PRACTICE PROJECTS

- **Excel**

- **SQL**

- **Python**

- **Advance Python**

- **Power BI**

Table of Contents

Project 1: Analyze Employee resignation data and derive summary statistics + visuals: Excel

Dataset source:
https://github.com/swapnilsaurav/projects/blob/main/Employee_Resignation_ProjectDataset.xlsx

Goal: Identify patterns and trends in employee resignations to improve retention strategies and HR planning.

Problem Statement:

An HR department wants to analyze past resignation data to understand:

- Which departments or job roles see higher attrition
- How resignation trends vary with age, tenure, gender, or performance
- What reasons are most cited for leaving
- How resignation trends have changed over time

The organization hopes to use this analysis to:

- Improve workforce planning
- Take preventive actions to reduce resignations
- Improve employee engagement and satisfaction

Solution Steps in Excel

Step 1: Data Cleaning

- Ensure all date columns are in Date format.
- Create a new column:
  - **Tenure (Years)** = (DateOfExit - DateOfJoining)/365

Step 2: Create Summary Statistics (in a new sheet)

Use Excel formulas such as:

Basic KPIs:

- **Total Resignations:**
  =COUNTA(EmployeeID)
- **Average Age at Resignation:**
  =AVERAGE(Age)
- **Average Tenure:**
  =AVERAGE(Tenure)
- **Resignations by Gender:**
  =COUNTIF(Gender,"Male"), =COUNTIF(Gender,"Female")
- **Resignations per Department (Pivot Table)**

Step 3: Use Pivot Tables for Analysis

a. Department-wise Resignations

- Rows: Department
- Values: Count of EmployeeID

b. Job Role vs Resignation Reason

- Rows: JobRole
- Columns: ResignationReason
- Values: Count of EmployeeID

c. Performance vs Resignation Rate

- Rows: PerformanceRating
- Values: Count of EmployeeID

d. Yearly Resignation Trend
- Create a column:
  - YearOfExit = YEAR(DateOfExit)
- Use Pivot Table:
  - Rows: YearOfExit
  - Values: Count of EmployeeID

Step 4: Create Visualizations

Use Excel charts based on pivot tables:
- **Bar Chart** – Resignations per Department
- **Line Chart** – Yearly Resignation Trend
- **Pie Chart** – Reasons for Resignation
- **Stacked Column Chart** – Gender distribution per Department
- **Scatter Plot** – Age vs Tenure

Step 5: Insights and Recommendations

Use conditional formatting and slicers to filter by:
- High-tenure exits
- Performance-based trends
- Department hot spots
- High attrition time periods

Sample insights:
- 45% of resignations are from IT and Sales
- Avg tenure of employees leaving due to "Better Opportunity" is < 3 years
- Performance Rating 2 has the highest resignation rate

Deliverables

You can structure your Excel file like this:

Sheet Name	Content Description
Employee_Resignations	Raw Data
Summary	Basic KPIs and formulas
Pivot_Dept	Pivot: Resignation count by Department
Pivot_Reason	Pivot: Reason vs Role
Charts	All visuals neatly arranged with titles

Project 2: Project: Analyze Employee resignation data and derive summary statistics + visuals: Python

Dataset source:
https://github.com/swapnilsaurav/projects/blob/main/Employee_Resignation_ProjectDataset.xlsx

Objective:

To analyze a company's employee resignation dataset using Python for:

- Summary statistics
- Trend analysis
- Departmental attrition insights
- Visualizations to identify key patterns and insights

Use Case Scenarios

The HR analytics team wants to:

1. Identify departments with the highest attrition
2. Understand age and performance patterns in resignations
3. Examine resignation trends over years
4. Analyze resignation reasons to drive retention strategies

Solution Using Python

Required Libraries:

```python
# Reading excel file in Python
# pandas / openpyxl
# pip install pandas openpyxl
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Step 1: Load the Dataset

```python
#Step 1: Load the Dataset
file = 'D:/datasets/Employee_Resignation_ProjectDataset.xlsx'
df = pd.read_excel(file)
print(df)
df.head()
```

Step 2: Data Preprocessing

```python
# Step 2: Data Preprocessing
# Convert dates to datetime
df['DateOfJoining'] = pd.to_datetime(df['DateOfJoining'])
df['DateOfExit'] = pd.to_datetime(df['DateOfExit'])

# Calculate tenure in years
df['TenureYears'] = (df['DateOfExit'] -
```

```
df['DateOfJoining']).dt.days / 365
df['YearOfExit'] = df['DateOfExit'].dt.year
```

Step 3: Summary Statistics

```
#Step 3: Summary Statistics
print("Total Resignations:", df.shape[0])
print("Average Age at Resignation:", df['Age'].mean())
print("Average Tenure (years):", df['TenureYears'].mean())
print("\nResignations by Gender:\n", df['Gender'].value_counts())
print("\nTop Resignation Reasons:\n",
df['ResignationReason'].value_counts())
```

Step 4: Department-Wise Resignations

```
# Step 4: Department-Wise Resignations
dept_counts = df['Department'].value_counts()
sns.barplot(x=dept_counts.index, y=dept_counts.values)
plt.xticks(rotation=45)
plt.title("Resignations by Department")
plt.ylabel("Count")
plt.tight_layout()
plt.show()
```

Step 5: Resignation Trends Over Years

```
# Step 5: Resignation Trends Over Years
yearly_counts = df['YearOfExit'].value_counts().sort_index()
yearly_counts.plot(kind='line', marker='o')
plt.title("Yearly Resignation Trend")
plt.xlabel("Year")
plt.ylabel("Resignation Count")
plt.grid(True)
plt.show()
```

Step 6: Reason vs Department Heatmap

```
# Step 6: Reason vs Department Heatmap
pivot = pd.pivot_table(df, values='EmployeeID',
                       index='Department',
                       columns='ResignationReason',
                       aggfunc='count', fill_value=0)
sns.heatmap(pivot, annot=True, fmt='d', cmap="YlGnBu")
plt.title("Department vs Resignation Reason")
plt.ylabel("Department")
plt.xlabel("Reason")
plt.tight_layout()
plt.show()
```

Step 7: Age vs Tenure Scatter Plot

```python
# Step 7: Age vs Tenure Scatter Plot
sns.scatterplot(data=df, x='Age', y='TenureYears', hue='Gender')
plt.title("Age vs Tenure of Resigned Employees")
plt.show()
```

Final Insights & Recommendations

1. Departments with the most resignations should be investigated for work culture, leadership, or growth path.
2. High performers (ratings 4–5) resigning frequently could indicate poaching or unmet expectations.
3. Frequent resignations due to **Better Opportunity** suggest market competitiveness and may require a compensation strategy revamp

Bonus: Save Insights to Excel

```python
summary = {
    "Total Resignations": df.shape[0],
    "Average Age": df['Age'].mean(),
    "Average Tenure": df['TenureYears'].mean(),
}
summary_df = pd.DataFrame([summary])
summary_df.to_excel("Resignation_Summary_Report.xlsx",
index=False)
```

Project 3: Project: Analyze Employee resignation data and derive summary statistics + visuals: Power BI

Dataset source:
https://github.com/swapnilsaurav/projects/blob/main/Employee_Resignation_ProjectDataset.xlsx

Problem Statement:

The HR leadership team of a mid-sized enterprise wants to understand the underlying trends and patterns in employee resignations. They are experiencing rising attrition in key departments and suspect a mix of age, performance, and tenure-based reasons.

They have collected historical resignation data and wish to use Power BI to:

- Visualize who is resigning (by department, gender, role)
- Understand when employees are most likely to resign (based on tenure, age, and exit year)
- Identify common reasons cited for resignation
- Track trends over time and across business units

Business Objectives:

1. Identify departments and job roles with the highest attrition
2. Understand tenure and age distribution of employees at the time of resignation
3. Track trends in resignation reasons year-on-year
4. Present data visually for stakeholder decision-making

Power BI Solution Approach

Step 1: Load the Dataset

- Import Employee_Resignation_Data_5000.xlsx into Power BI using Get Data → Excel.
- Load the sheet named Employee_Resignations.

Step 2: Data Transformation in Power Query

- Convert DateOfJoining and DateOfExit columns to Date type
- Add new calculated columns:

Tenure (Years) = DATEDIFF([DateOfJoining], [DateOfExit], YEAR)

Year of Exit = YEAR([DateOfExit])

Step 3: Build Key Visuals in Power BI Report View

Page 1: Executive Summary

- **Card Visuals**:
  - Total Resignations
  - Avg Age at Resignation
  - Avg Tenure
- **KPI Tile**: Resignations trend by Year of Exit

Page 2: Department-Level Insights
- **Clustered Column Chart**:
    - o X-axis: Department
    - o Y-axis: Count of EmployeeID
- **Stacked Column**:
    - o X-axis: Department
    - o Y-axis: Count of EmployeeID
    - o Legend: Gender

Page 3: Resignation Reasons & Performance
- **Stacked Bar Chart**:
    - o Axis: ResignationReason
    - o Legend: PerformanceRating
    - o Values: Count of EmployeeID
- **Pie Chart**:
    - o Values: ResignationReason
    - o Legend: ResignationReason

Page 4: Tenure & Age Analysis
- **Scatter Chart**:
    - o X-axis: Age
    - o Y-axis: Tenure (Years)
    - o Details: JobRole
    - o Color by: Gender
- **Histogram or Line Chart**:
    - o X-axis: Year of Exit
    - o Y-axis: Count of EmployeeID

Page 5: Interactive Filtering
- Add **Slicers**:
    - o Department
    - o Gender
    - o Year of Exit
    - o Resignation Reason

Step 4: Visual Design Best Practices
- Use consistent color themes (e.g., blues for males, pinks for females)
- Add tooltips for each visual
- Add page navigation buttons if required
- Add titles and subtitles to each visual for clarity

Power BI Measures (DAX Examples)

Total Resignations = COUNT(Employee_Resignations[EmployeeID])

Avg Age = AVERAGE(Employee_Resignations[Age])

Avg Tenure = AVERAGE(Employee_Resignations[Tenure (Years)])

High Performer Resignations =
CALCULATE(COUNT(Employee_Resignations[EmployeeID]),
    Employee_Resignations[PerformanceRating] >= 4)

Key Insights Delivered

Insight	Observation
High attrition	IT and Sales show 40%+ resignations
Reason pattern	"Better Opportunity" and "Work-life Balance" dominate
Tenure hotspots	Resignations mostly happen within 2–3 years
Age cluster	Common resignation ages: 28–35
Gender insight	55% of resignations are male employees

Power BI File Structure Suggestion

Page No	Report Page Name	Description
1	Executive Summary	Key KPIs and total resignation trends
2	Department Analysis	Who is resigning by department, gender
3	Reason & Performance	Visual link between reasons and ratings
4	Age & Tenure Analysis	Deeper analysis of demographic factors
5	Filters & Drilldowns	Slicers, filters for user exploration

Project 4: Analyze Employee resignation data and derive summary statistics + visuals: SQL

Problem Statement:

The HR leadership team of a mid-sized enterprise is witnessing increasing attrition in key departments. They aim to use historical resignation data to uncover patterns in who is leaving, why, and when. Key suspected drivers include age, performance ratings, and tenure.

The goal is to derive actionable insights that help HR take proactive retention measures using SQL analysis and Power BI visuals.

Business Objectives:

1. Identify departments and job roles with highest resignation rates.

2. Understand age and tenure distribution of employees at resignation.

3. Analyze reasons for resignation over the years.

4. Support data-driven decisions through clear visualizations.

Data Columns

Column Name	Description
EmployeeID	Unique ID of the employee
Name	Employee's name
Gender	Gender of the employee
Age	Age at time of resignation
Department	Department where they worked
JobRole	Job title
DateOfJoining	Joining date
DateOfExit	Exit/resignation date
PerformanceRating	Final performance rating (1–5)
ResignationReason	Self-reported reason for leaving
LastWorkingLocation	City/office they last worked in

SQL Table Creation

```
CREATE TABLE EmployeeResignations (
    EmployeeID INT PRIMARY KEY,
    Name VARCHAR(100),
    Gender VARCHAR(10),
    Age INT,
    Department VARCHAR(50),
    JobRole VARCHAR(50),
    DateOfJoining DATE,
```

```
    DateOfExit DATE,
    PerformanceRating INT,
    ResignationReason VARCHAR(100),
    LastWorkingLocation VARCHAR(50)
);
```

INSERT Queries

```
INSERT INTO EmployeeResignations VALUES
(1, 'Sneha Sharma', 'Female', 37, 'HR', 'HR Specialist', '2007-04-20', '2023-06-23', 5, 'Better
Opportunity', 'Chennai'),
(2, 'Ravi Sharma', 'Male', 27, 'HR', 'HR Specialist', '2016-05-02', '2024-01-31', 1, 'Work-life
Balance', 'Hyderabad'),
(3, 'Vikas Singh', 'Male', 50, 'Marketing', 'SEO Specialist', '2005-02-23', '2023-03-02', 2,
'Retirement', 'Bangalore'),
(4, 'John Patel', 'Male', 43, 'Sales', 'Sales Executive', '2013-07-10', '2015-08-10', 3, 'Low Job
Satisfaction', 'Chennai'),
(5, 'John Sharma', 'Female', 56, 'Sales', 'Sales Manager', '2006-10-08', '2020-02-24', 3, 'Work-life
Balance', 'Kolkata');
```

Key SQL Queries for Analysis

1. Department and Job Role with Highest Resignations

```
SELECT Department, JobRole, COUNT(*) AS ResignationCount
FROM EmployeeResignations
GROUP BY Department, JobRole
ORDER BY ResignationCount DESC;
```

2. Average Age and Tenure at Resignation

```
SELECT
    AVG(Age) AS AvgAge,
    AVG(DATEDIFF(DateOfExit, DateOfJoining)/365.25) AS AvgTenure
FROM EmployeeResignations;
```

3. Year-wise Trend in Resignations

```
SELECT YEAR(DateOfExit) AS ExitYear, COUNT(*) AS Resignations
FROM EmployeeResignations
GROUP BY YEAR(DateOfExit)
ORDER BY ExitYear;
```

4. Most Common Resignation Reasons

```
SELECT ResignationReason, COUNT(*) AS Count
FROM EmployeeResignations
GROUP BY ResignationReason
ORDER BY Count DESC;
```

5. Gender and Department-wise Resignation Split

```
SELECT Department, Gender, COUNT(*) AS ResignationCount
FROM EmployeeResignations
GROUP BY Department, Gender
ORDER BY Department, Gender;
```

6. Performance Rating at Exit Distribution

```
SELECT PerformanceRating, COUNT(*) AS EmployeesExited
FROM EmployeeResignations
GROUP BY PerformanceRating
ORDER BY PerformanceRating;
```