preprocess.sh

Preprocesses each file / does file conversion / normalizes headers columns / renames files to standard ones for further processing

Arguments

Inventory_file, stores_data, ecom_sales, ecom_returns, franch_sales,franch_returns

Sample usage

$../preprocess.sh inv_may.csv stores_may.csv Ecom_Sale_May2016.csv Ecom_Return_May2016.csv Franch_May_Sales.csv Franch_May_Return_Empty.csv

Details

Renames the arguments files to INV.csv, STORE_SALE.csv, ECOM_SALE.csv,ECOM_RETURN.csv, FRAN_SALE.csv, FRAN_RETURN.csv

sales_sum.sh

Computes the sales measures. No arguments

inv_sum.sh

Computes the inventory measures. No arguments.

merge.sh

Creates measures.csv for loading to **covers_new_view**. Takes date as argument

Sample Usage

$../merge.sh 30.04.2016

Article_mc.sh

Computes mapping from article-id to MerchaniseCategory. This is then merged with HIERARCHY.csv to create article_details.csv

article_measures.sh

Computes aggregate measures for each article id in the file article_measures.csv. This is loaded to the table **article_measures**

top_measures.sh

Computes the top 50 articles (dept wise, store wise) and top 10 article in each combination into files topm_dept.csv, topm_store.csv, topm_dept_store.csv. These are loaded respectively to tables top_dept, top_store,top_dept_store

HIERARCHY.csv

Contains the hierarchy mappings

Covers_new_view is created using the following command

```
USE `pentaho`;
CREATE
OR REPLACE ALGORITHM = UNDEFINED
DEFINER = `babyoye`@`%`
SQL SECURITY DEFINER
VIEW `covers_view_test` AS
SELECT
`measure_details1`.`mdate` AS `mdate`,
`measure_details1`.`Site` AS `site`,
`measure_details1`.`SLOC` AS `SLOC`,
`measure_details1`.`SOR` AS `SOR`,
`store_master`.`SOrg` AS `Organization`,
`store_master`.`SiteCategory` AS `siteCategory`,
`store_master`.`Name1` AS `StoreName`,
`store_master`.`City` AS `City`,
`article_details`.`Brand_Name_Code` AS `Brand_Name_Code`,
`article_details`.`MC_Desc_Code` AS `MC_Desc_Code`,
`article_details`.`Dept_Desc_Code` AS `Dept_Desc_Code`,
`article_details`.`SubDept_Desc_Code` AS `SubDept_Desc_Code`,
`article_details`.`Class_Desc_Code` AS `Class_Desc_Code`,
`article_details`.`SubClass_Desc_Code` AS `SubClass_Desc_Code`,
`measure_details1`.`Article` AS `Article`,
`article_details`.`Article_Desc` AS `Article_Desc`,
`measure_details1`.`Quantity` AS `Quantity`,
`measure_details1`.`NetSalesVal` AS `NetSalesVal`,
`measure_details1`.`COGS` AS `COGS`,
`measure_details1`.`Margin` AS `Margin`,
`measure_details1`.`TotalMAP` AS `TotalMAP`,
`measure_details1`.`TotalSOH` AS `TotalSOH`
FROM
((`measure_details1`
LEFT JOIN `article_details` ON ((`article_details`.`Article` = `measure_details1`.`Article`)))
LEFT JOIN `store_master` ON ((`measure_details1`.`Site` = `store_master`.`Site`)));
```

LOADING TABLES

1. measure_details1 - contains the sales/inventory measures along all dimensions like Article, Dept, SubDept, Class,SubClass, MerCategory,Store, SOR, SLoc etc

2. article_details - contains article hierarchy for every article-id

3. article_measures - contains sales/inventory measures along Article, Dept, SubDept, Class,SubClass, MerCategory. This doesn't contain Store related dimensions

4. date_details - dimension table for date, contains the dates at which measures are calculated

5. store_master - dimension table for store, contains store related information such as sitecategory, location etc

6. top_dept - contain top 50 articles (according to NSV) for each dept

7. top_store- contain top 50 articles (according to NSV) for each dept

8. top_dept_store - contains top 10 articles (according to NSV) for each dept,store combination
9. Covers_new_view

```
LOAD DATA LOCAL INFILE '/home/babyoye/union_measures.csv' INTO TABLE pentaho.measure_details1 FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA LOCAL INFILE '/home/babyoye/article_details.csv' INTO TABLE pentaho.article_details FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;


LOAD DATA LOCAL INFILE '/home/babyoye/article_measures.csv' INTO TABLE pentaho.article_measures FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA LOCAL INFILE '/home/babyoye/mdate.csv' INTO TABLE pentaho.date_details FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;
```

```
mysql> LOAD DATA LOCAL INFILE '/home/babyoye/store_data.conv.csv' INTO TABLE pentaho.store_master FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;
```

```
LOAD DATA LOCAL INFILE '/home/babyoye/union_topm_dept.csv' INTO TABLE pentaho.top_dept FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA LOCAL INFILE '/home/babyoye/union_topm_store.csv' INTO TABLE pentaho.top_store FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA LOCAL INFILE '/home/babyoye/union_topm_dept_store.csv' INTO TABLE pentaho.top_dept_store FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 LINES;
```