# Pakistan Airlines & Airports Data Analysis 🛩️

A comprehensive data analysis project exploring Pakistan's aviation industry through airlines and airports performance data using Python and SQL.

## 📋 Project Overview

This project analyzes Pakistan's aviation sector by examining passenger traffic, cargo operations, and airline performance patterns. The analysis provides insights into market dynamics, growth trends, and operational efficiency across Pakistani airlines and airports.

## 🎯 Objectives

* Identify Pakistan's busiest airports and top-performing airlines
* Analyze domestic vs international traffic patterns
* Track growth trends and market share distribution
* Examine cargo and mail operations efficiency
* Understand aircraft utilization patterns

## 🗂️ Project Structure

Pakistan-Aviation-Analysis/

│

├── datasets/

│ ├── airlines\_data.csv # Airlines performance data

│ └── airports\_data.csv # Airports traffic data

│

├── Pakistan\_Aviation\_Analysis.ipynb # Complete analysis notebook

└── README.md # Project documentation

## 📊 Datasets Description

### Airlines Dataset

* **Records**: [575 rows]
* **Time Period**: [From 2006 - 2024]
* **Key Metrics**: Passenger traffic, cargo operations, mail services, hub classifications

### Airports Dataset

* **Records**: [605 rows]
* **Time Period**: [From 2006 - 2024]
* **Key Metrics**: Aircraft movements, passenger volumes, cargo handling, domestic/international split

## 🔧 Technologies Used

* **Python Libraries**:
  + pandas - Data manipulation and analysis
  + numpy - Numerical computations
  + matplotlib - Data visualization
  + seaborn - Statistical visualizations
  + mysql-connector-python - Database connectivity
* **Database**: MySQL Workbench
* **Environment**: Jupyter Notebook

## 📈 Key Analysis Questions Explored

### **Airlines Analysis**

**.** Which airlines carry the most passengers and their market share?  
**.** Do airlines focus more on domestic vs international routes?  
**.** How do passenger embarked vs disembarked numbers compare?  
**.** Which airlines are most efficient in cargo operations?  
**.** Is Pakistan importing or exporting more goods by air?  
**.**  How are mail volumes trending over the years?

### Airports Analysis

**.** What are Pakistan's top 10 busiest airports?  
**.**  How do domestic and international passenger numbers compare over years?  
**.**  Which airports handle more cargo relative to passengers?  
**.** What are the aircraft utilization patterns (commercial vs non-commercial)?  
**.** Which airports show the highest year-over-year growth rates?  
**.** What airports are approaching capacity limits?

## 🚀 Getting Started

### Prerequisites

pip install pandas numpy matplotlib seaborn mysql-connector-python jupyter

### Setup Instructions

**Clone the repository**

1. git clone https://github.com/WardaIjaz/Pakistan-Aviation-Analysis.git

cd Pakistan-Aviation-Analysis

**Database Setup**

* + Install MySQL Workbench
  + Create a new database schema
  + Update database credentials in the notebook

**Run the Analysis**

jupyter notebook Pakistan\_Aviation\_Analysis.ipynb

## 📋 Analysis Workflow

1. **Database Connection** - Connect Python to MySQL Workbench
2. **Data Understanding** - Explore dataset structure and basic statistics
3. **Data Cleaning** - Handle missing values, outliers, and data type conversions
4. **Exploratory Data Analysis (EDA)** - Generate insights through visualizations
5. **Key Findings** - Summarize actionable insights and trends

## 🔍 Key Insights

### Market Leaders

* [Top airline by passenger volume]
* [Market share of leading airlines]
* [Busiest airports by passenger traffic]

### Growth Trends

* [Year-over-year growth patterns]
* [Domestic vs international traffic trends]
* [Cargo vs passenger operations growth]

### Operational Efficiency

* [Airlines with best cargo-to-passenger ratios]
* [Airports with highest aircraft utilization]
* [Import vs export balance findings]

## 📊 Sample Visualizations

The analysis includes various visualization types:

* Bar charts for ranking and comparisons
* Line charts for trend analysis
* Pie charts for market share distribution

## 🔮 Future Enhancements

* Interactive dashboard development with Power BI
* Predictive modeling for passenger demand forecasting
* Route optimization analysis
* Seasonal pattern analysis
* Comparison with regional aviation markets

## 📝 Conclusions

**.** Most of Pakistan's air travel happens through only 3 major airports, meaning these airports might face capacity problems while many smaller cities could attract much more passengers if properly developed .

****.** Pakistan has successfully built an international aviation hub but failed to develop its domestic market, representing a critical gap in aviation strategy and economic opportunity.**

****.** Pakistani airlines are doing well overall, but the country could attract many more foreign visitors by working with more airlines that focus on bringing people into Pakistan rather than just taking them out.**

****.** Pakistan's aviation sector demonstrates remarkable resilience and growth potential, making it an attractive market for long-term investment despite short-term volatilities**

****.**Pakistan International Airlines (PIA): **55.9% domestic**, **44.1% international** - slightly domestic-focused**

**Air Blue: **55.2% domestic,** **44.8% international** - similar domestic preference**

**Shaheen Air: **50.9% domestic,** **49.1% international** - most balanced operation\*\***

**These three airlines represent Pakistan's most resilient carriers, with diversified route portfolios that can weather both domestic economic challenges and international travel disruptions.**

## 📧 Contact

**Your Name** - wardaijaz2003@gmail.com

Project Link: [https://github.com/wardaijaz/Pakistan-Aviation-Analysis](https://github.com/yourusername/Pakistan-Aviation-Analysis)

⭐ **If you found this analysis helpful, please give it a star!** ⭐