

Jiteshwar Anjale

DSC680, Weeks 5-8

21-October-2023

Project 2: Milestone 3

Q&A (Answer the questions you proposed in Milestone 2)

Questions and Answers

Q: What is technical analysis?

Technical analysis is used to determine the features used to predict sales volume at a supermarket store.

Q: What is machine learning?

Machine learning is a type of artificial intelligence (AI) that involves training computer algorithms to recognize patterns in data and make predictions or decisions based on that data.

Q: How do graphics help?

Graphical visualization of the data enables us to recognize patterns.

Q. What are recurrent neural networks?

A recurrent neural network (RNN) is a type of neural network commonly used for processing sequential data, such as time series or natural language text.

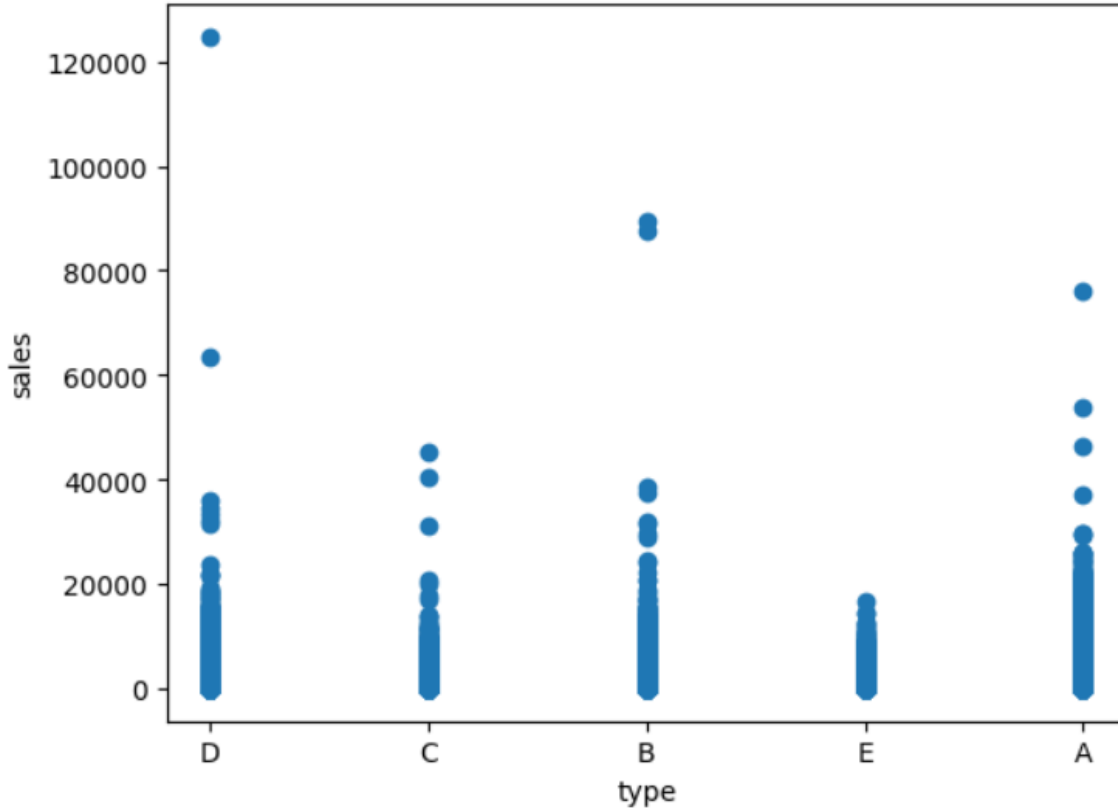
Q. Explain Autoregressive Integrated Moving Average Model (ARIMA) model algorithm.

An autoregressive integrated moving average (ARIMA) model is a statistical analysis model that leverages time series data to forecast future trends.

Q. Can the use of ML adversely affect the sales forecast of supermarket stores?

Automated, rapid transactions conducted by machine learning have been known to cause disruptions, but effects have been temporary and minimal.

Q. Does this type of stores affect the store sales?

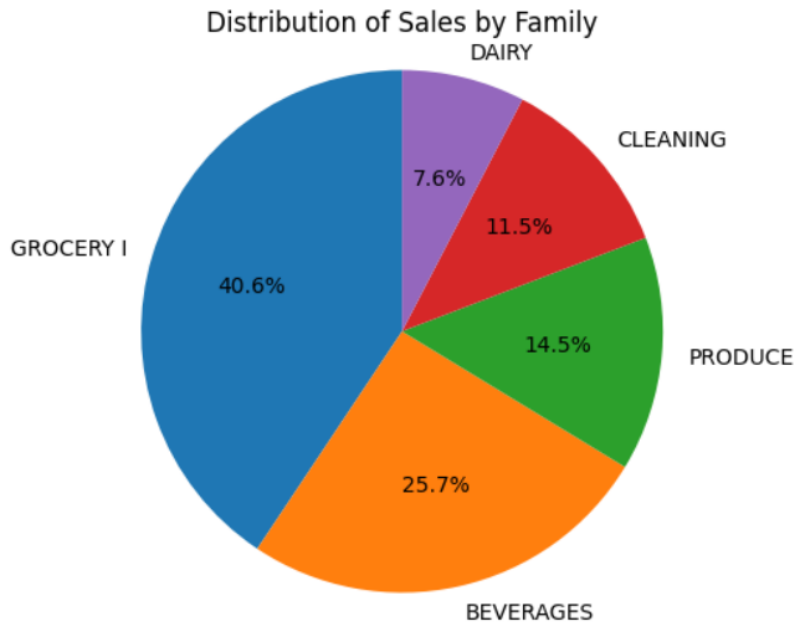


ANOVA (Analysis of Variance) is a statistical test used to determine whether there are significant differences between the means of two or more groups. It compares the variation between the groups (due to the different categories or factors) to the variation within the groups.

H_0 (>0.05)= The type of stores does not affect store sales. There is no significant difference in store sales between different types of stores.

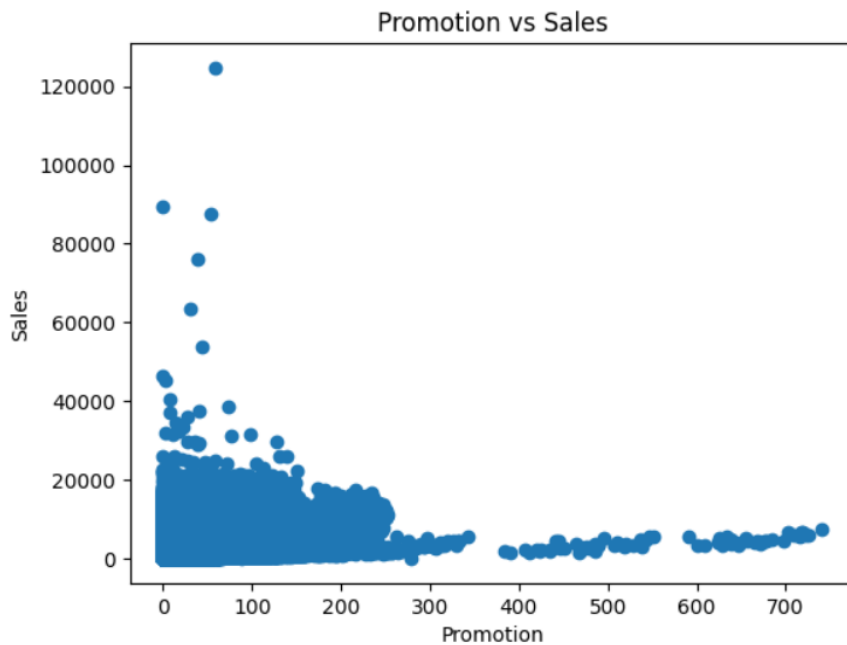
H_1 (<0.05)= The type of stores does affect store sales. There is a significant difference in store sales between different types of stores.

Q. Which family has the highest sales?



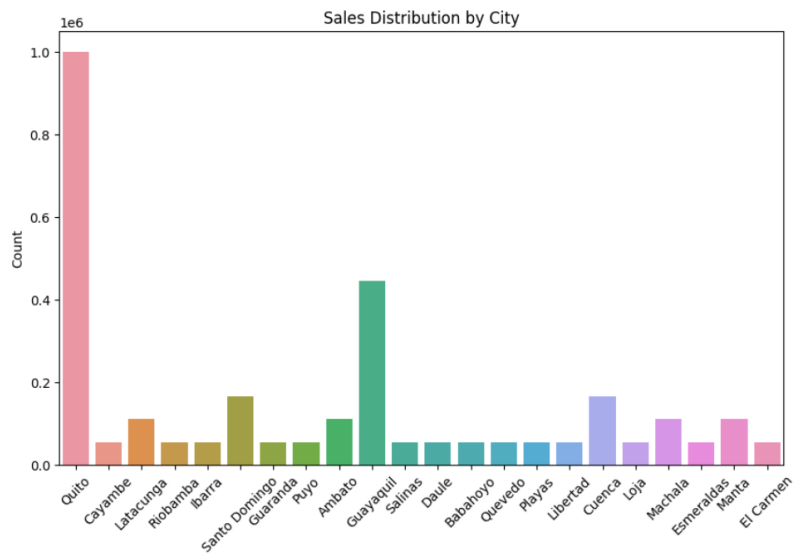
Based in the EDA, the GROCERY I is having the highest sales, and Beverages comes second highest.

Q. Does promotion able to improve the sales?



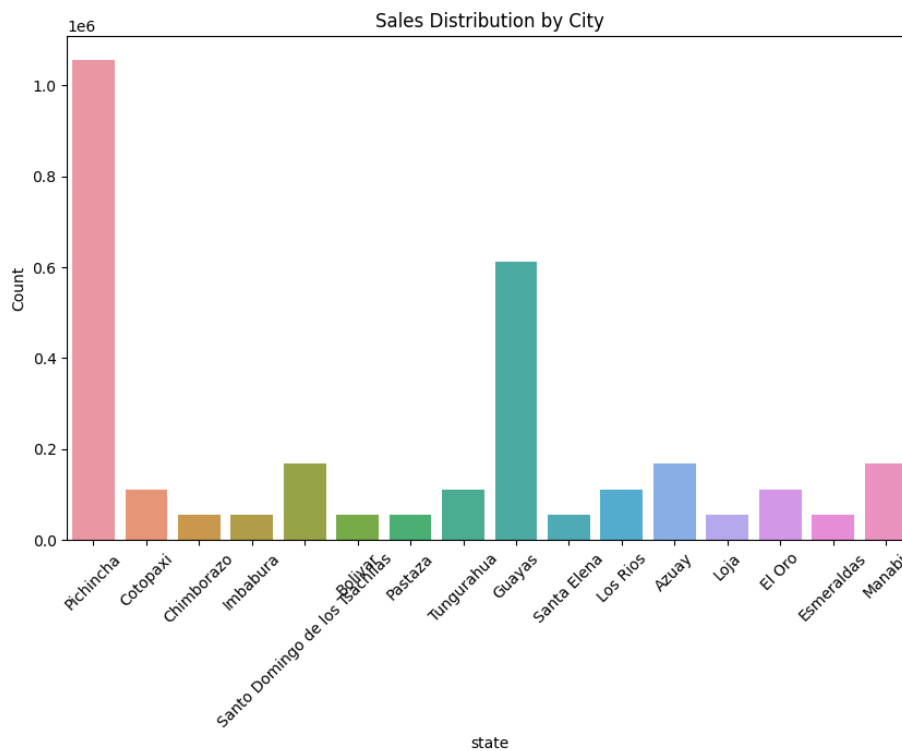
Based on the Pearson correlation coefficient of 0.4279 and the p-value of 0.0, we can reject the null hypothesis (H_0) and conclude that there is a significant relationship between promotion and store sales. Therefore, the promotion does affect store sales.

Q. Which city has the greatest number of customers?



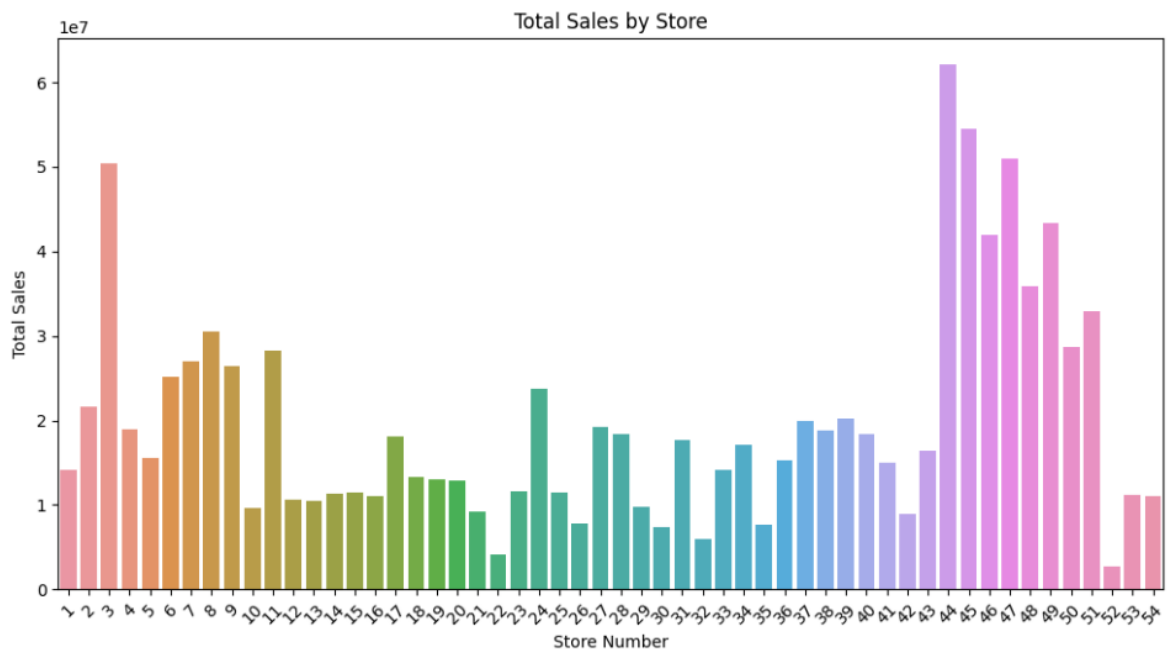
Based on the EDA, it shows that Quito has greatest number of customers.

Q. Which state has the greatest number of customers?



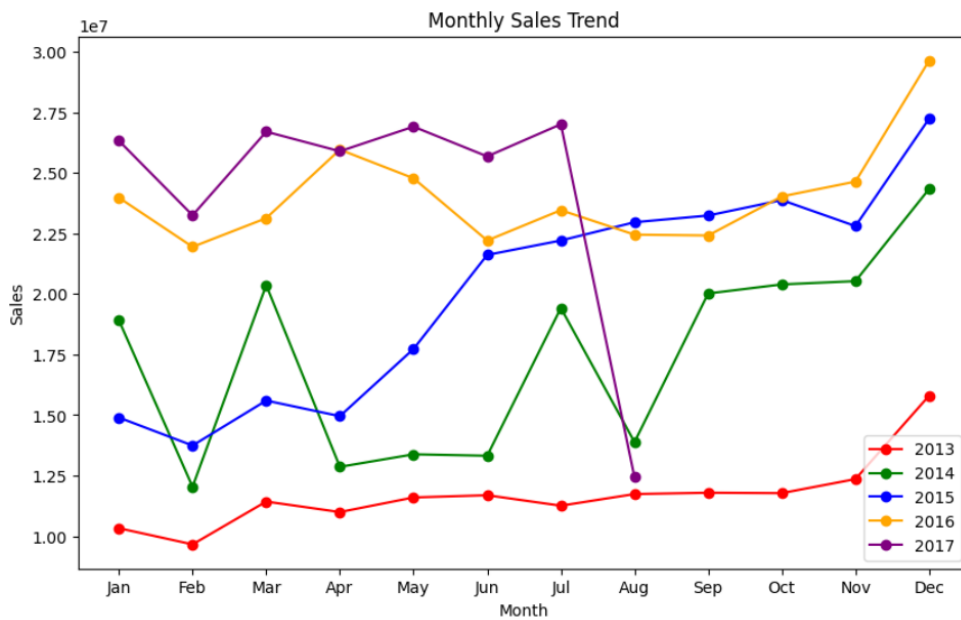
Based on the EDA, it shows that Pichincha has greatest number of customers.

Q. Which of the stores has the highest sales.



Based on the EDA, it shows that store number 44 has the highest sales.

Q. Which month has the most sales, and least sales.



Overall, the orange line which is 2016 has a stable high sales since January to Dec. Between the months in 2016, December had the most sales. In other hand, in comparing to other years, 2013 had an overall lowest sales achieved, especially during February.