

**Time Series Forecasting - Business Report**

**By: Aaryani Kadiyala**

**PGP-Data Science and Business Analytics**

**(PGPDSBA.O.JAN24.A)**

**Table of Contents**

1. Problem ----------------------------------------------------------------------------------------------------------------------- 5

**List of Figures:**

Figure 1: First 5 rows Dataset Sample --------------------------------------------------------------------------------------- 6

Figure 2: Last 5 rows Dataset Sample ---------------------------------------------------------------------------------------- 6

Figure 3: Data types of the dataset ------------------------------------------------------------------------------------------- 6

Figure 4: Statistical summary of the dataset ------------------------------------------------------------------------------- 7

Figure 5: Unique categories of categorical variables --------------------------------------------------------------------- 7

Figure 6: Univariant analysis of Age ------------------------------------------------------------------------------------------ 8

Figure 7: Univariant analysis of Work Experience ------------------------------------------------------------------------- 8

Figure 8: Bar Plot of the Gender ----------------------------------------------------------------------------------------------- 9

Figure 9: Bar Plot of the Transport -------------------------------------------------------------------------------------------- 9

Figure 10: Bar Plot between Transport and Gender ---------------------------------------------------------------------- 9

Figure 11: Transport and Gender in numbers ------------------------------------------------------------------------------ 9

Figure 12: Prediction with distance w.r.t Transport --------------------------------------------------------------------- 10

Figure 13: Prediction with Work Experience w.r.t Transport --------------------------------------------------------- 11

Figure 14: Prediction with Distance w.r.t Gender ------------------------------------------------------------------------ 12

Figure 15: Prediction with Age w.r.t Transport --------------------------------------------------------------------------- 13

Figure 16: Correlation of the dataset of the dataset -------------------------------------------------------------------- 13

Figure 17: Before Outlier Treatment ---------------------------------------------------------------------------------------- 14

Figure 18: After Outlier Treatment ------------------------------------------------------------------------------------------ 15

Figure 19: Shape of training & testing dataset ---------------------------------------------------------------------------- 16

Figure 20: Training Dataset --------------------------------------------------------------------------------------------------- 16

Figure 21: Testing Dataset ----------------------------------------------------------------------------------------------------- 16

Figure 22: Classification Report of Training Dataset --------------------------------------------------------------------- 16

Figure 23: Classification Report of Testing Dataset ---------------------------------------------------------------------- 16

Figure 24: Confusion Matrix of training Data Set ------------------------------------------------------------------------ 17

Figure 25: Confusion Matrix of testing Data Set -------------------------------------------------------------------------- 17

Figure 26: Classification Report of Training Dataset --------------------------------------------------------------------- 17

Figure 27: Classification Report of Testing Dataset ---------------------------------------------------------------------- 18

Figure 28: Confusion Matrix of training Data Set ------------------------------------------------------------------------ 18

Figure 29: Confusion Matrix of testing Data Set -------------------------------------------------------------------------- 18

Figure 30: Classification Report of Training Dataset --------------------------------------------------------------------- 19

Figure 31: Classification Report of Testing Dataset ---------------------------------------------------------------------- 19

Figure 32: Confusion Matrix of training Data Set ------------------------------------------------------------------------ 19

Figure 33: Confusion Matrix of testing Data Set --------------------------------------------------------------------------19

Figure 34: Classification Report of Training Dataset --------------------------------------------------------------------- 20

Figure 35: Classification Report of Testing Dataset --------------------------------------------------------------------- 20

Figure 36: Confusion Matrix of training Data Set ------------------------------------------------------------------------ 20

Figure 37: Confusion Matrix of testing Data Set -------------------------------------------------------------------------- 20

Figure 38: Classification Report of Training Dataset --------------------------------------------------------------------- 21

Figure 39: Classification Report of Testing Dataset ---------------------------------------------------------------------- 21

Figure 40: Confusion Matrix of training Data Set ------------------------------------------------------------------------ 21

Figure 41: Confusion Matrix of testing Data Set -------------------------------------------------------------------------- 21

Figure 42: Classification Report of Training Dataset --------------------------------------------------------------------- 22

Figure 43: Classification Report of Testing Dataset ---------------------------------------------------------------------- 22

Figure 44: Confusion Matrix of training Data Set ------------------------------------------------------------------------ 22

Figure 45: Confusion Matrix of testing Data Set -------------------------------------------------------------------------- 23

Figure 46 : Model Comparison for training data ------------------------------------------------------------------------- 24

Figure 47 : Model Comparison for testing data -------------------------------------------------------------------------- 24

Figure 48 : Important features of final model ---------------------------------------------------------------------------- 24

Figure 49: First 5 rows of the data ------------------------------------------------------------------------------------------- 26

Figure 50: Last 5 rows of the data ------------------------------------------------------------------------------------------- 26

Figure 51: Data types of the dataset ---------------------------------------------------------------------------------------- 26

Figure 52: Deal & Description in separate dataset ----------------------------------------------------------------------- 27

Figure 53: Deals Secured ------------------------------------------------------------------------------------------------------ 27

Figure 54: Deals Not Secured ------------------------------------------------------------------------------------------------- 27

Figure 55: No. of characters in each corpus ------------------------------------------------------------------------------- 27

Figure 56: After removing http links ---------------------------------------------------------------------------------------- 28

Figure 57: After words de-contraction of deals secured --------------------------------------------------------------- 28

Figure 58: After words de-contraction of deals not secured ---------------------------------------------------------- 28

Figure 59: After removing numbers of deals secured ------------------------------------------------------------------- 29

Figure 60: After removing numbers of deals not secured ------------------------------------------------------------- 29

Figure 61: After tokenizing of deals secured ------------------------------------------------------------------------------ 29

Figure 62: After tokenizing of deals not secured ------------------------------------------------------------------------- 29

Figure 63: After converting into lower case of deals secured --------------------------------------------------------- 30

Figure 64: After converting into lower case of deals not secured --------------------------------------------------- 30

Figure 65: After removing punctuations of deals secured ------------------------------------------------------------- 30

Figure 66: After removing punctuations of deals not secured -------------------------------------------------------- 30

Figure 67: After removing stopwords of deals secured ---------------------------------------------------------------- 31

Figure 68: After removing stopwords of deals not secured ----------------------------------------------------------- 31

Figure 69: After lemmatization of deals secured ------------------------------------------------------------------------- 31

Figure 70: After Normalization of deals secured ------------------------------------------------------------------------- 31

Figure 71: Wordcloud of deals secured ------------------------------------------------------------------------------------ 32

Figure 72: Wordcloud of deals not secured ------------------------------------------------------------------------------- 33

**Problem :**

### **Context**

### As an analyst at ABC Estate Wines, we are presented with historical data encompassing the sales of different types of wines throughout the 20th century. These datasets originate from the same company but represent sales figures for distinct wine varieties. Our objective is to delve into the data, analyze trends, patterns, and factors influencing wine sales over the course of the century. By leveraging data analytics and forecasting techniques, we aim to gain actionable insights that can inform strategic decision-making and optimize sales strategies for the future.

### **Objective**

### The primary objective of this project is to analyze and forecast wine sales trends for the 20th century based on historical data provided by ABC Estate Wines. We aim to equip ABC Estate Wines with the necessary insights and foresight to enhance sales performance, capitalize on emerging market opportunities, and maintain a competitive edge in the wine industry.

**Solution:-**