Introduction

In this project, we have performed an Exploratory Data Analysis on the Amazon Go Store dataset.

The link for the Dataset is: https://bit.ly/3i4rbWI

You can download it from the link.

In this dataset, we have many features like shipping method, Segment, country, City, State, Postal code, Region, product category, product sub-category, sales, Quantity, discount, and the Dependent variable is profit. The remaining are independent variables.

Tools, libraries and Languages Used:

- Jupyter Notebook
- Python
- Pandas, Numpy, plotly, matplotlib

Analysis from Heatmap

- Sales and Profit are Moderately Correlated
- Discount and Profit are Negatively Correlated

Overall Analysis

- The main reason which leads to loss is Discount as if some areas lead to loss due to more discounts, and some areas lead to fewer sales due to fewer discounts, hence it needs to be improved.
- It is better to give more discounts during festival seasons, additionally, that will result in more sales.

- The Home office segment needs better improvement.
- Some cities have fewer sales, lack of awareness can be the reason for this, hence advertising in those cities might help in more sales.