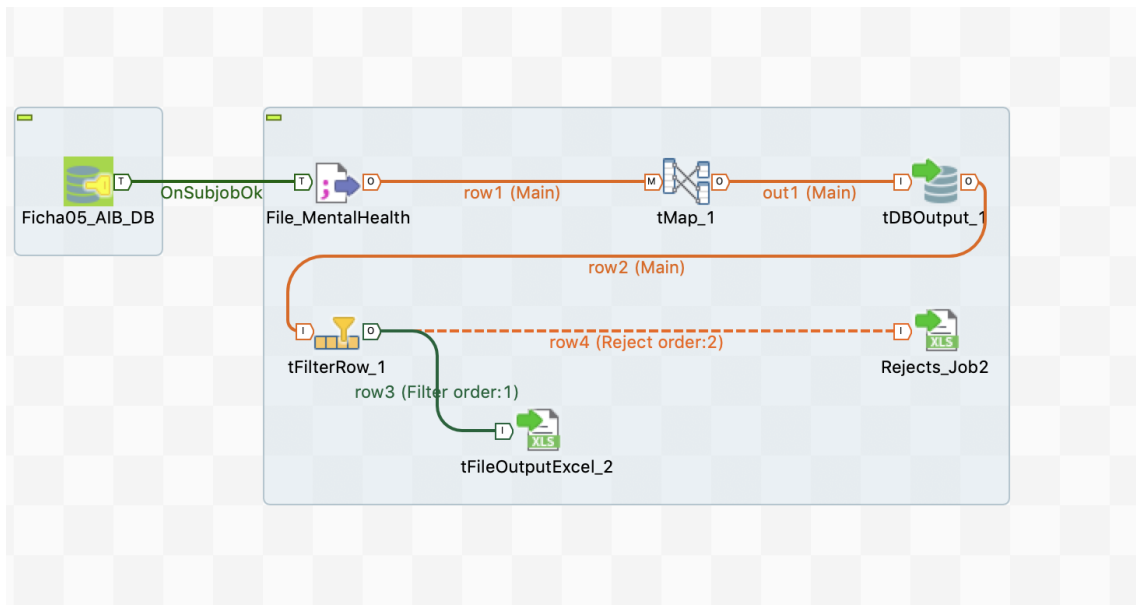


Figura 6: Job2- Design

Job(Job2 0.1) Contexts(Job2) Component Run (Job Job2)

Ficha05_AIB_DB(tDBConnection_1)(MySQL)

Basic settings

Database: MySQL Apply

Property Type: Repository DB (MYSQL):Ficha05_AIB

DB Version: Mysql 8

Host: "" Port: "3306"

Database: "Ficha05_AIB_Grupo2" Additional JDBC Parameters: "noDatetimeStringSync=true"

Username: "root" Password: "*****"

☐ Use or register a shared DB Connection

Data source

This option only applies when deploying and running in the Talend Runtime

☐ Specify a data source alias

Figura 7: Job2- DBConnection

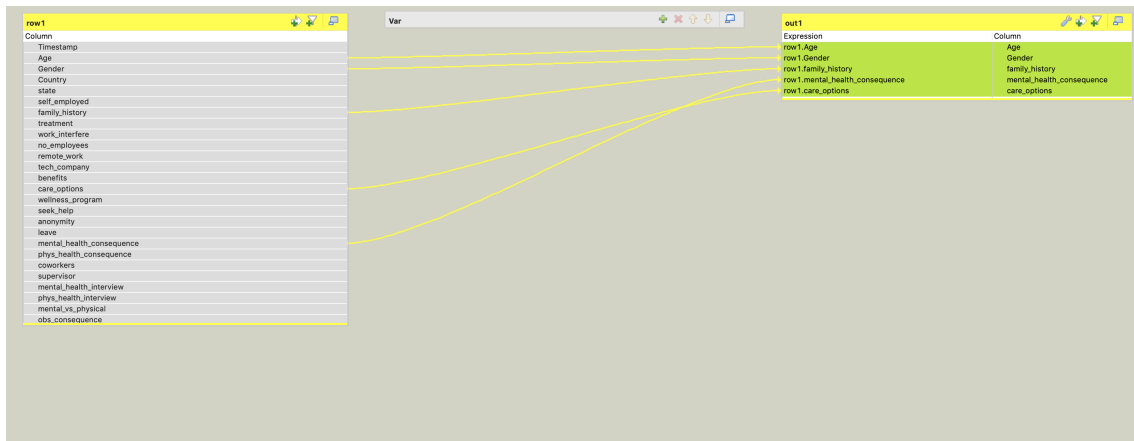


Figura 8: tMap_1

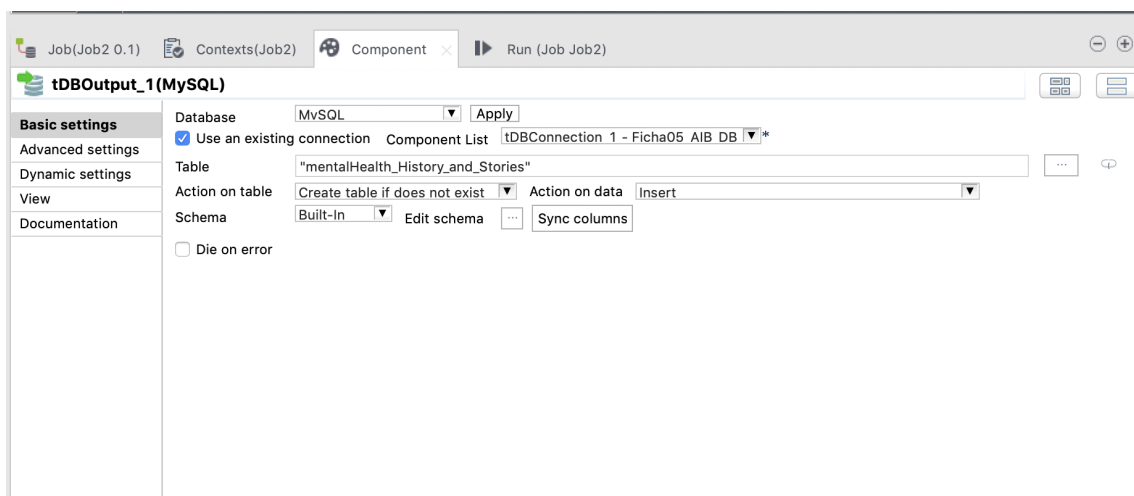


Figura 9: Criação de uma tabela denotada "mentalHealth_History_and_Stories" com as colunas selecionadas em tMap_1

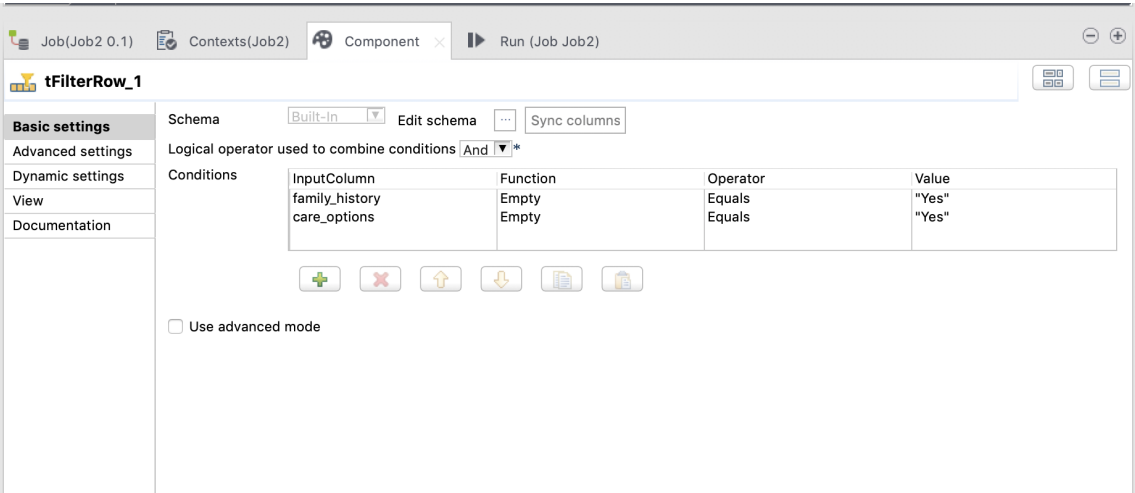


Figura 10: Filter aplicado

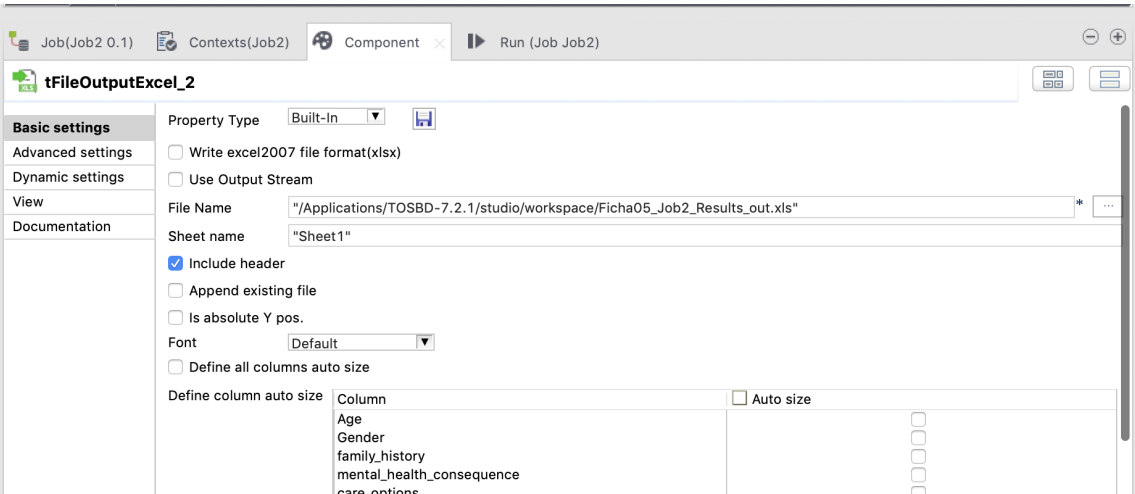


Figura 11: Resultados - Ficheiro de Output

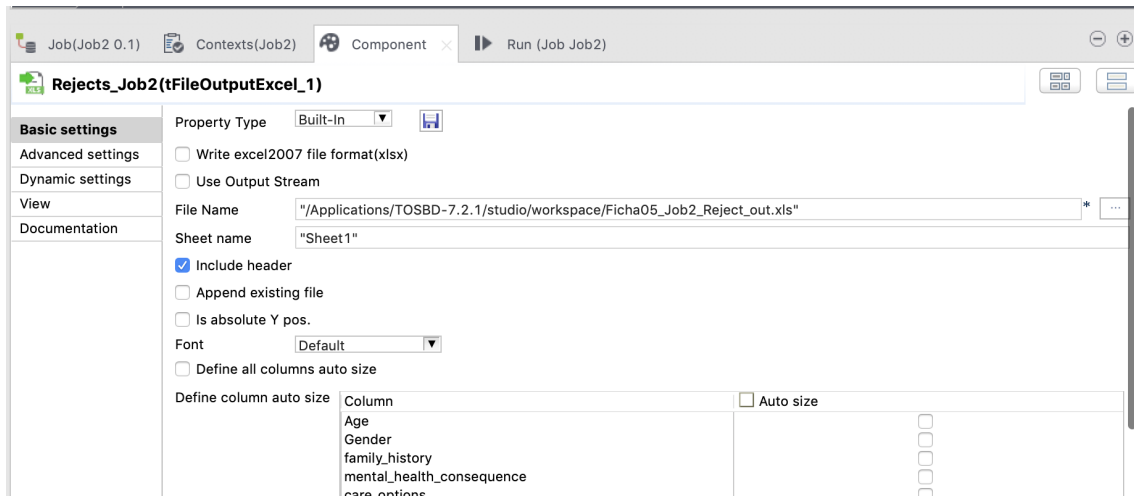


Figura 12: Rejects - Ficheiro de Output

Neste job abordamos o problema utilizando o tMap1 um out do dataset sendo tratado para filtrar a informação desejada. Depois filtra-mos por care_options e family_history positivos também para termos a noção da sua influencia nos dados relativamente a se achariam as pessoas que haveria mental_consequences de mencionar problemas desses no trabalho..

Os Dados obtidos da filtragem são então colocados num ficheiro Excel para os podermos avaliar. Os dados que não são aceites pelas condições de filtragem são colocados num ficheiro de Rejects.

2.3 Job3 - Criação e Execução

A terceira job centra-se no objectivo de criar tabelas que se distinguíssem no género do trabalhador e investigar sobre factores que teriam relativamente a doenças mentais dispondo-as assim. Assim apartir do dataset disponibilizado decidimos que as colunas de atributos essenciais à resolução do problema seriam as seguintes

- Age
- Genger
- self_employed
- no_employees
- teatment
- leave
- Country

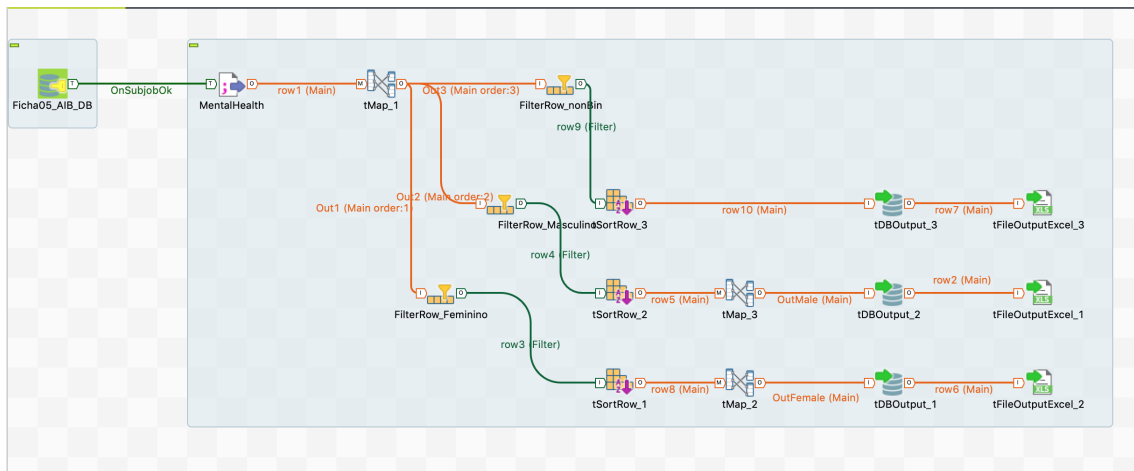


Figura 13: Job3 - Design



Figura 14: tMap1 - criação dos outs de sobre os quais trabalhamos

The screenshot shows the configuration window for 'FilterRow_Feminino(tFilterRow_2)'. The 'Basic settings' tab is selected. The 'Schema' is set to 'Built-in'. The 'Logical operator used to combine conditions' is set to 'Or'. The 'Conditions' table is as follows:

InputColumn	Function	Operator	Value
Gender	Empty	Equals	"Female"
Gender	Empty	Equals	"female"
Gender	Empty	Equals	"F"

There are buttons for adding (+), removing (-), and moving (up/down) conditions. A checkbox for 'Use advanced mode' is at the bottom.

Figura 15: tFilterRow - filtragem por género Feminino

The screenshot shows the configuration window for 'tSortRow_1'. The 'Basic settings' tab is selected. The 'Schema' is set to 'Built-in'. The 'Criteria' table is as follows:

Schema column	sort num or alpha?	Order asc or desc?
Age	num	asc

There are buttons for adding (+), removing (-), and moving (up/down) criteria. A green plus button is at the bottom right.

Figura 16: tSortRow - organização por idades

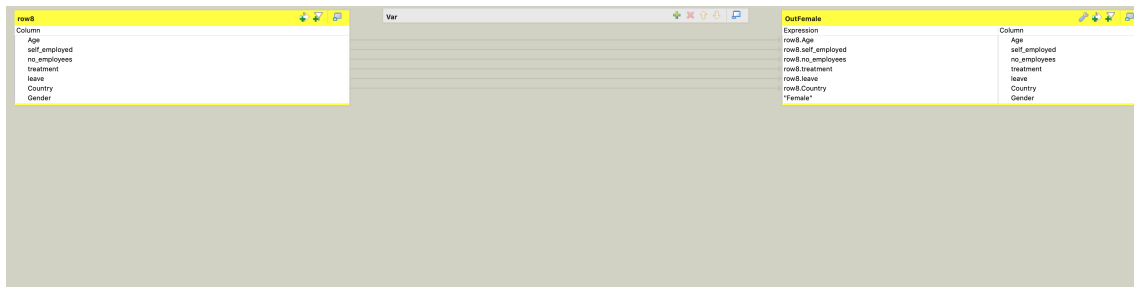


Figura 17: tMap2 - normalização do valor de género para Female

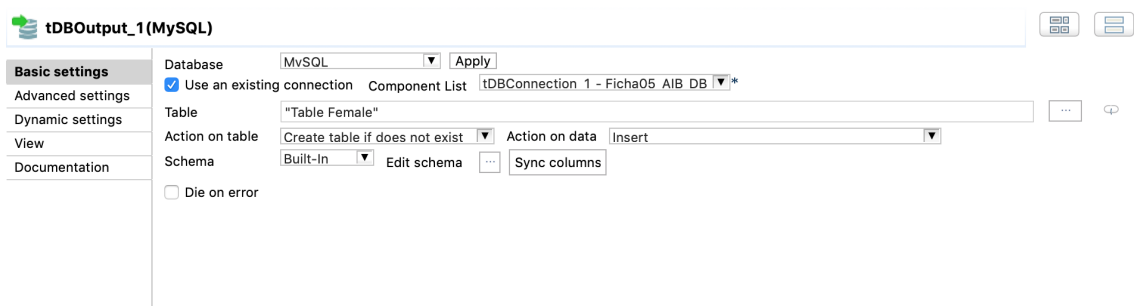


Figura 18: DBOutput para uma tabela "Table Female"



Figura 19: Resultados Feminino - Ficheiro de Output

Neste job abordamos o problema utilizando o tMap1 para criar três outs do dataset sendo que cada um ser tratado para filtrar a informação desejada por género. Depois ordenamos por idades também para termos a noção da sua influencia nos dados que procuramos debruçar nos.

Utilizamos ainda um tMap para normalizar os diferentes valores de um dado género e.g.: "Male", "male", "m" e "M" para "Male".

Por fim, para cada género, criamos uma tabela na Base de Dados inserindo nela os dados obtidos, e gerando um ficheiro Excel com os mesmos.

3 Conclusões

Acabada esta ficha, há que fazer uma retrospectiva de todo o trabalho efetuado. Com esta ficha conseguimos perceber a ferramenta Talend em paralelo com o MySQL, parceria esta que inicialmente trouxe alguns problemas mas que com o tempo foram diminuindo esses mesmos problemas.

Para finalizar, o grupo acha que concretizou os objectivos propostos inicialmente com a realização da 3 jobs.