1. **INTRODUCTION**

The **Diwali festival**, celebrated with grandeur across India, is not only a cultural and religious occasion but also a **major retail event** that drives significant economic activity. For retailers, this festive season represents an opportunity to maximize sales, attract new customers, and strengthen brand loyalty. During Diwali, consumers engage in high-volume purchases across various sectors, ranging from **electronics** and **fashion** to **jewelry**, **home decor**, and **groceries**. The influx of shoppers and the diverse range of products on offer creates a dynamic and complex retail environment, where understanding consumer behavior becomes paramount.

As the **e-commerce** industry continues to grow, driven by increased internet penetration and digital payment adoption, businesses are increasingly relying on **data-driven decision-making** to remain competitive. The sheer volume of data generated during Diwali sales—covering transaction details, customer demographics, product preferences, and regional buying patterns—can be overwhelming. However, the **effective analysis** of this data offers businesses the insights necessary to tailor their marketing strategies, optimize inventory management, and maximize customer satisfaction.

To tackle this challenge, the **Diwali Sales Dashboard** has been developed as an interactive and intuitive analytics platform aimed at providing retailers and analysts with a comprehensive tool for visualizing and understanding Diwali sales data. This dashboard is designed to simplify the process of gaining insights from the massive datasets that are generated during the festive period. By leveraging cutting-edge technologies such as **Python**, **Pandas**, **Seaborn**, and **Streamlit**, the dashboard allows users to perform deep analysis of Diwali sales data with minimal technical expertise.

The core idea behind this project is to create an accessible, **user-friendly platform** that can be used by anyone—whether they are small retailers, regional analysts, or large e-commerce platforms—to gain a clear understanding of **consumer spending patterns**. The ability to filter and segment data by factors such as **gender**, **age group**, **region**, **occupation**, and **product category** is crucial in uncovering insights that would otherwise be hidden in a sea of numbers.

Furthermore, this dashboard empowers retailers to **identify key trends** in real-time, such as the most popular product categories, the highest revenue-generating regions, and the customer segments that are most likely to engage with certain products. By offering **dynamic filtering options**, the tool enables businesses to analyze sales performance at a granular level, providing actionable insights that can drive targeted marketing efforts and **personalized campaigns**.

In essence, the **Diwali Sales Dashboard** transforms raw transactional data into clear, visual insights, helping businesses understand the **who, what, where, and when** of their Diwali sales. With the ability to explore sales trends and consumer behavior, businesses can make more informed decisions, enhance customer experience, and, ultimately, increase their sales performance during the Diwali season.

1. **LITERATURE SURVEY**

The use of business intelligence (BI) tools for data visualization and decision-making has seen widespread adoption across industries. Retail businesses, in particular, benefit from BI dashboards by transforming transactional data into actionable insights that enhance their strategic planning. A study by Gartner shows that businesses utilizing advanced data analytics tools have higher profitability and competitive advantage. Retailers who adopt these solutions can gain insights into consumer purchasing trends, inventory management, and marketing performance, resulting in improved profitability.

In the context of Diwali retail sales, seasonal trends are of immense significance, as sales during Diwali can contribute up to 40% of annual revenue for many businesses. Understanding these trends at a granular level—across demographics, product categories, and regions—is essential for optimizing product placement, personalized marketing, and targeted promotions.

However, the current retail industry lacks easily accessible and flexible tools to visualize and analyze data for specific events such as Diwali. While some major corporations rely on commercial BI tools such as Tableau and Power BI, these solutions are often expensive and require significant expertise. There is a growing demand for open-source, user-friendly solutions like the Diwali Sales Dashboard, which is designed to provide retailers with deep insights without the complexity of high-end BI platforms.

Existing research indicates that dashboards in retail analytics help businesses answer key questions, such as:

* Which products are performing best during the Diwali period?
* What demographic groups are spending the most?
* Which regions or states are contributing the highest revenue?
* How can businesses optimize their marketing campaigns based on customer segmentation?

The Diwali Sales Dashboard stands as a response to this need, providing a lightweight, customizable alternative to more expensive, commercial solution

1. **PROBLEM STATEMENT**

Retailers and analysts face several challenges when trying to analyze **Diwali sales data**, which can be complex and voluminous. These challenges include:

* **Lack of Interactive Filtering**: Traditional reports and dashboards may offer insights but often lack the ability to **dynamically filter data** across multiple variables. For example, businesses may want to compare sales across regions or age groups but are forced to manually adjust reports or perform cumbersome data manipulation.
* **Static Reporting Systems**: Many reporting tools are static, meaning they are not updated in real-time, or require manual interventions to adjust for new data. This results in outdated information that may not reflect the current sales trends.
* **Difficulty in Segmenting Customer Behavior**: Understanding consumer behavior during Diwali requires detailed segmentation, which can be difficult to achieve without the proper tools. Retailers often struggle to identify which **demographic groups** (such as age, gender, or occupation) are driving the most sales, making it challenging to tailor marketing and sales strategies effectively.
* **Underutilization of Data**: Many businesses collect vast amounts of transactional data but lack the tools or expertise to analyze it in a meaningful way. Without effective visualization, businesses miss out on identifying hidden insights that could drive improved decision-making.

**Core Problem**: The Diwali Sales Dashboard aims to provide a **solution to these challenges** by offering a **dynamic, user-friendly platform** that enables businesses to analyze sales data in real-time and across multiple dimensions, making it easier to uncover valuable insights.

**4. OBJECTIVE**

The key objectives of the **Diwali Sales Dashboard** are:

* **Dynamic Filtering**: Enable users to apply real-time filters based on key attributes such as **gender**, **region**, **age group**, **occupation**, and **product category**. This dynamic filtering provides a granular view of sales performance, helping businesses gain deeper insights into their customer base.
* **KPIs and Metrics**: Display critical **business metrics**, including **total revenue**, **average transaction value**, **unique customers**, and **category-specific performance**. These KPIs offer a quick overview of the retail performance during Diwali.
* **Customer Segmentation**: Allow businesses to explore the performance of different **customer segments** based on demographics and regions, identifying the most profitable groups for targeted marketing efforts.
* **Interactive Visualizations**: Provide **charts and graphs** that visually represent sales performance, trends, and customer behavior. These visualizations help users quickly grasp insights without delving into complex raw data.
* **Real-time Data Analysis**: Ensure that the dashboard updates dynamically as users apply filters, providing **real-time insights** that reflect current sales trends during the Diwali period.
* **Ease of Use**: Design the dashboard to be **intuitive and accessible** for non-technical users, eliminating the need for coding knowledge or advanced data analysis skills.

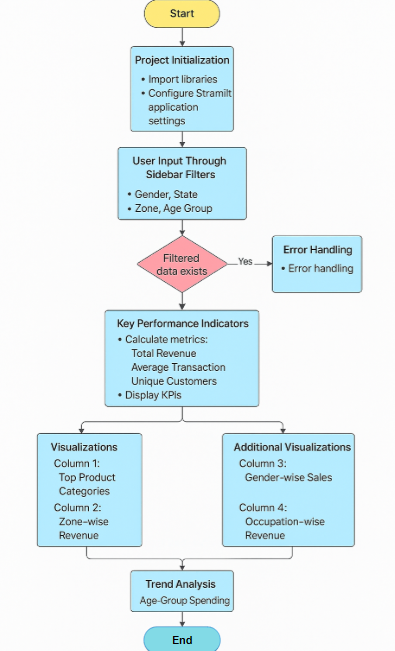
**5. PROPOSED SOLUTION**

The **Diwali Sales Dashboard** offers a comprehensive solution that addresses the challenges outlined above. Key features of the dashboard include:

* **Streamlit**: This framework allows for the rapid development of **interactive web applications** in Python, making it the ideal choice for building a responsive and user-friendly dashboard. Streamlit is designed for data scientists and analysts, making it easy to deploy Python code as an interactive web app with minimal effort.
* **Pandas**: Pandas is used to load, manipulate, and clean the **Diwali sales dataset**, which contains over **11,000 rows** of transactional data. With Pandas, data can be efficiently processed, aggregated, and transformed into the necessary format for analysis and visualization.
* **Seaborn and Matplotlib**: These Python libraries are leveraged to generate **dynamic and visually appealing charts**, including **bar graphs**, **pie charts**, and **KPI cards**. They are optimized for statistical data visualization and integrate seamlessly with **Streamlit**.
* **CSV Dataset**: The dataset used in this project contains key columns such as:
* **Gender**: Male, Female
* **Age Group**: Various age ranges (e.g., 18-25, 26-35, etc.)
* **Product Category**: Clothing, Electronics, Groceries, etc.
* **State and Region**: Geographical information, useful for regional sales analysis.
* **Orders and Amount**: Transaction data to calculate **Total\_Spend**.

The dashboard allows users to interactively explore the sales data and filter it based on different attributes such as **gender**, **age group**, **region**, and **product category**.

**6. FLOWCHART**



**7. ALGORITHM DETAILS**

**Data Preprocessing:**

* Null Handling: Remove rows with missing or irrelevant data (e.g., blank values in key columns like Amount).
* Column Calculation: Create a new column called Total\_Spend to calculate customer spending (Orders × Amount).
* Grouping Data: Aggregate sales data by region, gender, age group, and product category to generate insights.

**Dynamic Filtering:**

* Implement dynamic filters using Streamlit's multiselect widgets to allow the user to select values for gender, region, and other parameters. When the user selects a filter, the dataset is recalculated to reflect the filtered data.

**KPI Calculation:**

* Total Revenue: Sum of all Total\_Spend values for the filtered data.
* Average Transaction Value: Mean of the Amount column for the filtered data.
* Unique Customers: Count of distinct customers based on the Customer ID or equivalent identifier.

**Data Visualization:**

* Bar Charts: Display the total sales for each product category or region.
* Pie Charts: Show the percentage breakdown of sales by gender or age group.
* KPI Cards: Show key metrics like total revenue, unique customers, and average transaction value.

**8. IMPLEMENTATION**

import streamlit as st

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

**# Step 1: Set Page Configuration**

st.set\_page\_config(page\_title=" Diwali Sales BI", layout="wide")

**# Step 2: Load the Dataset**

df = pd.read\_csv("Enhanced\_Diwali\_Sales\_Data.csv")

**# Step 3: Sidebar Filters**

st.sidebar.header("Filter Data")

gender, state, zone, age\_group = (

st.sidebar.multiselect("Select Gender:", options=df['Gender'].unique(), default=df['Gender'].unique()),

st.sidebar.multiselect("Select State:", options=df['State'].unique(), default=df['State'].unique()),

st.sidebar.multiselect("Select Zone:", options=df['Zone'].unique(), default=df['Zone'].unique()),

st.sidebar.multiselect("Select Age Group:", options=df['Age Group'].unique(), default=df['Age Group'].unique()))

**# Step 4: Filter Data Based on User Selection**

filtered\_df = df[df['Gender'].isin(gender) & df['State'].isin(state) & df['Zone'].isin(zone) & df['Age Group'].isin(age\_group)]

**# Step 5: Key Performance Indicators (KPIs)**

st.title("Diwali Sales Dashboard")

st.metric("Total Revenue", f"₹{filtered\_df['Total\_Spend'].sum():,}")

st.metric("Avg. Transaction", f"₹{filtered\_df['Total\_Spend'].mean():.2f}"

if not filtered\_df.empty else "₹0")

st.metric("Unique Customers", filtered\_df['User\_ID'].nunique())

**# Step 6: Layout for Visualizations**

col1, col2 = st.columns(2)

# Subplot 1: Top Product Categories

with col1:

st.subheader("Top Product Categories")

top\_categories = filtered\_df['Product\_Category'].value\_counts().head(10)

if not top\_categories.empty:

sns.barplot(x=top\_categories.values,y=top\_categories.index,palette="magma").set(title="Top Product Categories")

st.pyplot()

else:

st.warning("No data for selected filters in Product Categories.")

# Subplot 2: Zone-wise Revenue

with col2:

st.subheader("Zone-wise Revenue")

zone\_sales = filtered\_df.groupby('Zone')['Total\_Spend'].sum().sort\_values()

if not zone\_sales.empty:

zone\_sales.plot(kind="barh", color="teal").set(title="Zone-wise Revenue")

st.pyplot() else:

st.warning("No data for selected filters in Zone-wise Revenue.")

**# Step 7: Additional Visualizations**

col3, col4 = st.columns(2)

# Subplot 3: Gender-wise Sales

with col3:

st.subheader("Gender-wise Sales")

gender\_sales = filtered\_df.groupby("Gender")["Total\_Spend"].sum()

if not gender\_sales.empty:

ax = gender\_sales.plot.pie(autopct="%1.1f%%", startangle=90, colors=["#FFA07A", "#20B2AA"], ylabel="")

ax.set\_title("Gender-wise Sales")

st.pyplot()

else:

st.warning("No data for selected filters in Gender-wise Sales.")

# Subplot 4: Occupation-wise Revenue

with col4:

st.subheader("Occupation-wise Revenue")

occupation\_sales = filtered\_df.groupby("Occupation")["Total\_Spend"].sum().sort\_values(ascending=False).head(10)

if not occupation\_sales.empty:

sns.barplot(x=occupation\_sales.values,

y=occupation\_sales.index, palette="cool").set(title="Occupation-wise Revenue")

st.pyplot()

else:

st.warning(" No data for selected filters in Occupation-wise Revenue.")

**# Step 8: Age Group Sales**

st.subheader("Age Group Spending Trend")

age\_group\_sales = filtered\_df.groupby("Age Group")["Total\_Spend"].sum()

if not age\_group\_sales.empty:

sns.barplot(x=age\_group\_sales.index, y=age\_group\_sales.values, palette="coolwarm").set(title="Age Group Spending")

st.pyplot()

else:

st.warning("No data for selected filters in Age Group Sales.")

**# Step 9: Footer Information**

st.caption("Built with Streamlit | Enhanced Dataset for Diwali Sales Intelligence")

**9. RESULT**

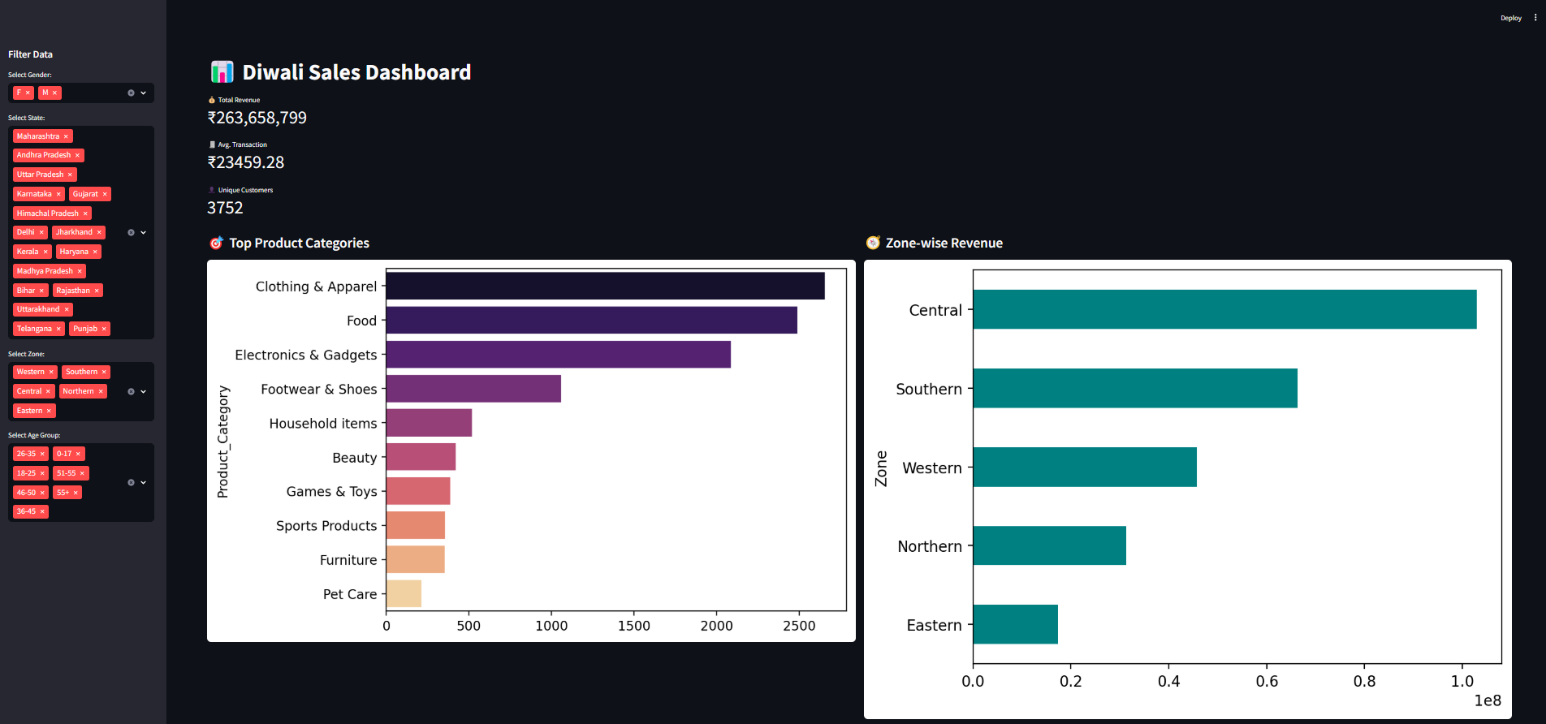
****

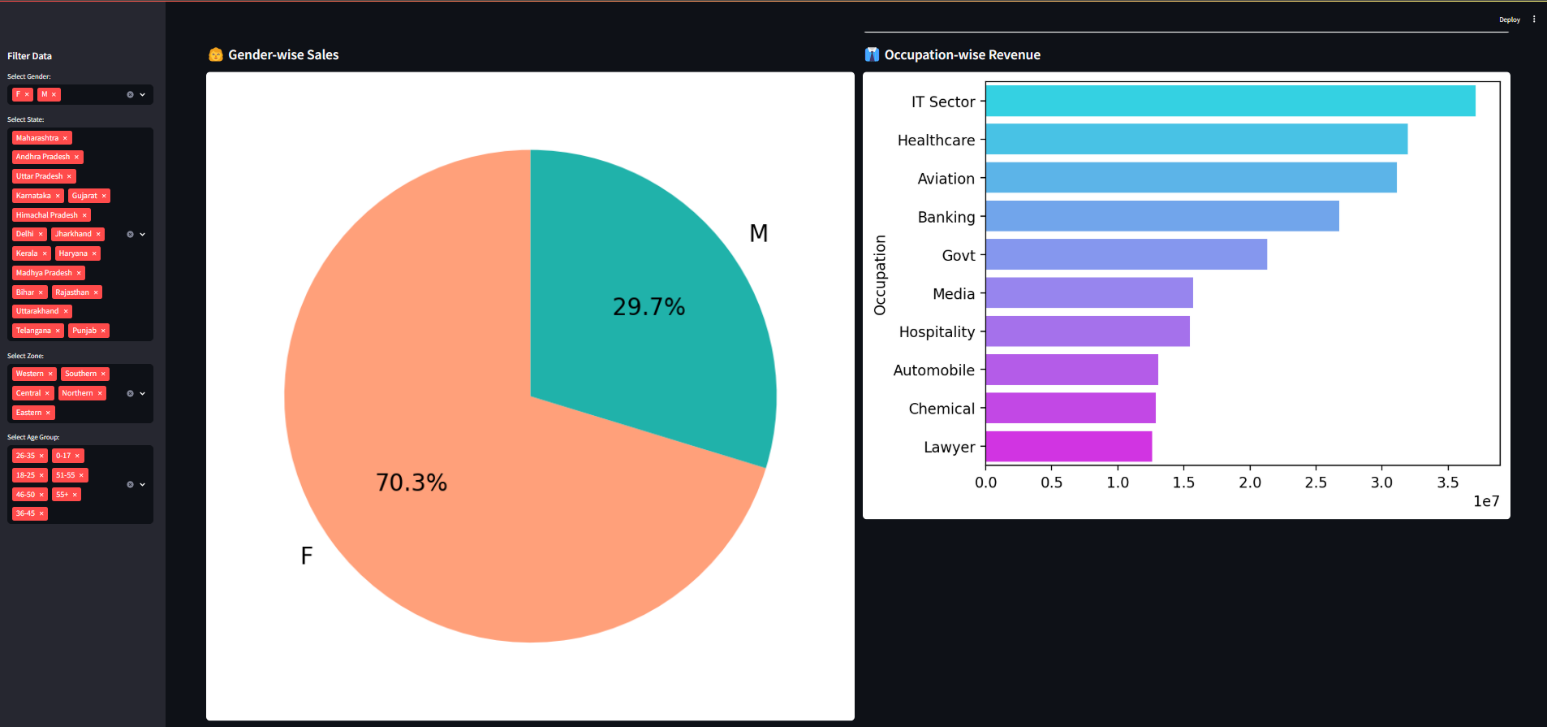
Fig 9.1: Diwali Sales Dashboard****

Fig 9.2: Diwali Sales Overview

**10. CONCLUSION**

The **Diwali Sales Dashboard** addresses the key challenges faced by retailers during the Diwali period by offering an **interactive, real-time data analysis tool**. It empowers businesses to **filter and segment** sales data, **track KPIs**, and **visualize trends** quickly and easily. By making this tool available, businesses can gain insights into customer behavior, optimize their product offerings, and improve their marketing strategies.

The dashboard is not only an open-source solution but also a **scalable tool** that can be adapted for other festive sales or retail events. It is designed to be easily customized, offering potential for future enhancements such as integrating predictive analytics or adding machine learning models for demand forecasting.