Sales and Temperature Data Analysis using Pandas

Introduction: -

➢ In this project we will have a closer look on the sales in Walmart stores and the changing in temperature degrees globally. We will try to answer some questions such as:
  ○ Which stores at Walmart score the highest revenue and lowest revenue?
  ○ What is the trend of sales for the store with the highest revenue over time?
  ○ What are the months that this specific store scored top 5 sales?
  ○ What is the average temperature in the USA for each month in 2021?
  ○ Is there a relationship between the temperature and the sales for Walmar? In other words, Are the sales affected when the temperature goes up?
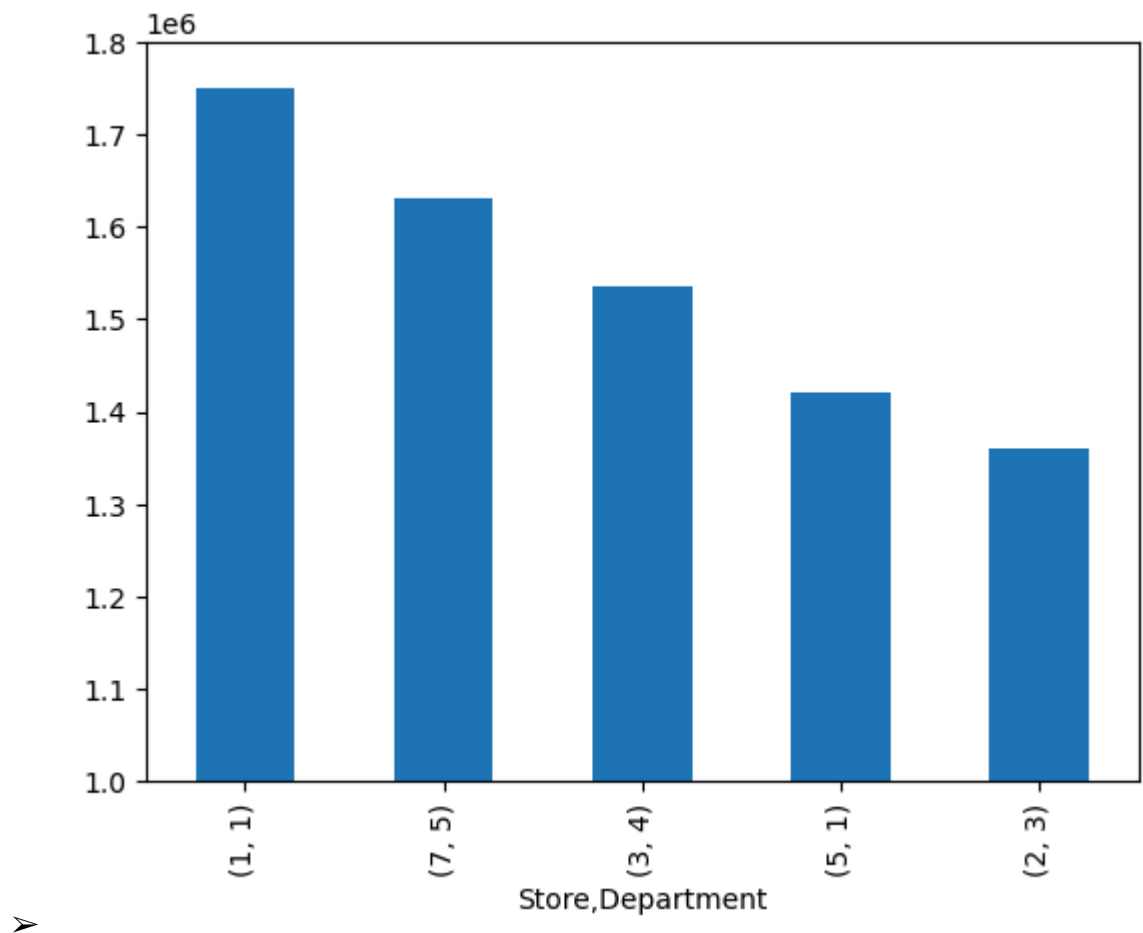
Methodology: -

➢ To apply EDA, Pandas and Matplotlib were used in this project
➢ Data cleansing was crucial as the data needed to be handled for further analysis and possible merging. Data cleansing steps applied:
  ○ Duplicated rows: Some rows were duplicated in both datasets.
  ○ Outliers: Some extreme values in the Sales and Global_Temperature columns existed as outliers. Handled but not dropped as they were needed for future statistics.
  ○ Inconsistent Casing: Upper and lower case variations in the City and Country columns.
  ○ String Paddings: Extra spaces before and after the city names in the global temperature data.
  ○ Inconsistent Date Formats: Different date formats within the same column.
  ○ Converting any date columns to DateTime objects for better manipulation.
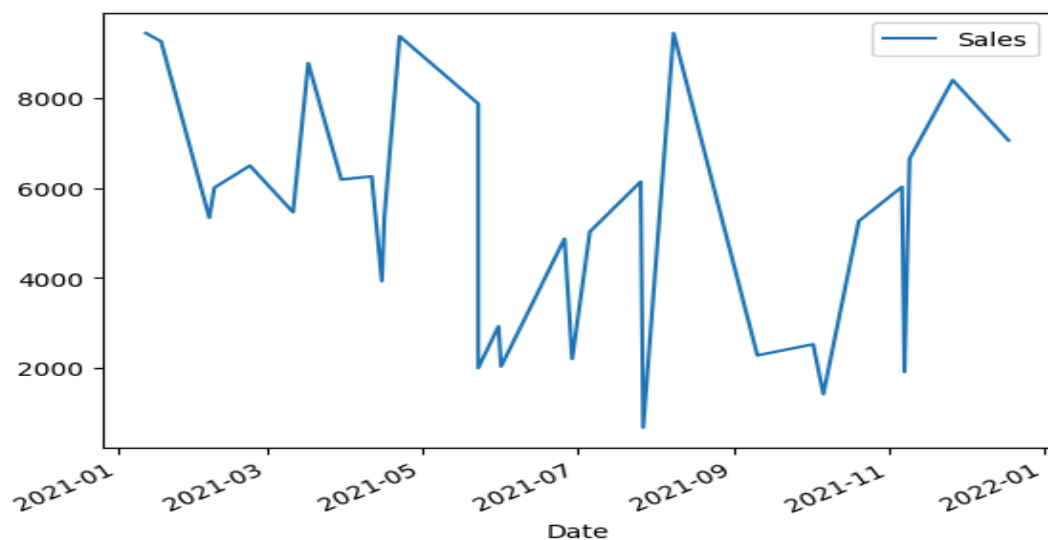 ○ Sorting Dates.

Results: -

Store 1 is our Best Store

❖ Our data analysis process shows that Store 1 has scored the highest total revenue in the year. (sum of all revenues across all months). **Let's call it our Best Store**
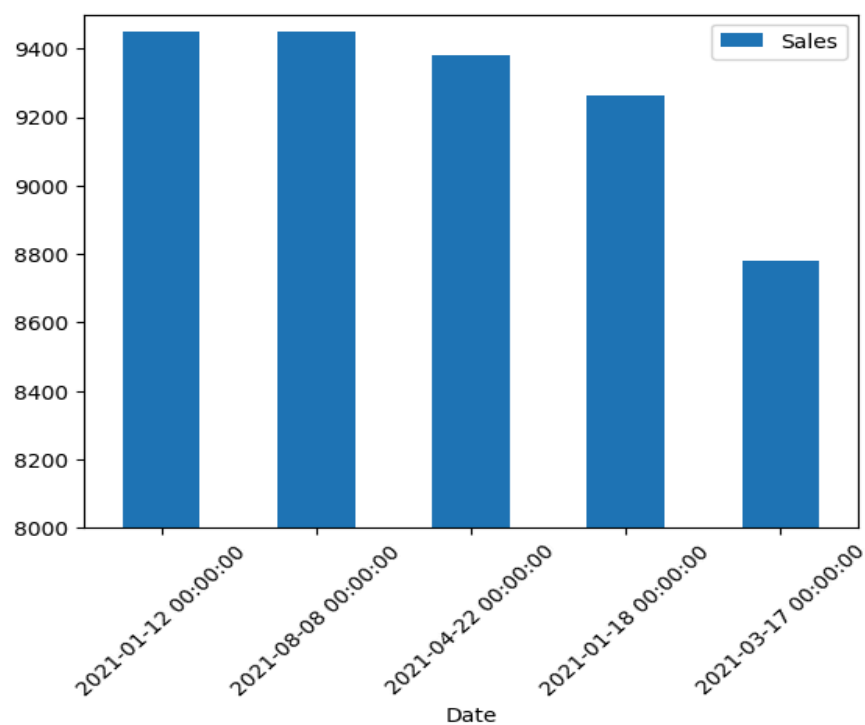

➢

Our Best Store Sales over the Year across Months

❖ For the sales over time, We can see that our best store kept scoring sales **not less than 4000** sales from **2021-Jan** till the end of **2021- April** reaching a **max** target of more than **9450** sales on **2021-Jan.** Noting that those months are Winter and Spring seasons.

❖ As we entered **June-2021** Walmart Store faced a significant **drop** in its sales scoring around **2000 sales**. Which finds it somehow challenging to go back to its high sales values until we reached **2021-August.** In August 2021, Walmart Scored Its **second highest sales** number which is exactly equal to **9450** like the **2021-Jan** number.

❖ Our best store scored its top 5 highest sales in January, August, April and March.

❖ **In conclusion, Walmart scored most of its top highest sales in the first 2 quarters of the year. Which indicates that it is more profitable in the first 6 months than the last 6 months.**
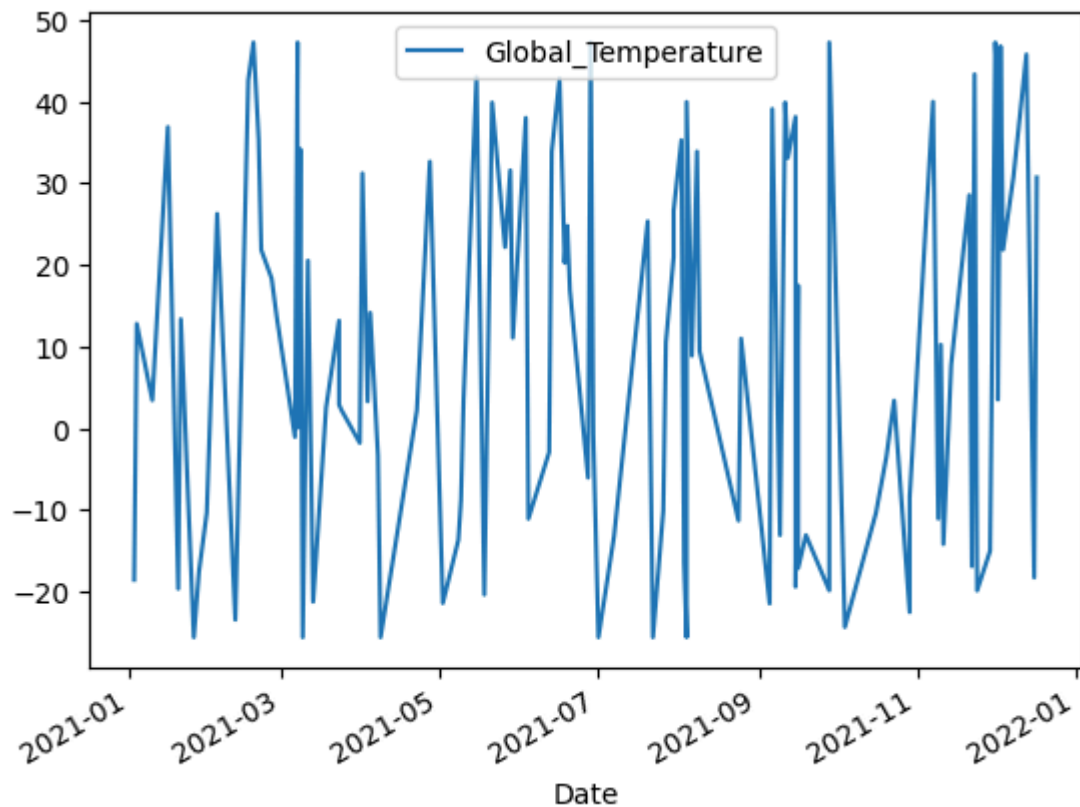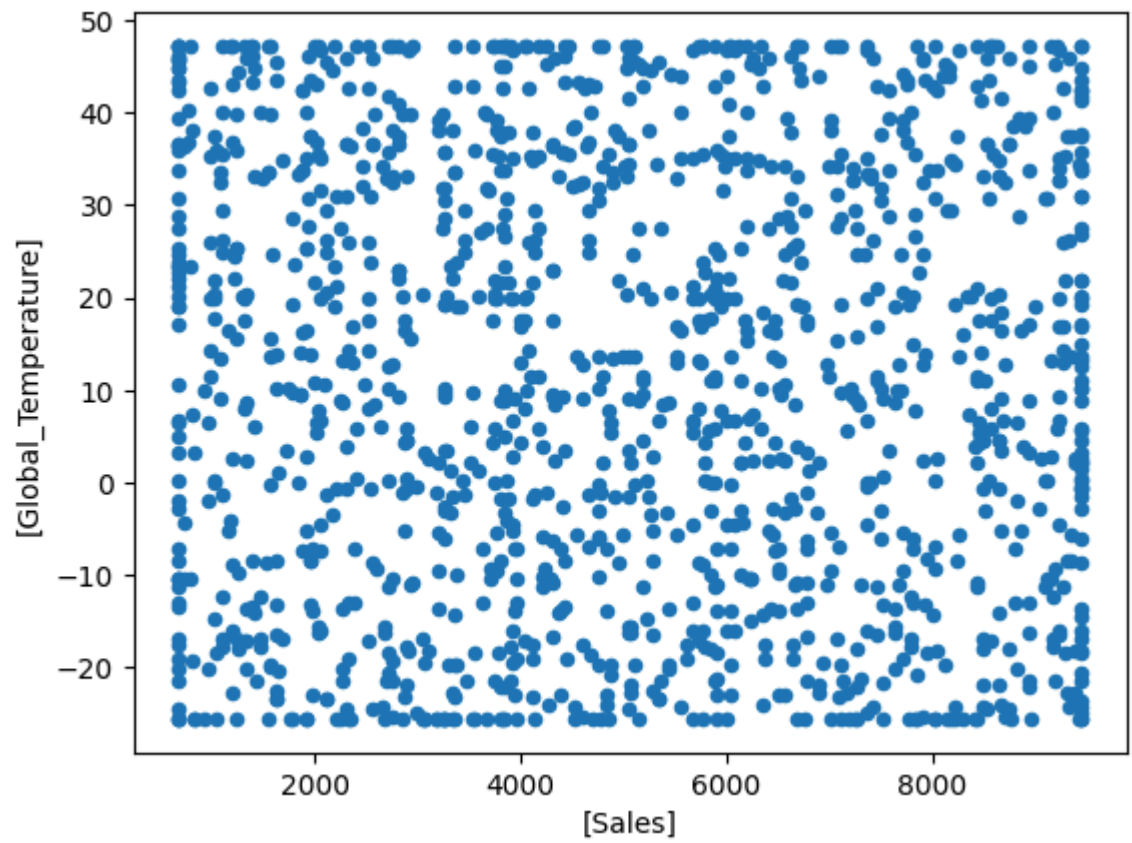
❖

❖

Average Temperature over the year by Month in USA

- ❖
- ❖ We can see that new york in the USA has a wide change in temperature during the whole year.

Are sales affected by the weather?

- ❖ By our EDA process, it showed that sales are not affected by change in temperatures with a correlation coefficient equals to -0.029. The coefficient shows a very weak negative correlation. Meaning that when the temperature rises sales decrease with very small amounts.

❖
❖ This plot shows how sales are not affected by the temperature values.