Transformations and logic explanation:

1. transformation on transactions.csv

1. Removed negative sign and moved to front

1. Create new column transaction\_amt, copy data of original column transaction\_amount data to transaction\_amt column

2. select column transaction\_amt

3. go in data section

4. select text to columns option

5. select next

6. Click next

7. Click advance

8. check trailing minus for negative number and click ok and then click finish.

<https://www.extendoffice.com/documents/excel/3521-excel-move-minus-sign-from-back-to-front.html>

2. To remove plus sign, apply formula =SUBSTITUTE(C2,"+",""), C2 is the original transaction\_amount column in csv file. C2 is my original transaction\_amount column, I am creating a new column transaction\_amt

2. Executed sql\_script.sql.

3. I used mysql workbench for importing and exporting the data. In step 6 I have mentioned links which I referred for the same.

After importing transactions and account\_info data:

Below is the data type I used for the columns in transactions table

Field Data type

account\_number int

transaction\_datetime text

transaction\_amount text

post\_date text

merchant\_number bigint

merchant\_description text

merchant\_category\_code int

transaction\_number int

transaction\_amt double

st text

Below is the data type I used for the columns in account\_info table

Field Data Type

last\_name text

first\_name text

street\_address text

unit text

city text

state text

zip int

dob text

ssn text

email\_address text

mobile\_number bigint

account\_number int

4. I executed the c++ program (main.cpp) for performing the Rule 1 and Rule 2 mentioned in part 2 of the test.

5. Logic for performing rule 1:

1. I made use of stddev() method which is related to standard deviation. I calculated standard deviation for each account and merchant number and checked if the difference between standard deviation value and maximum transaction amount value for respective account and merchant is greater than 500, then I am displaying such records as abnormally high transactions.

2. Standard deviation helps in understanding the dispersion of data relative to the mean.

6. For importing and exporting data I referred following website:

<https://dev.mysql.com/doc/workbench/en/wb-admin-export-import-table.html> <https://dev.mysql.com/doc/workbench/en/wb-admin-export-import-management.html>

7. I have exported MySQL database schema and data and stored in file dump\_file.sql

8. I had to include libmysql.dll in folder \mysqlproject\x64\Release in order to execute the program.

9. Below image shows the output of the program:

