



UNIVERSIDADE DO MINHO

MESTRADO INTEGRADO DE ENGENHARIA INFORMÁTICA

Aplicações Informáticas na Biomedicina

Ficha 5 - Grupo 4 Definição, Criação e Execução de Jobs no Talend

Autores:

António Lopes	A74357
Bruno Ferreira	A74155
Frederico Pinto	A73639
Pedro Lima	A61061

13 de Novembro de 2018

Conteúdo

1	Introdução	2
2	Jobs	4
2.1	Job 1	4
2.2	Job 2	5
2.3	Job 3	6
3	Conclusão	7

1 Introdução

A resolução desta ficha foi proposta na unidade curricular *Aplicações Informáticas na Biomedicina*, e resumidamente foi pedida a criação de três *Jobs* utilizando o *Talend*. Os *Jobs* por nós criados utilizam como referência o *dataset mental_health.csv* providenciado pelos docentes da disciplina.

Esse *dataset* contém dados reais de um estudo realizado em 2014 que avalia as atitudes dos trabalhadores em empresas de Tecnologias de Informação (TIs) relativamente à saúde mental, bem como a frequência de transtornos mentais em ambientes de trabalho.

Após analisar o *dataset*, verificamos que contém os vários campos que possuem valores distintos:

- Timestamp
- Age
- Country
- State
- Self_Employed
- Family_History
- Treatment
- Work_Interfere
- No_Employees
- Remote_Work
- Tech_Company
- Benefits
- Care_Options
- Wellness_Program
- Seek_Help
- Anonymity
- Leave
- Mental_Health_Consequence
- Phys_Health_Consequence
- Coworkers
- Supervisor
- Mental_Health_Interview
- Phys_Health_Interview

- Mental_vs_Physical
- Obs_Consequence
- Comments

2 Jobs

Nas próximas subsecções iremos apresentar os *Jobs* por nós criados. Nessa apresentação iremos falar sobre o problema que queremos resolver em cada um e toda a sua resolução.

2.1 Job 1

Neste primeiro *Job* queremos separar os utentes presentes no *dataset* por idade e se estão em tratamento.

Para resolver este problema optamos por guardar as seguintes informações dos utentes presentes no dataset:

- Age
- Gender
- State
- Self_Employed
- Country

Numa primeira instância, verificamos se a idade dos funcionários é superior a 30 anos e separamos aqueles que se encontram em tratamento daqueles que não se encontram em tratamento. Posteriormente dos funcionários que se encontram em tratamento, filtramos aqueles que têm mais de 50 anos, exportando o resultado para um ficheiro Excel. Contudo com esta filtragem, os funcionários que têm menos de 50 anos são guardados numa tabela da base de dados.

É de salientar que a informação dos funcionários que não são tratadas é guardada numa tabela da base de dados.

Apresentamos agora o *Job* criado no *Talend*:

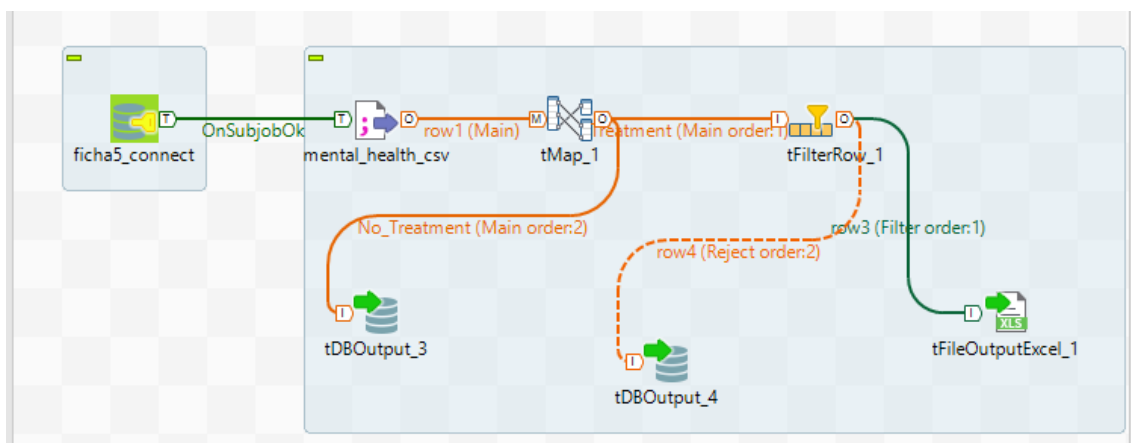


Figura 1: Job 1 Overview

2.2 Job 2

No *Job 2* temos como intuito separar os funcionários pelo seu estado de saúde mental e a facilidade que estes têm em falar com os seus colegas de trabalho/supervisor sobre saúde mental.

Para resolver o problema acima, selecionamos os seguintes campos do *dataset*:

- Age
- Gender (Processar gender)
- Country
- Work_Interfere
- Wellness_Program
- Seek_Help
- Mental_Health_Consequence
- Coworkers
- Supervisor
- Mental_Health_Interview
- Mental_vs_Physical
- Obs_Consequence

Após selecionarmos os campos acima, ordenamos-os pelo país, com isto que-remos facilitar a visualização da situação em cada país.

De seguida, apresentamos o *Job* criado:

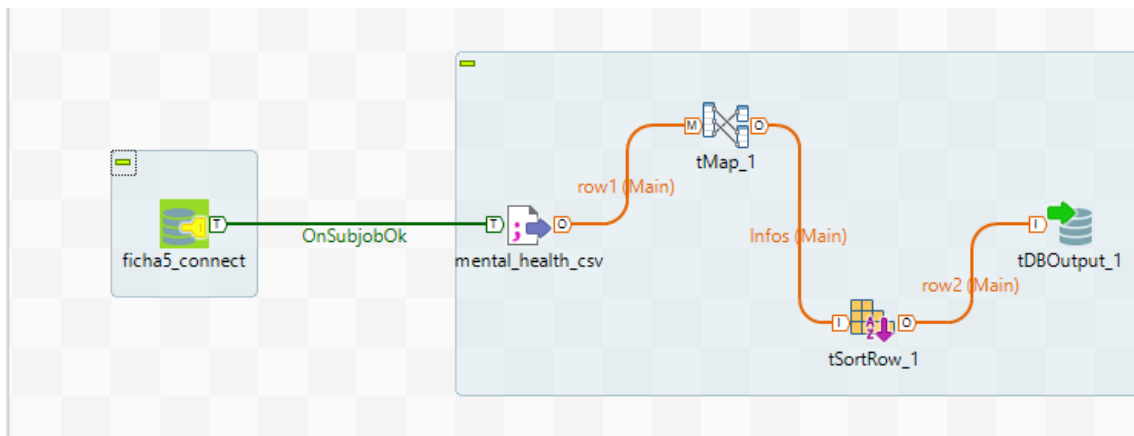


Figura 2: Job 2 Overview

2.3 Job 3

Com este *Job* pretendemos verificar a quantidade de pessoas que contêm empresa própria.

Para resolver este guardamos todos os dados dos funcionários.

No início começamos por ordenar os funcionários por idade, de seguida filtramos os funcionários que têm uma empresa própria. Após esta filtragem, guardamos os dados dos trabalhadores são guardados em dois ficheiros *Excel*. Um que irá armazenar os trabalhadores que têm empresa própria e outro que armazena os dados dos trabalhadores que não têm empresa própria.

Apresentamos agora o *Job* criado no *Talend*:

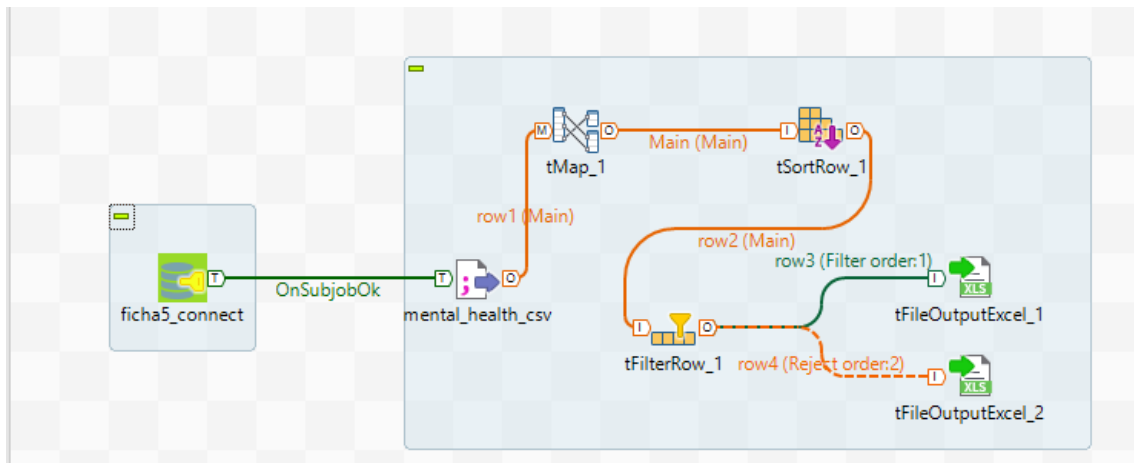


Figura 3: Job 3 Overview

3 Conclusão

Durante a resolução desta ficha fomos encontrando várias dificuldades. Essas dificuldades surgiram por ser a nossa primeira vez a utilizar o *Talend* sem suporte dos docentes. Contudo, com trabalho essas dificuldades desapareçam.

Apesar de tudo, conseguimos atingir o objetivo com sucesso, o que nos deixa bastante felizes.

A resolução desta ficha foi bastante importante do nosso ponto de vista, pois aumentou o nosso conhecimento sobre o *Talend* entre outras ferramentas, como *MySQL*.