**Sanchit Rokade**

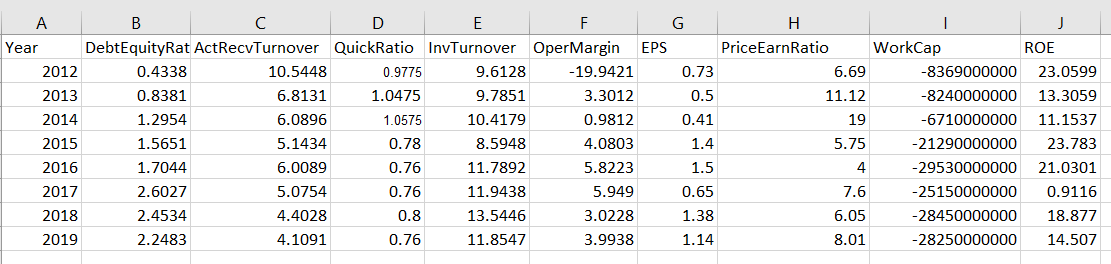
**UIN:677781611**

**UIC email: srokad3@uic.edu**

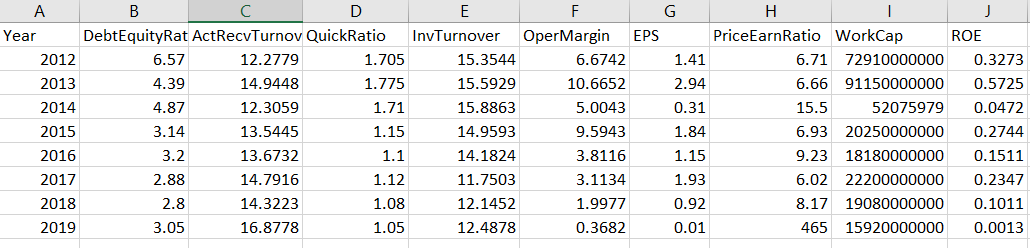
**Document for IDS 523 Midterm Exam:**

**1) (10 points) Financial Ratios:**

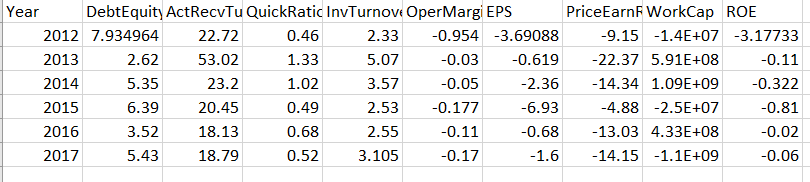
These are financial ratios for General Motors:



These are financial ratios for Ford Motors:



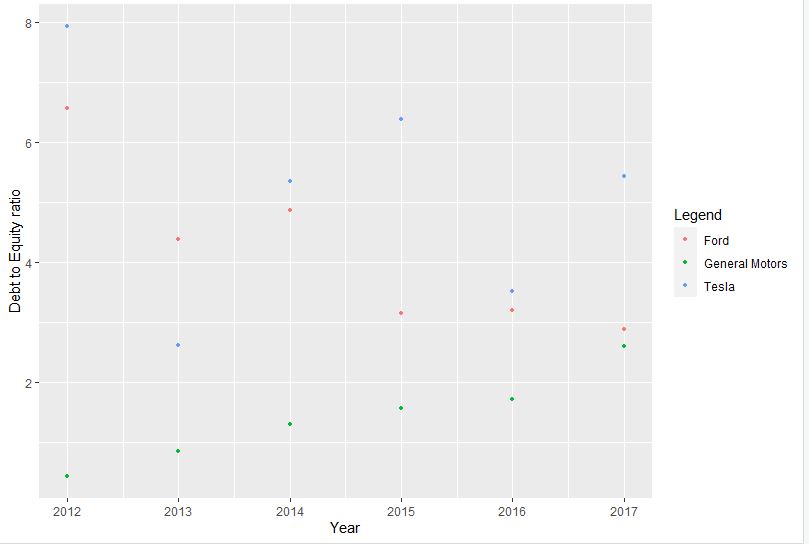
These are financial ratios for Tesla Motors:



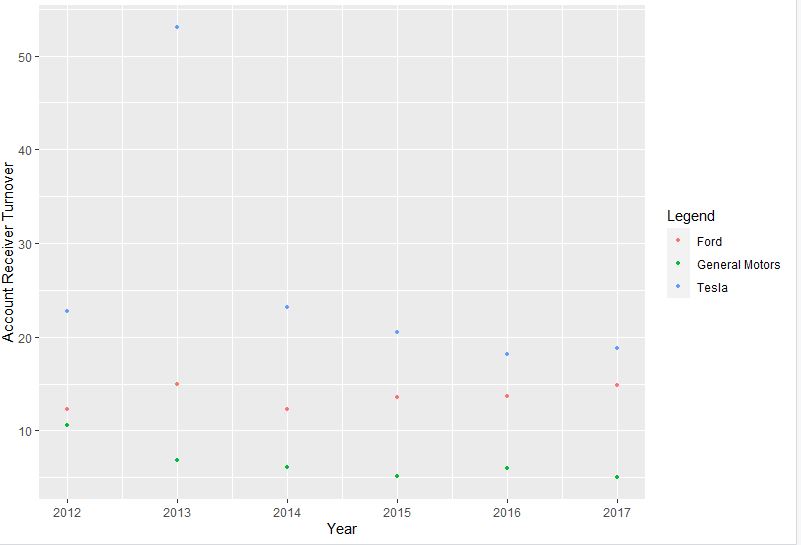
**2) (5 points) Visualization of Technical Metrics:**

I am using ggplot for visualization

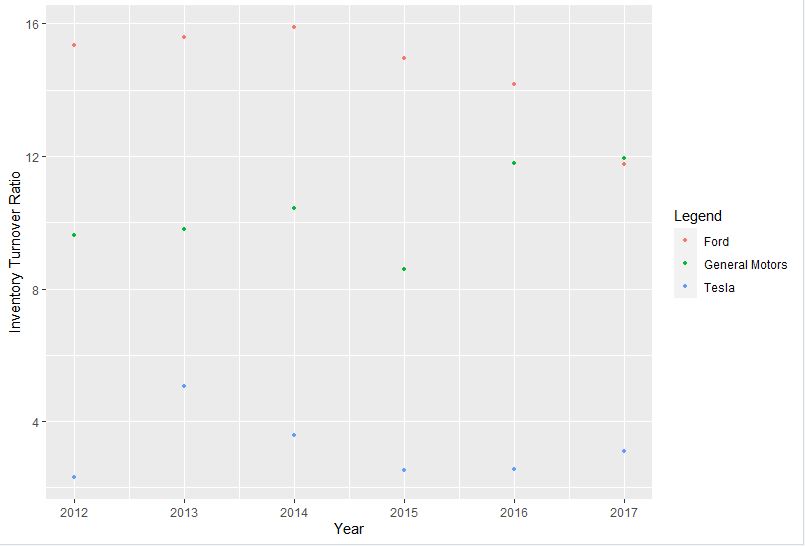
Plot for debt to equity ratios for the years 2012-2017



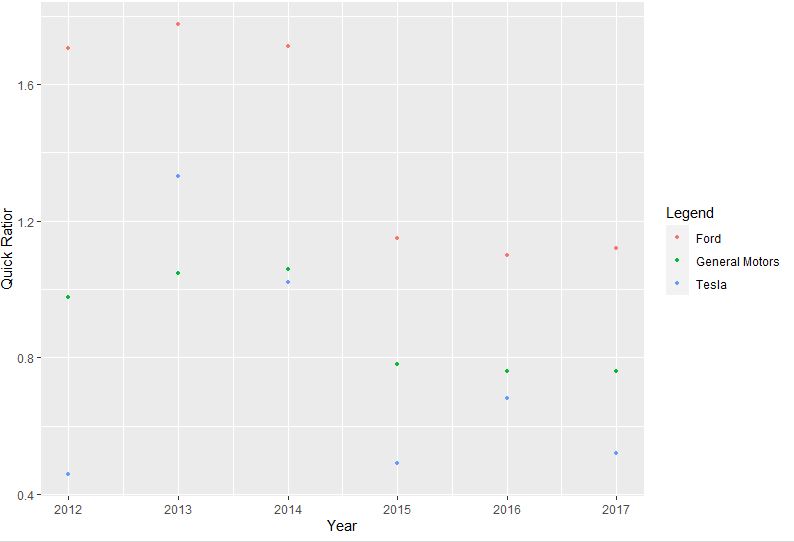
Plot for account receiver turnover ratios for the years 2012-2017



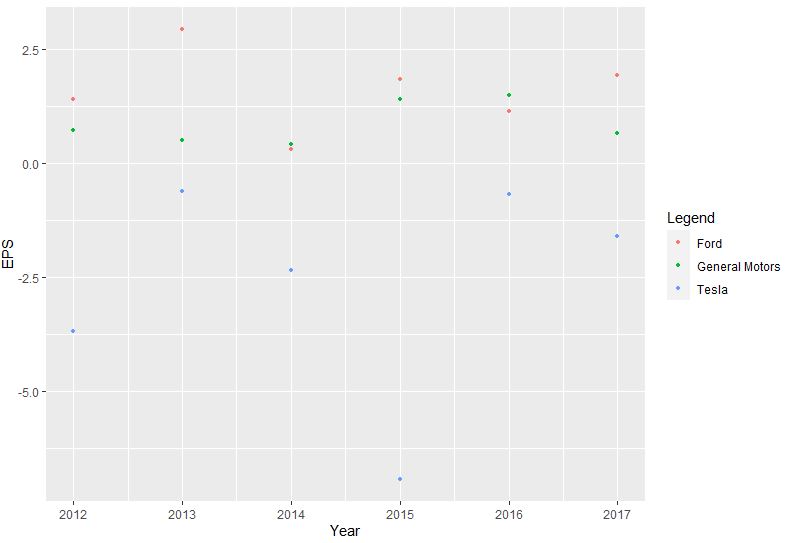
Plot for inventory turnover ratios for the years 2012-2017



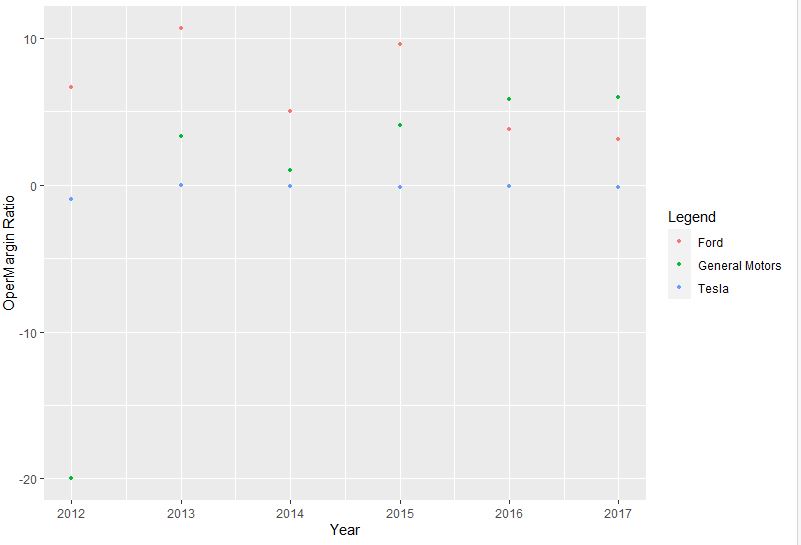
Plot for quick ratios for the years 2012-2017



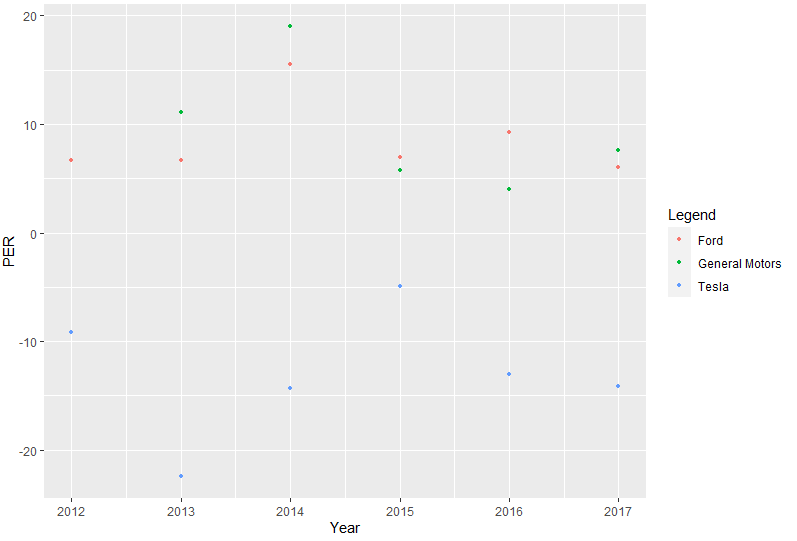
Plot for EPS ratios for the years 2012-2017



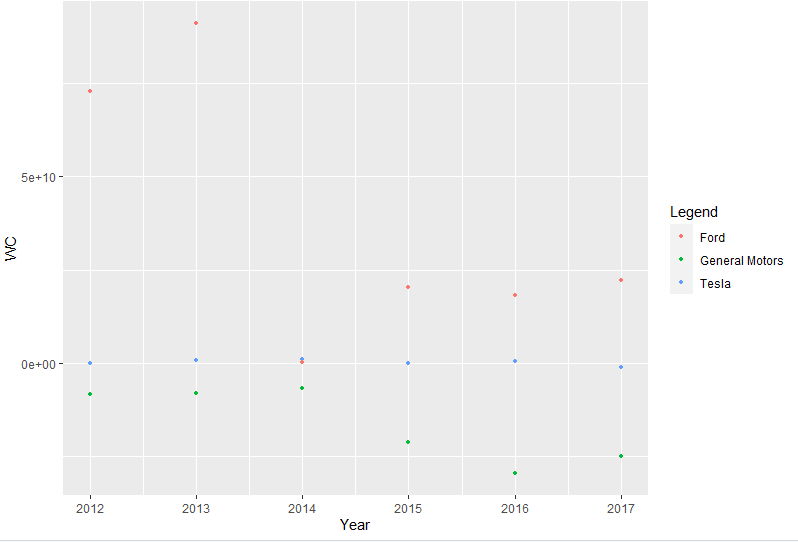
Plot for Operating Margin ratios for the years 2012-2017



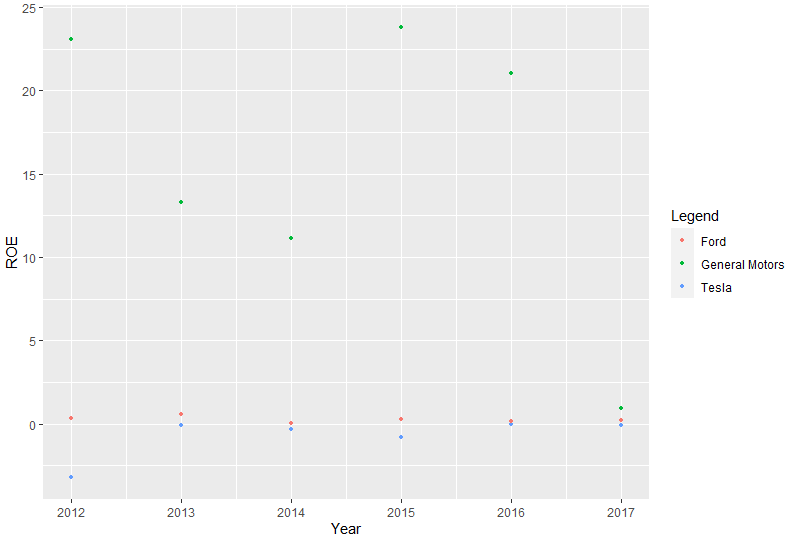
Plot for Price Earnings ratios for the years 2012-2017



Plot for Working Capital ratios for the years 2012-2017

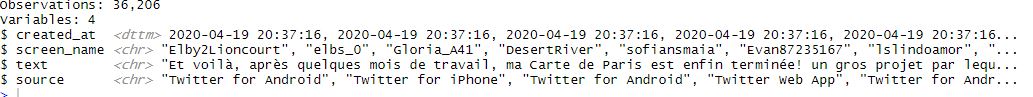


Plot for ROE ratios for the years 2012-2017

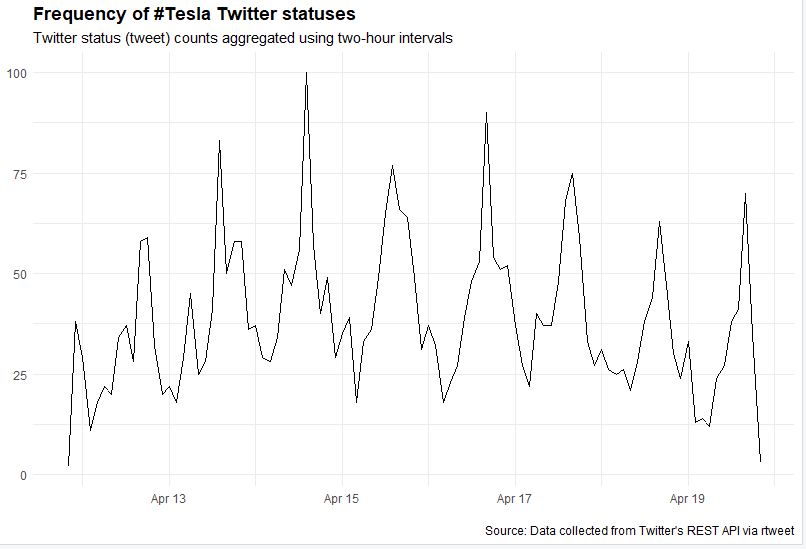


**3) (10 points) Intelligence Scanning with Twitter:**

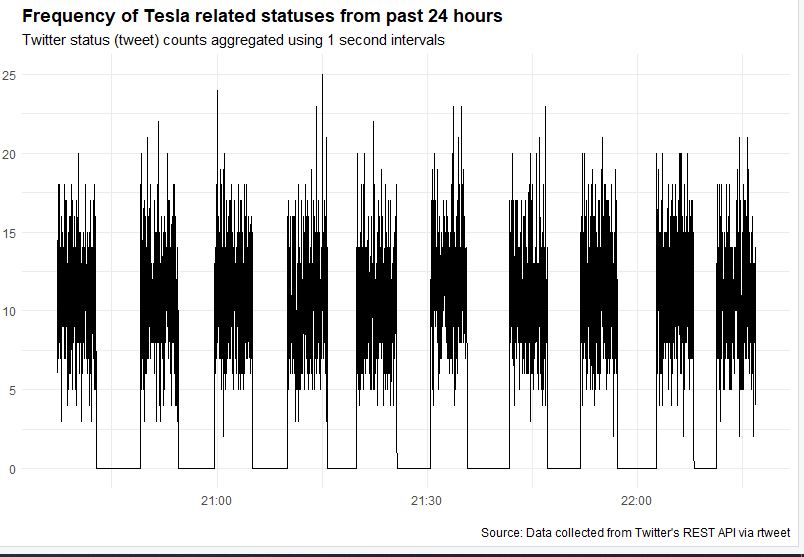
Glimpse of twitter data which includes source and text information along with time and name



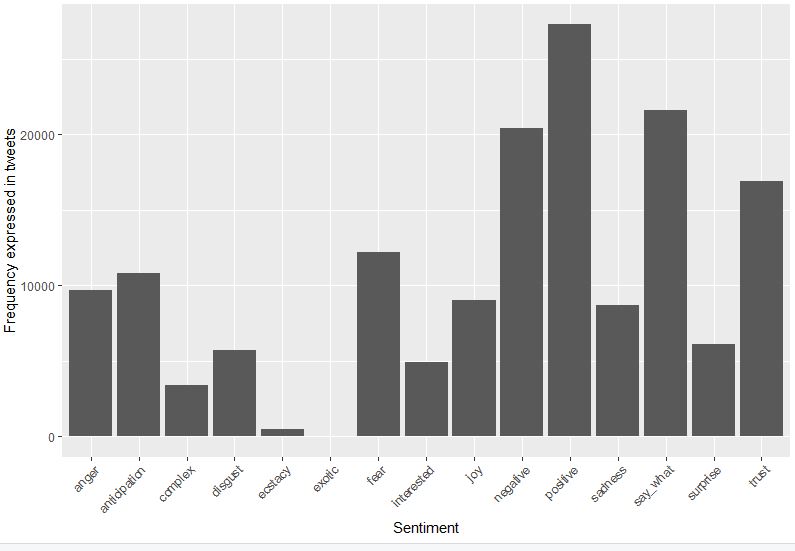
Frequency of tweets are high on 15 April and during weekends



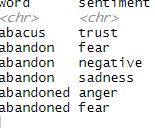
Frequency of tesla related statuses from past 24 hours using 1 second interval where the peak is between 21:00 and 22:00



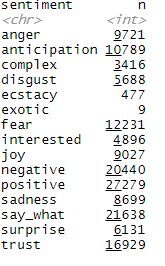
Frequency of words expressed in tweets where mostly it is positive



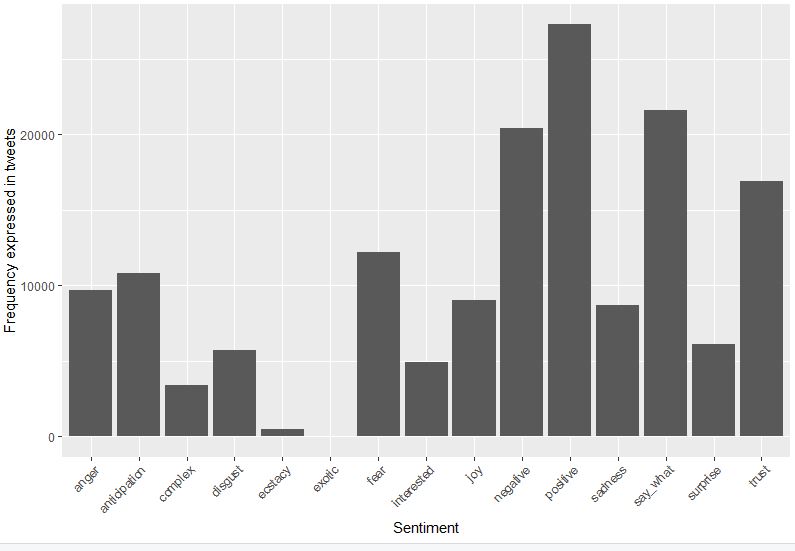
Snapshot of sentiment lexicon



Count for words on sentiment analysis



Adhoc lexicon have grouped the words related to competitors of Tesla into "complex", interesting words related to electric motor vehicle into "interested". From histogram, most popular emotions would be positive vs negative



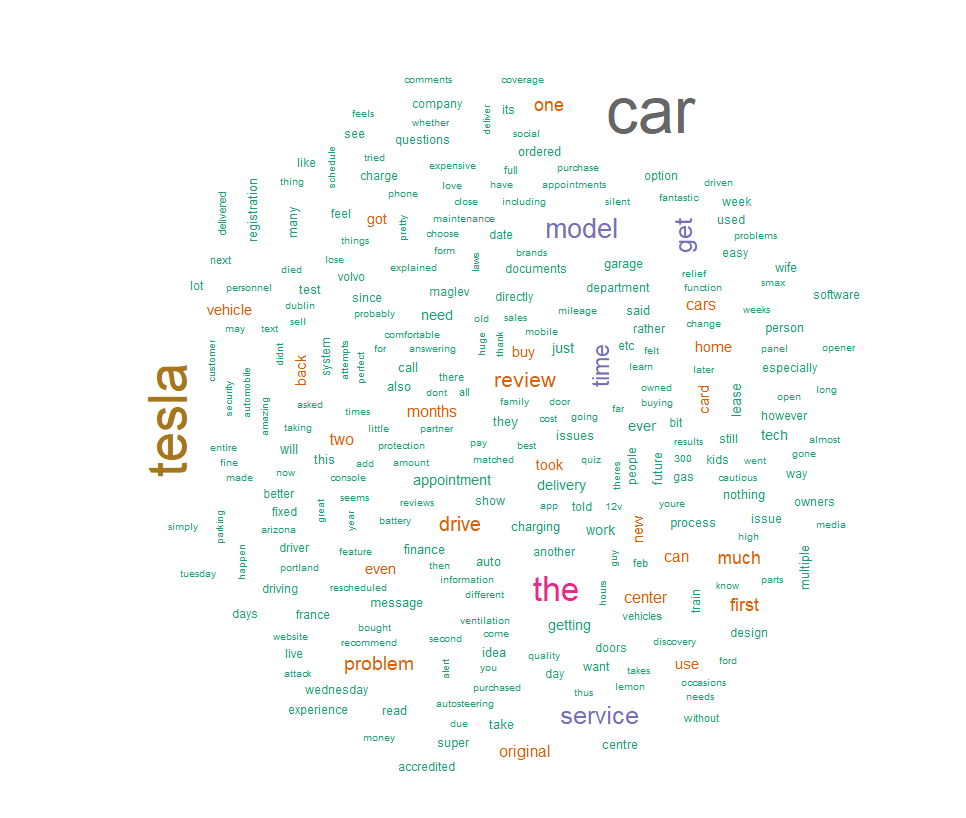
**4) (10 points) Extract Intelligence from Product User Forums and Visualize using Wordcloud**

Used the rvest package to inject structure into web scraping for intelligence scanning from User Forums.

**Wordcloud Plot:**

These are the reviews received for Tesla motors from consumers.

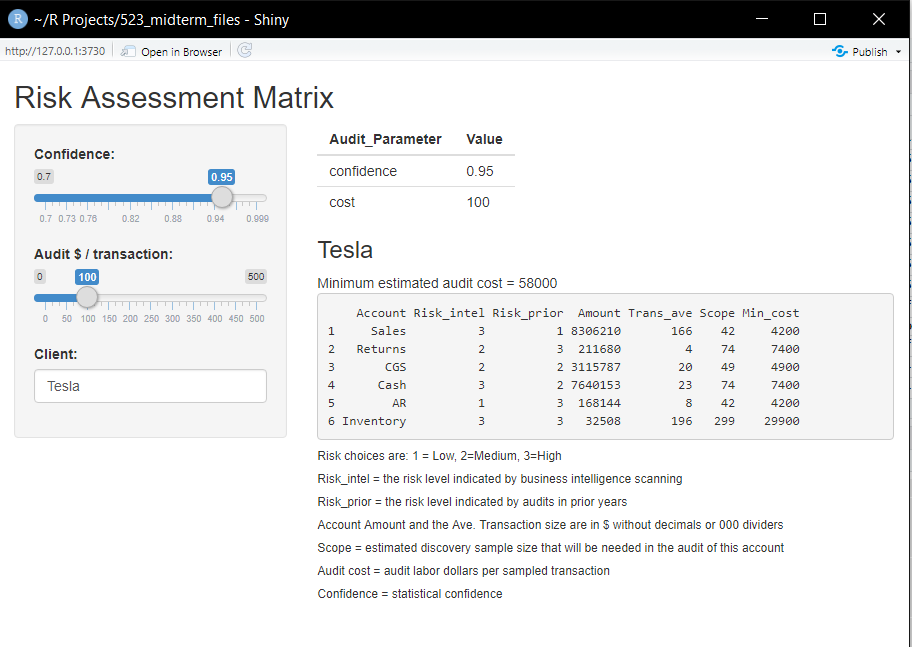
The wordcloud highlights the most popular words being used in the user forums. Here, users are talking mostly about the Tesla models, their performances and vehicle parameters which are mostly positive. There are also some mentions of other related attributes like "console", "garage", "tech" and has a good experience with respect to Tesla motors.



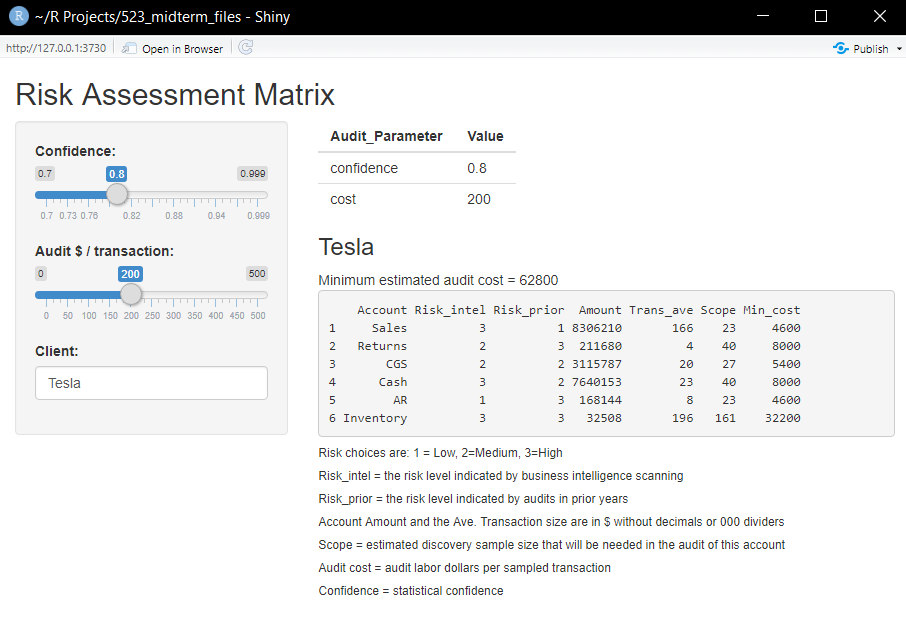
**5) (10 points) Planning and Risk Assessment:**

Created risk assessment matrix for tesla by calculating amounts for various account and updated the values given by professor and the csv file is included in zip folder.

**6) (5 points) Generate the Audit Budget from the Risk Assessment Matrix app**



As seen above for Confidence = 0.95 and cost = $100 the minimun estimated audit cost is $58000.



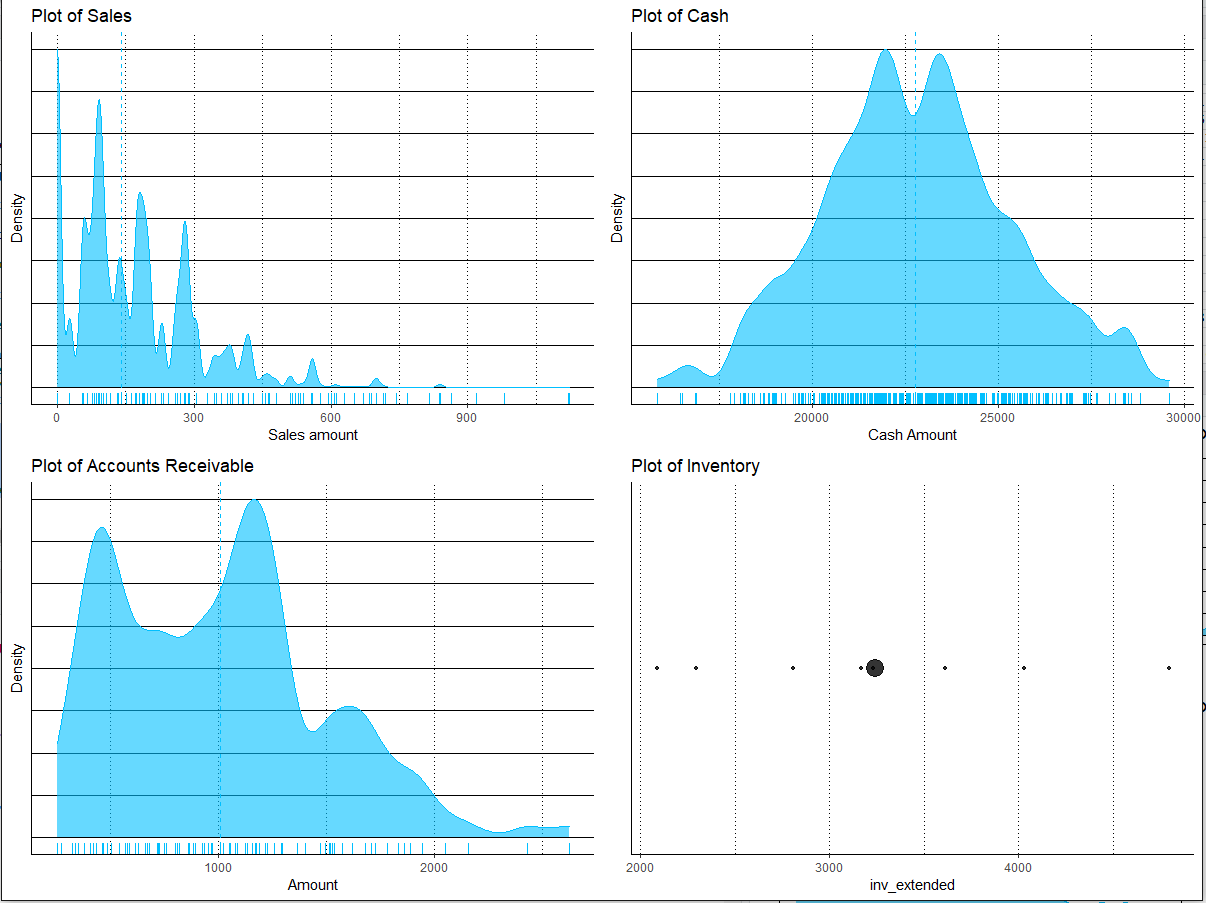
As seen above for Confidence = 0.80 and cost = $200 the minimun estimated audit cost is $62800.

**7) (5 points) Preliminary Tests of Transactions Controls**

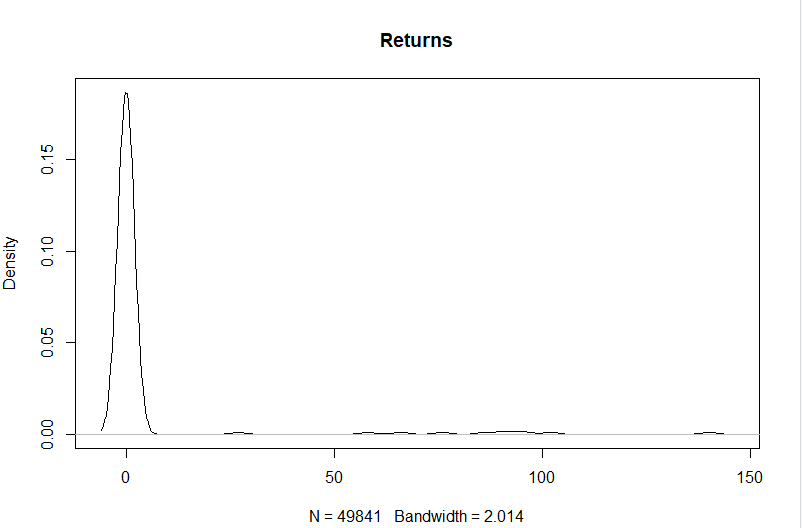
Used ggplot for visualizing plot and ggarrange function under pubrr package for plotting multiple ggplots.

Our general assumption is that auditing has normal distribution for the different transactions throughout the company's activities.

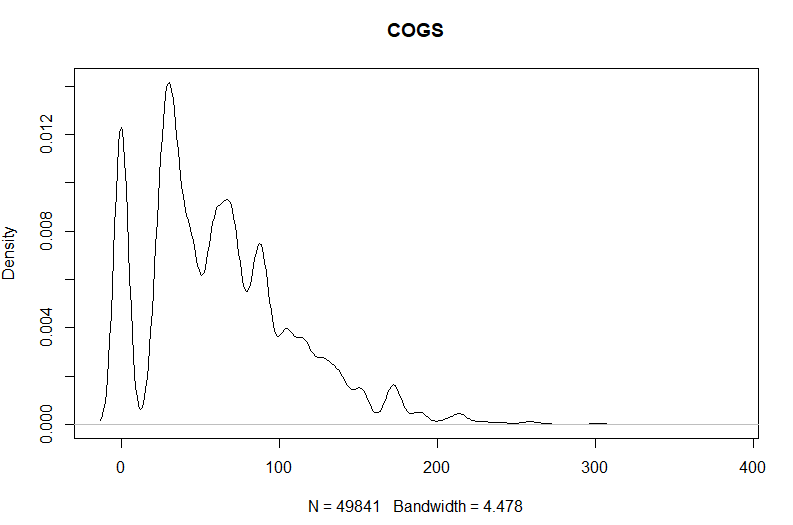
Plotting the transaction distributions:



Here we can see that sales are not normally distributed and is highly right-skewed meaning mean is greater than median. Cash shows a normal distribution for our assumptions. Also, Accounts Receivable distribution has a multimodal right skewed distribution.



As we can see that the graph is skewed so my take is that we need to change risk factor with respect to Returns as an improvement.



As we can see that the graph is multi-modal right skewed, so my take is that we need to change risk factor with respect to COGS as an improvement.

Here are the final conclusions based on AICPA's aspects to sampling risk when performing tests of controls:

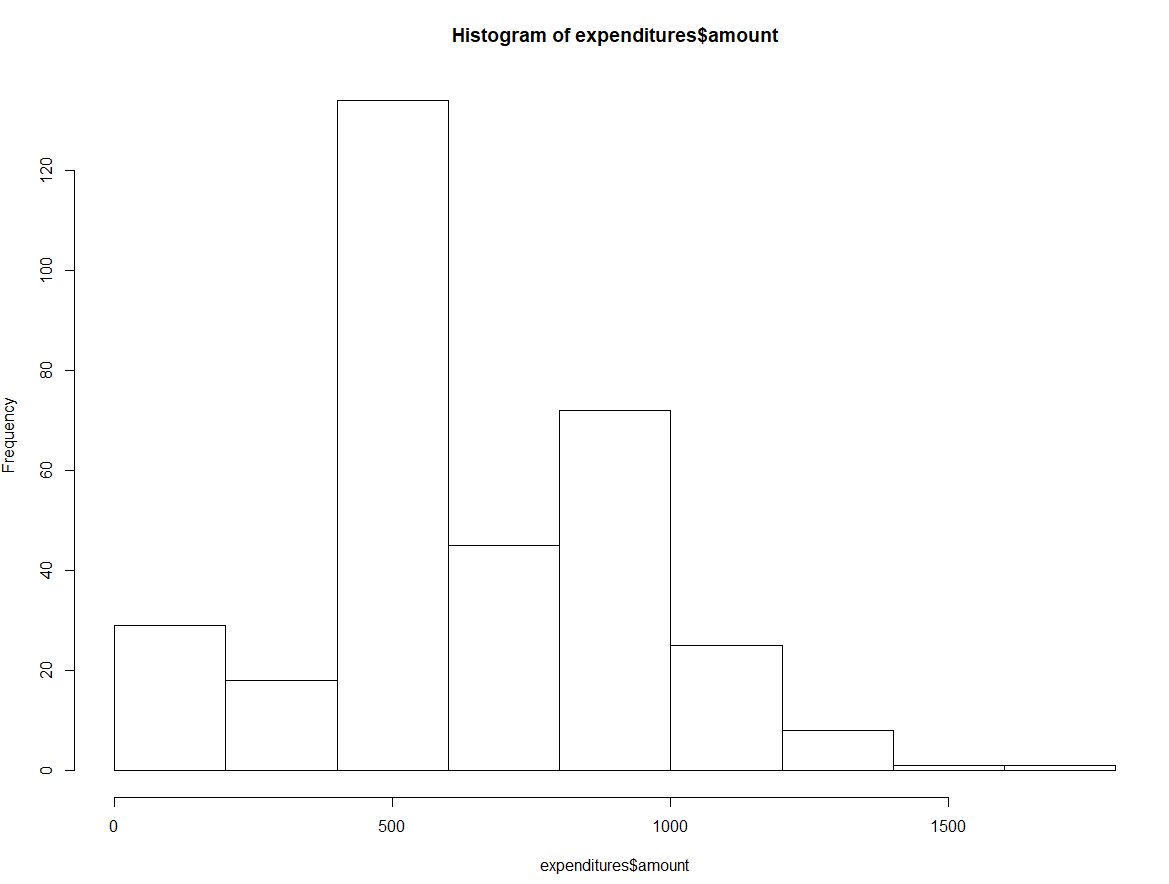
1. Assessing control risk is too low represents the risk that an audit sample supports the conclusion that the design and operation of an internal control is effective when in fact it is not.

2. Assessing control risk too high represents the risk that an audit sample supports the conclusion that the design and operation of an internal control is not effective when in fact it is effective.

Attribute estimation will help us decide on actual error rate of system that the process is transaction stream and attribute sampling used for correcting samples. If discovery sampling gives that particular transaction stream is out of control.

**8) (5 points) Employee Expenditures Audit**

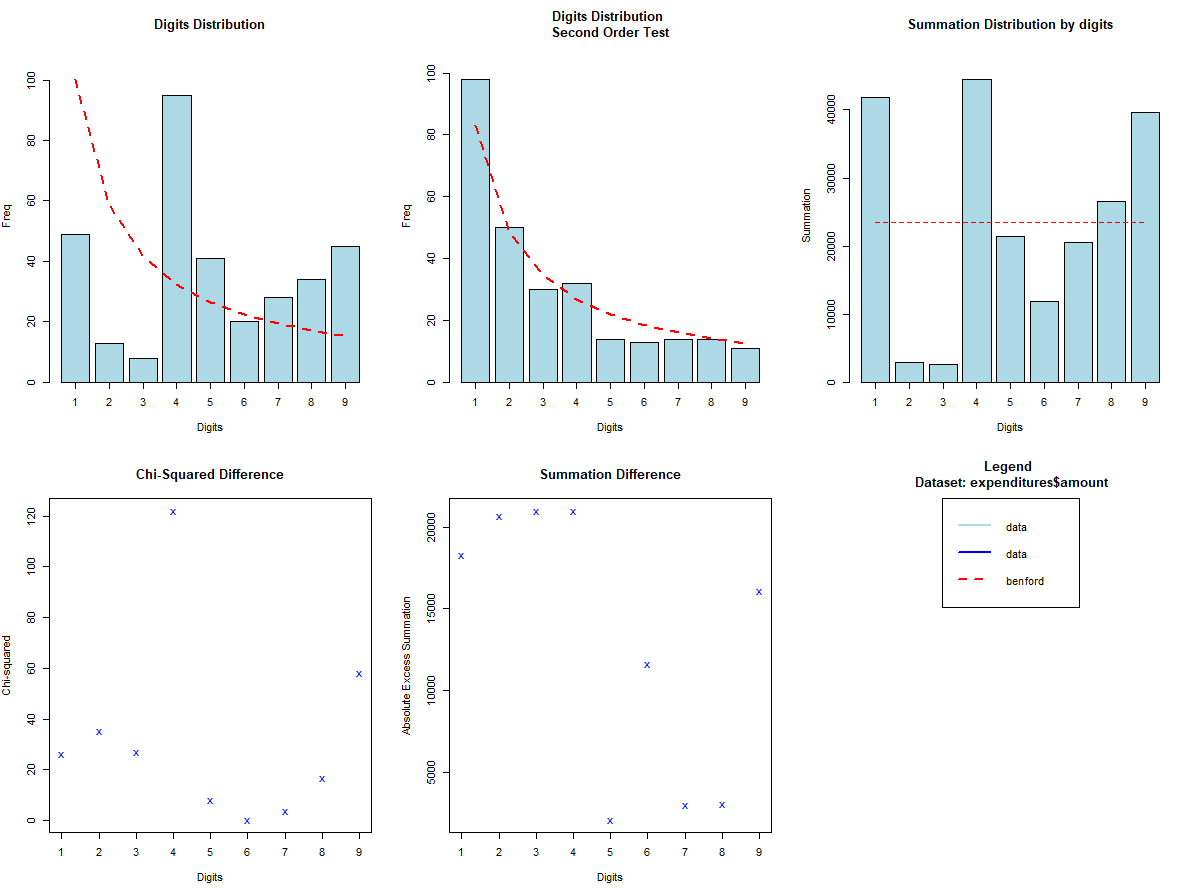
Reading expenditure data to get idea to do Benford test



Here we can see that histogram of expenditure is rightly skewed meaning mean is greater than median

Benford test is conducted for employee expenditure auditing

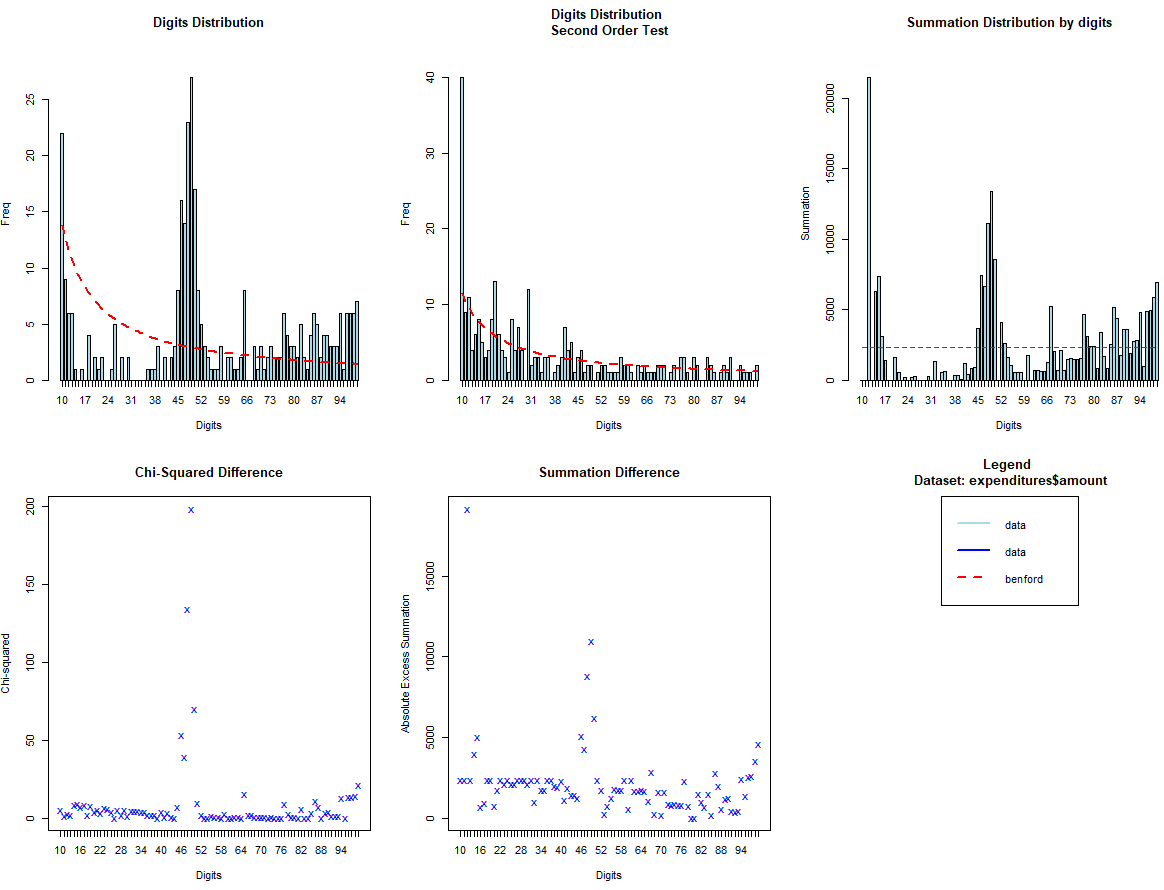
For Workers and Staff, the authorized limit is $50 and therefore has taken no of digits as 1 for first test.

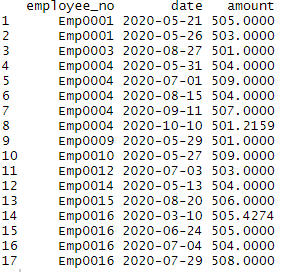


Result suggests that digits distribution second order test for 5 remains under controlled observations of Benford Test and therefore suggests that the number 50 occurred well within its limits.

For Line supervisors, the authorized limit is $500

Here, we have taken the number of digits as 2 being consistent to the Benford Testing logic.

Results suggests that the digits 50 were in the data more often than expected and hence, classified on the employees for all contracts with 50 as the first two digits to find several line supervisors who submitted an excessive number of expense reports for $500.

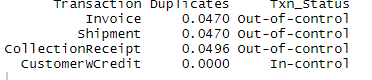


Displays the amount for each employee no with respective to date

9) **(10 points) Sales Duplicates and Omissions Error Rates**

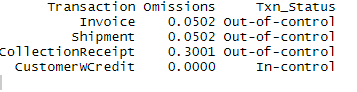
Summary for following transaction

If transaction duplicate percentage > 1% then it is out-of-control else, it is in-control.



Summary for Omissions of both transaction:

If transaction omission percentage >1% then it is out-of-control else, it is in-control.



**10) (10 points) Error Rates in Sales Amounts**



Performing tests with discovery sampling, we see that the error rates are 0%, which indicates tolerable error amounts in sales and to be sure of that have performed attribute sampling t-tests which is usually performed when the error rates are identified to be 'out-of-control'.



The discovery sample size was 59, an additional $224- 58 = 166 records are required for attribute tests of occurrence rate, and $636 - 58 = $578 records are required for attribute tests of amount rate.



The out-of-control critical value with a confidence of 95% is 0.05 error rate, thus both in occurrence and amount auditor determines sales system to be in-control.

**11.**

**A part**

**(5 points) Inventory Omissions and Duplicates**

Summary for Duplicates of both transaction:

If transaction duplicate percentage > 1% then it is out-of-control else, it is in-control.



Summary for Omissions of both transaction:

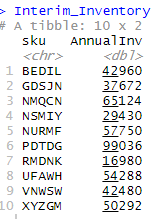
If transaction omission percentage >1% then it is out-of-control else, it is in-control.



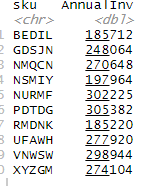
**B part**

**(10 points) Inventory Stock on Hand with a Turnover of Less than 5 times.**

Here is the annual inventory for sku under interim inventory for 2020:

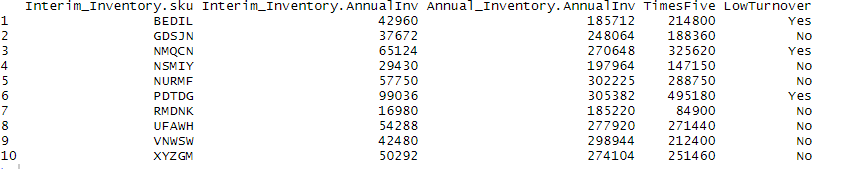


Here is the sum of inventory for sku under annual inventory for 2020:



Below is table where the condition is if TimesFive > Annual inventory then it is LowTurnOver or else it is not.

LowTurnOver is bad which means that it has excess inventory



**12) (5 points) Inventory Stock on Hand Lower of Cost or Market**

Here is the table for inventory count differences



Total Inventory Markdown is sum of Inventory count difference and inventory markdown



Here is the market and cost of inventory by sku mentioning total markdown for LOCOM rule



**13) (5 points) Write the SAS 115 Letter to Management**

I have qualified opinion that is the Tesla Motors data was fairly presented with minor errors in some parameters which are as follows:

* There were duplicates and omissions in inventory number, collections, shipping number for question 9 which makes data difficult to interpret and analyze
* There were duplicates and omissions in purchase orders for question 11 which makes data difficult to interpret and analyze
* Also, the risk associated with returns and cogs was incorrectly interpreted so we need to change risk factor for both
* The plot for cash, sales, accounts receivable for question 7 doesn’t give much indication for analysis as much of the data is skewed
* The error rate for Sales amount with respect to sampling
* There was excess inventory that is LowTurnOver for some sku like BEDIL, NMQCN, PDTDG