



Curso de Engenharia de Sistemas e Informática - 5º Ano
Análise Inteligente de Dados
Ficha de Trabalho N.º 9

Objectivo: Familiarização com o ambiente do Clementine e utilização de fontes de dados.

I - Introdução ao Clementine a Familiarização com o Ambiente

1. Inicialize o Clementine e familiarize-se com o respectivo ambiente:

- menus e facilidades de ajuda em linha
- procure ajuda para qualquer tópico que lhe seja familiar, relativamente a Data Mining
- localize o tópico de ajuda relativo ao **Nó Select** e familiarize-se com a operação

2. Pratique a colocação de Nós no Painel Stream

3. Seleccione o **Nó Var.File** da Paleta Sources (Fontes) e coloque-o no painel Stream.

4. Seleccione o **Nó Table** da paleta de **Output** (Saída) e coloque-a a seguir ao **Nó Var.File**.

5. Ligue esses dois nós, utilizando o botão do meio (ou se o rato não o tiver, utilize os dois botões simultaneamente para emular o 3.º botão).

6. Edite o **Nó Var.File** e visualize as opções respectivas.

7. Desligue os dois nós.

8. Apague um dos nós do Stream.

9. Saia do Clementine sem salvar o Stream.

II - Leitura de Ficheiros de Dados

Vamos utilizar dois ficheiros

charity.csv (campos delimitados por virgulas)

charity.sav (ficheiro SPSS)

cujo download para sua área de trabalho pode ser efectuado por si dos ficheiros em anexo a esta ficha.

O ficheiro charity contém detalhes acerca da resposta dos indivíduos à campanha, o seu comportamento de donativos p/ caridade e dados base do tipo demográfico (idade, sexo e grupo).

1. Inicie o Clementine, ou se já estiver inicializado, assegure-se que o painel de Stream está limpo.

2. Seleccione o **Nó Var-File** do painel Fontes e coloque-o no painel Stream.

3. Edite o nó e indique que o ficheiro a ler é o charity.csv. O ficheiro contém os nomes dos campos na 1.ª linha. Dessa forma coloque o check na opção que indica ao Clementine para ler os campos da 1.ª linha.

4. Retorne ao painel Stream fazendo **OK**.

5. Coloque um **Nó SPSS Import** no painel Stream.

6. Edite esse nó e indique que o ficheiro a ler é charity.sav. O fich. SPSS contém valores e labels das variáveis. Verifique as opções para utilizar ambos .

7. Retorne ao painel Stream fazendo **OK**.

8. Para assegurar que ambos os nós fonte estão a funcionar correctamente, ligar-lhe um **Nó Table** a cada um.

9. Execute ambos os streams individualmente, utilizando a opção **Execute** do menu contexto (menu pop-up que surge com um clique no botão direito do rato), ou simultaneamente através do click no botão EXECUTE colocado por baixo da paleta Generated Models.
10. Inspeccione as tabelas e familiarize-se com os campos respectivos e analise alguns dos registos. Quando tiver acabado a operação, feche as tabelas (através do botão de fechar a janela ou através do Table... Dismiss).
11. Como vamos utilizar um desses streams nos exercícios seguintes, escolha qual dos Nós Fonte prefere e apague o Stream que contém o outro.
12. Guarde o stream restante numa pasta onde vai guardar o seu trabalho de AID com o nome ExerTp9.