

Lab_07 - AMD_N_04_Lab_07

1. - Case study: “the e-commerce”

2. - Create a database and populate it with the e-commerce dataset

Ficheiros completos:

- 00_script_CREATE_DB.sql
- 01_script_CREATE_SCHEMA.sql
- 02_script_IMPORT_POPULATE_SCHEMA.sql

Os dados foram importados corretamente.

3. - Build views to assist dataset analyzes (for the “market basket”)

Ficheros completos:

- 03_script_CREATE_VIEW.sql

O ficheiro "z_VIEW_expected.txt" tem os mesmos resultados que "z_VIEW_result.txt".

4. - Export a basket for the extraction of “association rules”

Fichero completo:

- 04_script_EXPORT_DATA.sql

O ficheiro "z_dataset_sample_OUT_expected.txt" tem os mesmos resultados que "z_dataset_sample_OUT_result.txt".

5. - ... info about the “.basket” format and “string normalization”

Ao aplicar os passos de "string normalization" permite reduzir a variedade de formatos e potenciais erros no conjunto de dados, tornando-o mais uniforme e fácil de analisar, especialmente para um conjunto de dados de maior dimensão.

6. - Generate the “.basket” format for sparse matrix representation

Fichero completo:

- `_goPy_transform_v02.py`

```
#for transactionID in basket.keys():
for transactionID in basket:
    items = basket[transactionID]
    line = ", ".join(items.keys())
    f.write( line + '\n' )
...

# set the number of items at each transaction
for itemID, count in basket[transactionID].items():
    itemID_index = all_itemset[ itemID ]
    line_list[itemID_index] = count
```

Os ficheiros "zz_dataset_2012_01.basket" e "zz_dataset_2012_01.tab" tem os mesmos resultados que "_expected".

7. - Build your “market basket analysis” (e.g., with Orange workflows)

Ficheiro Orange:

- `AssociationRules.ows`