WALMART SALE ANALYSIS REPORT

1.Imported necessary Libraries for the analysis.

2. Imported the CSV file using Pandas.

3.Checked the Data types, verified the columns , Duplicated data and Null values in the Data set.

4. Verified the sales of Stores by grouping them and Analysed that Store 20 has the Highest sale of them all.

5. Represented the Bar visualisation of Sales by Each store with respective to the Months in year.

5.Calculated the Standard deviation and Mean of sales of each store out of all the Store 20 has

Highest STD.

6. Converted the Date column object variable to datetime64[ns] and imported a new variable days to the Date column and Splits the months , year from Date column for better analysis .

7. Analysed that on the Q3 2012 the Highest sale on the 8th month and Store 3 has the Highest sale on the month.

8. Analysed that total of 220 holidays has good sale than the mean sales in non-holiday season for all stores together

9.Cerated data for Store 1 to Hypothesize if CPI, unemployment, and fuel price have any impact on sales.

10.Wrangling the Store 1 data and Displayed the Bar Visualization for sale on each year by the Store 1.

11. Visualised the Skewness graphs, Box plot to find outliers and Scattered plot to find out relation among the variables.

12.Visualized the Heat Map showing the correlation among the variables.

13. Performed the train test using sklearn.model_selection and Concluded that CPI , Temperature has Impact on sales rather than the Unemployment.

Report by

Muniyappan Gopal.