**Retail Analysis with Walmart Data**



* In this report we are going to discuss the retail analysis with Walmart data of the historical data that covers sales from 2010-02-05 to 2012-11-01 using python language and the process will be as follow:

1. Importing Data

2- display data

3- visualize quantitative variables distributions

4- perform data cleaning

* First:

We will import our packages which are NumPy, pandas, matplotlib and seaborn to help us in performing our data in a suitable way and in making our calculations and at the end visualizing our data will direct us to be aware of the end goal we have mind.

Graphical user interface, text, application

Description automatically generated

* Second:

We will display our data: -

Table

Description automatically generated

* Third:

We will do some operations on our data in order to visualize data much easier as wanted in the assignment.

A picture containing graphical user interface

Description automatically generated

* Fourth:

Looking at the data we can see that its already a clean dataset.

* For example: there is no rows that contains NAN values to be dropped.
* **Now we are going to perform some needed calculations on the dataset:**

a-Identifying the store that has the maximum sales.

Text

Description automatically generated with low confidence

b- Identifying the store that has the maximum standard deviation.

Graphical user interface

Description automatically generated with low confidence

c-

**First**: We are going to calculate the mean value of the non-holiday’s weeks: -

A picture containing text

Description automatically generated

**Second**: we are going to find out holidays that have higher sales than the mean sales in the non-holiday season for all stores together.

A picture containing table

Description automatically generated

d- Providing a monthly and semester view of sales in units, here is some sample rows:

Graphical user interface, application

Description automatically generated

e-1-we have scatter graph between Weekly sales and Stores

Chart, bar chart

Description automatically generated

We will represent the Unemployment of DataChart, box and whisker chart

Description automatically generated

2- We will represent the CPI of Data

Chart, box and whisker chart

Description automatically generated

Now we will Show comparison between the sales in years Chart, pie chart

Description automatically generated

3-we have a graph between Weekly sales and Fuel Price

Chart, scatter chart

Description automatically generated

**Fuel Price has no effect in the weekly sales**

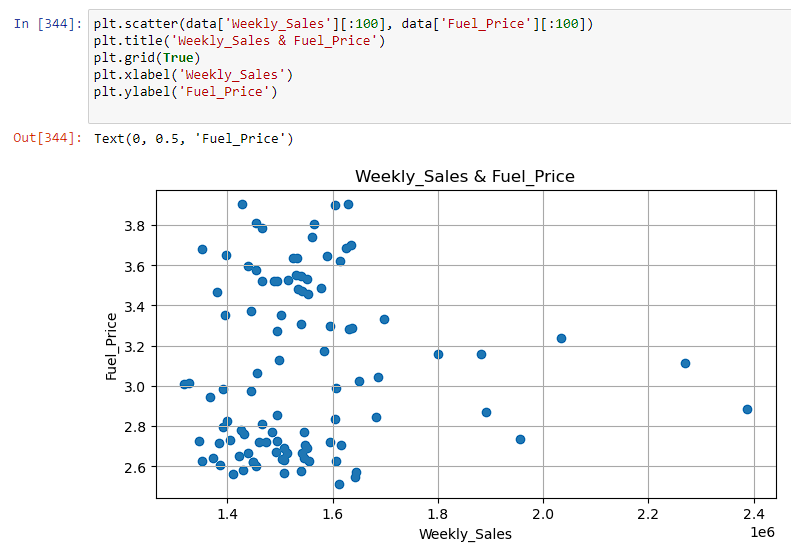
3-we have a graph between Weekly sales and temperature

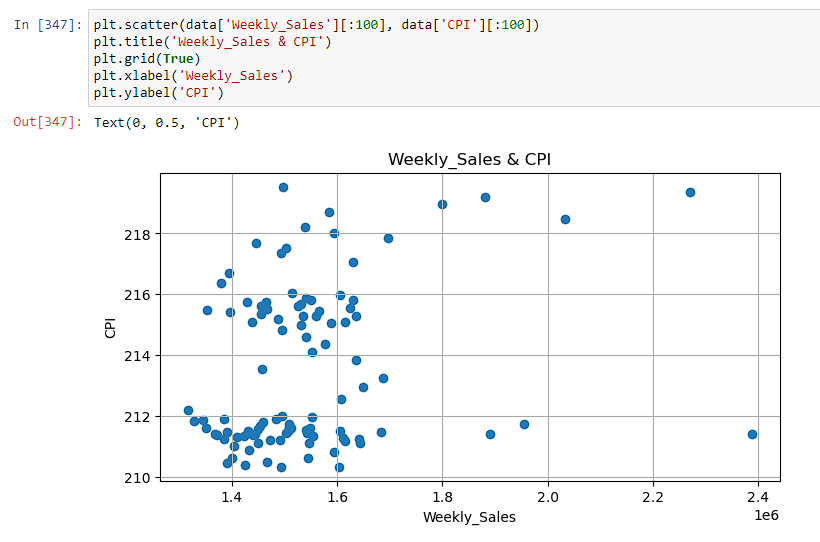
**Chart, scatter chart

Description automatically generated**

**temperature has no effect in the weekly sales**

second, we have a graph between Weekly sales and Fuel\_Price

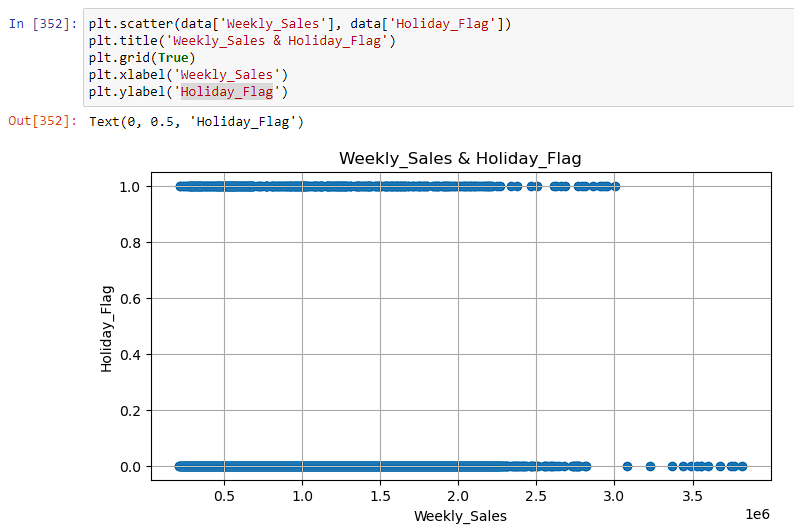


Third, we have a graph between Weekly sales and CPI

Fourth, we have a graph between Weekly sales and UnemploymentChart, scatter chart

Description automatically generated

fifth we have a graph between Weekly sales and Holiday\_Flag



* And the last plot we have is to summarize all these relations which is a heatmap that identifies positive and negative correlations between weekly sales and all other numeric variables.

Chart, treemap chart

Description automatically generated

**And that’s everything … We apologize for elongating, and lengthiness.**

**Hope we met all the requirements. 😊**

**Thanks…**