**SQL Sales Data Project**

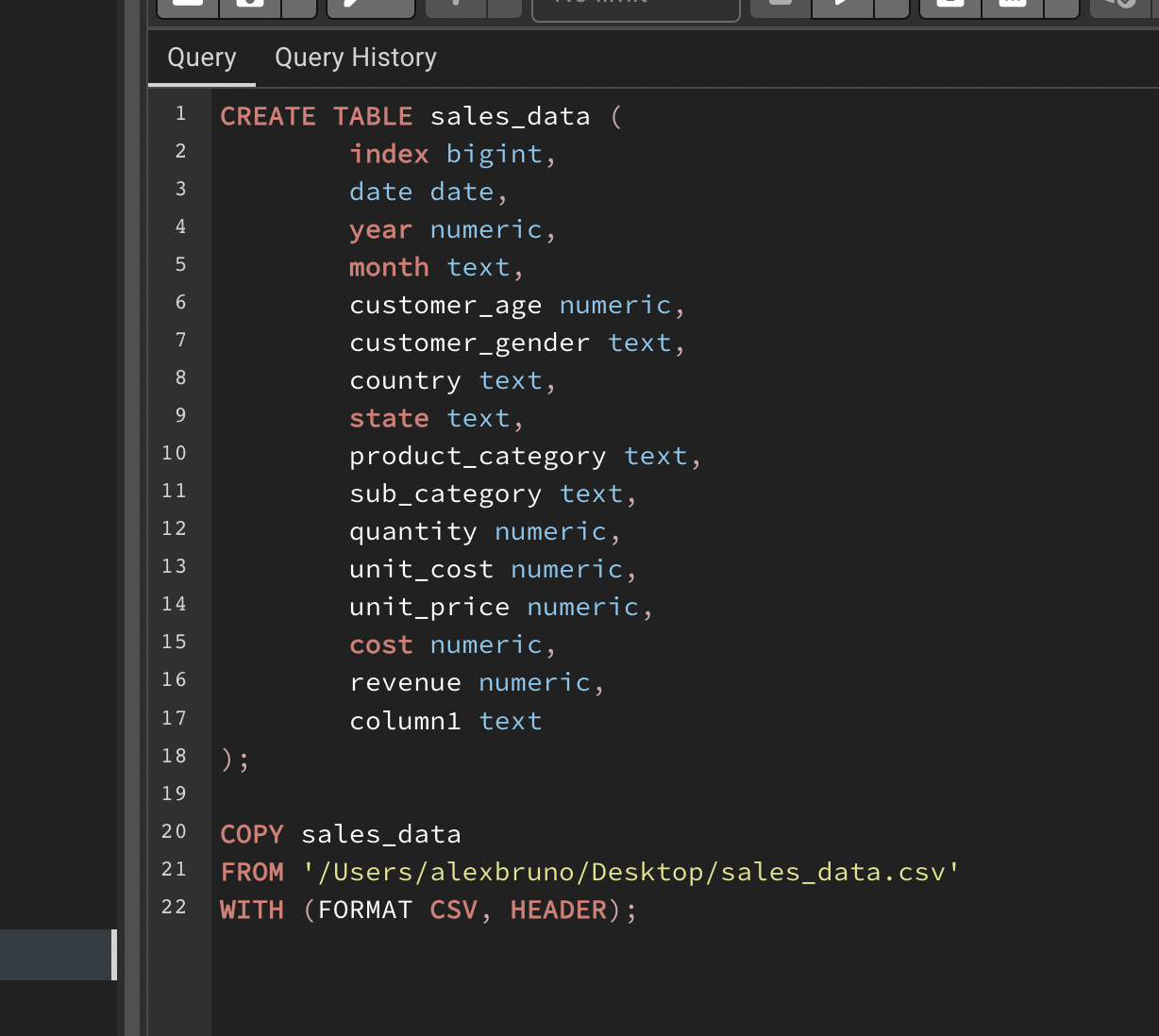
Source: found sample sales data from Kaggle.com

SQL type: PostgreSQL

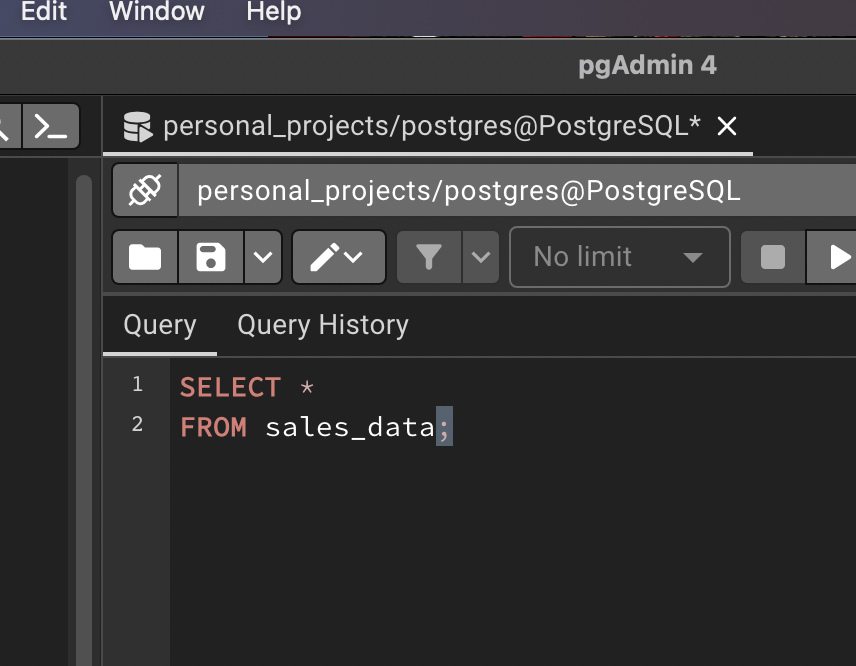
Admin Tool: pgAdmin4

Creating table within database and loading into pgAdmin4:

\*Some columns needed to be loaded as numeric to port over correctly or else an error message appeared. The columns were unnecessarily input with a decimal format. Will have to clean that later to perform calculations.



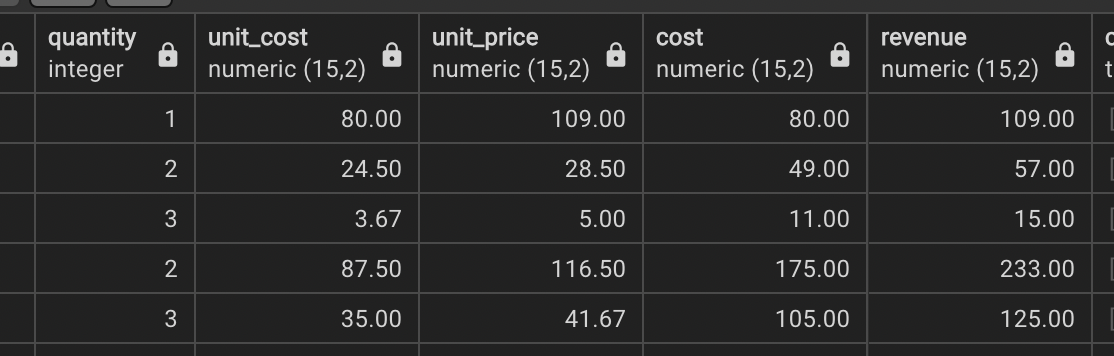
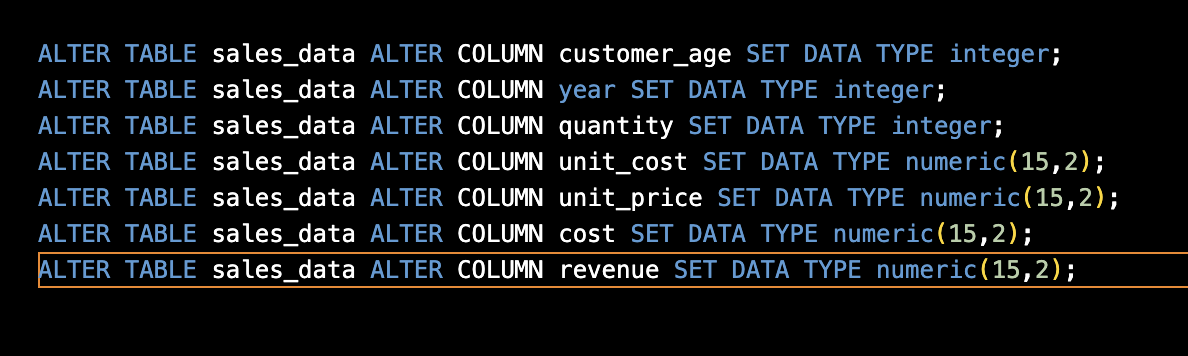
Inspecting initial data:



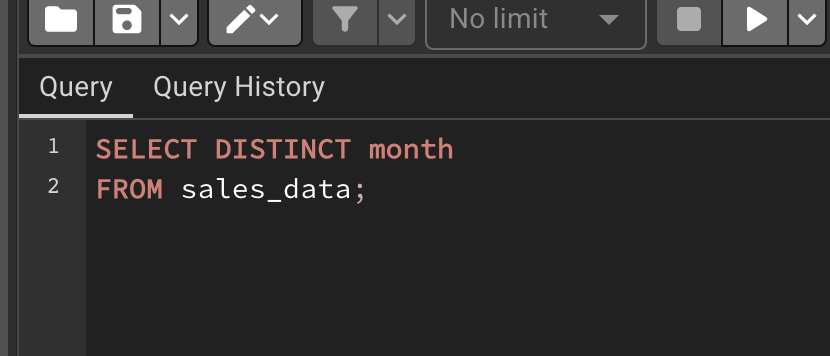
\*Data loaded correctly with 34,867 total rows.

**Data Cleaning**

Running ALTER TABLE statements to convert column data types to the correct type with good format:



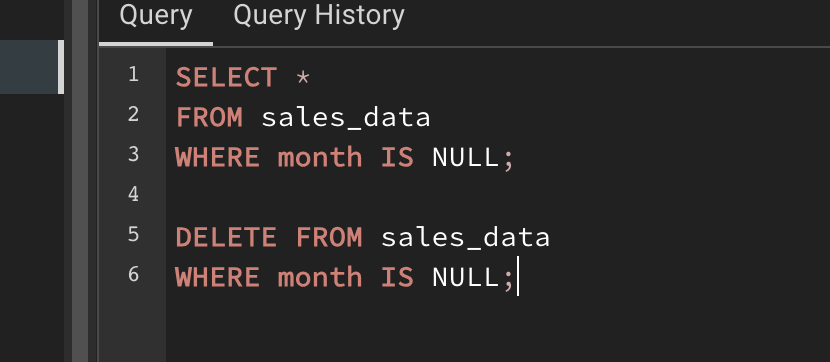
Running SELECT DISTINCT statement over every text column to find any inconsistent or missing values:



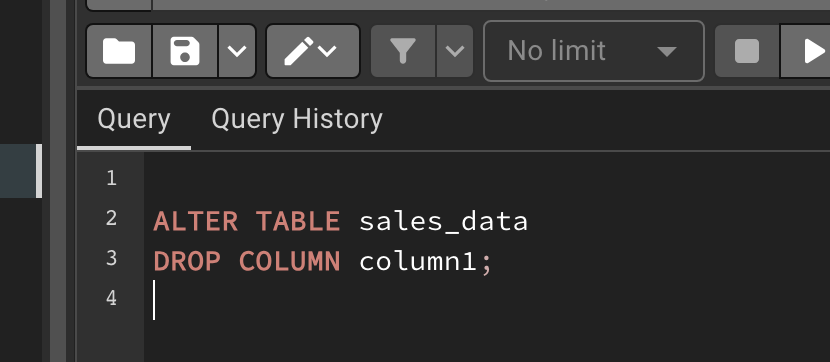
Cleaning tasks found:

1. Eliminate bottom row that only had a value in one column. Must have been a column that ran a total for some reason.
2. Eliminate empty column at the end of the columns titled column1.

Task 1 code:

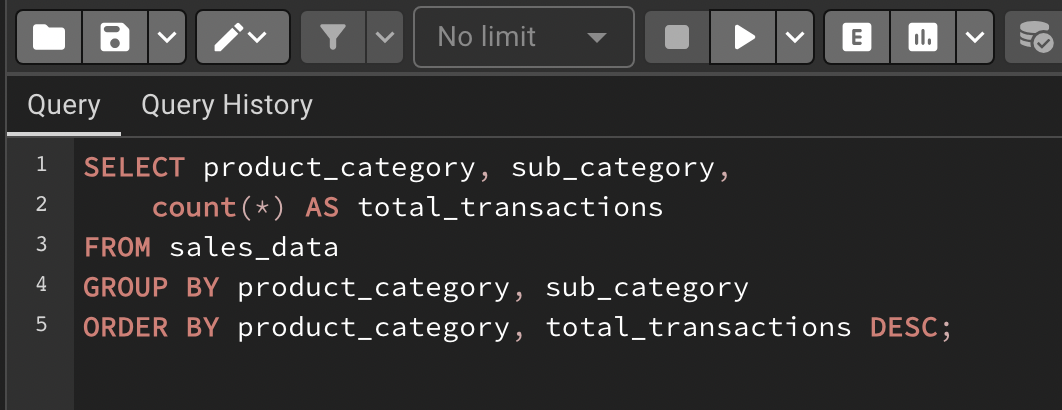


Task 2 code:



**EDA (Exploratory Data Analysis)**

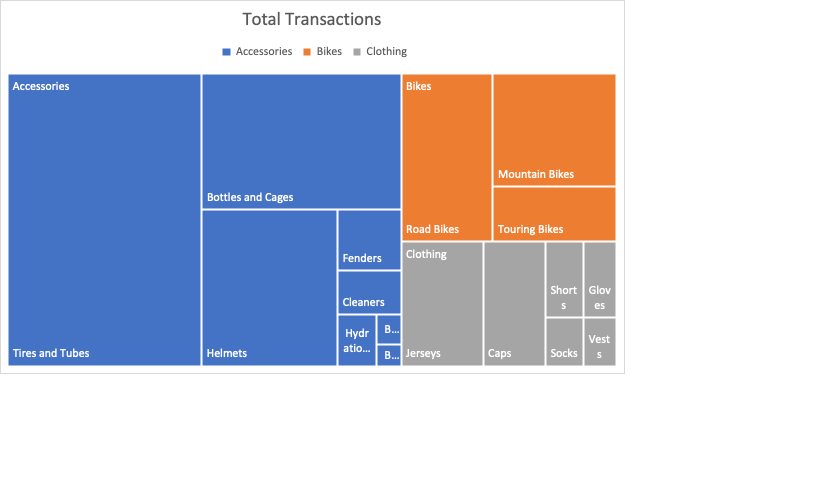
1. Most frequently bought items in each category/sub-category:



Results:



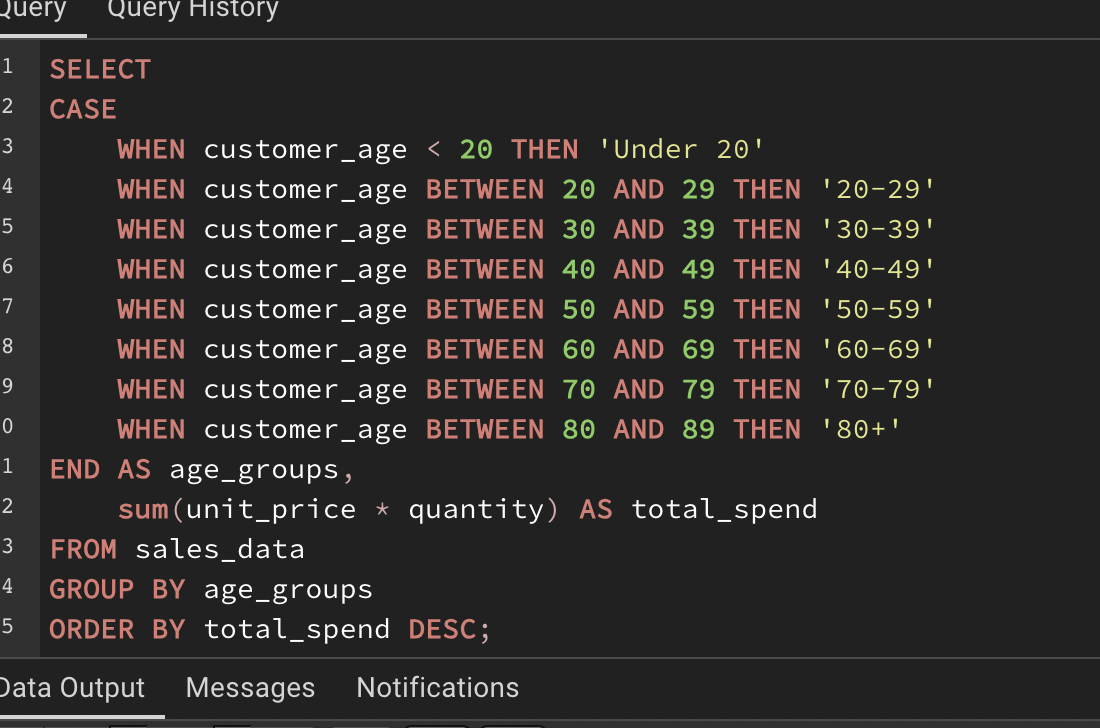
Exported to Excel to visualize results with a treemap.



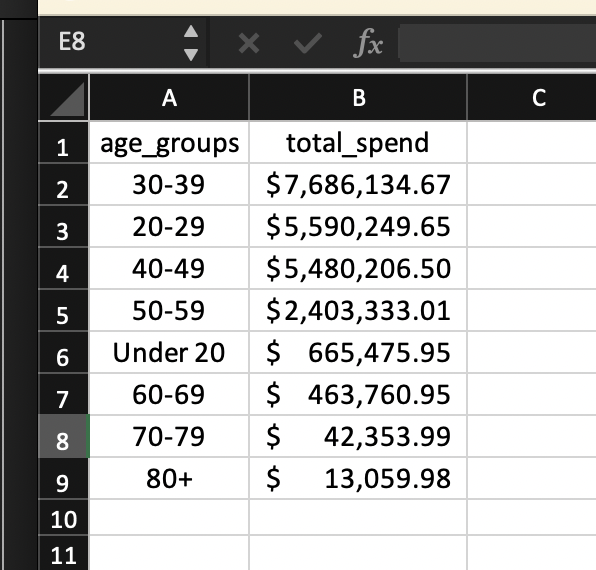
Notes:

* Accessories clearly holds the most total transactions out of all the categories.
* Tires and tubes are ¼ of ALL sales. The store should focus marketing on those two products.

1. Total spend by age category:



Results:

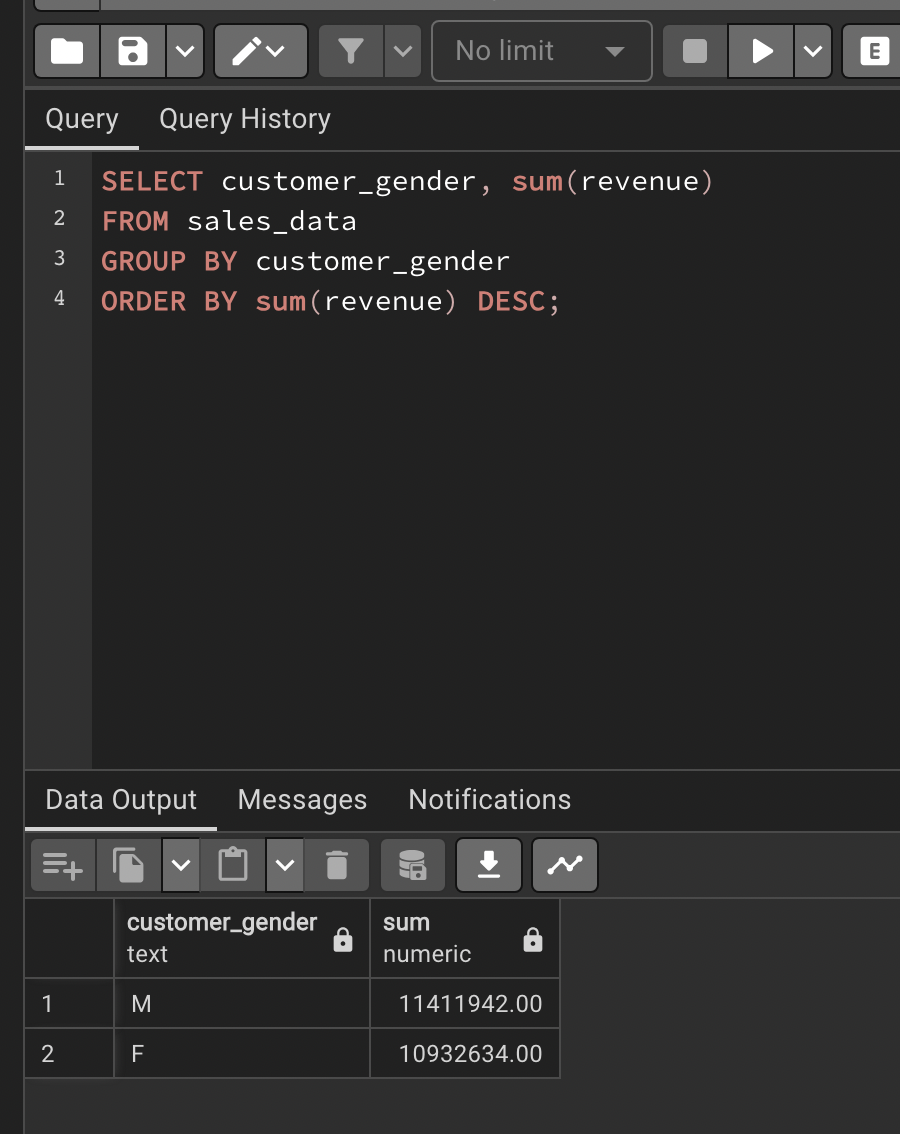


Visualized with Bar Chart:

Notes:

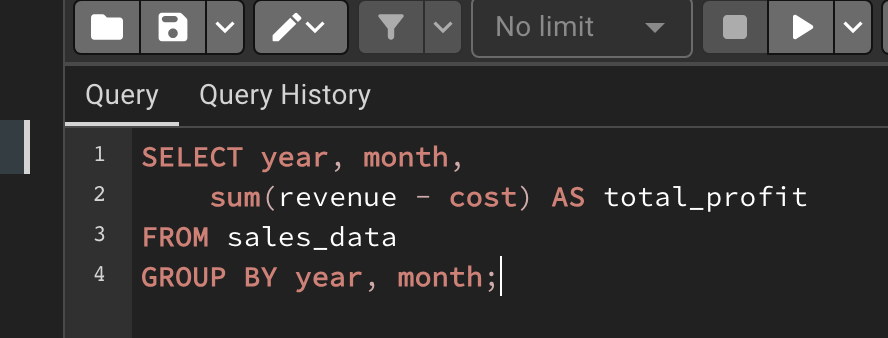
* Their largest spending groups, by far, are people in their 20’s, 30’s, and 40’s, which makes sense for a biking company
* Marketing/sales teams should be focused on these groups as customer targets
* People in their 50’s would be a good focus age group if the company is looking to expand their reach and acquire more market shares

1. Revenue by customer gender:

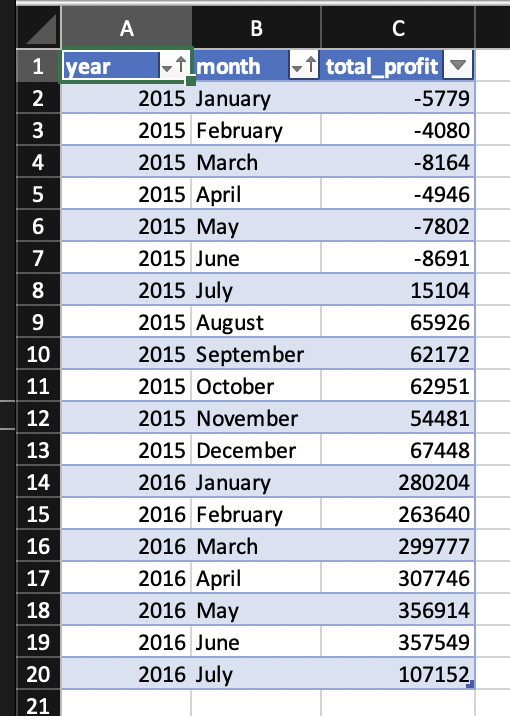


Notes: surprisingly even amount of revenue between genders. No further recommendations.

1. Total monthly profit/loss calculation:



Results:



Visualized with line graph:

Notes:

* Seems to me like January 2015 was the beginning of the company and it took them a while to turn a profit
* Peaked in April/May of 2016 then profits went down sharply in July 2016
* Would need more data/a different data set to determine what caused the increases/decreases in business