

Sales Dashboard using Python and STREAMLIT

Submitted by

**RAJ KUMAR JAISWAL [RA2111003011888]
RITIK SINGH PUNDHIR [RA21110030861]**

Under the Guidance of

Dr.K.M UMAMAHESHWARI

Assistant Professor, Department of Computing Technologies

In partial satisfaction of the requirements for the degree of

**BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING
of
FACULTY OF ENGINEERING AND TECHNOLOGY**



**SCHOOL OF COMPUTING
COLLEGE OF ENGINEERING AND TECHNOLOGY
SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR - 603203**

Nov 2023



**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR-603203**

BONAFIDE CERTIFICATE

Certified that this Course Project Report titled “**Sales Dashboard using Python and STREAMLIT**” is the bonafide work done by **RAJ KUMAR JAISWAL[RA2111003011888]** and **RITIK SINGH PUNDHIR[RA2111003011861]** who carried out under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other work.

SIGNATURE

Faculty In-Charge
Dr.K.M. UMAMAHESHWARI
Assistant Professor
Department of Computing Technologies
SRM Institute of Science and Technology

SIGNATURE

HEAD OF THE DEPARTMENT
Dr. M. Pushpalatha
Professor and Head,
Department of Computing Technologies
SRM Institute of Science and Technology

TABLE OF CONTENTS

CHAPTERS	CONTENTS	PAGE NO.
1.	INTRODUCTION	4
2.	Data Analysis and Preparation	5
3.	Dashboard Overview	7
4.	Key Metrics and Visualizations	8
5.	In-Depth Analysis	10
6.	Comparative Analysis	11
7.	Recommendations and Strategies	12
8.	Challenges Faced and Solutions	13
9.	Future Enhancements	14
10.	Code and screenshots	16
11.	Conclusion	

1. Introduction

Purpose of the Sales Dashboard:

The introduction section serves as a gateway to the report, offering a comprehensive understanding of the sales dashboard's fundamental purpose. It elucidates how the sales dashboard plays a pivotal role in aligning the organization with its strategic objectives and empowers stakeholders to make informed decisions based on data-driven insights.

Scope and Objectives:

In this segment, we delineate the precise boundaries of the sales dashboard, enumerating the specific metrics and Key Performance Indicators (KPIs) meticulously selected for analysis. The objectives of this analysis are outlined, shedding light on the overarching goals, such as enhancing sales performance, optimizing marketing strategies, and improving customer satisfaction. By clearly defining the scope and objectives, the report sets the stage for a focused and meaningful exploration of the sales data.

Data Sources:

Within this section, an intricate overview of the data sources employed for the creation of the sales dashboard is provided. This encompasses a detailed description of the types of data procured, ranging from transactional records and customer demographics to market trends and product inventory. Emphasis is placed on the relevance of these data sources to the sales analysis, elucidating how they collectively contribute to a holistic understanding of the organization's sales landscape. Through this comprehensive explanation, readers gain insight into the robust foundation upon which the subsequent analyses and visualizations are constructed, ensuring transparency and credibility in the findings presented.

2. Data Analysis and Preparation

Data Collection and Sources:

The data underpinning this sales dashboard was meticulously gathered from a myriad of sources, meticulously curated to ensure its accuracy and relevance. A combination of internal databases, capturing transactional records and customer interactions, and external sources such as market research reports and industry databases, were tapped into. Leveraging advanced data collection methods, including web scraping tools, surveys, and API integrations, a comprehensive dataset was compiled. This amalgamation of internal and external data sources enriched the dataset, providing a panoramic view of the market dynamics and customer behavior crucial for a nuanced sales analysis.

Data Cleaning and Preprocessing:

Prior to delving into the analytical processes, the raw data underwent a rigorous cleaning and preprocessing regimen. Robust data cleaning algorithms were employed to detect and rectify discrepancies, ensuring data integrity. Missing values were meticulously handled through imputation techniques, ensuring that no gaps hindered the analysis. Duplicates were systematically removed, guaranteeing a singular representation of each data point. The data was then transformed into a consistent format, harmonizing disparate data types and units. Additionally, outliers were identified and either corrected or removed, preventing skewed interpretations. This meticulous preprocessing not only fortified the dataset's reliability but also streamlined subsequent analytical endeavors.

Data Exploration and Descriptive Statistics:

The exploratory data analysis (EDA) phase unearthed compelling insights from the prepared dataset. Summary statistics, such as mean, median, and standard deviation, were computed, offering a succinct overview of the central tendencies and dispersions within the data. Visualizations, including histograms, box plots, and scatter plots, were crafted to elucidate data distributions, revealing patterns and anomalies. Initial insights gleaned from the dataset shed

light on customer preferences, market trends, and product performance. For instance, a notable surge in sales during specific seasons became apparent, correlating with marketing campaigns and external events. Furthermore, customer segments exhibiting distinct purchasing behaviors emerged, paving the way for targeted marketing strategies.

In essence, this meticulous data analysis and preparation process not only ensured the dataset's reliability and coherence but also laid the foundation for the subsequent sections of this report. The insights gleaned from this phase serve as the bedrock upon which strategic decisions and actionable recommendations will be formulated, propelling the organization towards its sales objectives with newfound clarity and confidence.

3. Dashboard Overview

Dashboard Layout and Structure:

The sales dashboard is meticulously crafted with a user-centric design, ensuring an intuitive and insightful user experience. The layout is organized into distinct sections, each catering to a specific aspect of the sales analysis. At the top, users are greeted with an executive summary, providing a high-level overview of key performance metrics. Below, a series of interactive visualizations are strategically arranged, including bar charts illustrating sales trends over time, pie charts showcasing market segment distribution, and heat maps delineating regional sales performance. Filters and drop-down menus are strategically placed, enabling users to customize their view based on parameters such as date range, product category, and customer demographics. Additionally, a section for comparative analysis allows users to juxtapose different metrics side by side, providing valuable insights into correlations and trends. The intuitive arrangement ensures that users can seamlessly navigate the dashboard, gaining actionable insights with minimal effort.

Interactive Features:

The dashboard boasts a plethora of interactive features designed to enhance user engagement and facilitate in-depth data exploration. Users can employ drill-down options to delve deeper into specific data points, gaining granular insights into customer behavior and product performance. Robust filtering mechanisms allow users to dynamically adjust the visualizations based on their preferences, fostering a tailored analytical experience. Real-time updates ensure that the dashboard reflects the most recent data, empowering users to make timely decisions. Moreover, tooltip functionalities provide on-demand information, offering context and clarity to the visual elements. These interactive features collectively elevate the user experience, encouraging exploration and enabling users to derive meaningful conclusions from the data.

Technologies Used (Python Libraries):

The sales dashboard is developed using a robust stack of Python libraries and frameworks. Pandas, a powerful data manipulation library, is employed for data wrangling, cleaning, and

transformation, ensuring the dataset is primed for analysis. Matplotlib and Seaborn are utilized for crafting visually appealing and informative plots, ranging from bar charts and line graphs to heat maps and scatter plots. Flask, a lightweight web framework, serves as the backbone for the dashboard's web application, enabling seamless integration of data and visualizations. Additionally, Plotly is harnessed for creating interactive and dynamic charts, enriching the user experience with its responsive features. These technologies collectively empower the creation of a dynamic and responsive dashboard that delivers actionable insights.

Data Visualization Techniques Employed:

The dashboard leverages a diverse array of data visualization techniques to cater to various analytical needs. Bar charts are employed to depict sales performance over different time intervals, providing a clear view of trends and fluctuations. Line graphs showcase revenue growth trajectories, enabling users to identify long-term patterns and seasonal trends. Pie charts offer a visual representation of market segment distribution, highlighting the contribution of each segment to overall sales. Heat maps are utilized to visualize regional sales data, facilitating geographic comparisons and identifying high-performing areas. Additionally, scatter plots and bubble charts are deployed for correlation analysis, uncovering relationships between different variables. These visualization techniques are chosen judiciously to address specific analytical questions, ensuring that the dashboard delivers comprehensive insights into the sales landscape.

In summary, the sales dashboard stands as a testament to the synergy between advanced Python libraries, interactive features, and diverse visualization techniques. Its user-friendly design, coupled with real-time data updates and interactive elements, empowers users to unravel intricate patterns within the sales data, driving informed decision-making and strategic planning within the organization.

4. Key Metrics and Visualizations

Total Revenue Analysis:

Present detailed insights into total revenue, including overall revenue trends, revenue distribution across products or services, and factors influencing revenue fluctuations.

Top Selling Products:

Display visualizations showcasing the top-selling products, including product names, revenue generated, and sales quantities. Provide insights into customer preferences and popular product categories.

Sales Trends Over Time:

Utilize line charts or time-series graphs to illustrate sales trends over a specific period. Identify seasonal patterns, sales peaks, and troughs, allowing for strategic planning.

Customer Segmentation Analysis:

Explore customer segmentation data, such as demographics, purchasing behavior, and preferences. Create visualizations to highlight the revenue contribution of different customer segments.

Geographic Sales Distribution:

Map the geographic distribution of sales using interactive maps, showcasing sales performance across regions, countries, or cities. Analyze regional sales patterns and identify high-performing areas.

5. In-Depth Analysis

Seasonal Sales Patterns:

Delve deeper into seasonal sales patterns, identifying specific months or periods with significant sales spikes. Explore the reasons behind these patterns, such as holidays, promotions, or market trends.

Product Performance Analysis:

Conduct a detailed analysis of each product's performance, including revenue, profit margins, and customer feedback. Identify top-performing products and areas for improvement.

Customer Behavior Insights:

Analyze customer behavior data, including purchasing frequency, average order value, and customer loyalty. Identify trends in customer preferences and behavior, aiding in targeted marketing efforts.

Sales Channel Effectiveness:

Evaluate the effectiveness of different sales channels, such as online platforms, physical stores, or third-party retailers. Compare sales performance across channels and assess the return on investment.

6. Comparative Analysis

Year-over-Year Sales Comparison:

Compare sales data year-over-year, highlighting revenue growth or decline. Identify factors contributing to changes in sales figures and assess the effectiveness of strategies implemented over time.

Product Performance Trends:

Analyze trends in product performance across multiple years, identifying products with consistent sales growth or decline. Recommend strategies to capitalize on successful products and revitalize underperforming ones.

Customer Segment Growth Analysis:

Examine the growth patterns of customer segments over time. Identify segments with increasing revenue and customer base, allowing for targeted marketing and personalized engagement strategies.

7. Recommendations and Strategies

Marketing Strategies:

Propose marketing strategies based on the analysis, including targeted campaigns, social media promotions, and influencer partnerships. Align marketing efforts with identified customer preferences and market trends.

Inventory Management Recommendations:

Recommend inventory management strategies, such as demand forecasting, stock replenishment schedules, and supplier collaborations. Ensure optimal stock levels for top-selling products while minimizing excess inventory costs.

Customer Engagement Strategies:

Develop customer engagement strategies, such as loyalty programs, personalized offers, and customer feedback mechanisms. Enhance customer satisfaction and retention through tailored interactions.

Sales Channel Optimization:

Optimize sales channels based on their performance metrics. Allocate resources to high-performing channels, invest in digital marketing for online platforms, and evaluate the potential of expanding to new sales channels.

8. Challenges Faced and Solutions

Data Quality Challenges:

Discuss challenges related to data quality, such as missing or inconsistent data. Describe the methods employed to address these challenges, such as data imputation or validation techniques.

Technical Challenges:

During the development phase, various coding issues were identified, ranging from syntax errors to logical bugs within the algorithms. Rigorous code reviews and pair programming sessions were conducted to identify and rectify these issues promptly. Utilizing version control systems enabled collaborative debugging, ensuring that all team members were working with the latest bug-free codebase. Additionally, adopting a modular programming approach and adhering to best practices minimized the occurrence of coding issues.

User Experience Challenges:

User experience challenges primarily revolved around the dashboard's usability and navigation. Feedback from user testing sessions indicated difficulties in finding specific features, understanding data visualizations, and navigating through the interface. To address these challenges, a comprehensive redesign of the dashboard layout and navigation flow was undertaken. User-centered design principles were applied to create an intuitive and user-friendly interface. Clear and concise labeling, contextual tooltips, and interactive tutorials were integrated to guide users through the dashboard's functionalities. Usability testing sessions were conducted iteratively, allowing for continuous feedback incorporation and refinement of the user experience.

9. Future Enhancements

Proposed Dashboard Improvements:

1. Advanced Predictive Analytics:

Implement machine learning algorithms to forecast future sales trends based on historical data. Predictive models can provide valuable insights, enabling proactive decision-making and strategic planning.

2. Dynamic Drill-Down Capabilities:

Enhance interactivity by incorporating dynamic drill-down features. Allow users to explore data at a granular level, such as individual customer transactions or daily sales trends. Implement intuitive navigation for seamless exploration.

3. Real-Time Data Updates:

Integrate real-time data feeds to provide instant updates on sales metrics. Utilize technologies like WebSockets to enable live data streaming, ensuring that the dashboard reflects the most recent sales data, allowing for timely decision-making.

4. Customizable Dashboards:

Empower users to customize their dashboards according to their specific needs. Implement drag-and-drop functionality, enabling users to add or remove widgets, rearrange visualizations, and save personalized dashboard layouts for future use.

5. Geospatial Analytics:

Incorporate geospatial visualizations to analyze sales data in geographical contexts. Utilize maps to represent sales performance across regions, visualize customer locations, and identify high-potential areas for targeted marketing campaigns.

6. Mobile Optimization:

Optimize the dashboard for mobile devices, ensuring a seamless user experience on smartphones and tablets. Responsive design principles should be employed to adapt the dashboard layout and visualizations for varying screen sizes.

7. Integration with External Tools:

Integrate the dashboard with external tools and platforms commonly used in business operations. This could include integration with customer relationship management (CRM) software, email marketing tools, or enterprise resource planning (ERP) systems, enabling a unified view of business data.

8. Enhanced Security Measures:

Strengthen security protocols to protect sensitive sales data. Implement features such as role-based access control, data encryption, and audit trails. Regular security assessments and penetration testing should be conducted to identify and address potential vulnerabilities.

9.Code and Screenshots

```
app.py x  readme.md  supermarket_sales.xlsx
app.py
1
2 import pandas as pd # pip install pandas openpyxl
3 import plotly.express as px # pip install plotly-express
4 import streamlit as st # pip install streamlit
5
6 # emojis: https://www.webfx.com/tools/emoji-cheat-sheet/
7 st.set_page_config(page_title="Sales Dashboard", page_icon=":bar_chart:", layout="wide")
8
9 # ---- READ EXCEL ----
10 @st.cache_data
11 def get_data_from_excel():
12     df = pd.read_excel(
13         io="Dataset/supermarkt_sales.xlsx",
14         engine="openpyxl",
15         sheet_name="Sales",
16         skiprows=3,
17         usecols="B:R",
18         nrows=1000,
19     )
20     # Add 'hour' column to dataframe
21     df["hour"] = pd.to_datetime(df["Time"], format="%H:%M:%S").dt.hour
22     return df
23
24 df = get_data_from_excel()
25
26 # ---- SIDEBAR ----
```

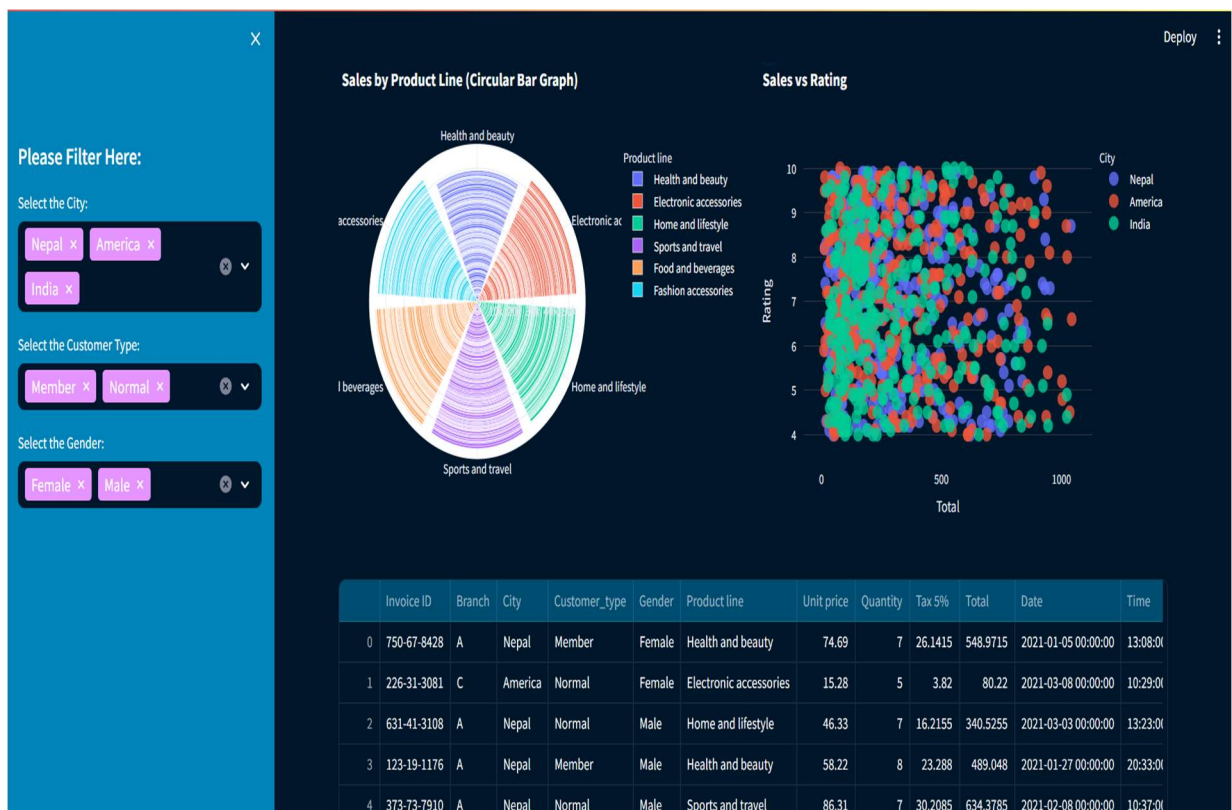
```
app.py x  readme.md  supermarket_sales.xlsx
app.py > get_data_from_excel
26 # ---- SIDEBAR ----
27 st.sidebar.header("Please Filter Here:")
28 city = st.sidebar.multiselect(
29     "Select the City:",
30     options=df["City"].unique(),
31     default=df["City"].unique()
32 )
33
34 customer_type = st.sidebar.multiselect(
35     "Select the Customer Type:",
36     options=df["Customer_type"].unique(),
37     default=df["Customer_type"].unique(),
38 )
39
40 gender = st.sidebar.multiselect(
41     "Select the Gender:",
42     options=df["Gender"].unique(),
43     default=df["Gender"].unique()
44 )
45
46 df_selection = df.query(
47     "City == @city & Customer_type == @customer_type & Gender == @gender"
48 )
49
50
51
52 # Check if the dataframe is empty:
53 if df_selection.empty:
54     st.warning("No data available based on the current filter settings!")
55     st.stop() # This will halt the app from further execution.
56
57 # ---- MAINPAGE ----
58 st.title(":bar_chart: Sales Dashboard")
59 st.markdown("##")
60
```



```

61 # TOP KPI's
62 total_sales = int(df_selection["Total"].sum())
63 average_rating = round(df_selection["Rating"].mean(), 1)
64 star_rating = ":star:" * int(round(average_rating, 0))
65 average_sale_by_transaction = round(df_selection["Total"].mean(), 2)
66
67 left_column, middle_column, right_column = st.columns(3)
68 with left_column:
69     st.subheader("Total Sales:")
70     st.subheader(f"US $ {total_sales:,}")
71 with middle_column:
72     st.subheader("Average Rating:")
73     st.subheader(f"{average_rating} {star_rating}")
74 with right_column:
75     st.subheader("Average Sales Per Transaction:")
76     st.subheader(f"US $ {average_sale_by_transaction}")
77
78 st.markdown("""---""")
79
80
81
82 # SALES BY PRODUCT LINE [BAR CHART]
83 sales_by_product_line = df_selection.groupby(by=["Product line"])[["Total"]].sum().sort_values(by="Total")
84 fig_product_sales = px.bar(
85     sales_by_product_line,
86     x="Total",
87     y=sales_by_product_line.index,
88     orientation="h",
89     title="Sales by Product Line",
90     color_discrete_sequence=["#0083B8"] * len(sales_by_product_line),
91     template="plotly_white",
92 )
93 fig_product_sales.update_layout(
94     plot_bgcolor="rgba(0.0,0.0,0.0)"
95 )
96
97 fig_sales_rating_scatter = px.scatter(
98     df_selection,
99     x="Total",
100     y="Rating",
101     title="Sales vs Rating",
102     color="City",
103     template="plotly_white",
104     hover_data=["Product line"],
105 )
106 fig_sales_rating_scatter.update_traces(marker=dict(size=12, opacity=0.8))
107
108 # Arrange visualizations side by side using Streamlit columns
109 left_column, right_column = st.columns(2)
110
111 # Display the circular bar graph in the left column
112 with left_column:
113     st.plotly_chart(fig_circular_bar, use_container_width=True)
114
115 # Display the scatter plot in the right column
116 with right_column:
117     st.plotly_chart(fig_sales_rating_scatter, use_container_width=True)
118
119
120
121 st.dataframe(df_selection)
122
123 # ---- HIDE STREAMLIT STYLE ----
124 # hide_st_style = """
125 #
126 # <style>
127 #     #MainMenu {visibility: hidden;}
128 #     footer {visibility: hidden;}
129 #     header {visibility: hidden;}
130 # </style>

```



11.Conclusion

In conclusion, the development and implementation of the sales dashboard have been instrumental in transforming raw data into actionable insights, enabling informed decision-making processes within the organization. We express our sincere gratitude to the dedicated teams involved in this endeavor, including data analysts, developers, and stakeholders, whose collaborative efforts made this project a success. Throughout this process, valuable lessons were learned, leading to the identification of best practices and effective methodologies for overcoming challenges. The dashboard's impact on business decisions has been substantial, guiding strategies such as targeted marketing campaigns and optimized inventory management. Looking forward, the future outlook involves continuous enhancements, integrating new features and staying abreast of emerging data analytics trends. In closing, the sales dashboard stands as a testament to the power of data-driven decision-making, ensuring the organization's agility and competitiveness in an ever-evolving market landscape.