Project Title:

Data Analytics on IndiGo Business Model

Index:

|  |  |
| --- | --- |
| Topics | Page Number |
| 1. Problem Statement | 2 |
| 2. Problem Description | 2 |
| 3. Objectives | 3 |
| 4. Datasets Used | 3 |
| 5. Key Analysis of Project | 4 |
| 6. Outcome and Benefits | 5 and 6 |
| 7. Data Sources | 6 |
| 8. Tools and Technologies | 7 |
| 9. File Structure of the project | 7 |
| 10. Code Module Structure | 7 |
| 11. Implementation | 8 |
| 12. Output Screenshots | 8,9 and 10 |
| 13. Codes Used in the Project | 10 - 14 |
| 14. Closure of The Project | 14 |

Data Analytics on IndiGo Business Model

* Problem Statement:

Analize the data of Indian Airlines and bring out the following: (DATA ANALYTICS ON **IndiGo** BUSINESS MODEL)

1. Top 3 sectors

2. Which sector in future will see more demand

3. Which airline is more likely to grow by market share

* Problem Description:

As India witnesses rapid growth in air travel, especially with the expansion of regional airports and rising middle-class demand, it becomes essential for airlines to strategically analyse operational data to stay competitive. **IndiGo Airlines**, India’s largest low-cost carrier, has maintained profitability and market leadership by leveraging data analytics across its operations. From optimizing ticket pricing and managing high-yield routes to forecasting demand in emerging sectors, data-driven decisions have helped IndiGo improve efficiency and expand strategically. By analysing passenger trends, load factors, and route performance, the airline identifies profitable routes and adjusts its network accordingly.

This project focuses on analysing the business model(datasets) of **IndiGo Airlines** through data analytics, with the objective of understanding route performance, market trends, and future growth opportunities in the Indian aviation sector. Using real-world air traffic data, the project identifies the **top-performing sectors (routes)** by passenger volume, and applies forecasting techniques to predict **emerging high-demand routes**.

The project also evaluates **market share trends** among major Indian carriers to determine which airline is best positioned for future growth. By simulating and visualizing route-level data, the analysis demonstrates how IndiGo and similar airlines can make **data-driven decisions** to improve profitability, expand strategically, and enhance operational efficiency.

This work showcases the role of data analytics in aviation and offers practical insights into how airlines can align their route planning with actual demand patterns using historical data, predictive modelling, and visual storytelling tools.

* Objectives:

1. Identify the Top 3 Sectors (Routes) for IndiGo:
2. Identifying the Busiest months of every Routes
3. Year-wise Growth of IndiGo's Market Share
4. Seasonal Travel Peaks
5. How Airlines are coming up with New Routes? What Data and What Data Analysis is helping them to arrive at the introduction of the new Routes. (Top 5 growing routes or in demand routes)

* Datasets Used:

|  |  |
| --- | --- |
| **\***flight\_data.csv | - Monthly passenger traffic on each route (Sector- wise) |

The datasets used in this project were sourced from the **Indian Aviation Traffic** repository on GitHub, which compiles and shares publicly available data released by the **Directorate General of Civil Aviation (DGCA), Government of India**. This repository includes monthly records of airline-wise passenger traffic, sector-level city pair movements, and market share data of domestic airlines. From this repository, two datasets — city.csv and carrier.csv — were extracted, cleaned, and transformed into flight\_data.csv and market\_share.csv to align with the project’s analysis requirements.

|  |  |
| --- | --- |
| **\***market\_share.csv | * Monthly market share (%) of all airlines including IndiGo |

* Key Analysis of The Project:

**\* Top Performing Routes (Sectors):**

* Identified the **top 3 busiest flight routes** based on total passenger volume.
* These routes represent high-demand travel corridors which are essential for airline route planning and prioritization.

**\* Fastest Growing Sectors:**

* Calculated the **growth percentage in passenger traffic** over time for each sector.
* Revealed the **top 5 routes with the highest growth**, indicating emerging demand and future expansion opportunities.
* **IndiGo uses** such analytics to predict such **Tier 2 city routes** with growing demand, as there won’t be **any other competition** in such routes and thus, would help in expecting full **Profitability** in such routes

**\* Busiest Month per Route:**

* Found the **month with the highest passenger count** for every individual route.
* Useful for understanding **route-specific peak travel seasons** to optimize fleet utilization and scheduling.

**\* Seasonal Travel Peaks:**

* Aggregated all flight data by **calendar month** to discover consistent **seasonal travel patterns** (e.g., summer or festive spikes).
* Supports **dynamic pricing**, promotional campaigns, and **staff allocation planning**.

**\* IndiGo Market Share Trends:**

* Filtered and analyzed IndiGo’s monthly market share.
* Visualized its **market dominance and fluctuation** over time, which helps assess brand positioning and business strategy.
* Outcome and Benefits:

**Outcomes:**

This project successfully transformed raw flight and market data into actionable business insights. By leveraging Python-based data analysis and visualization techniques, we achieved the following:

1. **Identified Top Performing Routes**
   * Using passenger data, the most popular flight sectors (routes) were determined based on total passenger volume.
   * These routes represent the most heavily trafficked air corridors in India, particularly relevant for IndiGo and other airlines for optimizing scheduling and promotions.
2. **Analysed Fastest Growing Routes**
   * By computing percentage growth in passengers over time, the analysis revealed emerging sectors with increasing demand.
   * This helps identify investment opportunities or potential for introducing new services or increasing frequency.
   * This helps airlines like IndiGo to predict new routes (between smaller cities) which has high demand).
3. **Detected Busiest Months for Each Route**
   * The busiest travel month for each flight sector was extracted, providing insight into peak seasonality on a per-route basis.
   * Airlines can use this data to dynamically adjust fleet allocation, staffing, and maintenance schedules.
4. **Uncovered Seasonal Travel Peaks (Across All Routes)**
   * Aggregated data across years highlighted high-demand travel months such as May, October, and December.
   * This is critical for dynamic pricing strategies, marketing campaigns, and airport resource planning.
5. **Tracked IndiGo’s Market Share Over Time**
   * A clear view of how IndiGo’s share in the domestic market changed monthly.
   * Useful for competitive benchmarking and evaluating performance trends.

**Business Benefits:**

1. **Strategic Route Planning**
   * Airlines can focus on profitable and high-growth routes for expansion, while re-evaluating underperforming ones.
2. **Optimized Resource Allocation**
   * Ground crew, flight crew, aircraft, and airport gates can be assigned based on peak months and busy routes to reduce operational bottlenecks.
3. **Dynamic Pricing & Promotions**
   * Airlines can implement seasonal pricing models based on demand patterns and busiest travel periods.
4. **Customer-Centric Planning**
   * Understanding when and where passengers travel most allows for tailored offers, frequent flyer benefits, and targeted communications.
5. **Market Intelligence for Competitive Advantage**
   * IndiGo and competing carriers can benchmark their market performance and adjust strategies in response to shifts in demand or competitive pressure.
6. **Data-Driven Decision Making**
   * The project emphasizes the role of analytics in airline management, supporting informed decisions at both strategic and operational levels.

* Data Sources:

Data used in this project were extracted from the following sources:

* **Indian Aviation Traffic** repository from GitHub(<https://github.com/Vonter/india-aviation-traffic/tree/main/aggregated/domestic>)
* **DGCA** India (Directorate General of Civil Aviation)
* data.gov.in
* **Airline websites** (IndiGo, Air India, etc.)
* Tools & Technologies Used:

Programming Language: Python

Coding Platform Used: VS Code

Libraries Used: pandas, matplotlib, seaborn, NumPy, scikit-learn/statsmodels (for forecasting)



* File Structure:

indigo\_project\_hackathon\_zipfile/

│

├── code/

│ ├── main.py # Main script to run the project

│ ├── analysis.py # Contains functions for extracting insights

│ └── visualization.py # Contains functions for plotting charts

│

├── datasets/

│ ├── flight\_data.csv # Dataset with sector-wise passenger count per month

│ └── market\_share.csv # Dataset showing monthly market share of airlines

### Code **Module Structure Diagram**

┌────────────────┐

│ main.py │

│ (entry point) │

└──────┬─────────┘

│

┌───────────┴────────────┐

▼ ▼

┌──────────────┐ ┌────────────────┐

│ analysis.py │ │ visualization.py │

│ - Business │ │ - Graphs & Plots │

│ logic │ │ - Aesthetics │

└──────────────┘ └──────────────────┘

* Implementation:

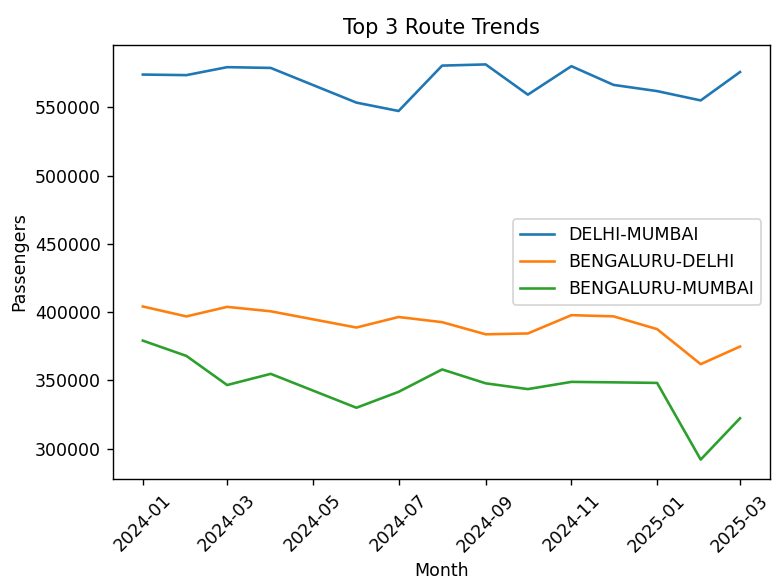
This project was implemented using **Python** with a modular and object-oriented approach. It is divided into three main files:

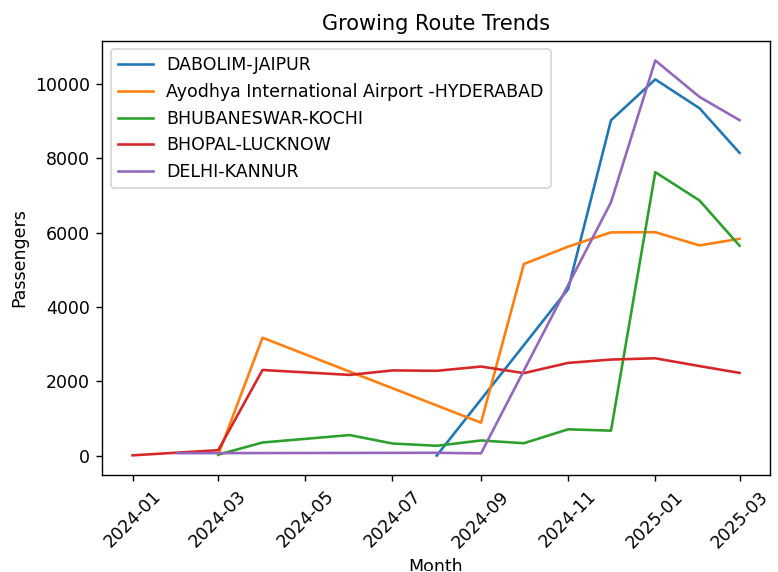
* **main.py**: A CLI-based interface that allows users to select various analysis options. It connects to analysis and visualization modules.
* **analysis.py**: Contains two classes – FlightDataAnalyzer for flight route analysis and IndiGoAnalyzer for market share analysis.
* **visualization.py**: Includes plotting functions using Matplotlib for visual insights such as bar charts and line graphs.

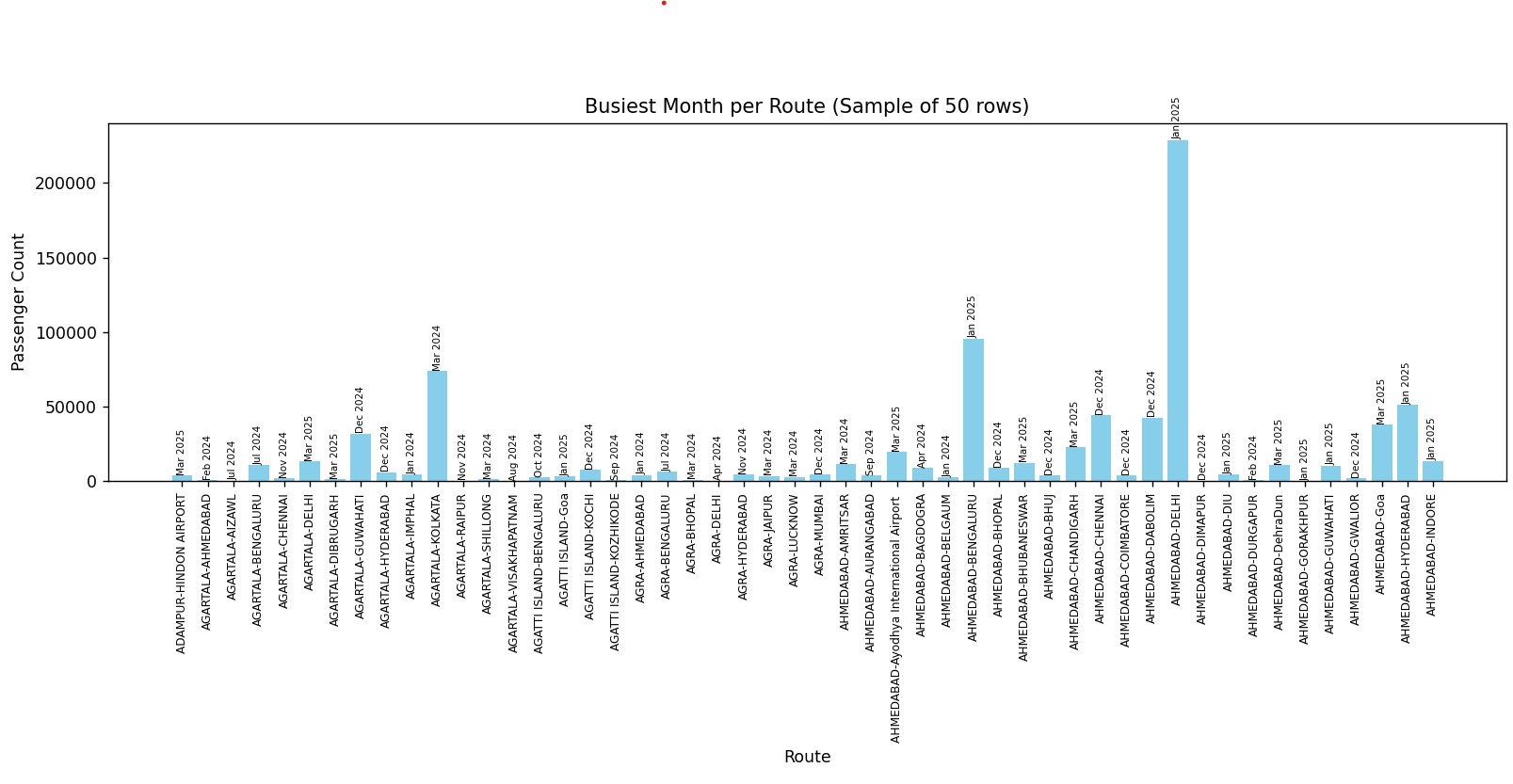
The code loads flight and market share datasets, processes them with pandas, and displays insights using matplotlib.

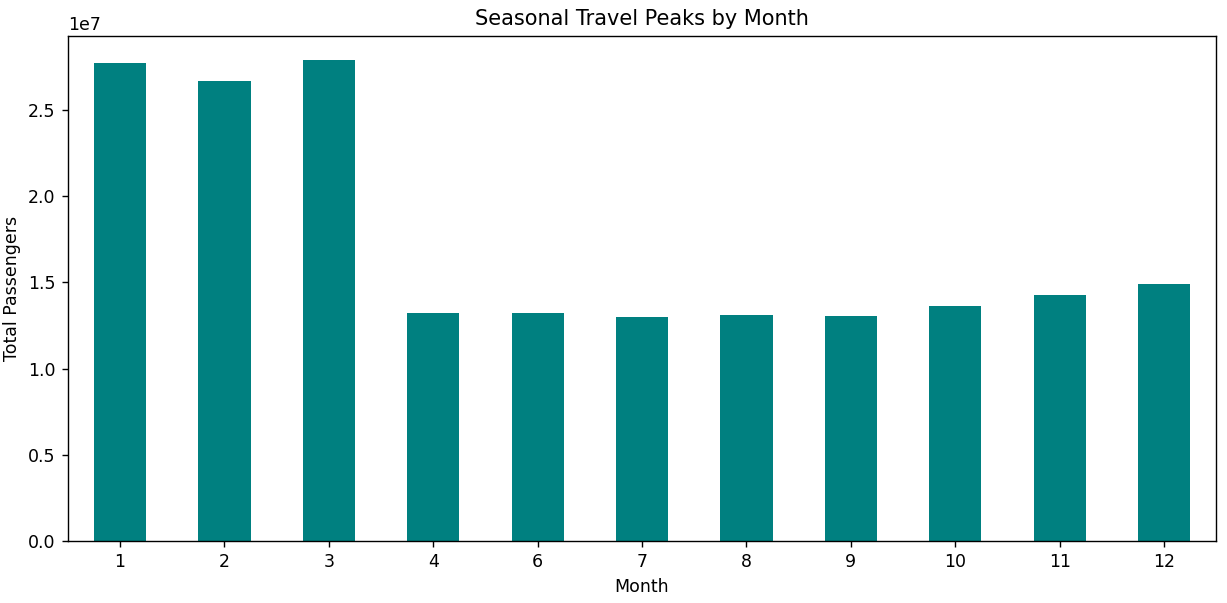
The menu lets users choose actions like viewing top routes, growing sectors, peak travel months, and IndiGo’s market trends.

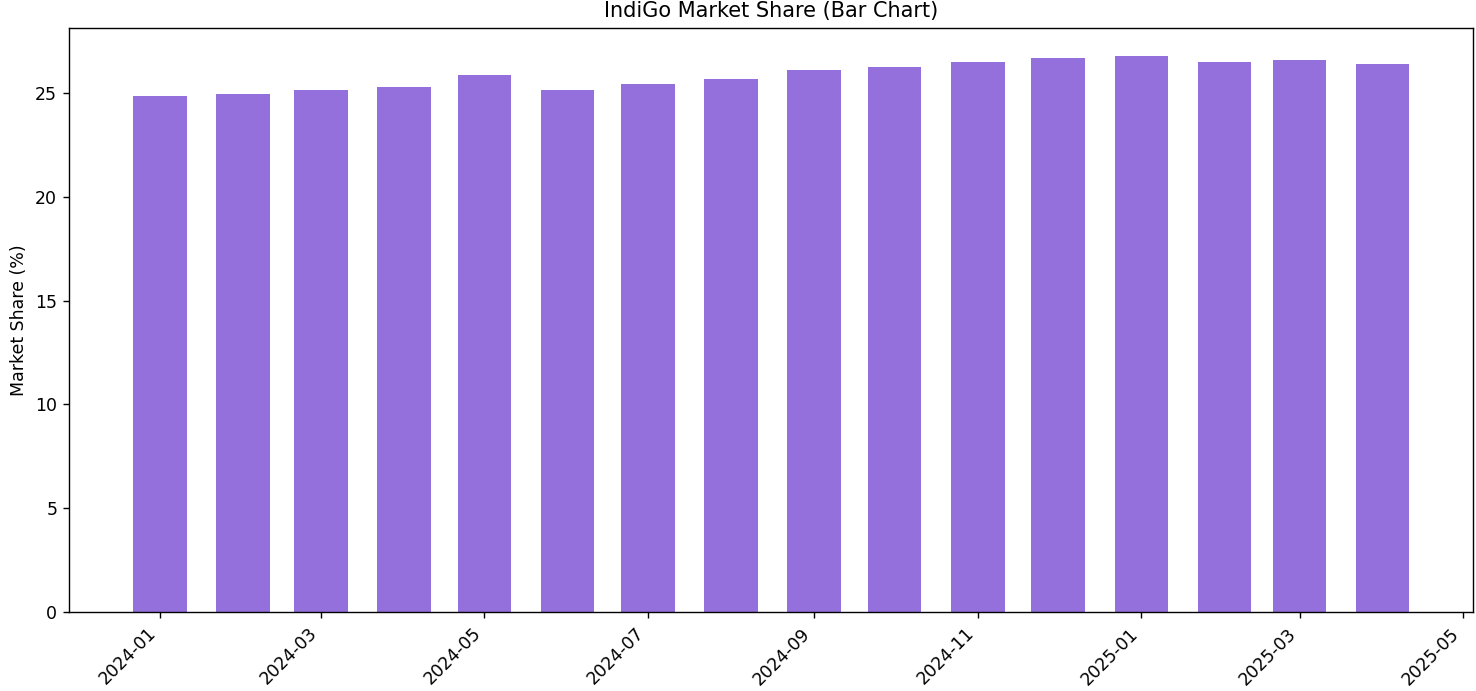
* Output Screenshots:











* Codes used In the Project

1 .Main  
from analysis import FlightDataAnalyzer, IndiGoAnalyzer

from visualization import (

    plot\_top\_sectors,

    plot\_growing\_sectors,

    plot\_market\_share,

    plot\_busiest\_months,

    plot\_seasonal\_peaks

)

def main():

    flight = FlightDataAnalyzer("datasets/flight\_data.csv")

    indigo = IndiGoAnalyzer("datasets/market\_share.csv")

    while True:

        print("\n--- Flight Analysis ---")

        print("1. Top 3 Routes with Most Passengers")

        print("2. Top 5 Growing Routes")

        print("3. Busiest Month per Route")

        print("4. Seasonal Travel Peaks")

        print("5. IndiGo Market Share Trend")

        print("0. Exit")

        choice = input("Enter your choice: ")

        if choice == "1":

            top\_sectors = flight.get\_top\_sectors()

            plot\_top\_sectors(flight.df, top\_sectors)

        elif choice == "2":

            growing = flight.get\_growing\_sectors()

            plot\_growing\_sectors(flight.df, growing)

        elif choice == "3":

            busiest\_df = flight.get\_busiest\_months()

            plot\_busiest\_months(busiest\_df)

        elif choice == "4":

            seasonal\_data = flight.get\_seasonal\_peaks()

            plot\_seasonal\_peaks(seasonal\_data)

        elif choice == "5":

            plot\_market\_share(indigo.df)

        elif choice == "0":

            print("Exiting.")

            break

        else:

            print("Invalid option. Try again!")

if \_\_name\_\_ == "\_\_main\_\_":

    main()

2. Analysis Block

import pandas as pd

class FlightDataAnalyzer:

    def \_\_init\_\_(self, filepath):

        self.df = pd.read\_csv(filepath)

        self.df['Month'] = pd.to\_datetime(self.df['Year'].astype(str) + '-' + self.df['Month'].astype(str) + '-01')

        self.df = self.df[['Month', 'Sector', 'Passengers']]

    def get\_top\_sectors(self):

        top = self.df.groupby("Sector")["Passengers"].sum().sort\_values(ascending=False).head(3)

        print("Top 3 routes with most passengers:")

        print(top)

        return top.index.tolist()

    def get\_growing\_sectors(self):

        growth = []

        for s in self.df["Sector"].unique():

            data = self.df[self.df["Sector"] == s].sort\_values("Month")

            if len(data) >= 6:

                start = data.iloc[0]["Passengers"]

                end = data.iloc[-1]["Passengers"]

                if start > 0:

                    change = ((end - start) / start) \* 100

                    growth.append((s, change))

        growth.sort(key=lambda x: x[1], reverse=True)

        print("Top 5 growing routes:")

        for s in growth[:5]:

            print(f"{s[0]} - {round(s[1], 2)}%")

        return [s[0] for s in growth[:5]]

    def get\_busiest\_months(self):

        busiest = self.df.loc[self.df.groupby("Sector")["Passengers"].idxmax()]

        print("Busiest month for each route:")

        print(busiest[["Sector", "Month", "Passengers"]])

        return busiest

    def get\_seasonal\_peaks(self):

        self.df["Month\_Num"] = self.df["Month"].dt.month

        seasonal = self.df.groupby("Month\_Num")["Passengers"].sum()

        return seasonal

class IndiGoAnalyzer:

    def \_\_init\_\_(self, filepath):

        df = pd.read\_csv(filepath)

        df['Month'] = pd.to\_datetime(df['Month'])

        self.df = df[df['Airline'].str.lower() == 'indigo'].sort\_values("Month")

    def print\_recent\_share(self):

        print("Recent IndiGo market share:")

        print(self.df[["Month", "Market\_Share"]].tail(10))

3. Visualization Block

import matplotlib.pyplot as plt

def plot\_top\_sectors(df, sectors):

    for s in sectors:

        data = df[df["Sector"] == s]

        monthly = data.groupby("Month")["Passengers"].sum()

        plt.plot(monthly.index, monthly.values, label=s)

    plt.title("Top 3 Route Trends")

    plt.xlabel("Month")

    plt.ylabel("Passengers")

    plt.legend()

    plt.xticks(rotation=45)

    plt.tight\_layout()

    plt.show()

def plot\_growing\_sectors(df, sectors):

    for s in sectors:

        data = df[df["Sector"] == s]

        monthly = data.groupby("Month")["Passengers"].sum()

        plt.plot(monthly.index, monthly.values, label=s)

    plt.title("Growing Route Trends")

    plt.xlabel("Month")

    plt.ylabel("Passengers")

    plt.legend()

    plt.xticks(rotation=45)

    plt.tight\_layout()

    plt.show()

def plot\_market\_share(df):

    df["Share"] = df["Market\_Share"].str.replace('%', '').astype(float)

    plt.figure(figsize=(12, 6))

    plt.bar(df["Month"], df["Share"], color="mediumpurple", width=20)

    plt.title("IndiGo Market Share (Bar Chart)")

    plt.xlabel("Month")

    plt.ylabel("Market Share (%)")

    plt.xticks(rotation=45, ha='right')

    plt.tight\_layout()

    plt.show()

def plot\_busiest\_months(df):

    df = df.head(50)

    sectors = df["Sector"]

    passengers = df["Passengers"]

    months = df["Month"].dt.strftime("%b %Y")

    plt.figure(figsize=(13, 6))

    bars = plt.bar(sectors, passengers, color="skyblue")

    plt.title("Busiest Month per Route (Sample of 50 rows)")

    plt.xlabel("Route")

    plt.ylabel("Passenger Count")

    plt.xticks(rotation=90, ha='center', fontsize=7)

    for bar, month in zip(bars, months):

        height = bar.get\_height()

        plt.text(bar.get\_x() + bar.get\_width() / 2, height + 200, month,

                 ha='center', va='bottom', fontsize=6, rotation=90)

    plt.tight\_layout()

    plt.show()

def plot\_seasonal\_peaks(seasonal\_series):

    plt.figure(figsize=(10, 5))

    seasonal\_series.plot(kind='bar', color='teal')

    plt.title("Seasonal Travel Peaks by Month")

    plt.xlabel("Month")

    plt.ylabel("Total Passengers")

    plt.xticks(rotation=0)

    plt.tight\_layout()

    plt.show()

### **Closure of the Project:**

This project successfully delivers a data-driven analysis of the Indian aviation sector with a specific focus on flight route performance and IndiGo's market trends. Through the integration of Python, pandas, and matplotlib, the system offers clear insights into top-performing routes, seasonal travel patterns, and the growth trajectory of airlines. The modular, menu-driven CLI interface makes it user-friendly and interactive.

In summary, the project combines practical data analysis with effective visual communication, helping stakeholders make informed decisions about route planning, seasonal trends, and competitive positioning in the airline industry