"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."

Final Problem: Predict Credit Card Spending

In the era where people create quintillion bytes of data on daily basis, companies compete to gather insights on their customers. One of the primary goals the business sector strives to achieve is personalization on individual level. In order to achieve such precise level, each customer's lifestyle has to be well understood. Data scientists need to know what is needed, what is not, and how to respond to the gained insights. Great challenges await all challengers of Data Track in TechJam 2017: Final Round to fully utilize huge amount of data in limited amount of time.

You, as a data scientist, are tasked to utilize credit card information and transactions to determine buying behavior of all the customers. The given data dates from January 2016 to June 2017. The goal is to predict amount of spending and buying frequency of five designated categories in July 2017.

Provided Files: techjam_final_data.zip

- File name: cc_log.csv
  - Dummy credit card transaction
  - Period: 01/01/2016 - 31/07/2017
  - Size: 6,011,450 Rows

Field Name	Data Type	Description
card_no	STRING	Dummy Credit Card Number (Link to cc_info)
txn_dt	DATETIME	Transaction Date
txn_tm	TIMESTAMP	Transaction Time
bill_amt	DECIMAL	Transaction amount (THB)
card_acpt_cty	STRING	Country Accept Credit Card
mrch_tp_cd	INTEGER	Merchant category code (link to Final_categories.csv)

"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."

card_type	STRING	Visa or Master

- File name: cc_info.csv
  - Dummy credit card information
  - Size: 99,954 rows

Field Name	Data Type	Description
card_no	STRING	Dummy Credit Card Number (Link to cc_log)
card_type	STRING	Visa or Master
opn_dt	TIMESTAMP	Card issue date.
exp_dt	STRING	Card expiry date.
cr_lmt_amt	DECIMAL	Credit limit amount.
prev_cr_lmt_amt	DECIMAL	Previous credit limit.
main_zip_cd	STRING	Dummy zip code of home address.
cr_line_amt	DECIMAL	Total available credit amount.
incm_amt	INTEGER	income amount per month.
brth_estb_yr	INTEGER	Birthday Year (A.D.)
gnd_ind	STRING	Gender (0 : Female, 1: Male)

- File name: Final_categories.csv
  - Categories of Merchant category code
  - Size: 1,186 Rows

Field Name	Data Type	Description

"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."

Categories	STRING	Categories of Merchant category code
MCC	INTEGER	Merchant category code (link to cc_log.csv)

Expected Output

- File name: Team_xx.csv
  - Description: Predicted frequency and total spending of each category (Automobiles and Vehicles, Clothing Stores, Service Providers, Transportation, and Utilities) in July, 2017
  - Each line contains frequency and total spending for each category
  - The order of card numbers must follow the card numbers in cc_info.csv
  - Each line must have predicted values corresponded to the ordered columns as shown on the following table

Field Name	Data Type	Description
Automobiles and Vehicles - Frequency	INTEGER	Frequency of Transaction in Automobiles and Vehicles Category
Clothing Stores - Frequency	INTEGER	Frequency of Transaction in Clothing Stores Category
Service Providers - Frequency	INTEGER	Frequency of Transaction in Service Providers Category
Transportation - Frequency	INTEGER	Frequency of Transaction in Transportation Category
Utilities - Frequency	INTEGER	Frequency of Transaction in Utilities Category
Automobiles and Vehicles – Total Spending	INTEGER	Total Spending of Transaction in Automobiles and Vehicles Category

"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."

Clothing Stores – Total Spending	INTEGER	Total Spending of Transaction in Clothing Stores Category
Service Providers – Total Spending	INTEGER	Total Spending of Transaction in Service Providers Category
Transportation – Total Spending	INTEGER	Total Spending of Transaction in Transportation Category
Utilities – Total Spending	INTEGER	Total Spending of Transaction in Utilities Category

Examples

File: cc_log.csv

```
4410123456098153,2017-03-21 00:00:00,10:22:19,5660,TH,4722,visa

4410123456018545,2017-03-21 00:00:00,23:30:43,100,US,5818,visa
```

File: cc_info.csv

```
4410123456000001,visa,1997-10-14 00:00:00,1017,50000,0,73120,50000,21000,1953,0

4410123456000002,master,2010-06-29 00:00:00,0620,146000,0,43000,146000,72000,1965,0
```

File: Final_categories.csv

```
Clothing Stores,5611

Automobiles and Vehicles,5511
```

Output File: Team_01.csv

```
23,4,2,0,0,870,650,5000,0,0

0,0,2,0,0,0,0,10000,0,0
```

"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."