7-Day Action Plan: HR Dataset Analysis

Day 1 — Choose Dataset & Exploration

- Download dataset from Kaggle (IBM HR Attrition or HR Analytics).
- Load into Pandas (pd.read_csv).
- Check head(), info(), describe().
- Record: number of rows, columns, and data types (categorical/numeric).

Day 2 — Data Cleaning

- Check missing values (isnull().sum()).
- Handle or drop missing values.
- Clean categorical values (Yes/No → 1/0).
- Convert date columns to datetime (if any).
- Remove duplicates.

Day 3 — Exploratory Data Analysis (EDA): General Overview

- Number of employees by department.
- Average age and salary.
- Gender distribution.
- Stacked bar chart: Department × Gender.

Day 4 — Attrition Analysis (Main Part)

- Calculate attrition (Yes/No).
- Global attrition percentage.
- Breakdown by department, age group (20–30, 30–40…), salary level.
- Visualizations: pie chart (overall), bar chart (by department).

Day 5 — Correlation & Advanced Insights

- Correlation matrix (Salary ↔ YearsAtCompany ↔ JobSatisfaction).
- Heatmap visualization.
- Attrition vs YearsAtCompany (line chart).
- Attrition vs JobSatisfaction (bar chart).

Day 6 — Visual Storytelling

- Create 4–5 key visualizations:
- - Attrition % by department
- - Age distribution histogram
- - Average salary by job role
- - Correlation heatmap
- Add titles and short descriptions to each chart.

Day 7 — Final Polishing & Documentation

- Save notebook as HR_Attrition_Analysis.ipynb.
- Write a short intro: 'This project analyzes HR data based on IBM dataset…'
- Add markdown explanations for each step.
- Conclusion example: 'Analysis shows highest attrition in Sales, among young employees with low salary.'
- Upload to GitHub (folder → dataset, notebook, README.md).