

SMA_AIRLINE Project Documentation Report

Project Title: SMA_AIRLINE - Exploratory and Analytical Insight on Airline Data

Author: [Your Name]

Date: [Insert Date]

Overview:

This project conducts an in-depth Exploratory Data Analysis (EDA) on airline passenger data. It leverages time-series techniques and statistical visualizations to derive actionable insights, primarily focusing on the Simple Moving Average (SMA) as a forecasting mechanism.

Section-by-Section Breakdown:

1. Library Imports

Relevant Python libraries were imported, including:

- pandas for data manipulation
- matplotlib.pyplot and seaborn for visualizations
- warnings to suppress unnecessary warnings
- numpy for numerical operations

2. Data Loading and Initial Examination

- The dataset was imported using `pd.read_csv()` from a local `.csv` file.
- `head()`, `shape`, and `info()` were used to understand the size and structure of the data.
- Column names and types were validated.
- Null values were checked and confirmed to be absent or minimal.

SMA_AIRLINE Project Documentation Report

3. Data Cleaning & Preprocessing

- Date Conversion: The Month column was converted to datetime format to facilitate time-series operations.
- Indexing: The Month column was set as the dataframe index to align with time-based plots.
- Validation: Ensured all rows were chronologically ordered and no date gaps were present.

4. Exploratory Data Analysis (EDA)

- Line Plot of Passengers over Time: Visualized overall trends and seasonal effects.
- Rolling Statistics: Calculated and plotted the Simple Moving Average (SMA) using a 12-month window.
- Visual Comparison: Original data and SMA were plotted on the same graph to illustrate smoothed trends.

5. Insights & Conclusions

- The passenger count demonstrates a consistent upward trend over the years.
- The 12-month SMA offers a clear view of long-term patterns, helping to reduce noise in the data.
- Seasonality is evident in the form of repeating peaks and valleys, likely tied to travel seasons.

Deliverables:

- Cleaned, structured, and indexed airline passenger dataset.
- SMA-enhanced time-series visualization.
- Visual and statistical summary of trends and insights.

Key Learnings:

- Use of SMA for trend analysis in time-series data.
- Importance of data indexing in time-based analytics.
- Power of simple visualizations in extracting domain-level insights.

SMA_AIRLINE Project Documentation Report

Next Steps:

- Extend analysis using Exponential Moving Average (EMA) or ARIMA.
- Build a dashboard using Plotly or Streamlit for interactive visualization.
- Evaluate forecasting accuracy with metrics like MAPE or RMSE.

End of Report