Problem Statement :

Find the top selling item department category in each store on every day and load this information to mysql table

Tables :

**Dim\_item** (already in hive from previous exercise)

**Dim\_store** (already in hive from previous exercise)

**Fact\_sales** (Table is in mysql DB. Need to load it to a hive table)

Mysql table name : labuser\_database.fact\_sales

**Sales\_amount** (Data is present in a file and need to load it to a hive table.

File location : /mnt/bigdatapgp/edureka\_719925/data\_prep/sales\_amount/sales\_amount.txt

**Import column description in the tables :**

Fact\_Sales:

|  |  |  |
| --- | --- | --- |
| column name | data type | desc |
| item\_id | int | unique item number |
| store\_nbr | int | unique store number |
| sales | array | list of all the sales ids of a given item in a store |

Sales\_amount

|  |  |  |
| --- | --- | --- |
| column name | data type | desc |
| sales\_id | int | unique sales\_id |
| sales\_amount | float | amount of the sale |
| transaction\_date | string | sales transaction timestamp |

Dim\_item

|  |  |  |
| --- | --- | --- |
| column name | data type | desc |
| mds\_fam\_id | int | unique item number |
| dept\_catg\_desc | string | Item department category name |

Dim\_store

|  |  |  |
| --- | --- | --- |
| column name | data type | desc |
| store\_nbr | int | unique store number |
| store\_nm | string | store name |