Maheshvari A S

Week-wise Implementation Plan

Milestone 1: Data Foundation and Cleaning

Week 1: Project Initialization and Dataset Setup

1. Define goals, KPIs, and workflow:

Goal :

**Project Goals – Air Fly Insights**

* Analyse airline flight data to identify patterns in delays and cancellations.
* Measure overall performance of airlines and airports using clear KPIs.
* Discover seasonal and time-based trends affecting flight operations.
* Compare airlines and routes to find the most reliable and the most delay-prone.
* Understand key reasons behind delays and cancellations.
* Present insights in an easy-to-read **Power BI dashboard**.
* Provide a final report and presentation that summarize findings for decision-makers.

KPIs :

**High-Level KPIs (Executive Dashboard)**

These are the “must-have” top metrics:

1. **Total Flights Operated**
2. **% Delayed Flights** = (Flights with ArrDelay > 0) ÷ Total Flights
3. **% Cancelled Flights**
4. **Average Delay (min)** (overall + arrival/departure separately)
5. **Top 5 Airlines by On-Time Performance**
6. **Top 5 Airports by Congestion (avg delay minutes)**
7. **Seasonal Trend of Delays** (Monthly trend line)
8. **Cancellation Reasons Share** (Carrier, Weather, NAS, Security)

**Detailed Operational KPIs (Analyst/Manager view)**

**Flight Performance**

* **On-Time % by Airline** (Airline vs Industry Avg)
* **On-Time % by Route (Origin → Destination)**
* **Average Taxi-In & Taxi-Out Time** (indicator of airport congestion)
* **Top 10 Delay-Prone Routes**

**Delay Cause Analysis**

* **Carrier Delay %** per airline
* **Weather Delay %** per month/season
* **NAS Delay (Air Traffic Control) %** by airport
* **Late Aircraft Delay Impact** (connecting flights effect)

**Cancellations**

* **Cancellation % by Cause**
* **Top Airlines with Highest Cancellations**
* **Peak Cancellation Months (Seasonality)**

**Time-Based**

* **Delays by Time of Day** (Morning vs Afternoon vs Evening)
* **Day-of-Week Performance** (Weekend vs Weekday delays)
* **Holiday Effect Analysis** (Nov/Dec spikes, etc.)

**Route & Airport Level**

* **Busiest Airports (by flight count)**
* **Worst Performing Airports (avg delay & cancellations)**
* **Top Origin-Destination Pairs by Traffic & Delay**
* **Geographic Delay Map** (using Origin/Destination + delay avg)

**Advanced/Bonus KPIs**

* **Delay Cost Estimation** (Avg Delay mins × Industry cost per min)
* **Flight Diversions %**
* **Flight Utilization Metrics** (Actual vs Scheduled Elapsed Time)
* **Airport Congestion Index** = (TaxiIn + TaxiOut times combined)
* **Season-Wise Carrier Performance**

Work flow :

**Project Workflow – AirFly Insights**

1. **Project Initialization**
   * Define goals and key performance indicators (KPIs).
   * Set up project files and structure.
2. **Data Preparation**
   * Import dataset into Power BI.
   * Check columns, data types, and missing values.
   * Clean and transform data (dates, routes, delays, cancellations).
3. **Feature Creation**
   * Derive useful fields such as Month, Day of Week, Hour, Route, and Delay Causes.
4. **KPI Development**
   * Build measures for flight counts, delays, cancellations, average times, and airline/airport performance.
5. **Visualization & Analysis**
   * Create charts and maps to explore:
     + Flight trends over time
     + Delay and cancellation patterns
     + Route and airport performance
     + Seasonal and time-based insights
6. **Dashboard Design**
   * Organize visuals into clear pages (Executive Summary, Delay & Cancellation, Route & Airport Insights).
   * Add filters and interactivity for easy exploration.
7. **Insights & Recommendations**
   * Summarize findings into key observations.
   * Provide recommendations for airlines/airports to improve efficiency.
8. **Final Reporting**
   * Prepare Power BI dashboard, final PDF/README, and presentation slides.
   * Submit project deliverables with documentation.

2. Load CSVs using pandas

1. Explore schema, types, size, and nulls
2. Perform sampling and memory optimizations