

# **PROJECT REPORT**

## **DIWALI SALES ANALYSIS**

**Submitted by**

**NAME: ANUSHREE**

**ROLL NUMBER: RK21UWB65**

**REGISTRATION NUMBER: 12113879**

**SECTION: K21UW**

**COURSE CODE –  
INT353**

**Under the Guidance of**

**(Shivangini Gupta)**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING  
LOVELY PROFESSIONAL UNIVERSITY,  
JALANDHAR, PUNJAB, INDIA**



**L** OVELY  
**P** ROFESSIONAL  
**U** NIVERSITY

## **TABLE OF CONTENTS**

EXPLORATORY DATA ANALYSIS PROJECT REPORT .....	(i)
TABLE OF CONTENTS .....	(ii)
INTRODUCTION .....	(iii)
DOMAIN KNOWLEDGE .....	(iv)
WHY I CHOOSE THIS DATASET... ..	(v)
LIBRARIES USED AND APPROACHES .....	(vi)
DATA DESCRIPTION.....	(vii)
DATA CLEANING .....	(viii)
DATA EXPLORATION... ..	(ix)
UNIVARIATE ANALYSIS.....	(x)
BIVARIATE ANALYSIS.....	(xi)
MULTIVARIATE ANALYSIS... ..	(xii)
HYPOTHESIS TESTING.....	(xiii)
FINDINGS AND INSIGHTS.....	(xiv)
LIMITATIONS.....	(xv)
RECOMMENDATIONS.....	(xvi)
CONCLUSION.....	(xvii)
REFERENCES.....	(xviii)
ACKNOWLEDGMENT.....	(xix)

## **INTRODUCTION**

Diwali, also known as the Festival of Lights, is one of the most celebrated and widely observed festivals in India and other parts of the world. Beyond its cultural and religious significance, Diwali also marks a significant period for businesses, with a surge in consumer spending during the festive season. The Diwali Sales Analysis Exploratory Data Analysis (EDA) Project aims to shed light on the trends, patterns, and insights within the sales data during this festive period.

As businesses gear up for Diwali, the project focuses on leveraging data analytics techniques to analyze sales data and derive actionable insights. The analysis will delve into various aspects, including product preferences, geographical variations, purchasing behaviors, and the impact of marketing strategies on sales performance. By examining historical sales data, the project aims to uncover hidden patterns and correlations, providing valuable information for businesses to optimize their strategies for future Diwali seasons.

The primary focus of this Diwali Sales Analysis EDA Project is to unravel invaluable insights from the dataset, honing in on key aspects such as pinpointing geographical regions where sales are at their peak. By scrutinizing the data, we aim to discern not only the geographic hotspots but also the specific products that garner heightened consumer interest. Additionally, a nuanced exploration into the gender-based purchasing patterns will be conducted to identify which demographic contributes significantly to the festive sales.

The EDA process involves exploring, cleaning, and visualizing the data to uncover meaningful trends. Various statistical and machine learning techniques will be employed to extract actionable insights from the dataset. The project's findings will not only aid businesses in refining their Diwali marketing strategies but also contribute to a deeper understanding of consumer behavior during festive seasons.

In summary, the Diwali Sales Analysis EDA Project serves as a beacon, illuminating the nuances of consumer behavior and market dynamics during the festive season. Through a thorough exploration of sales data, businesses can gain valuable insights to enhance their decision-making processes and capitalize on the opportunities presented by the vibrant Diwali market.

## **DOMAIN KNOWLEDGE**

The domain of "Diwali Sales Analysis" resides at the intersection of business analytics, marketing strategy, and cultural festivities. Diwali, or the Festival of Lights, is a major cultural and religious celebration in India and various other parts of the world. Beyond its traditional significance, Diwali has evolved into a significant economic event, marked by increased consumer spending and vibrant market activities.

Understanding the dynamics of Diwali sales requires a grasp of both cultural aspects and business strategies. Businesses often witness a surge in sales during the Diwali season due to increased consumer purchasing power, gifting traditions, and the festive spirit that encourages people to indulge in shopping. Consequently, companies strategize to capitalize on this festive fervor by launching special Diwali-themed marketing campaigns, promotions, and product offerings.

Analyzing Diwali sales data involves exploring patterns, trends, and consumer behaviors specific to this festive period. Geographical variations in sales, popular product categories, and demographic preferences become critical factors in shaping marketing and sales strategies. Additionally, understanding the impact of cultural elements, such as traditional customs and rituals associated with Diwali, is vital for businesses aiming to create resonant and culturally relevant marketing campaigns.

In summary, the domain of Diwali Sales Analysis requires a nuanced understanding of the cultural significance of Diwali, coupled with a keen awareness of business strategies, market dynamics, and consumer behavior during festive seasons. This interdisciplinary approach is crucial for extracting meaningful insights from sales data and guiding businesses in making informed decisions to enhance their Diwali-specific marketing and sales endeavors.

## **WHY I CHOOSE THIS DATASET**

Choosing the Diwali Sales dataset for analysis offers several compelling reasons:

- 1. Cultural and Economic Significance:** Diwali is a major cultural and religious festival in various regions, especially in India. It holds immense significance, not only in the lives of individuals but also in the economic landscape. Analyzing Diwali sales data provides insights into consumer behavior during a period of heightened spending and festivities.
- 2. Seasonal Variation:** Diwali is characterized by a noticeable surge in consumer activity, making it an ideal period for businesses to boost their sales. Studying the dataset allows you to understand how consumer behavior changes during this festive season, helping businesses tailor their strategies accordingly.
- 3. Marketing Opportunities:** Diwali presents businesses with unique marketing opportunities. By delving into the dataset, you can identify successful marketing strategies, popular product categories, and effective promotional campaigns that have historically resonated with consumers during this festive period.
- 4. Geographical Insights:** The dataset can offer valuable information about regional variations in sales during Diwali. Understanding which areas experience the highest sales can assist businesses in targeting specific regions more effectively and customizing their marketing approaches accordingly.
- 5. Target Customer Identification:** Analyzing gender-based and demographic preferences during Diwali can aid in identifying target customer segments. This knowledge is crucial for businesses looking to personalize their offerings and marketing messages to better connect with their audience.
- 6. Strategic Decision-Making:** The insights derived from the dataset can guide businesses in making informed strategic decisions for future Diwali seasons. Whether it's optimizing product offerings, adjusting pricing strategies, or refining marketing campaigns, the data can serve as a valuable resource for planning and decision-making.

In essence, choosing the Diwali Sales dataset aligns with the goal of understanding and leveraging the unique dynamics of the festive season, providing a rich source of information for businesses seeking to enhance their market presence and capitalize on the opportunities presented during this culturally significant period.

## **LIBRARIES USED AND APPROACHES**

In the Diwali Sales Analysis EDA project, several key libraries are utilized for data manipulation, analysis, and visualization. The primary libraries include:

**1. Pandas:** Used for handling and organizing structured data, Pandas facilitates tasks such as loading datasets, data cleaning, and basic exploratory data analysis.

**2. NumPy:** Essential for numerical operations and computations, NumPy is employed to work with numerical data efficiently, providing support for arrays and mathematical functions.

**3. Matplotlib:** This library is instrumental for creating a diverse range of static visualizations, enabling the representation of trends and patterns in the Diwali sales dataset.

**4. Seaborn:** Built on top of Matplotlib, Seaborn offers a high-level interface for generating aesthetically pleasing statistical graphics. It is employed to enhance the visual interpretation of the data.

**5. SciPy:** This library was used for hypothesis testing.

### **Approaches:**

#### **1. Data Handling and Cleaning:**

- Pandas is used for loading the dataset and addressing any data cleaning tasks, ensuring that the dataset is prepared for analysis by handling missing values and outliers.

#### **2. Descriptive Analysis:**

- Basic statistical summaries are generated using Pandas to understand the central tendencies and distributions of the Diwali sales data.

#### **3. Data Visualization:**

- Matplotlib and Seaborn are employed for creating visualizations such as line plots, bar charts, and scatter plots. These visualizations help in revealing insights and patterns in the dataset.

#### **4. Correlation Analysis:**

- Seaborn is utilized to generate correlation matrices and heatmaps, allowing for the identification of relationships between different variables in the dataset.

## **DATA DESCRIPTION**

### **Data Description:**

1. **User\_ID:** A unique identifier assigned to each user.
2. **Cust\_name:** The name of the customer.
3. **Product\_ID:** A unique identifier for each product.
4. **Gender:** The gender of the customer (e.g., Male, Female).
5. **Age\_Group:** The age group to which the customer belongs.
6. **Age:** The specific age of the customer.
7. **Marital\_Status:** Indicates whether the customer is married or not.
8. **State:** The state in which the customer resides.
9. **Zone:** The geographical zone to which the state belongs.
10. **Occupation:** The occupation of the customer.
11. **Product\_Category:** The category to which the product belongs.
12. **Orders:** The number of orders placed by the customer.
13. **Amount:** The total amount spent by the customer.

### **Source of the Data:**

The dataset was downloaded from Kaggle, a popular platform for data science competitions and datasets. Kaggle provides a diverse range of datasets contributed by the community and data providers, fostering a collaborative environment for data analysis and machine learning projects.

### **Data Preprocessing:**

1. **Handling Missing Values:** Checking for and addressing any missing values in the dataset to ensure completeness.
2. **Data Type Validation:** Verifying and correcting data types to ensure that each variable is appropriately represented (e.g., numeric, categorical).
3. **Outlier Detection:** Identifying and handling any outliers that may affect the analysis and visualizations.
4. **Data Transformation:** Converting and transforming data as needed for analysis, such as encoding categorical variables or scaling numeric features.
5. **Exploratory Data Analysis (EDA):** Conducting an initial exploration of the dataset to understand the distribution of variables, identify patterns, and gain insights into potential relationship

## **DATA CLEANING**

Certainly, here is an expanded explanation of the data cleaning steps, incorporating the additional information you provided:

### **1. Handling Missing Values:**

- Identified columns with missing values, including 'Status' and 'Unnamed1,' and recognized that these columns contained null values.
- Decided to drop the 'Status' and 'Unnamed1' columns due to the presence of null values, as these columns were not deemed essential for the analysis.
- Furthermore, dropped the rows containing null values in the 'Amount' column

### **2. Outlier Detection and Treatment:**

- Conducted an outlier analysis on relevant numerical columns such as 'Age' to identify data points significantly deviating from the typical distribution.
- Implemented outlier treatment techniques, such as IQR(Inter Quartile Range), to remove the outliers.

### **3. Data Type Validation and Correction:**

- Validated the data types of each variable to ensure correct representation.
- Identified a column, potentially 'Amount,' where the data type required modification.
- Corrected the data type of the 'Amount' column to align it with the appropriate numeric format for analysis.

### **4. Data Transformation:**

- Transformed variables if necessary for analysis, including converting categorical variables into a suitable format.
- Altered the name of a column as per project requirements.

The data cleaning process also involved a strategic decision to drop specific columns, change the data type of a column, and rename a column. These decisions were made to enhance the dataset's suitability for analysis and to address specific requirements outlined in the Diwali Sales Analysis EDA project.



## **DATA EXPLORATION**

### **Initial Summary Statistics**

#### Overview:

- We initiated the analysis by loading the dataset into a Pandas DataFrame.
- A thorough review of the dataset structure, consisting of the initial and final rows, along with the dataset's shape and size, provided a comprehensive understanding of the data's breadth and scope.
- We examined column details, scrutinizing their names, data types, and non-null counts, shedding light on the dataset's composition.

#### Handling Data Integrity:

- Our data cleaning process involved the identification and removal of missing values and duplicates.
- A conscious decision was made to drop rows with null values to ensure data completeness and accuracy.

### **Trends, Patterns, and Observations:**

- **Gender-Based Analysis:** Visual exploration into the gender demographics and their purchasing habits exposed a significant trend. Females constituted the majority of buyers, demonstrating a higher purchasing power compared to males.
- **Age Group Analysis:** A detailed exploration into different age groups revealed that the age group of 26-35, particularly females, dominated the purchasing landscape.
- **State-Wise Analysis:** Geographic distribution was explored, uncovering top states for orders and purchases. This highlighted consumer concentration and spending patterns in various states.
- **Zone-Based Analysis:** Understanding consumer distribution across different zones illuminated insights into consumer behavior and spending patterns in distinct geographical areas.

## **UNIVARIATE ANALYSIS**

### **Q1. WHAT IS THE NUMBER OF PURCHASE DONE BY DIFFERENT GENDER?**

Through the analysis we have found out that the number purchase done by females is '7832' and that by male is '3407'. This clearly show that most of the buyers are females.

### **Q2. WHICH AGE GROUP HAS HEIGHEST NUMBER OF PURCHASES?**

Through the visualization we got to know the maximum no of buyers are from the age group of 26-35.

### **Q3. WHAT IS THE DISTRIBUTION OF CUSTOMERS ACROSS DIFFERENT STATES?**

This visualization show that the maximum no of costumers are in Uttar Pradesh followed by Maharashtra and Karnataka.

### **Q4. WHAT IS THE DISTRIBUTION OF CUSTOMERS ACROSS DIFFERENT ZONES?**

This visualization shows how many customers are there in different zones. And we through this we found that the central zone in having highest no of customers.

### **Q5. WHAT IS THE DISTRIBUTION OF CUSTOMERS' MARITAL STATUS IN THE DATASET**

Through this visualization we got to know that the maximum no of customers are married

### **Q6. WHAT IS THE NUMBER OF PURCHASE MADE BY DIFFERENT OCCUPATION**

This visualization shows that the maximum no of purchase are made by the people of IT sector followed by that of healthcare and aviation

### **Q7. WHAT IS THE DISTRIBUTION OF PRODUCTS ACROSS DIFFERENT PRODUCT CATEGORIES?**

This visualization shows which product category has how much sale and we can see that the clothing and 'Apparel and product' category has height sales followed by 'Food' and 'Electronic and Gadgets'.

### **Q8. Which 'Product\_ID' is the most popular or has the highest number of purchases in the dataset?**

Through the analysis we found out that the most popular Product\_ID is: P0026524

## **BIVARIATE ANALYSIS**

### **Q9. WHAT IS THE AMOUNT SPEND BY DIFFERNT GENDERS ON PURCHACHING?**

Through this visualization we got to know that females have spent more money on purchasing than man

### **Q10. HOW DOES THE DISTRIBUTION OF 'AGE GROUP' VARY BETWEEN DIFFERENT GENDERS?**

This visualization shows us how the distribution of 'Age Group' varies between different genders and from the visualization we can clearly observe that the females in the age group of 26-35 are having the highest no of purchase.

### **Q11. WHAT IS THE AMOUNT SPEND BY DIFFERNT AGE GROUP ON PURCHACHING?**

This visualization shows about the amount spend by different age group and we can clearly see that the maximum amount is spend by the people in the age group of 26-35.

### **Q12 WHICH STATES HAVE THE HIGHEST NUMBER OF ORDERS?**

This visualization show which state has maximum no of orders and we can clearly, we that Uttar Pradesh has maximum no of orders.

### **Q13 WHAT IS THE AMOUNT SPEND BY DIFFERNT STATES ON PURCHACHING?**

This visualization show which state has spent how much money on purchasing and we can clearly, we that Uttar Pradesh has spent maximum amount on purchasing. But here we also get to know that although the number of order is more in Kerala but the amount spent on these orders is very less.

### **Q14. WHAT IS THE AMOUNT SPEND BY PEOPLE OF DIFFERNT ZONES ON PURCHACHING?**

Through this visualization we get to know that the customers of central zone has spent the maximum money on purchasing

### **Q15. WHICH COMBINATIONS OF MARITAL STATUS AND GENDER CONTRIBUTE THE MOST TO THE TOTAL SALES AMOUNTS?**

Through this visualization we got to know that the maximum no of customers are married and are females

### **Q16. WHAT IS THE AMOUNT SPEND BY DIFFERNT OCCUPATIONS ON PURCHACHING?**

This visualization shows that the maximum amount spent on purchase is by the people of IT sector followed by that of healthcare and aviation

### **Q17 WHICH PRODUCT CATEGORIES CONTRIBUTE THE MOST TO THE TOTAL SALES AMOUNTS?**

This visualization shows which product category maximum amount has been spent. And out that although the number of purchases are highest in the 'Apparel and product' category but the amount spent is more in the 'Food' category

## **MULTIVARIATE ANALYSIS**

### **Q18.WHAT IS THE DISTRIBUTION OF AVERAGE SPENDING ACROSS VARIOUS PRODUCT CATEGORIES IN DIFFERENT ZONES?**

The 'Auto' category exhibits consistently high average spending across all zones, suggesting that it is a popular and high-spending product category regardless of the customer's zone. This indicates that 'Auto' products have a broad customer appeal and strong demand in all regions.

### **Q19. WHAT CORRELATIONS CAN BE OBSERVED AMONG THE NUMERICAL COLUMNS 'AGE,' 'ORDERS,' AND 'AMOUNT' IN THE DATASET**

Through this visualization we got to know that there exists a correlation between the 'Age' and 'Amount' columns.

### **Q20. WHAT IS THE OVERALL CORRELATION AMONG THE NUMERICAL COLUMNS IN THE DATASET?**

This visualization shows the correlations among the numerical columns of the data set

## **HYPOTHESIS TESTING**

In the analysis of our dataset, we conducted a correlation hypothesis test to investigate whether there is a significant correlation between the variables "Age" and "Amount." This test aimed to determine whether there is a linear association between the age of customers and the amount they spent on products. The null and alternative hypotheses were defined as follows:

**Null Hypothesis (H0):** - There is no correlation between the two variables "Age" and "Amount." In statistical terms, the correlation coefficient is equal to zero.

**Alternative Hypothesis (Ha):** - There is a significant correlation between the two variables "Age" and "Amount."

### **Test Results: -**

Upon performing the correlation hypothesis test, we obtained the following results:

Pearson Correlation Coefficient: 0.0309

P-Value: 0.0010

The Pearson Correlation Coefficient (r) measures the strength and direction of the linear relationship between "Age" and "Amount." In this case, the coefficient was calculated to be **0.0309**, indicating a positive but very weak linear correlation.

The P-Value represents the probability of observing a correlation as extreme as the one calculated, assuming the null hypothesis is true. In this test, the p-value was found to be 0.0010, which is significantly smaller than the commonly chosen significance level (alpha) of 0.05.

### **Hypothesis Conclusion**

Since the p-value (0.0010) is less than the chosen significance level (alpha), we reject the null hypothesis (H0). This means that there is a statistically significant correlation between the age of customers and the amount they spent on products.

.

## **FINDINGS AND INSIGHTS**

After a comprehensive analysis of the Diwali sales dataset, several key findings and insights have emerged, shedding light on distinct patterns and preferences among consumers. The primary focus centers on married women aged 26-35 years residing in Uttar Pradesh (UP), Maharashtra, and Karnataka, engaged in occupations within the IT, Healthcare, and Aviation sectors.

### **1. Demographic Insights:**

- Marital Status: Married women exhibit a notable inclination toward Diwali purchases, suggesting that this demographic group actively participates in festive shopping.
- Age Group: The age group of 26-35 years appears to be a significant segment, indicating a prime period for heightened consumer engagement during the Diwali season.

### **2. Geographical Trends:**

- Regional Preferences: Consumers from Uttar Pradesh, Maharashtra, and Karnataka emerge as key contributors to Diwali sales, suggesting potential regional trends or cultural influences on purchasing behavior.
- Varied Regional Impact: The dataset implies that these regions significantly contribute to the overall sales, emphasizing the importance of tailoring marketing strategies to regional preferences and customs.

### **3. Occupational Patterns:**

- Occupational Influence: Women employed in the Information Technology (IT), Healthcare, and Aviation sectors demonstrate a higher likelihood of engaging in Diwali purchases. This insight unveils the impact of occupation on consumer behavior during the festive season.
- Diverse Professional Backgrounds: The diverse occupational backgrounds suggest that Diwali shopping is not limited to specific industries, highlighting the universality of festive spending across various professional sectors.

### **4. Product Category Preferences:**

- Preferred Product Categories: The analysis indicates a distinct preference for products falling within the Food, Clothing, and Electronics categories among the identified demographic group.
- Multifaceted Consumer Interests: The diversity of product categories suggests that these consumers are not limited to specific types of products, showcasing a multifaceted interest in various goods during the Diwali period.

## **LIMITATIONS**

1. Data Quality Issues: The dataset had instances of inaccurate, incomplete, and noisy data, which posed challenges in ensuring data quality and drawing meaningful insights.
2. Data Bias: The data exhibited potential bias, particularly in terms of representation. It didn't fully reflect the diversity of the target population, possibly due to limitations in the data sources or collection methods.
3. Data Collection Challenges: The data collection methods used may have introduced biases, making it difficult to provide a completely objective analysis. Additionally, a lack of transparency in data collection procedures limited our understanding of the data's origin.
4. Limited Variables: The analysis was constrained by the available variables in the dataset. Not having access to additional variables limited the scope of the analysis.
7. Interpretation Challenges: The analysis was based on patterns and correlations within the data, but causation could not be definitively established. This posed challenges in interpreting the findings accurately.
8. Scope Constraints: The project was constrained to the available dataset and the research questions at hand. Due to project limitations, certain variables or research questions couldn't be explored in-depth.
9. Data Privacy and Security: Adhering to data privacy and security standards was a critical concern, especially when dealing with sensitive information.
10. Time Constraints: The project had time constraints that limited the depth and thoroughness of the analysis. Some areas of investigation remained unexplored due to these time limitations.

## **RECOMMENDATIONS**

Building on the insights derived from the Diwali Sales Analysis, several actionable recommendations can be proposed for businesses aiming to optimize their strategies for the festive season:

### **1. Targeted Marketing:**

- Action: Tailor campaigns for married women aged 26-35 in UP, Maharashtra, and Karnataka.
- Why: Directly address preferences of a key consumer group for increased campaign effectiveness.

### **2. Regional Customization:**

- Action: Customize strategies based on the significant contribution of UP, Maharashtra, and Karnataka to Diwali sales.
- Why: Leverage regional variations for tailored promotions and offerings.

### **3. Occupation-Centric Offers**

- Action: Collaborate with IT, Healthcare, and Aviation sectors for occupation-centric promotions.
- Why: Tailor promotions to the higher likelihood of purchases from these professional sectors.

### **4. Consumer Engagement Initiatives:**

- Action: Implement initiatives to actively engage married women aged 26-35.
- Why: Recognize the significance of this demographic and enhance customer engagement through targeted initiatives.

### **5. Continuous Monitoring:**

- Action: Regularly monitor and analyze sales data during subsequent Diwali seasons.
- Why: Adapt strategies to changing trends and preferences for sustained success.

### **6. Customer Feedback Mechanism:**

- Action: Implement a feedback mechanism to gather insights on product satisfaction.
- Why: Direct feedback guides product refinement and strategy improvement based on customer preferences.



## **CONCLUSION**

In summary, the exploratory data analysis of Diwali sales data has uncovered valuable insights into consumer behavior and market dynamics. The analysis highlighted the significance of married women aged 26-35 as active participants in Diwali purchases. Regional trends emphasized the importance of tailoring strategies for key states like Uttar Pradesh, Maharashtra, and Karnataka. The influence of occupation, particularly in IT, Healthcare, and Aviation sectors, on purchasing behavior was notable. Consumer preferences spanning Food, Clothing, and Electronics categories showcased the multifaceted interests of the target demographic.

The findings suggest opportunities for businesses to refine marketing approaches, customize strategies for specific regions, and offer diverse product assortments. Continuous monitoring and adaptation to evolving trends are crucial for sustained success in the dynamic Diwali sales market. The analysis sets the stage for future insights with potential data augmentation, including variables such as customer feedback and promotional campaign data. Overall, the EDA serves as a valuable tool for businesses seeking to optimize their strategies during the Diwali season.

## **REFERENCES**

### 1. Dataset:

- Specify the source of the Diwali sales dataset, including details such as the platform (e.g., Kaggle) and the dataset name or URL.

### 2. Libraries:

- Mention the Python libraries used in your project, such as Pandas, NumPy, Matplotlib, and Seaborn. You can provide the versions used if deemed necessary.

### 3. Analysis Techniques:

- If you applied specific analysis techniques or methodologies inspired by external sources, it's good practice to reference those materials or publications.

### 4. Documentation and Tutorials:

- If you followed any specific documentation, tutorials, or online resources for coding or analysis, acknowledge those sources.

## **ACKNOWLEDGMENT**

I would like to extend my sincere gratitude to all those who have been instrumental in the successful completion of this Diwali Sales Analysis EDA project.

The dataset used for this analysis was sourced from Kaggle, and I want to express my appreciation for the collaborative environment that Kaggle provides for data scientists, making diverse datasets accessible for analysis.

The project heavily relied on the use of open-source Python libraries, including Pandas, NumPy, Matplotlib, and Seaborn. These libraries played a pivotal role in facilitating data manipulation, analysis, and visualization, contributing significantly to the project's success.

I would like to express my heartfelt thanks to everyone who has been a part of this journey. Their active guidance, help, cooperation, and encouragement were indispensable, ensuring the timely completion and presentation of the project.

A special note of appreciation goes to my teacher, Miss Shivangini Gupta, for her consistent support and guidance throughout the project's duration. I am also thankful to all the faculty and staff members of our department for their kind cooperation and assistance.

My gratitude extends to my family and friends for their unwavering support and encouragement. Their motivational words have been a driving force, keeping me focused on completing the project to the best of my abilities.

I would also like to acknowledge my classmates for providing moral support and encouragement, fostering a positive collaborative environment.

In conclusion, I appreciate the collective efforts of everyone involved, directly or indirectly, in contributing to this project. I hope that this Diwali Sales Analysis will serve as a valuable resource for those who engage with it.

Link for inynb file: -

[https://drive.google.com/file/d/1EF30LnGeQGRHUZJo\\_WPE7n9w3cOzfWig/view?usp=drive\\_link](https://drive.google.com/file/d/1EF30LnGeQGRHUZJo_WPE7n9w3cOzfWig/view?usp=drive_link)

link for ppt:

[https://docs.google.com/presentation/d/1DsuiJbtWk1LNs\\_cvRdbTd8V3\\_xhh7H AH/edit?usp=drive\\_link&ouid=112547457027320482644&rtpof=true&sd=true](https://docs.google.com/presentation/d/1DsuiJbtWk1LNs_cvRdbTd8V3_xhh7H AH/edit?usp=drive_link&ouid=112547457027320482644&rtpof=true&sd=true)