



D MART SALES ANALYSIS

Between (2019-2022)
Predicted Profit Or Loss

**PREPARED BY
AAKASH R**

CONTENTS

Points for discussion



01 Introduction

02 Objective

03 Collecting The data

04 Preprocessing The Data

05 EDA Process



Time line Series

06 Model Selection

07 Predicting

08 Conclusion

An aerial photograph of a city skyline, likely New York City, taken at dusk. The sky is a mix of purple, blue, and orange. A semi-transparent dark rectangle is centered over the image, containing the text 'DATA SOURCE' in large, bold, orange capital letters.

DATA SOURCE

https://docs.google.com/spreadsheets/d/165vRoweWpk_kUR0WNNuZZti1L4AbmLsVxGc5Pzjvxuk/edit#gid=0

INTRODUCTION

Bussiness Problem

As a Business manager, To find the product sales, loss or profit for the last 4 years. According to the predicted report the level of marketting is to be increased in the state of increasing the profit. Derive business problems by exploring the data



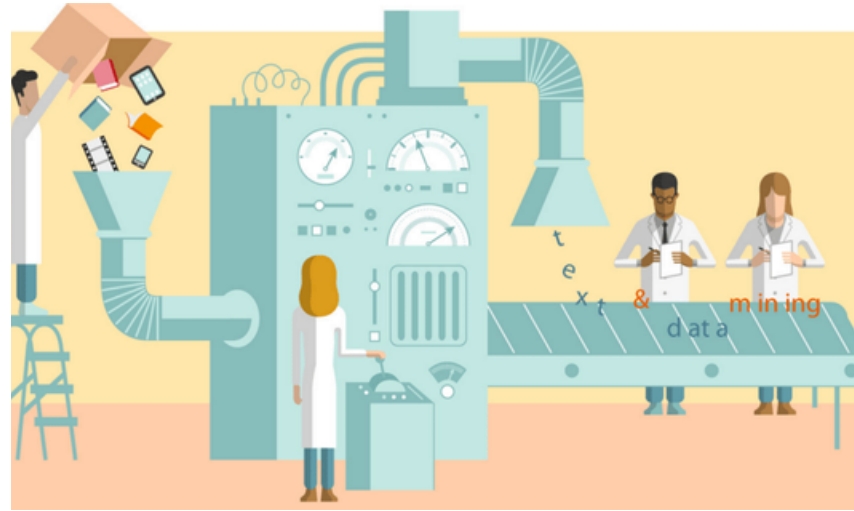
OBJECTIVES



Bussiness Solution

1. Identifying the sales percentage and evaluating the loss according to different states and promoting the marketing in particular states.
2. To identify more product sell by state and less product sell by state.
3. To identify more Product sold and less product sold.

Machine Learning Process



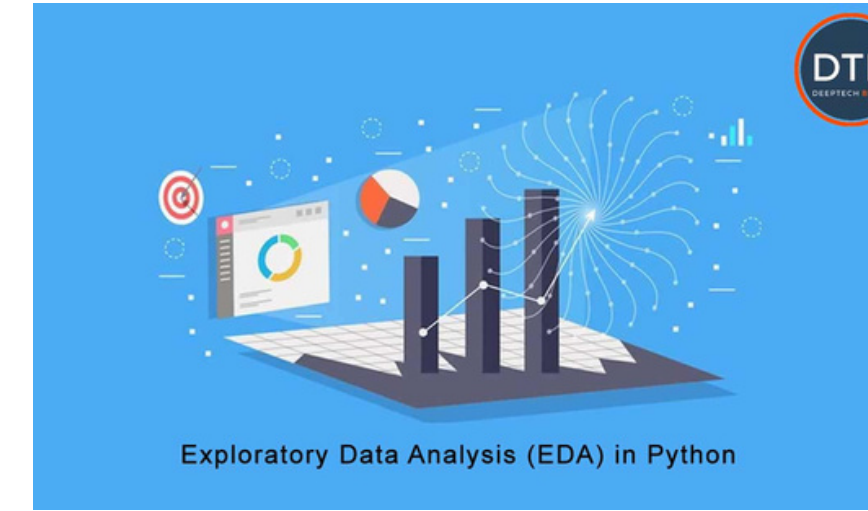
Collecting the Data

Collecting data for training the ML model is the basic step in the machine learning pipeline. The predictions made by ML systems can only be as good as the data on which they have been trained.



Preprocessing The Data

Real-world raw data and images are often incomplete, inconsistent and lacking in certain behaviors or trends. They are also likely to contain many errors. So, once collected, they are pre-processed into a format the machine learning algorithm can use for the model.

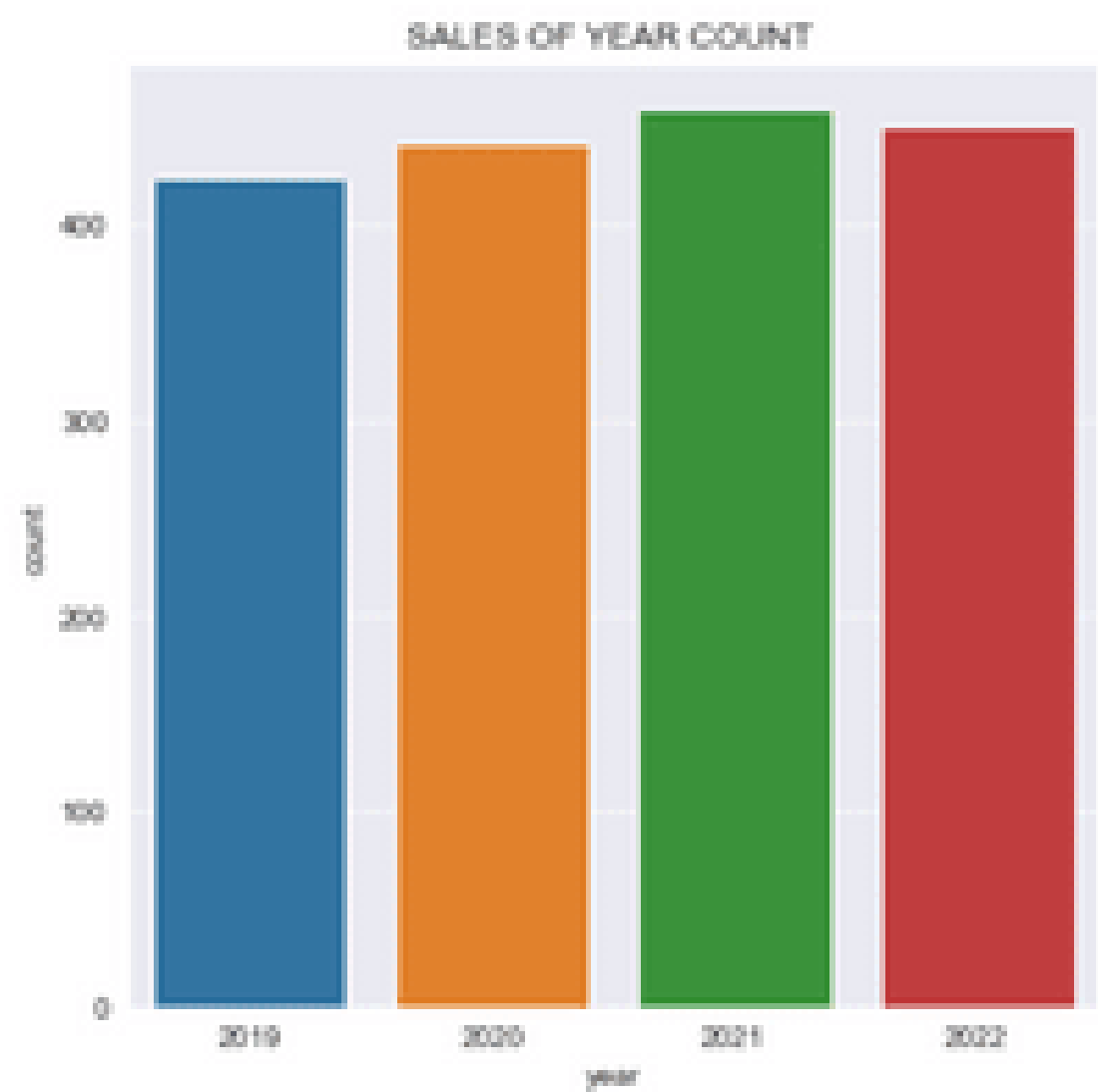


EDA Process

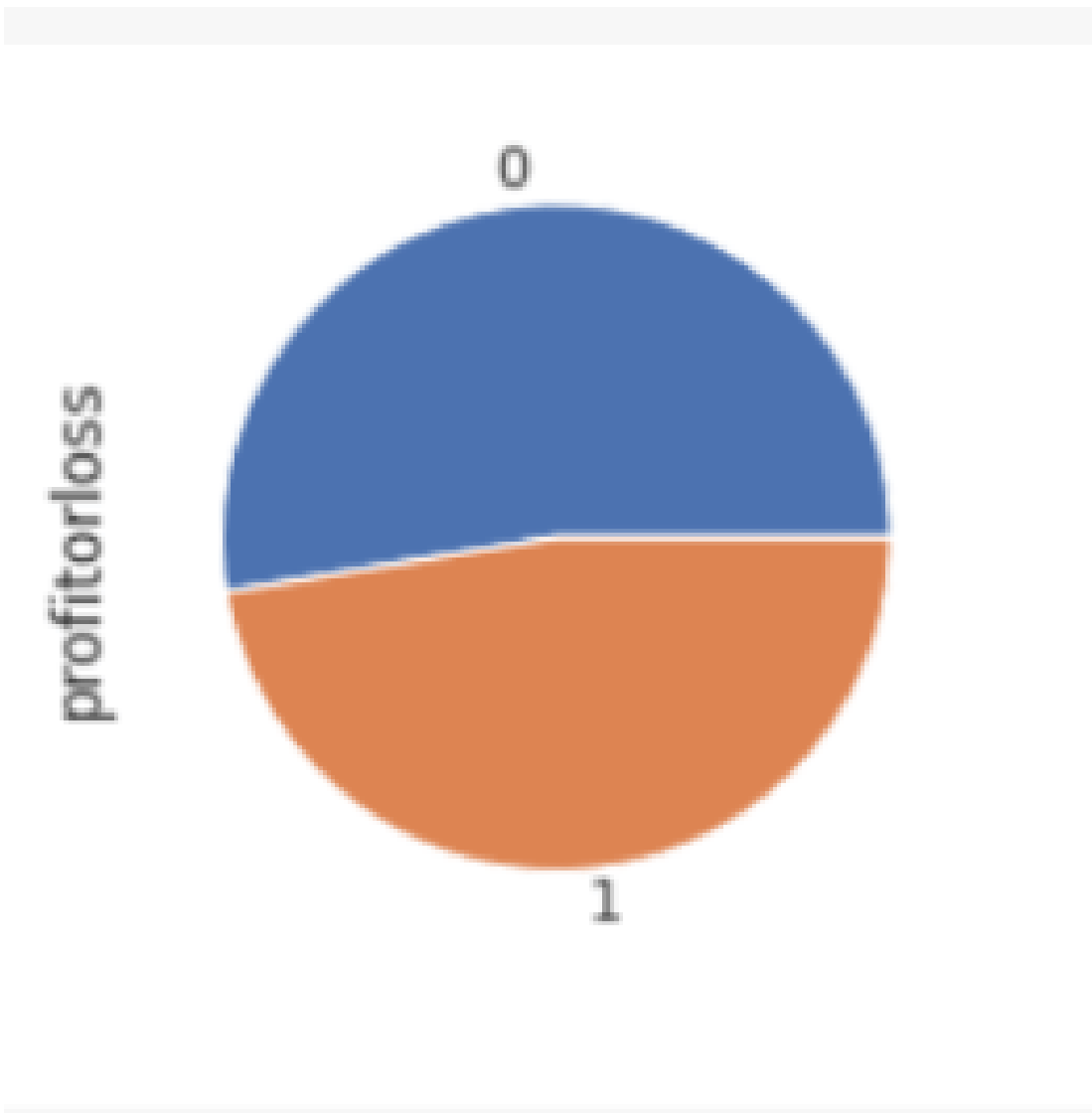
Exploratory Data Analysis (EDA) is an approach to analyze the data using visual techniques. It is used to discover trends, patterns, or to check assumptions with the help of statistical summary and graphical representations.

DATA VISUALIZATION

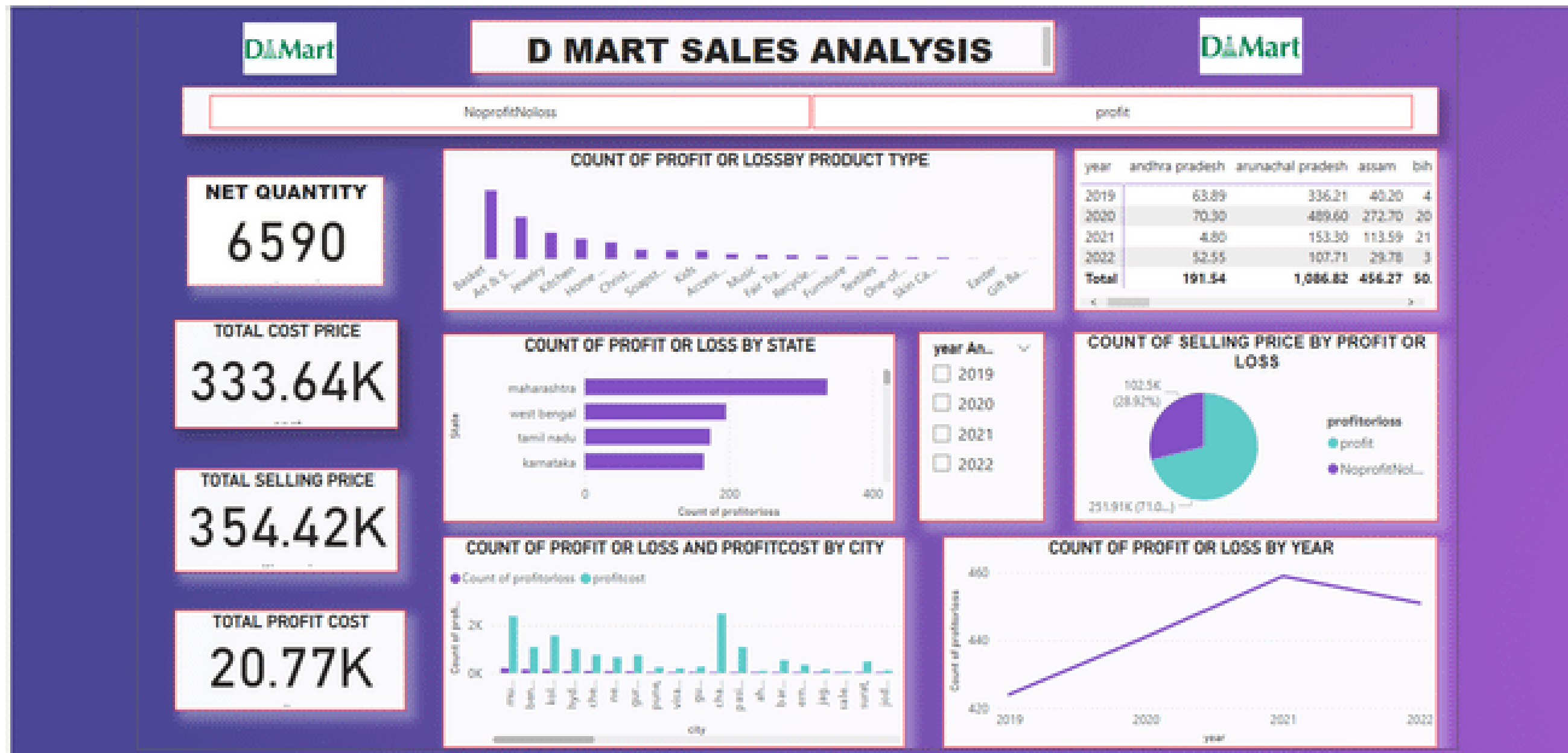
SALES COUNT BY YEAR



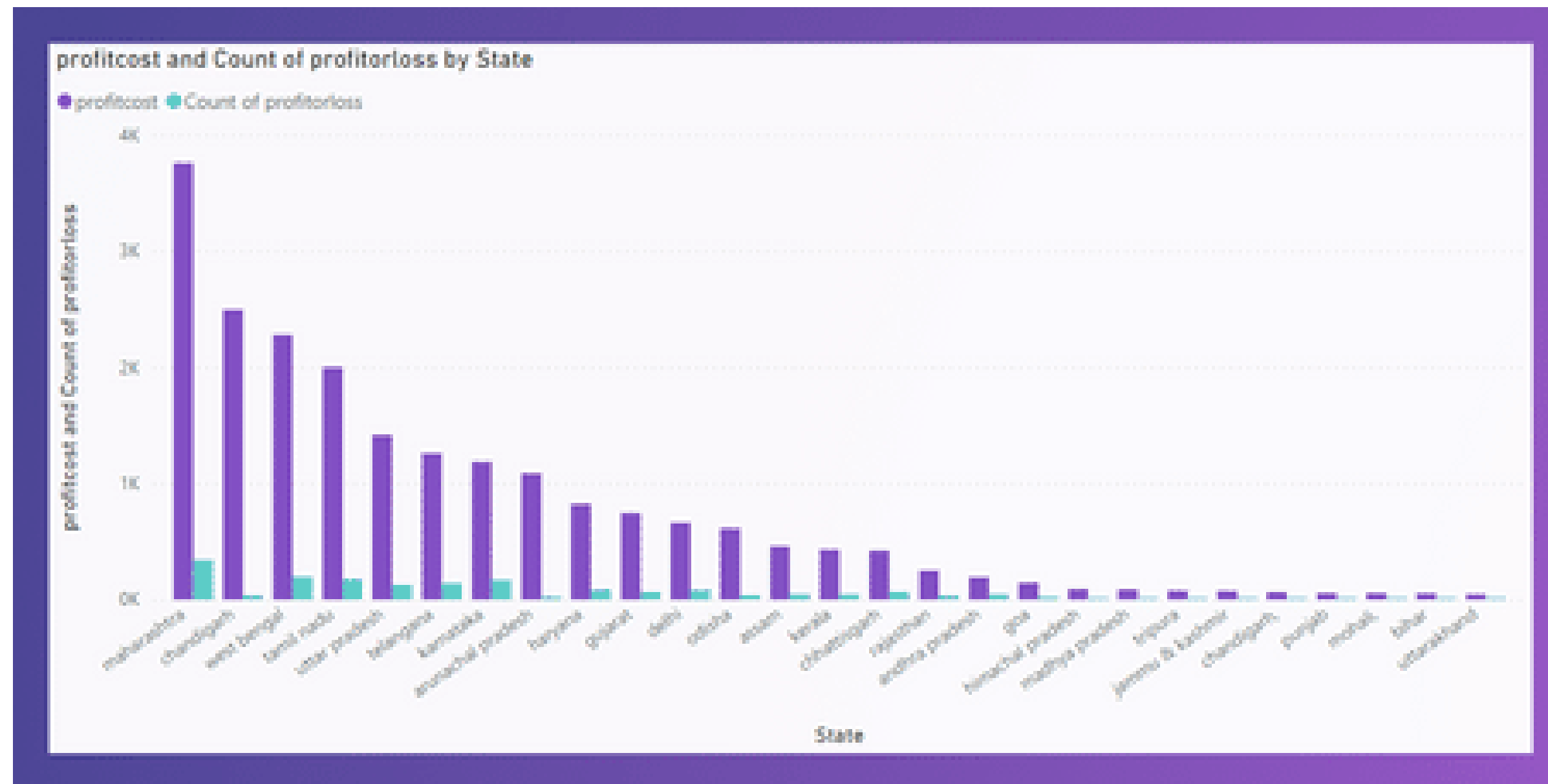
SALES COUNT BY PROFIT OR LOSS



SALES ANALYSIS DASHBOARD USING POWER BI



COUNT OF PROFIT COST AND PROFIT OR LOSS BY STATE

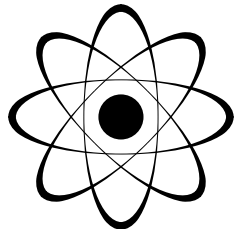


MODEL SELECTION

01 **DECISION TREE
REGRESSOR**

02 **RANDOM FOREST
REGRESSOR**

03 **ARIMA MODEL**

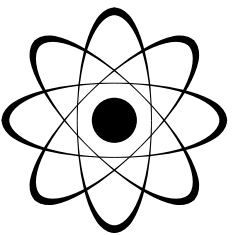


ARIMA MODEL

ARIMA is a form of regression analysis that indicates the strength of a dependent variable relative to other changing variables.

DECISION TREE REGRESSOR

It is used for regression problems where you are trying to predict something with infinite possible answers such as the price of a selling price.



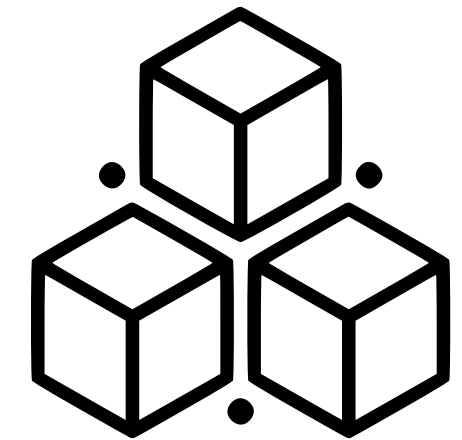
RANDOM FOREST REGRESSOR

Random Forest Regression is a supervised learning algorithm that uses ensemble learning method for regression. Ensemble learning method is a technique that combines predictions from multiple machine learning algorithms to make a more accurate prediction than a single model.



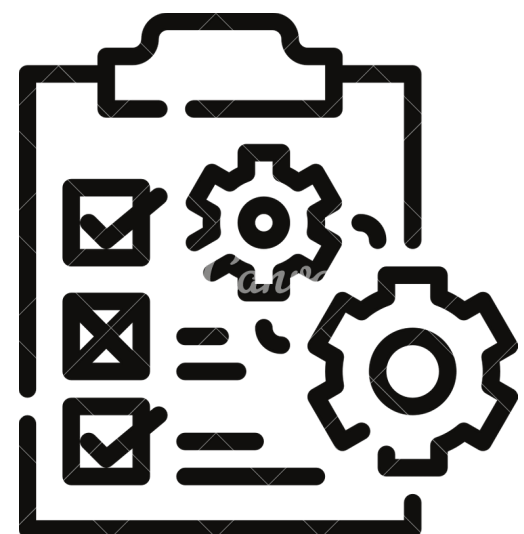
PREDICTING MODEL

1.To predict model have been r^2 score and mean squared error.



2.Time series forecasting occurs when you make scientific predictions based on historical time stamped data.

3.Making predictions about the future is called extrapolation in the classical statistical handling of time series data.



CONCLUSION:



1.To Analyze The datasets target Attribute is Profit or loss and To analyze between past 4 years "No profit No loss".

2.Major Our product Sold by Mumbai State and At uttarakhand was a sold By Less number of Products.

3.In Year 2020 Our Product are sold By More and In 2019 Product are sold By Less Numbers.Current Year Our product are sold More compare between last year.

4.The Product of Basket are More sold and Gift Basket Product are sold less.



Thank you for your time!

**Feel free to reach out to us for any
questions or concerns.**