Report
On
Industry Internship

Submitted as Partial Fulfillment of
Master of Computer Applications
Semester – IV

Developed at
INFOLABZ I.T. Services PVT. LTD.

Developed by
Janki Panchal (202300719010042)

Under Guidance of
Mr. Meet Thakar, FCAIT
Mr. Kirit Suthar INFOLABZ I.T. Services Pvt. Ltd.

Faculty of Computer Applications & Information Technology
(FCAIT)

GLS University
Ahmedabad-380 006

<table>
<tr><td>Role:</td><td>Data Science Intern</td></tr>
<tr><td>Description:</td><td>This AI-driven system forecasts future revenue by analyzing historical data and market trends. It detects anomalies, captures seasonal patterns, and provides actionable insights. The solution supports financial planning and strategic decision-making across industries like retail and e-commerce.</td></tr>
<tr><td>Tools Used / Required:</td><td><u>Development side</u>
Hardware: NA

Software:
• Operating System: Windows 10
• Front End: Streamlit
• Back End: Python, Machine Learning Models, AI models
• Framework: Streamlit, Scikit-learn
• Report Generation Tool: Microsoft Word

<u>Deployment side</u>

Hardware:
• Processor: Intel i5 or higher
• RAM: 8GB minimum
• Storage: 250GB SSD or more
• Internet Connectivity: Stable broadband connection
Software:
• Operating System: Windows 10 or above
• IDE/Tools: PyCharm, Python 3.7+
• Required Libraries</td></tr>
<tr><td>Developed at:</td><td>INFOLABZ I.T. Services PVT. LTD., Ahmedabad</td></tr>
<tr><td>Developed By:</td><td><table><tr><th>Enroll No.</th><th>Name</th><th>Institute</th><th>Contribution</th></tr><tr><td>202300719010042</td><td>Janki Panchal</td><td>FCAIT</td><td>Contributed to data preprocessing and visualization using Pandas, NumPy, and Matplotlib. Assisted in building an interactive Streamlit UI for trend analysis and supported project documentation.</td></tr></table></td></tr>
<tr><td>Guided By:</td><td><table><tr><th>Internal at Institute</th><th>External at company</th></tr><tr><td>Name: Mr. Meet Thakar
Designation: Assistant Professor</td><td>Name: Mr. Kirit Suthar
Designation: Software Developer</td></tr></table></td></tr>
</table>

Declaration of Originality and non-plagiarized content

I am declaring that the work presented in this project report titled **"AI-Driven Revenue Forecasting and Trend Analysis for Business Growth"** is my original creation and has been carried out independently. All the coding, design, and implementation have been done by me and are the result of my own effort, understanding, and research.

This project has not been copied or submitted elsewhere, in part or full, for the award of any degree or certification. The content used from external sources such as online libraries, documentation, tutorials, or open-source repositories has been appropriately referenced and acknowledged in the references section.

I have strictly adhered to ethical and academic standards, ensuring that the codebase is free from plagiarism. The development of this project reflects my own learning, problem-solving approach, and application of technical knowledge.

INDEX

Acknowledgment

I would like to express my sincere gratitude to all those who supported me throughout the development of this project, titled **"AI-Driven Revenue Forecasting and Trend Analysis for Business Growth."**

I am especially thankful to my internal mentor, **Mr. Meet Thakar**, for his invaluable guidance, continuous support, and insightful feedback throughout the project. His mentorship played a crucial role in shaping the direction of my work.

I also extend my heartfelt appreciation to the instructors at **InfoLabz IT Services Pvt. Ltd.**, whose practical insights and expert advice significantly enriched my understanding and approach.

My special thanks to **Dr. Devarshi Mehta**, Project Coordinator, and **Dr. Harshal Arolkar**, Head of PG Programme, for their academic support and encouragement throughout this journey.

I am also grateful to **Dr. Savita Gandhi**, Dean, PG Programme, for providing the opportunity and platform to undertake this industry internship and project work, which has immensely contributed to my academic and professional development.

This project has greatly enhanced my knowledge and skills in **Python programming, machine learning**, and the design of **user-centric applications using Streamlit**. The entire experience has been a significant milestone in my learning journey.

Introduction of company and Guide

Infolabz IT Services Pvt. Ltd. is a forward-looking technology solutions provider committed to delivering innovative, scalable, and customized digital services. Founded with the goal of transforming how businesses operate in the digital age, Infolabz specializes in a wide range of IT services including software development, mobile app development, web design, data science, machine learning, cloud computing, IoT solutions, and cybersecurity.

With a strong emphasis on practical learning and industry exposure, Infolabz also plays an active role in nurturing talent through internship programs, live projects, and mentorship initiatives. These opportunities allow students and budding professionals to gain hands-on experience in real-world projects and emerging technologies.

Guide Name: – Mr. Kirit Suthar

Mr. Kirit Suthar, a Senior Developer at Infolabz IT Services Pvt. Ltd., serves as the technical guide and mentor. With extensive experience in software development and project implementation, he provides valuable support and direction throughout the engagement.

Abstract giving brief introduction about role/system

In today's data-driven world, the role of a Data Analyst is crucial in turning complex data into meaningful insights that support better business decisions. In the project titled ***AI-Driven Revenue Forecasting and Trend Analysis for Business Growth***, the Data Analyst is responsible for building and analyzing predictive models to forecast future revenue trends. This helps businesses plan more effectively and grow with confidence.

Using machine learning techniques—specifically the Random Forest algorithm—the Data Analyst creates accurate and reliable forecasting models based on historical data. Random Forest is known for its high accuracy and does not require heavy data preprocessing, making it a practical choice for this task. The Data Analyst's work includes gathering and cleaning data, performing exploratory data analysis (EDA), and selecting important features to improve model performance.

In addition to the modeling, the Data Analyst designs a user-friendly interface using Streamlit. This allows business users to easily upload data, visualize trends, detect anomalies, and view revenue forecasts. With this system, business stakeholders can make informed decisions based on clear and accessible data insights.

Overall, the Data Analyst plays a key role in helping businesses use their data for smarter decision-making, better risk management, and sustainable growth. By leveraging tools like Random Forest, the project provides accurate revenue forecasts that lead to improved financial planning and long-term success.

.

Team details with clear work distribution

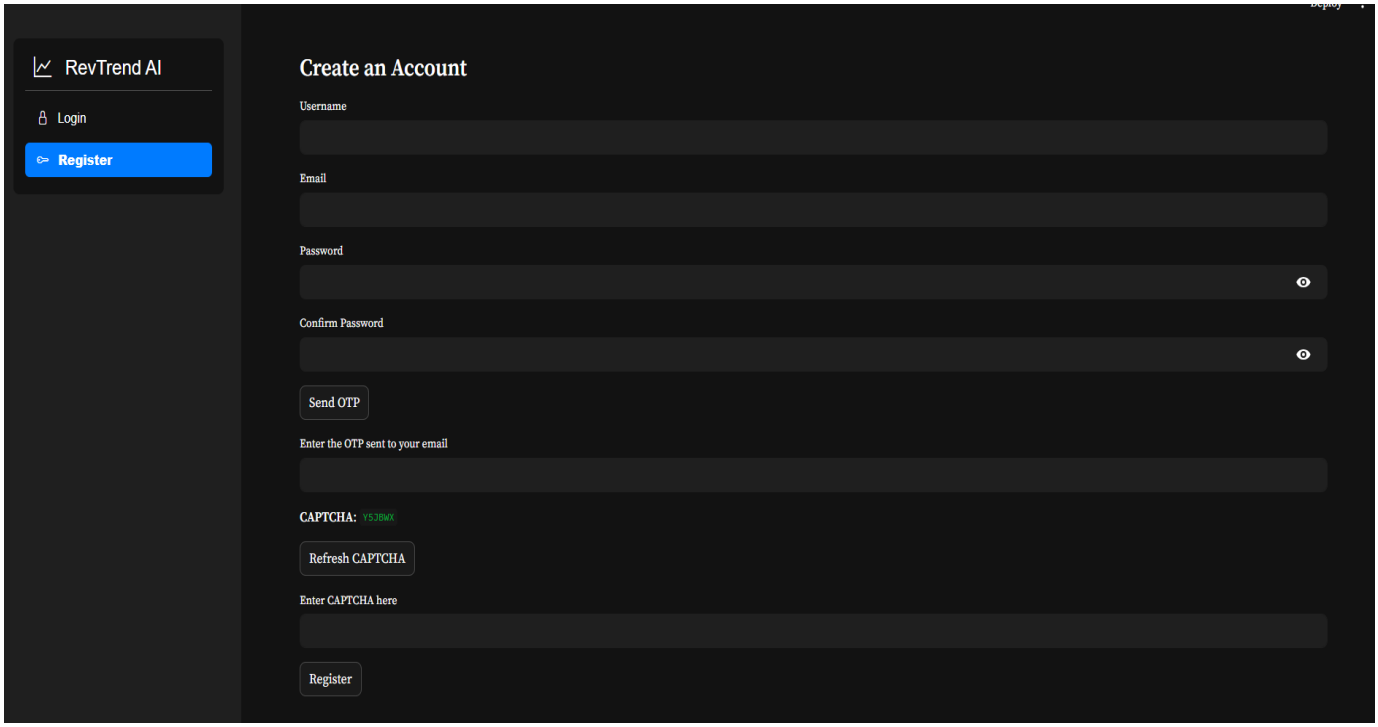
- **Project Title:** *AI-Driven Revenue Forecasting and Trend Analysis for Business Growth*
- **Technologies Used:** Python, Streamlit, Pandas, NumPy, Matplotlib, Scikit-learn, Random Forest, CSV
- **Key Tasks:**
  This project focuses on building a smart, user-centric platform for forecasting business revenue and analyzing trends to support strategic decision-making. The system includes an interactive, role-based login and registration module enhanced with OTP and CAPTCHA verification for secure access. A robust data ingestion and preprocessing pipeline was implemented to handle validation and cleaning of uploaded datasets. Using dynamic visualizations and statistical summaries, users can explore revenue trends interactively. A machine learning model based on the Random Forest algorithm was developed to accurately forecast future revenue, complete with performance evaluation metrics and CSV export functionality. The entire application is designed with a modern user interface and streamlined through a multi-tab navigation system using Streamlit. Contributions also include enhancing the user interface and preparing comprehensive project documentation to ensure usability and clarity for end users.

Task with description and screenshots

Task Title: AI-Driven Revenue Forecasting and Trend Analysis for Business Growth

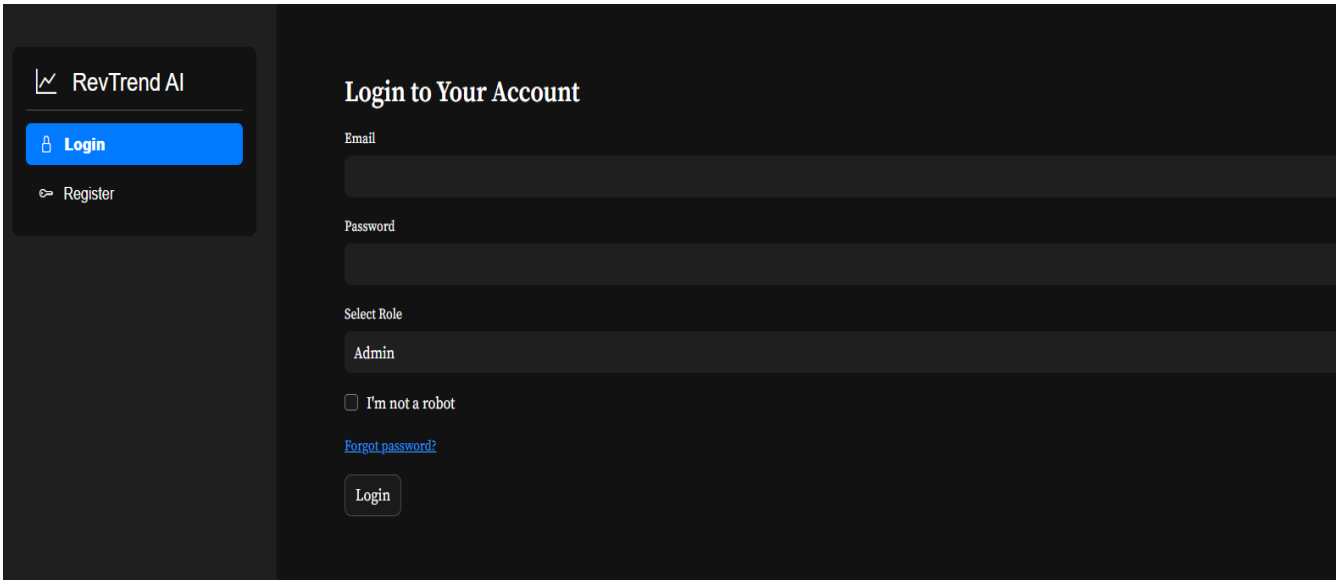
Register Page

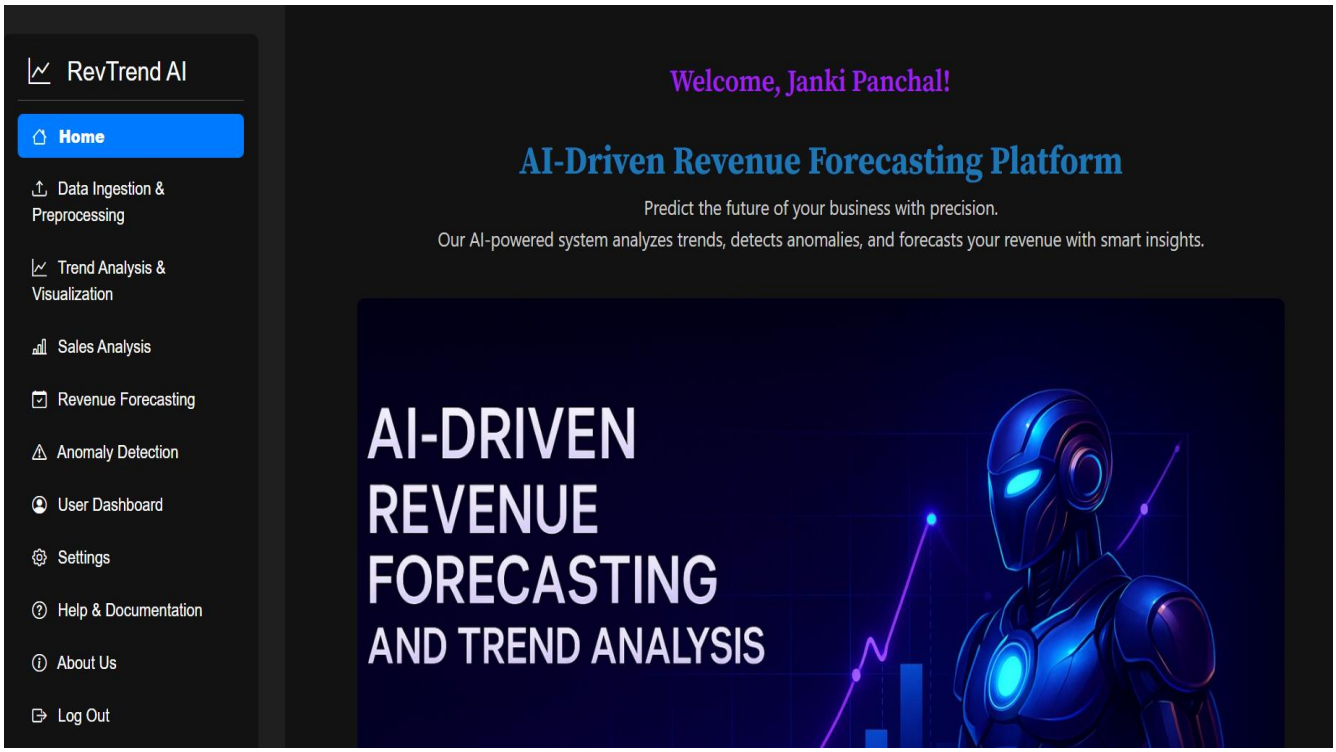
The Register page enables new users to create an account by providing a username, email, password, and confirming the password. Users must verify their email via an OTP sent to their email address, enter CAPTCHA text for security, and complete the registration process. After successful registration, the user's data is saved, and they are notified of a successful account creation, ready to log in.



Login Page

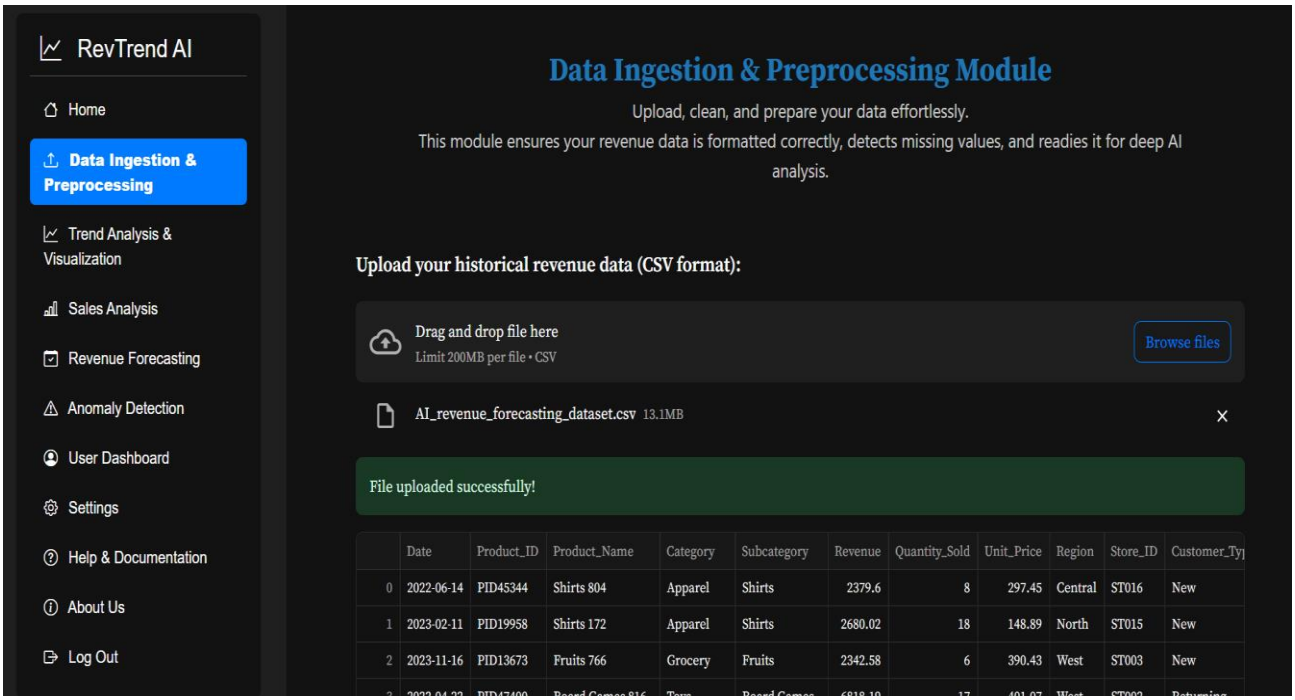
The Login page allows users to securely access their accounts by entering their email, password, and selecting their role (Admin, Analyst, or Manager). After successful login, the user is redirected to the "Home" tab, and their session is initialized with the selected role and username.

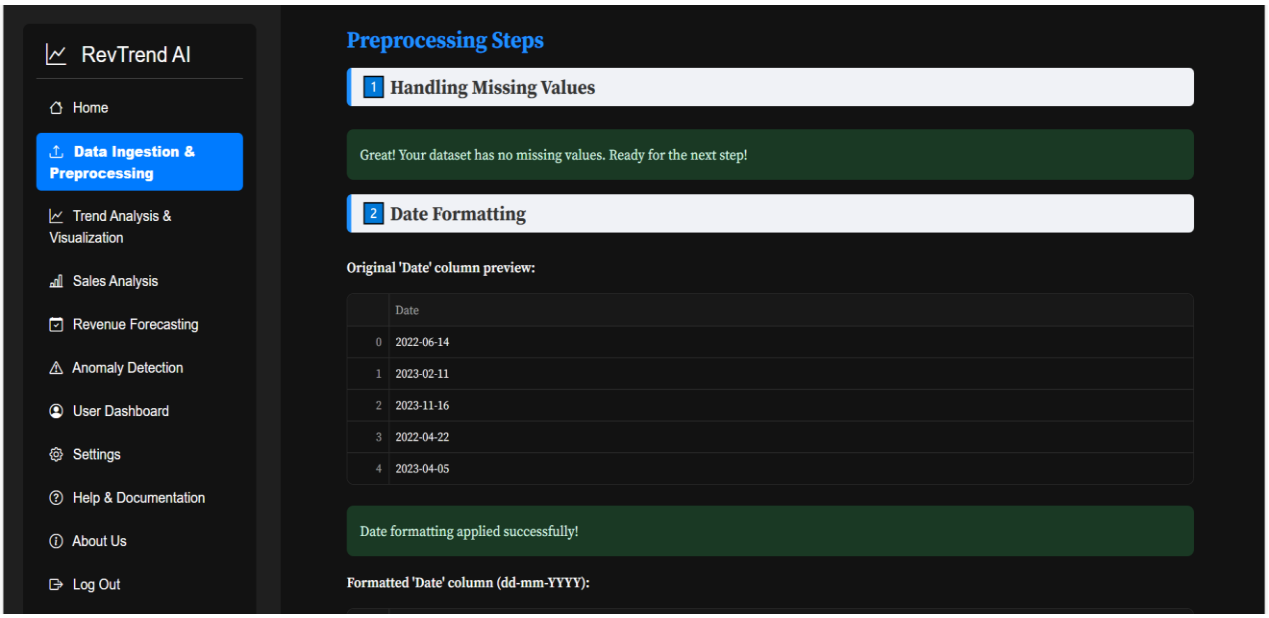


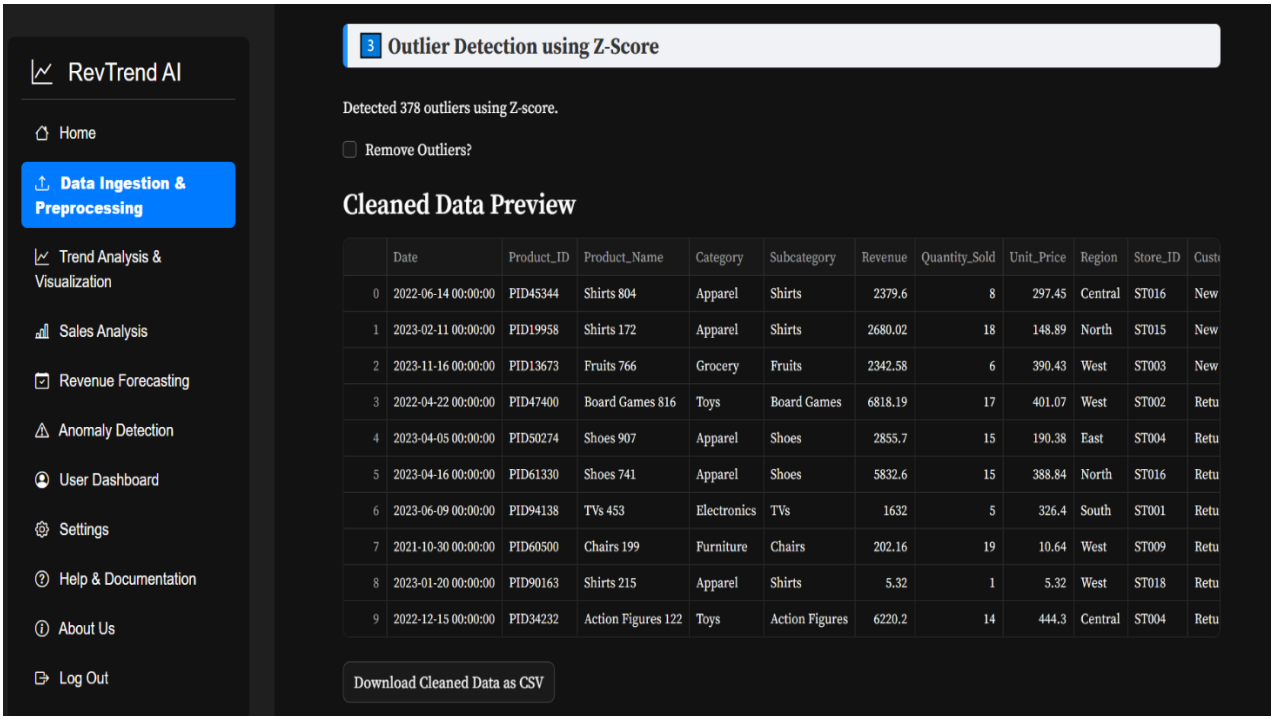


Data Ingestion & Preprocessing

This module allows users to seamlessly upload their historical revenue data in CSV format for further analysis. It performs essential preprocessing steps such as handling missing values, standardizing date formats, and detecting outliers using Z-score techniques. The goal is to ensure that your dataset is clean, consistent, and structured correctly before feeding it into forecasting models. Through this intuitive interface, users can preview raw and cleaned data, choose to remove anomalies, and download the processed dataset for use in future modules. This step forms the foundation for accurate AI-driven revenue predictions and business insights.





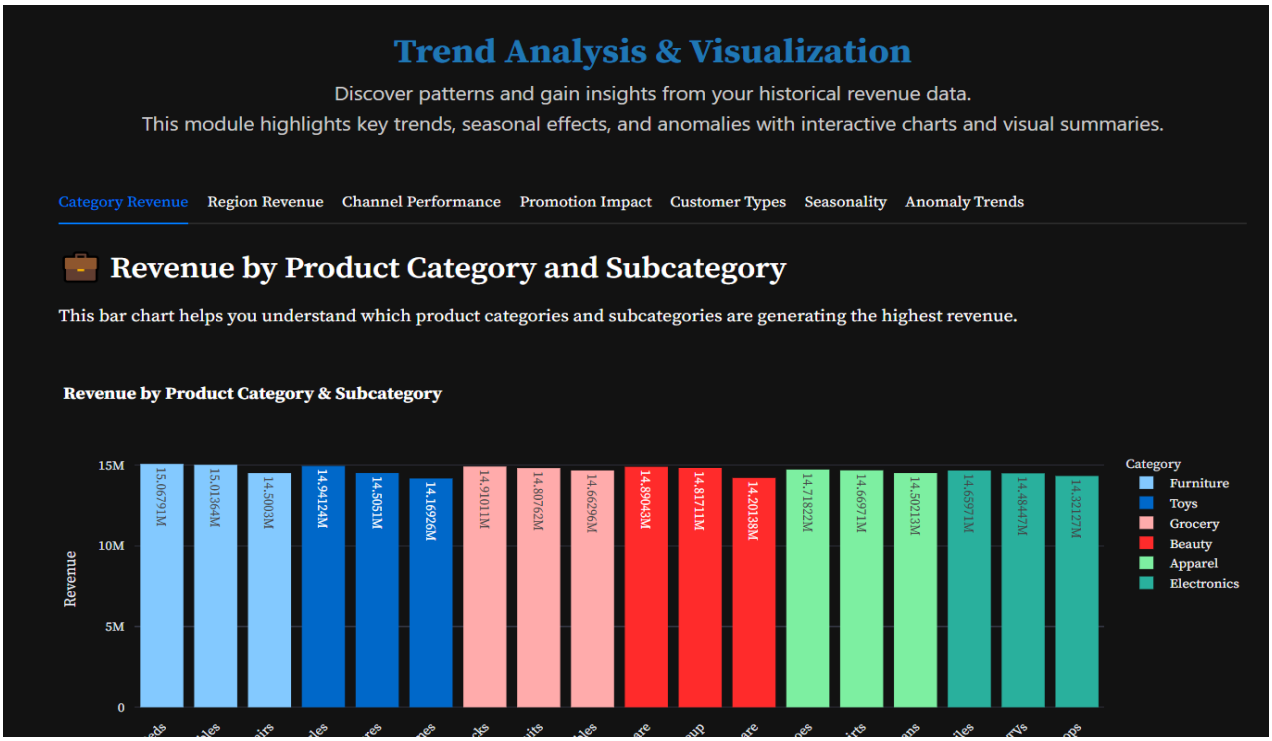


	Date	Product_ID	Product_Name	Category	Subcategory	Revenue	Quantity_Sold	Unit_Price	Region	Store_ID	Custo
0	2022-06-14 00:00:00	PID45344	Shirts 804	Apparel	Shirts	2379.6	8	297.45	Central	ST016	New
1	2023-02-11 00:00:00	PID19958	Shirts 172	Apparel	Shirts	2680.02	18	148.89	North	ST015	New
2	2023-11-16 00:00:00	PID13673	Fruits 766	Grocery	Fruits	2342.58	6	390.43	West	ST003	New
3	2022-04-22 00:00:00	PID47400	Board Games 816	Toys	Board Games	6818.19	17	401.07	West	ST002	Retu
4	2023-04-05 00:00:00	PID50274	Shoes 907	Apparel	Shoes	2855.7	15	190.38	East	ST004	Retu
5	2023-04-16 00:00:00	PID61330	Shoes 741	Apparel	Shoes	5832.6	15	388.84	North	ST016	Retu
6	2023-06-09 00:00:00	PID94138	TVs 453	Electronics	TVs	1632	5	326.4	South	ST001	Retu
7	2021-10-30 00:00:00	PID60500	Chairs 199	Furniture	Chairs	202.16	19	10.64	West	ST009	Retu
8	2023-01-20 00:00:00	PID90163	Shirts 215	Apparel	Shirts	5.32	1	5.32	West	ST018	Retu
9	2022-12-15 00:00:00	PID34232	Action Figures 122	Toys	Action Figures	6220.2	14	444.3	Central	ST004	Retu

Trend Analysis & Visualization

This module enables users to uncover valuable insights from their historical revenue data using interactive and dynamic visualizations. It provides multiple tabs for in-depth exploration of key performance indicators such as revenue by product category, region, customer type, and sales channel. Users can analyze the impact of promotions, track seasonal patterns, and detect anomalies with heatmaps.

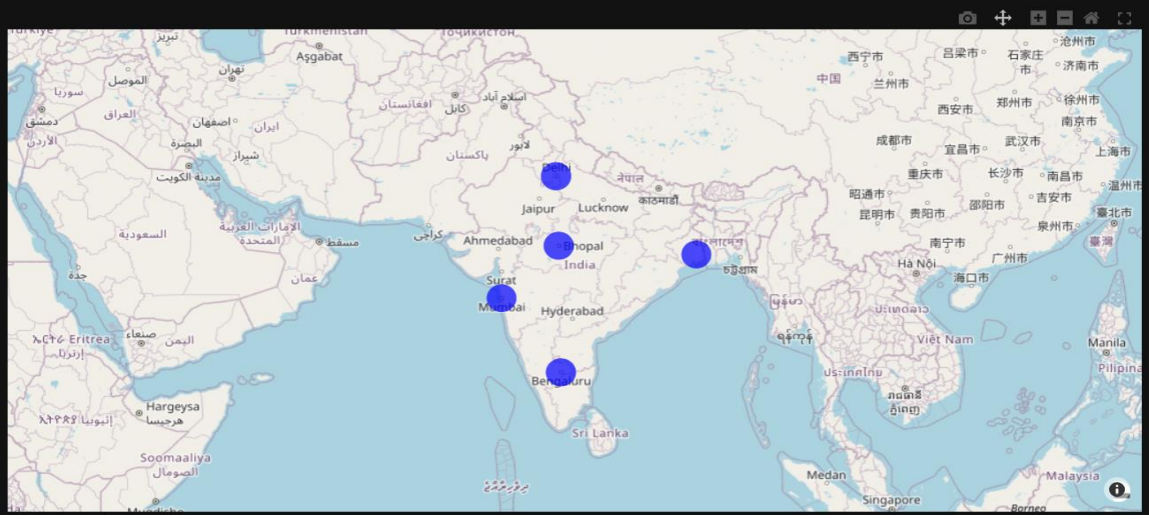
From insightful bar charts to geographic map visualizations and donut charts, each visualization helps businesses understand their strengths, bottlenecks, and opportunities. Interactive summary tables accompany each graph to offer quick numerical insights. This module forms a core part of the forecasting pipeline by offering visual clarity into past revenue trends, allowing for informed, data-driven decisions.



Category Revenue | Region Revenue | Channel Performance | Promotion Impact | Customer Types | Seasonality | Anomaly Trends

🌍 Revenue by Region

(Map-based visualization to identify regional performance.

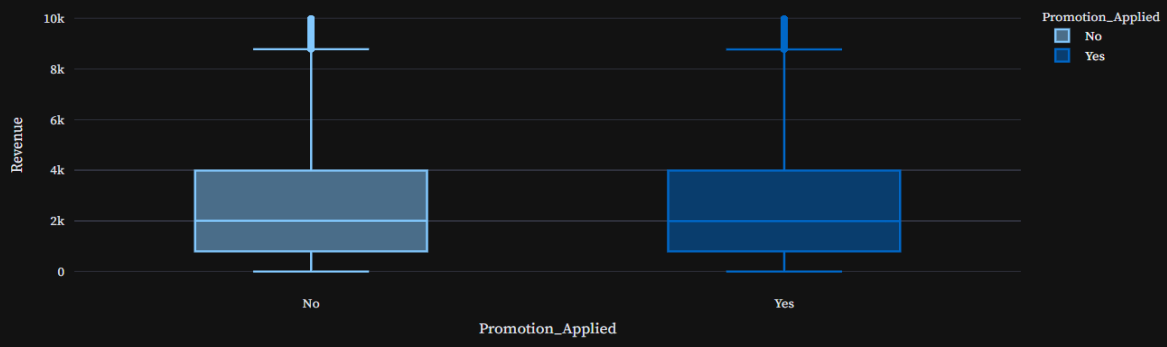


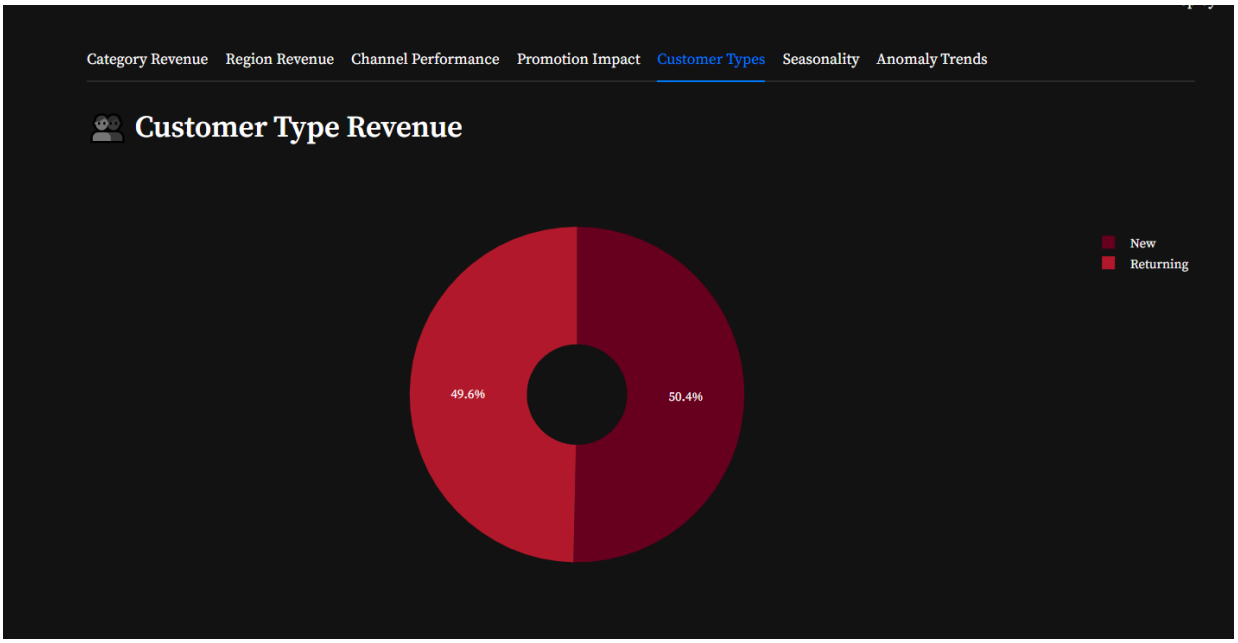
Category Revenue | Region Revenue | Channel Performance | Promotion Impact | Customer Types | Seasonality | Anomaly Trends

🎯 Promotion vs Non-Promotion Revenue

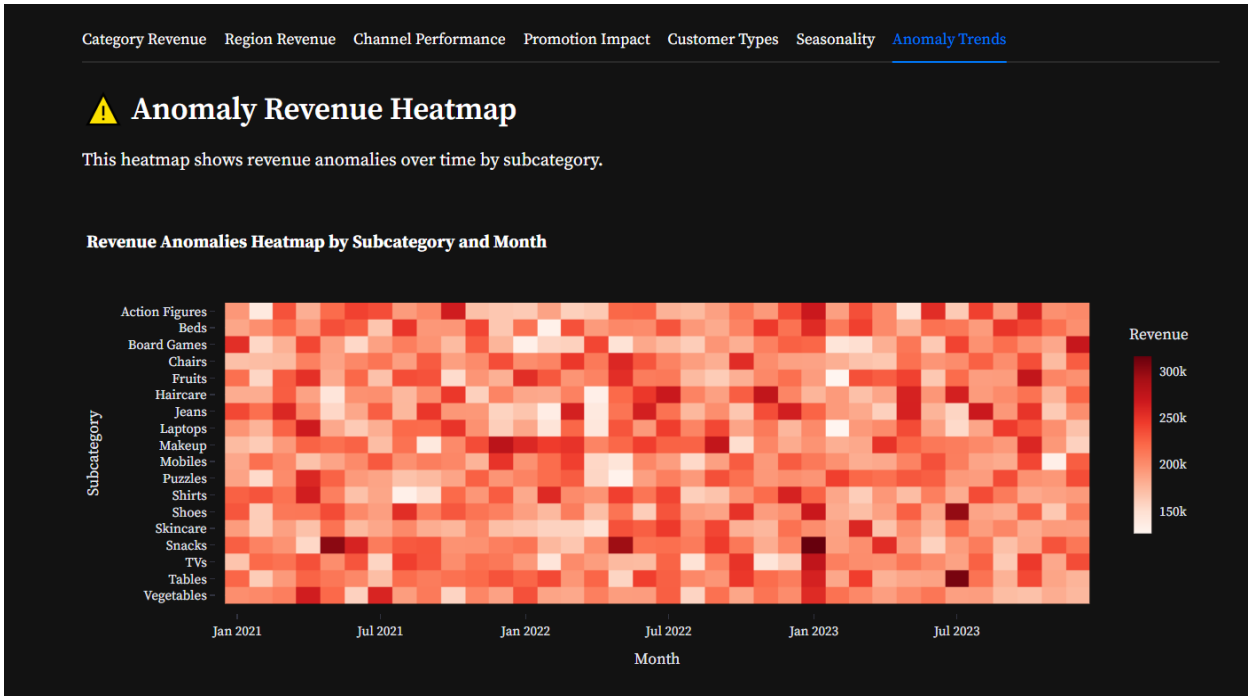
This box plot shows the revenue distribution for transactions with and without promotions.

Revenue Distribution Based on Promotion Application




Category Revenue
Region Revenue
Channel Performance
Promotion Impact
Customer Types
Seasonality
Anomaly Trends
Customer Type Revenue
New
Returning
49.6%
50.4%


Category Revenue
Region Revenue
Channel Performance
Promotion Impact
Customer Types
Seasonality
Anomaly Trends
Seasonal Revenue Trends
Revenue
50M
40M
30M
20M
10M
0
53.63485M
52.70329M
52.62849M
52.56185M
52.31409M
Spring
Winter
Summer
Monsoon
Autumn
Season
Season
Spring
Winter
Summer
Monsoon
Autumn

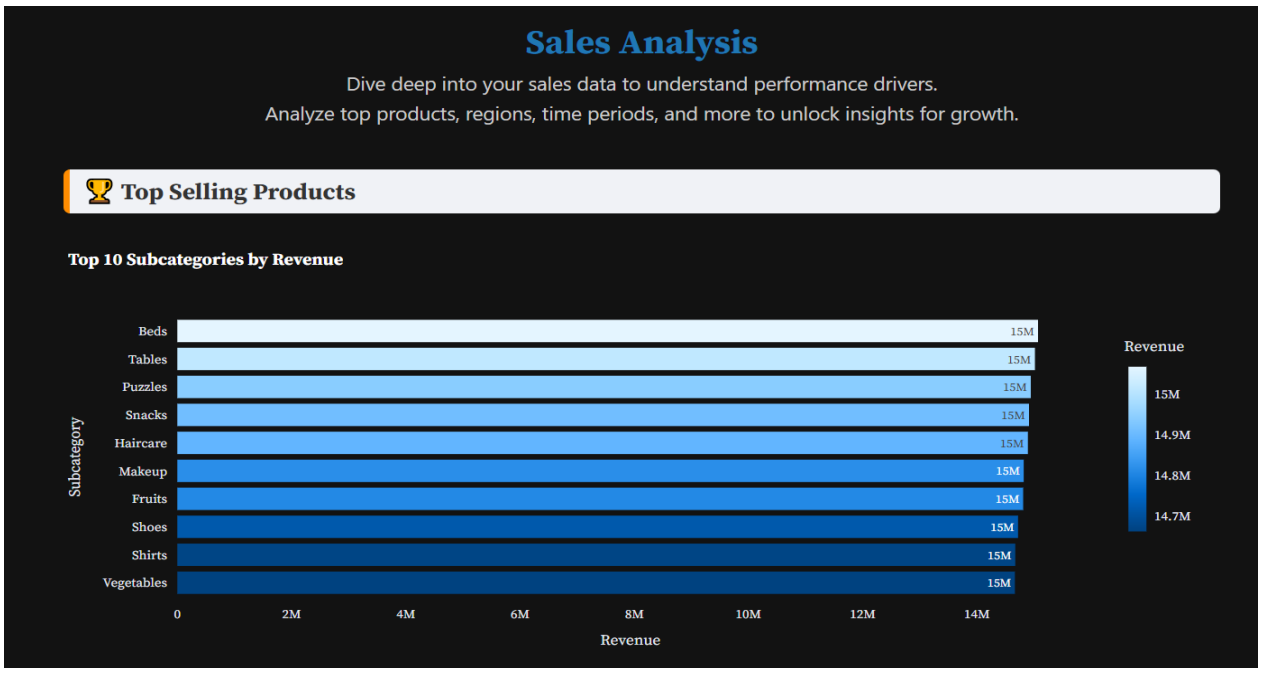


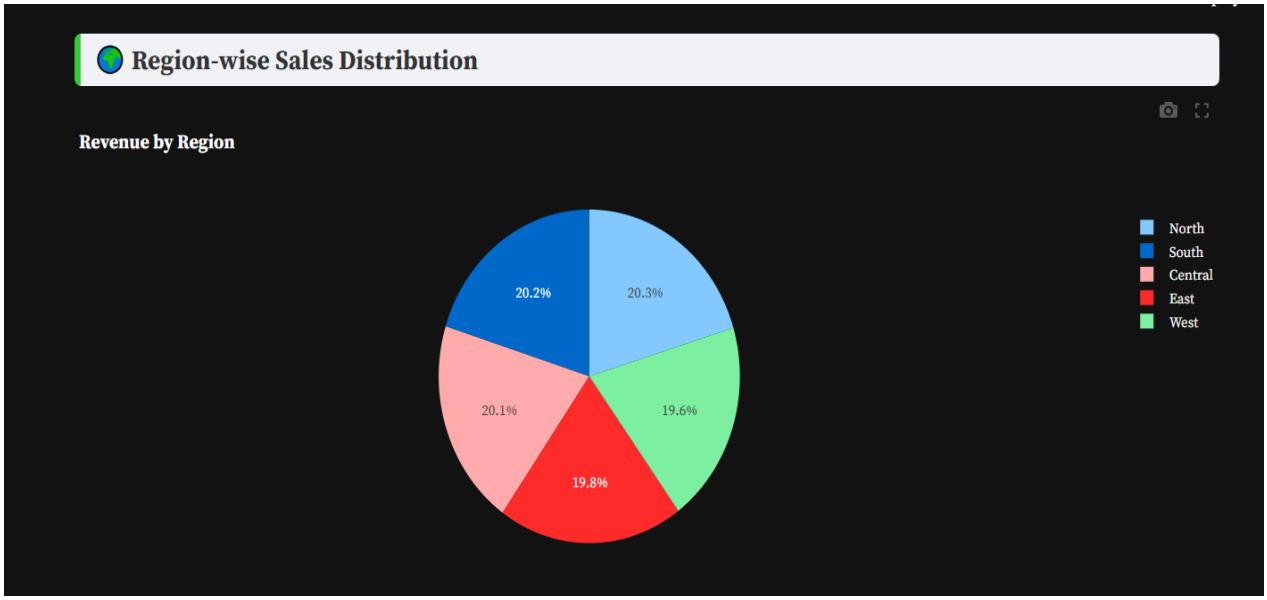
Sales Analysis Module Description

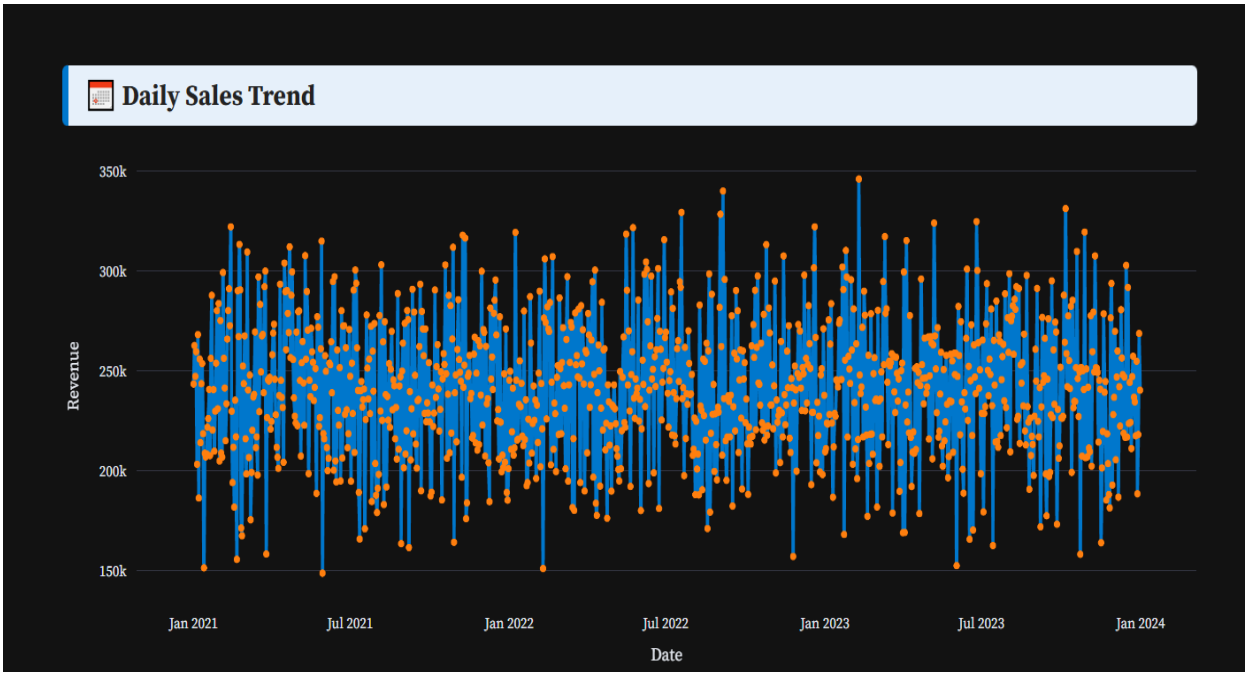
The **Sales Analysis** tab empowers users to explore key sales performance metrics and discover meaningful trends from their data. This module includes a set of insightful visualizations and summaries designed to answer core business questions and highlight opportunities for growth. Here's a breakdown of what it offers:

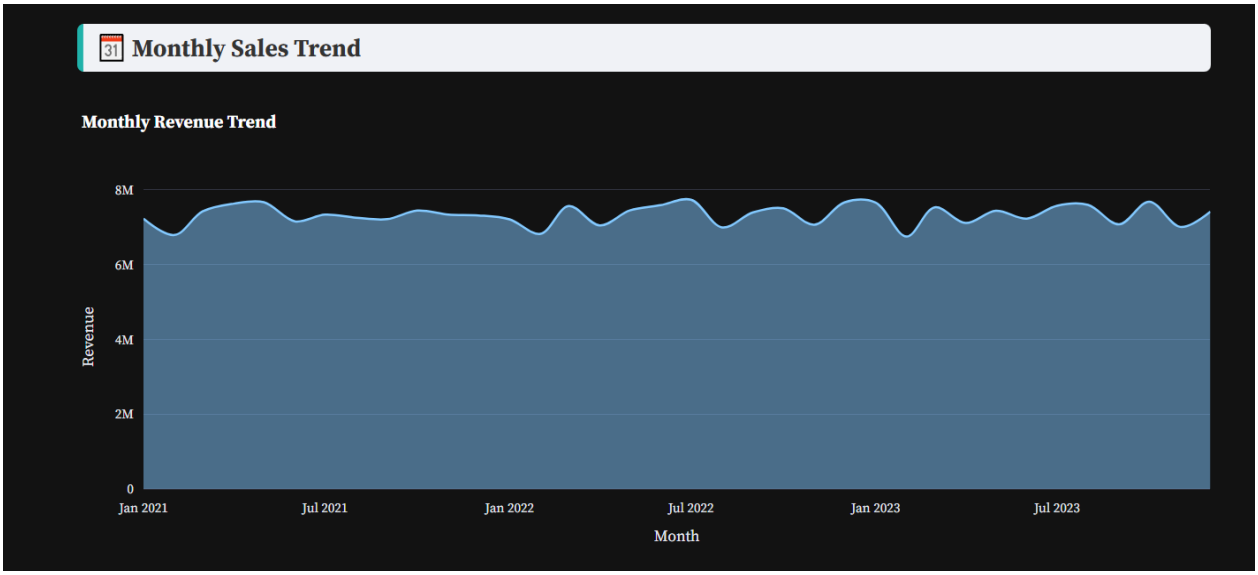
- **Top Selling Products:**
  Quickly identify your highest revenue-generating subcategories to understand what's driving your business.
- **Region-wise Sales Distribution:**
  Explore how your sales are distributed geographically with an interactive pie chart showing revenue contributions from each region.
- **Revenue by Customer Type:**
  Analyze which customer segments (e.g., Retail, Corporate, etc.) are contributing most to your revenue, helping tailor marketing and service strategies.

- **Daily Sales Trend:**
  Visualize sales performance over time to detect daily fluctuations, seasonal spikes, or dips using an engaging line chart.
- **Monthly Sales Trend:**
  Track broader patterns in monthly revenue to make informed strategic decisions and identify long-term growth trends.





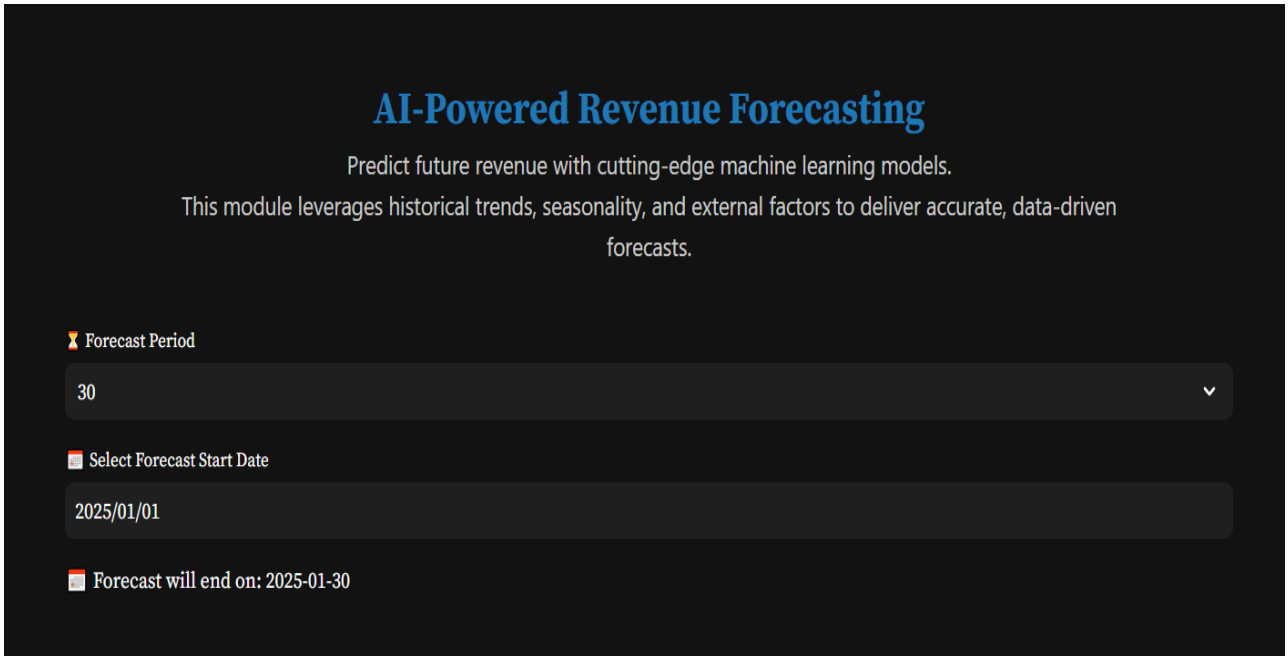

Daily Sales Trend
350k
300k
250k
200k
150k
Revenue
Jan 2021
Jul 2021
Jan 2022
Jul 2022
Jan 2023
Jul 2023
Jan 2024
Date


Monthly Sales Trend
Monthly Revenue Trend
8M
6M
4M
2M
0
Revenue
Jan 2021
Jul 2021
Jan 2022
Jul 2022
Jan 2023
Jul 2023
Month

Revenue Forecasting – AI-Driven Predictions for Business Growth

The **Revenue Forecasting** module empowers businesses to anticipate future revenue with precision by leveraging historical sales trends, seasonal patterns, and machine learning techniques. This section provides a user-friendly interface to:

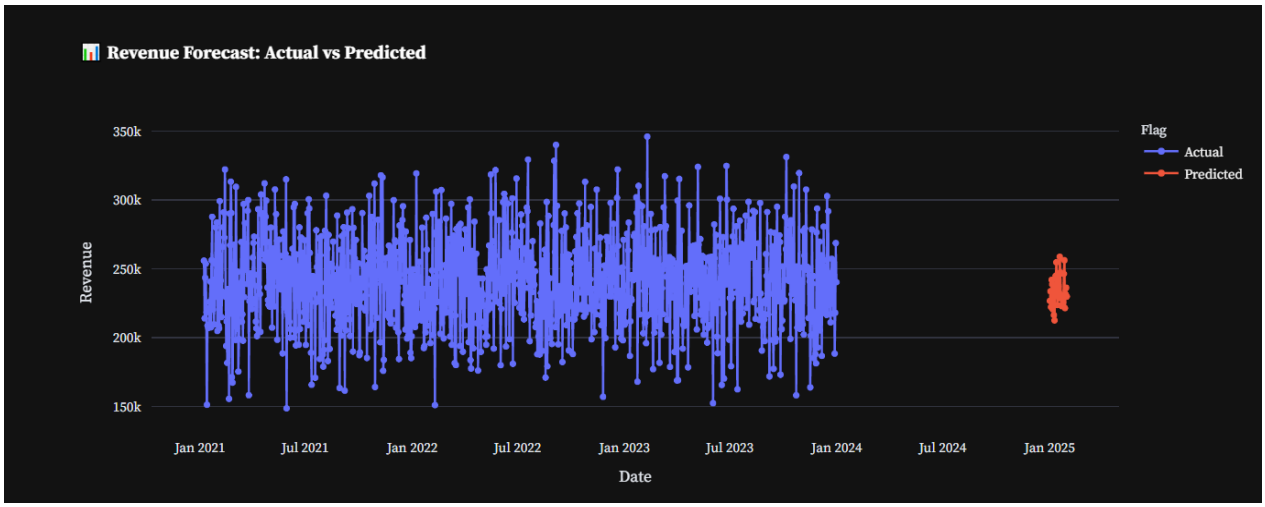
- **Upload and preprocess data** including handling categorical variables and generating useful time-based features (month, day, weekday, lags).
- **Train a Random Forest Regressor** model on historical daily revenue data to capture key patterns and dependencies.
- **Forecast future revenue** for 30, 60, or 90 days ahead, starting from a user-selected forecast date.
- **Visualize predicted vs. actual revenue trends** using interactive line charts built with Plotly for better decision-making.
- **Download forecasted results** in CSV format for further analysis and reporting.



Forecasted Data

	Date	Revenue	Flag
1088	2025-01-01	226710.1602	Predicted
1089	2025-01-02	233669.5138	Predicted
1090	2025-01-03	222231.6837	Predicted
1091	2025-01-04	242017.3801	Predicted
1092	2025-01-05	238742.5083	Predicted
1093	2025-01-06	220571.6427	Predicted
1094	2025-01-07	233877.6454	Predicted
1095	2025-01-08	216330.3992	Predicted
1096	2025-01-09	212542.5647	Predicted
1097	2025-01-10	235551.1819	Predicted

Download Forecasted Data (CSV)

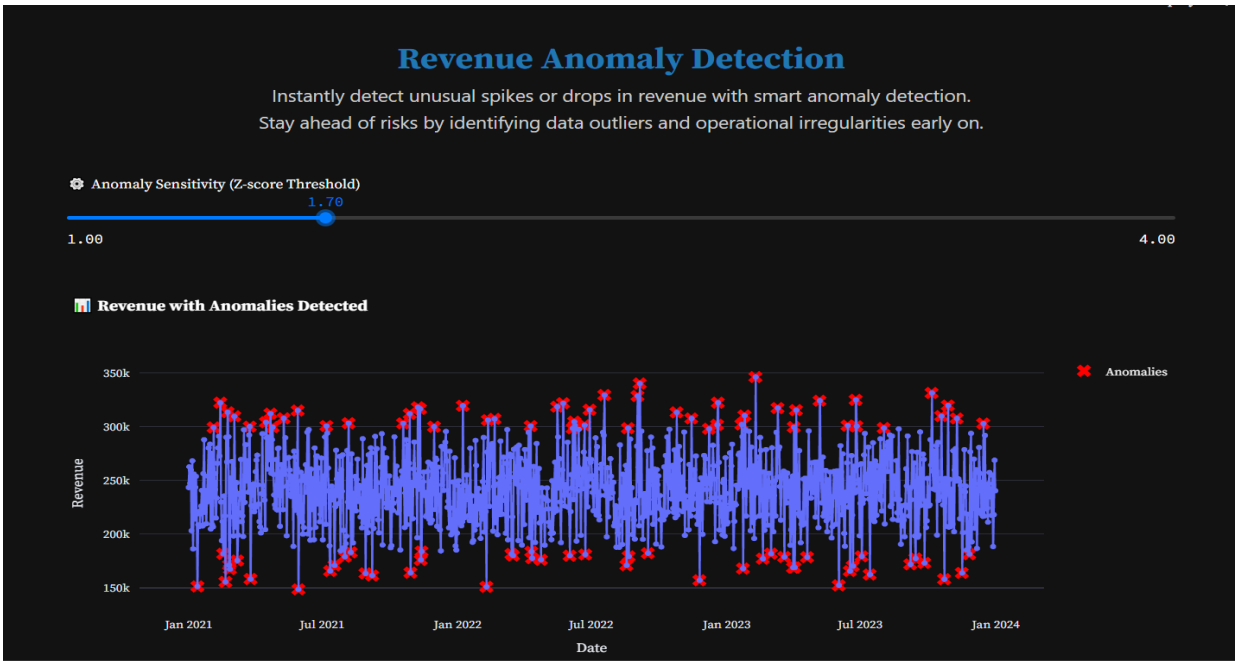


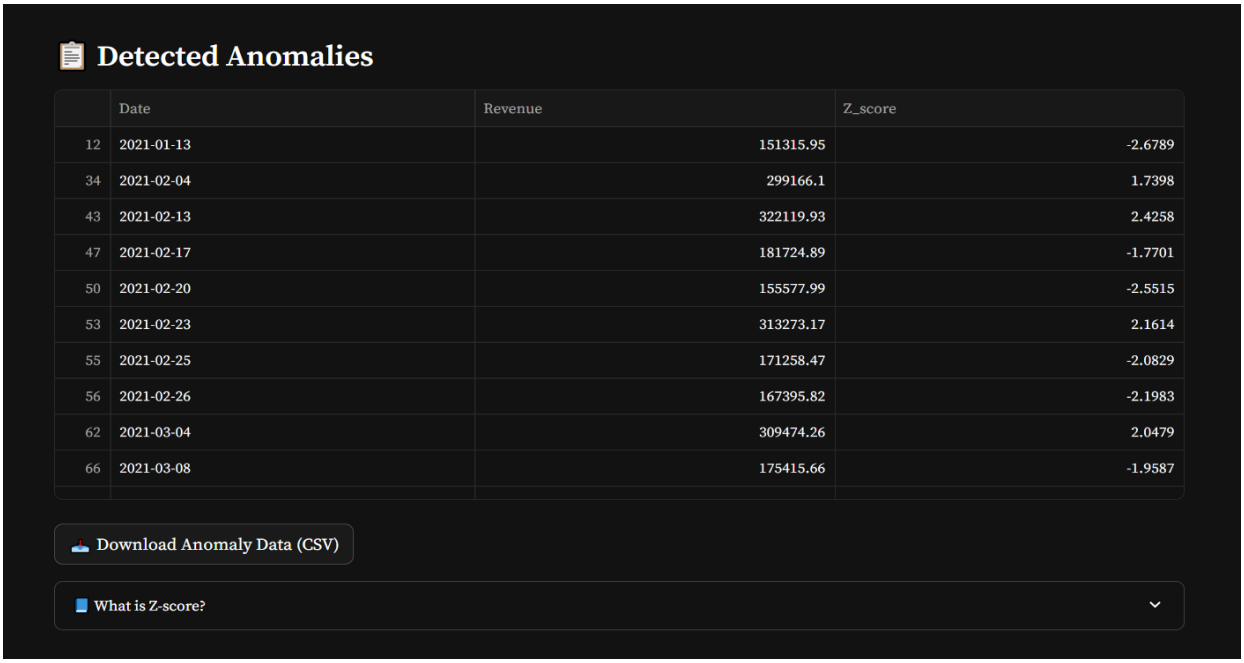
The **Anomaly Detection** module helps businesses quickly identify irregularities in daily revenue patterns using statistical outlier detection. Leveraging **Z-score analysis**, it flags significant deviations—whether spikes or drops—that may indicate operational issues, fraud, or unexpected market behavior.

Users can interactively adjust the **Z-score sensitivity threshold** to control how strict or lenient the detection is, tailoring it to different business scenarios. A clean, interactive line chart overlays red markers on anomalous revenue points, allowing for quick visual insight.

Key features include:

- Daily revenue trend chart with highlighted anomalies
- Adjustable sensitivity using Z-score threshold slider
- Tabular view of detected anomalies with dates and scores
- Downloadable CSV of anomalous records
- An in-app explainer to demystify Z-score-based anomaly detection





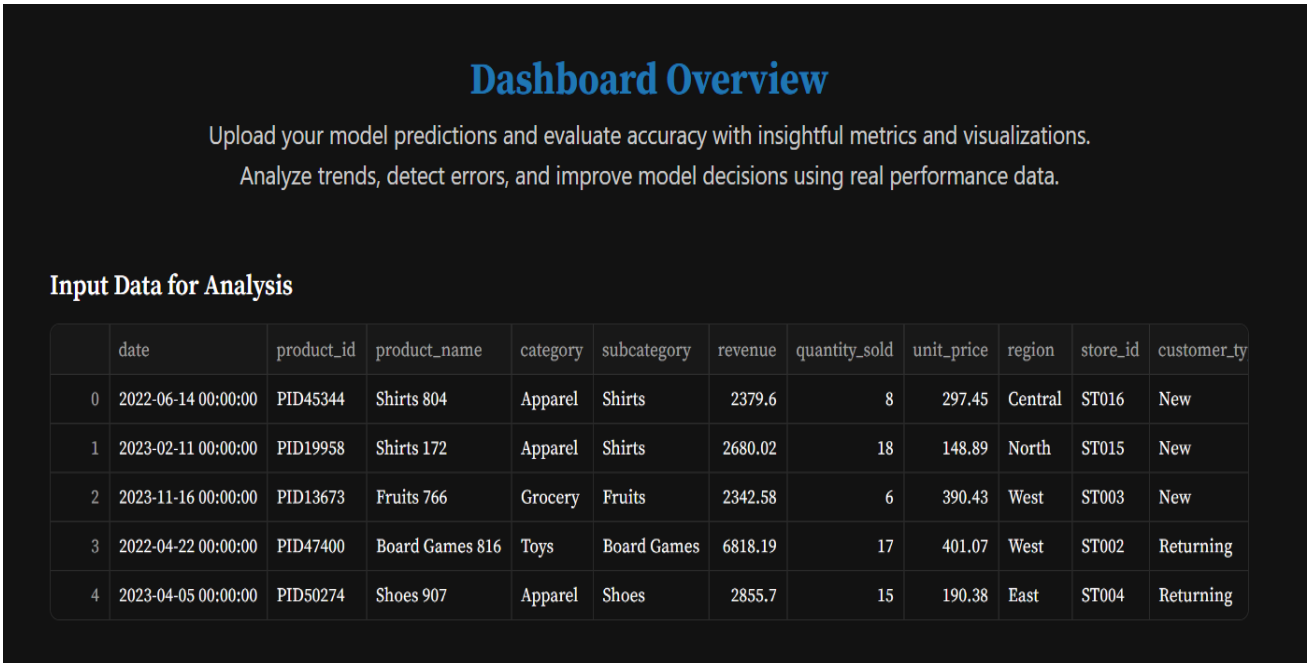
	Date	Revenue	Z_score
12	2021-01-13	151315.95	-2.6789
34	2021-02-04	299166.1	1.7398
43	2021-02-13	322119.93	2.4258
47	2021-02-17	181724.89	-1.7701
50	2021-02-20	155577.99	-2.5515
53	2021-02-23	313273.17	2.1614
55	2021-02-25	171258.47	-2.0829
56	2021-02-26	167395.82	-2.1983
62	2021-03-04	309474.26	2.0479
66	2021-03-08	175415.66	-1.9587

User Dashboard Module Description

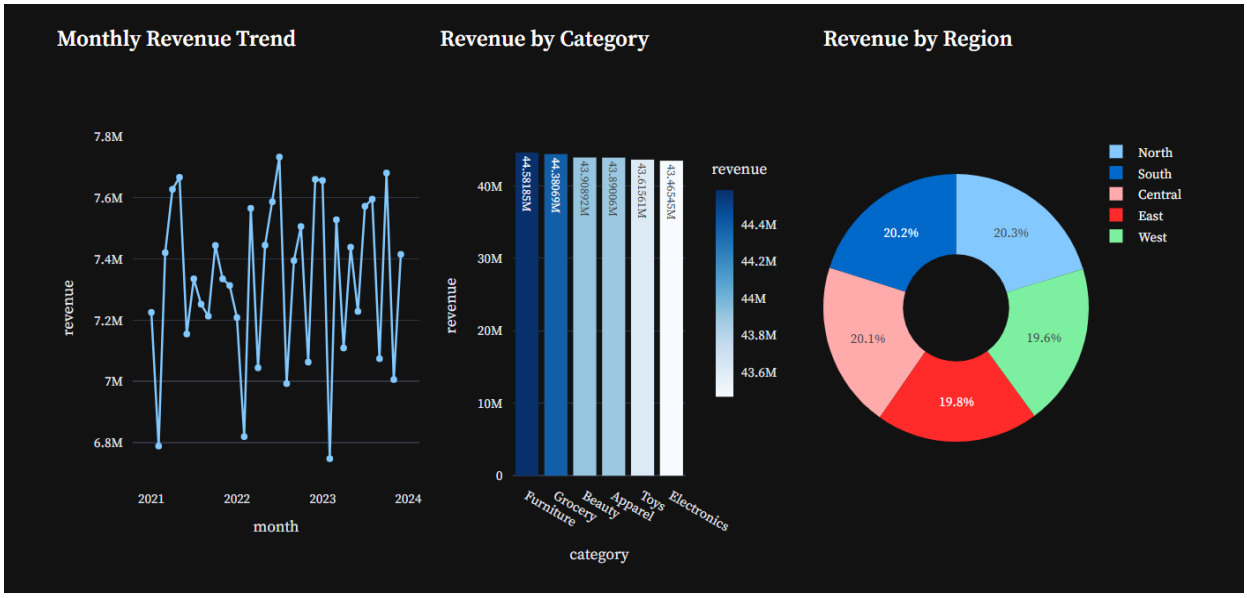
The **User Dashboard** provides a comprehensive view of your uploaded forecasting data, offering powerful visual analytics to evaluate performance and derive actionable insights. Designed for business analysts and decision-makers, this interactive dashboard helps you monitor trends, assess model predictions, and detect anomalies.

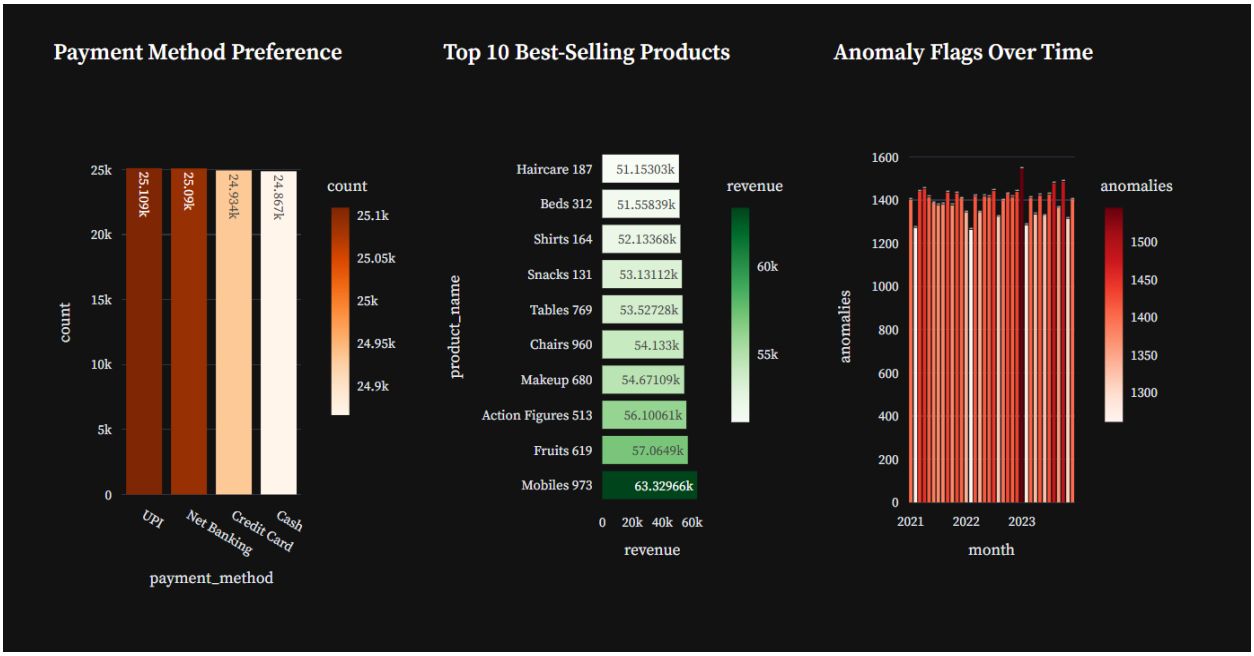
Key Features:

- **Monthly Revenue Trend:** Visualizes how revenue has changed over time, helping to identify seasonal effects and growth patterns.
- **Revenue by Product Category:** Highlights which categories contribute most to your revenue, aiding product-level decision-making.
- **Regional Revenue Distribution:** Displays geographic revenue breakdown with an intuitive pie chart.
- **Payment Method Preference:** Analyzes customer payment trends to optimize checkout strategies and support operations.
- **Top 10 Best-Selling Products:** Lists your highest-revenue products, helping you focus on top performers.
- **Anomaly Flags Over Time:** Shows how often anomalies were detected, helping assess model health and spot unusual months.

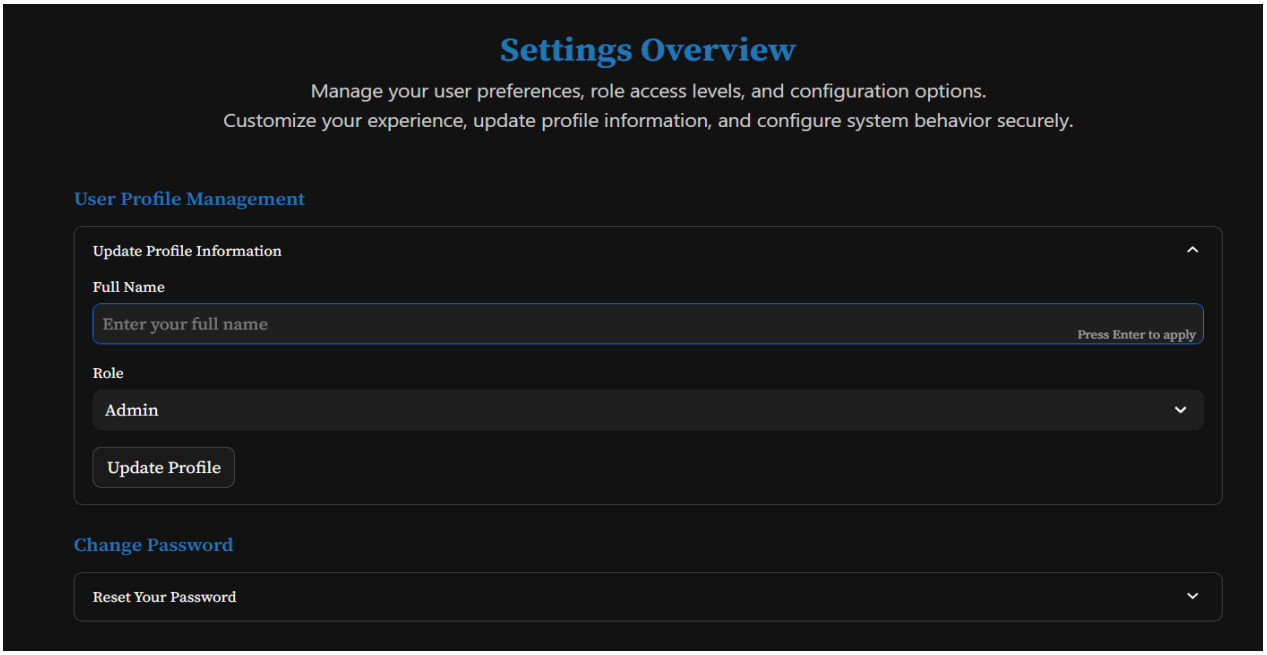


	date	product_id	product_name	category	subcategory	revenue	quantity_sold	unit_price	region	store_id	customer_ty
0	2022-06-14 00:00:00	PID45344	Shirts 804	Apparel	Shirts	2379.6	8	297.45	Central	ST016	New
1	2023-02-11 00:00:00	PID19958	Shirts 172	Apparel	Shirts	2680.02	18	148.89	North	ST015	New
2	2023-11-16 00:00:00	PID13673	Fruits 766	Grocery	Fruits	2342.58	6	390.43	West	ST003	New
3	2022-04-22 00:00:00	PID47400	Board Games 816	Toys	Board Games	6818.19	17	401.07	West	ST002	Returning
4	2023-04-05 00:00:00	PID50274	Shoes 907	Apparel	Shoes	2855.7	15	190.38	East	ST004	Returning


Monthly Revenue Trend
7.8M
7.6M
7.4M
7.2M
7M
6.8M
revenue
2021
2022
2023
2024
month
Revenue by Category
40M
30M
20M
10M
0
revenue
44.5818M
44.3806M
43.9089M
43.8906M
43.6156M
43.4654M
Furniture
Grocery
Beauty
Apparel
Toys
Electronics
category
revenue
44.4M
44.2M
44M
43.8M
43.6M
Revenue by Region
20.2%
20.3%
20.1%
19.6%
19.8%
North
South
Central
East
West


Payment Method Preference
25k
20k
15k
10k
5k
0
count
25.109k
25.09k
24.934k
24.867k
UPI
Net Banking
Credit Card
Cash
payment_method
count
25.1k
25.05k
25k
24.95k
24.9k
Top 10 Best-Selling Products
Haircare 187
Beds 312
Shirts 164
Snacks 131
Tables 769
Chairs 960
Makeup 680
Action Figures 513
Fruits 619
Mobiles 973
51.15303k
51.55839k
52.13368k
53.13112k
53.52728k
54.133k
54.67109k
56.10061k
57.0649k
63.32966k
product_name
0
20k
40k
60k
revenue
revenue
60k
55k
Anomaly Flags Over Time
1600
1400
1200
1000
800
600
400
200
0
anomalies
2021
2022
2023
month
anomalies
1500
1450
1400
1350
1300

The **Settings** tab allows users to manage their profile, update roles, and securely change their password. Customize your account preferences to enhance your experience and maintain account security.



Logout Tab Description:

The **Logout** tab allows users to securely log out of their account, ending the current session. This ensures that sensitive information is protected and that users are required to log in again for future access. It's a simple but crucial step for maintaining privacy and security.

Session State Description:

Session state in Streamlit helps store information across different user interactions within a session. It maintains variables like username, role, and preferences during the user's time in the app. When a user logs in or performs other actions, their session state holds and manages this data until they log out or the session expires, providing a seamless user experience.

Learning and Work Experience

During the development of my project, *AI-Driven Revenue Forecasting and Trend Analysis for Business Growth*, I had the opportunity to explore various aspects of data analytics, machine learning, and application development. The project focused on building an AI-based forecasting system that predicts future revenue trends and provides actionable business insights using historical data.

Key Learnings:

- **AI and Machine Learning:** Gained hands-on experience in implementing machine learning models like **Random Forest Regressor** for accurate revenue forecasting and trend analysis across different business domains.
- **Data Preprocessing:** Developed a robust data pipeline to clean, transform, and prepare time-series business data using **Pandas** and **NumPy**, enabling efficient model training and analysis.
- **Revenue Forecasting System Development:** Built an intelligent forecasting engine that evaluates historical revenue data and provides future revenue predictions, enabling proactive business decision-making.
- **Web Application Development:** Used **Streamlit** to design and deploy a user-friendly web interface, allowing business users (analysts, managers, and decision-makers) to interact with the forecasting system easily.
- **Data Visualization:** Integrated **Matplotlib** and **Plotly** to generate dynamic charts and graphs, offering clear insights into historical trends, predicted values, and anomalies in revenue patterns.

Work Experience:

- **Collaboration:** Worked with business analysts and mentors to ensure the forecasting system aligned with real-world financial planning and strategic growth requirements.
- **Model Development:** Contributed to the development, tuning, and validation of the AI-driven revenue forecasting model to improve prediction accuracy and business relevance.
- **UI Design:** Designed an intuitive and visually appealing interface using **Streamlit**, ensuring an accessible and seamless user experience across all roles.

Salient Features and Achievements

Salient Features:

- **Role-Based Login:** Implemented secure, role-based dashboards for Admins, Analysts, and Managers, each with customized access and features aligned to their responsibilities.
- **Revenue Forecasting Engine:** Developed an AI-powered forecasting system using the **Random Forest algorithm** to accurately predict future revenue based on historical trends.
- **User Authentication & Registration:** Integrated a secure login and registration system with **OTP verification**, **CAPTCHA**, and role selection to ensure data privacy and authorized access.
- **Data Upload & Preprocessing:** Enabled users to upload datasets in various formats (CSV, Excel), with automated preprocessing including handling missing values, date formatting, and outlier detection.
- **Trend Analysis & Visualization:** Incorporated dynamic data visualizations using **Matplotlib** and **Plotly** to display historical revenue trends, anomaly detection, and future forecasts for better decision-making.

Achievements:

- **Improved User Experience:** Designed an intuitive, professional interface using **Streamlit**, ensuring a smooth and engaging experience for business users, analysts, and stakeholders.
- **Optimized Revenue Forecasting:** Refined the forecasting algorithm using **Random Forest**, resulting in enhanced prediction accuracy and reliable financial insights.
- **Scalable System Design:** Built a modular and scalable architecture to support the addition of advanced models, new data sources, and expanded role-based features in the future.
- **Comprehensive Testing:** Performed extensive testing across modules—including data upload, trend analysis, and forecasting—to ensure accuracy, reliability, and seamless functionality.
- **Collaborative Development:** Worked closely with mentors and domain experts to align the forecasting system with real-world business needs and strategic planning.

Future Enhancements

- **Enhanced Data Management:** Improve the current CSV-based data handling by organizing file structures, automating data validation, and optimizing performance for larger datasets.
- **Model Performance Optimization:** Refine the Random Forest forecasting model with advanced hyperparameter tuning, additional feature engineering, and improved evaluation techniques to enhance accuracy.
- **Automated Data Ingestion Pipeline:** Develop an automated pipeline to manage data uploads, preprocessing, and cleaning processes, reducing manual tasks and ensuring a smooth user experience.
- **Advanced Analytics Dashboard:** Expand the existing dashboard with interactive visualizations, such as product-wise revenue trends, growth comparisons, and filterable charts using Plotly and Matplotlib.
- **User Customization Options:** Introduce **customizable user preferences** for different forecast scenarios, such as regional breakdowns, product-specific forecasts, and adjustable prediction periods.
- **Cross-Platform Compatibility:** Enhance the **Streamlit app's compatibility** by ensuring it works seamlessly across various devices (desktop, tablet, mobile) for a wider user base.
- **Integration with External APIs:** Integrate with external data sources, such as **market trends or financial news**, to enrich the forecasting model with up-to-date information for more accurate predictions.

References

- **Streamlit Documentation:**
  Streamlit. *Streamlit Documentation*. https://docs.streamlit.io
- **Plotly Documentation:**
  Plotly. *Plotly Documentation*. https://plotly.com/python/
- **Python Documentation:**
  Python Software Foundation. *Python Official Documentation*. https://docs.python.org/3/
- **Random Forest Algorithm:**
  Understanding Random Forest: The Complete Guide. Towards Data Science. https://towardsdatascience.com/
- **Data Preprocessing Techniques:**
  Data Preprocessing in Python: A Guide to Data Cleaning and Analysis. Towards Data Science. https://towardsdatascience.com/
- **Machine Learning Model Evaluation Metrics:**
  Model Evaluation Metrics for Machine Learning. Towards Data Science. https://towardsdatascience.com/
- **Pandas Documentation:**
  Pandas. *Pandas Documentation*. https://pandas.pydata.org/pandas-docs/stable/
- **Matplotlib Documentation:**
  Matplotlib. *Matplotlib Documentation*. https://matplotlib.org/
- **NumPy Documentation:**
  NumPy. *NumPy Documentation*. https://numpy.org/doc/stable/
- **Time Series Forecasting:**
  Introduction to Time Series Forecasting with Random Forest. Analytics Vidhya. https://www.analyticsvidhya.com/

Appendix

Tools Used:

- **Streamlit** – For UI with tabs and role-based access
- **Pandas** – For data cleaning and processing
- **Scikit-learn (Random Forest)** – For revenue prediction
- **Matplotlib & Plotly** – For graphs and trend visualization

Domain Terms:

- **Revenue Forecasting** – Predicting future income
- **Time-Series** – Data over time
- **Random Forest** – ML model for prediction
- **Preprocessing** – Cleaning and preparing data

Coding Conventions:

- Followed **PEP 8** Python coding style for clean and readable code
- Added **docstrings** to explain functions and modules
- Used **modular approach**: separate functions for each task (upload, preprocess, visualize, forecast)

Sample Dataset Structure:

- **sales_data.csv** – Includes historical sales data, marketing spends, region-wise sales, etc.
- **forecasting_results.csv** – Contains model predictions, historical vs. predicted revenue data, and performance metrics
- **user_data.csv** – Stores user login and role data for admin and other role-based access management

Technologies Used:

- **Frontend & UI** – Streamlit (multi-tab navigation, interactive UI, role-based access)
- **Data Handling** – Pandas, CSV files (cleaning, merging, filtering)
- **Machine Learning** – Random Forest (used for revenue prediction)
- **Visualization** – Matplotlib, Plotly (trend graphs, forecast results)
- **File Handling** – CSV upload and download for datasets and results
- **Authentication** – OTP-secured registration, CAPTCHA, and role-based login system