CV Writeup

Sale Insight - Electronics Store [Python]

(.) Extracted valuable insights from a given data set over a one-year period to understand strong sales trends.

(.) The final report was effective in displaying monthly sales and the reasons for such patterns, as well as the

most popular selling city and product. This report also discusses the optimum times to run adverts in order

to increase traction.

(.) This report will assist business owners in making critical decisions regarding their company's growth.

Data: https://github.com/KeithGalli/Pandas-Data-Science-Tasks (Youtube)

This was a kaggel data set where I use Python Pandas & Python Matplotlib to analyse and answer business questions about 12 months’ worth of sales data.

The data contains hundreds of thousands of electronics store purchases broken down by month, product type, cost, purchase address, etc. So I used this data set to performed Exploratory Data Analysis to bring out Key Insights.

I had 12 different excel files containing sales details of each months. So I had to combine all those excel file in to one excel file so that I can do the analysis of the whole year. To merge the excel files I took the help of Stack Overflow and merged all the data in one data frame.

Then I performed some cleaning operation on our data by Drop NaN values from Data Frame, removing rows based on a condition and Change the type of columns (to\_numeric, to\_datetime, astype).

Once I had cleaned up our data a bit, I moved the data exploration section. In this section I explore 5 high level business questions related to the data.

1. I found out What was the best month for sales? How much was earned that month? and while doing so I added a Month column in the data frame by extracting the month from the order date and Used group by to perform aggregate analysis. I also plotted a Bar chart of Sales Vs Month to visualize the data in a better manner. I found that December had the highest sales as it is the Festivel season .

2. I then found out What city sold the most product. To find this I extracted the city name along with the states using lambda function and then used group by to perform aggregate analysis. Here also I plotted a Bar chart of City vs Sales to visualize the highest selling city. I found that San Francisco had the highest sales as it is an Electronic Hub in US and Silicon Valley has a huge demand of electronics.

3. I also found that What time should the company should display advertisements to maximize the likelihood of customer’s buying product? For this I used

date time function to extract the Hour of purchase from the order date and then used group by function to find the best time according to the sales of that hour.

Here I used Line Chart to plot the hourly sales to understand the peak time of sales. I found that 11th Hour and 19th Hour are the best time for advertisement

as most people order at that time of the day.

4. I found What product sold the most? Why do you think it sold the most? For this I firstly grouped the data based on product to find the quantity sold for each

product and then plotted the data with respect to product name in a bar chart to understand the highest selling product. Then to find the reason behind the sales behaviour, I found the average price of each product and plotted the avg price on the same chart using subplot. Here I found that AAA Battery pack had the highest sales as the average price of the product is the lowest, while the Mac Book Laptop has the lowest sales as the average price of the product is high.

5. I found What products are most often sold together? and for this I created a Grouped column by finding duplicates of order id and then transforming those by join function. Then I with the help of stack overflow used the combination and counter functions to create a dictionary containing 2 product names and the count of times those were sold together. Here I found that iPhone and Lightning Charging Cable were sold the most together.