Team / Alpha

- محمد فاروق
- احمد رجب
- صفاء احمد
- حسام رفعت
- هاجر الفقي
- مصطفى كارم

Project Idea: Employee Attrition Analysis in a Virtual Company Using Excel and Power BI

Introduction

This project focuses on analyzing employee data from a virtual company (inspired by the well-known HR Analytics dataset), with a particular emphasis on factors influencing employee attrition, especially among the younger age group (18–25 years), as reflected in the provided sample. The dataset includes information about 12 young employees, such as age, salary, distance from home, job satisfaction, and several other variables.

The goal is to identify patterns and key drivers of attrition, helping HR management make informed decisions to improve employee retention.

Project Objectives

1. **Understand attrition rates**: Calculate the overall attrition percentage and break it down by categories (e.g., age, department, gender).
2. **Identify key factors**: Explore relationships between attrition and variables such as salary, distance from home, job satisfaction, and overtime.
3. **Provide recommendations**: Suggest strategies to reduce attrition, such as salary adjustments or job satisfaction programs.
4. **Build interactive dashboards**: Present the results visually in an engaging, easy-to-understand format.

Tools Used

- **Excel**: For data cleaning, basic statistical analysis, and simple visualizations (available free or via Microsoft 365).
- **Power BI**: For advanced analytics, data modeling, and interactive dashboards (Power BI Desktop is free, with optional web publishing).

Project Steps

1. Data Collection and Cleaning (in Excel)

- **Import the data**: Copy the provided dataset (tab-delimited format) into a new Excel sheet. Use the *Text to Columns* tool to separate fields.
- **Data cleaning**:
  - Check for missing or incorrect values (e.g., the "Over18" column is always "Y"; remove if not useful).
  - Convert categorical values into numeric or binary formats when needed (e.g., "Attrition": Yes = 1, No = 0).
  - Create new calculated columns, such as "Annual Income" = MonthlyIncome × 12.
  - Use Excel filters to detect duplicates or fix data entry errors.
- **Basic statistics (using PivotTables)**:
  - Calculate average age, salary, and distance from home by department.
  - Compute attrition rate: `=COUNTIF(Attrition,"Yes") / COUNT(Attrition)` → In the sample: 58% (7 out of 12).

---

2. Exploratory Data Analysis (EDA) in Excel

- **Charts**:
  - Bar chart: Attrition rate by *AgeGroup* or *Department* (e.g., most attrition in Sales).
  - Pie chart: Distribution of gender or marital status among employees who left.
  - Line chart: Relationship between *DistanceFromHome* and attrition (e.g., higher distances like 22 km appear linked to attrition in some cases).
- **Correlations**: Use `CORREL` to check relationships, such as between overtime and attrition (e.g., 3 of 7 who left had overtime = Yes).
- **Advanced analysis**: Use `VLOOKUP`/`XLOOKUP` to link fields, or *Solver* to model scenarios (e.g., would a 10% salary increase reduce attrition?).

---

3. Advanced Analysis and Visualization in Power BI

- **Import data**: Load the cleaned Excel file into Power BI (*Get Data* > *Excel*).
- **Data modeling**:
  - Create relationships (e.g., link *JobRole* to *Department*).
  - Use Power Query for additional cleaning and calculated fields (e.g., "Attrition Rate" = DIVIDE(COUNTROWS(FILTER(Table, [Attrition]="Yes")), COUNTROWS(Table))).
- **Visualizations**:
  - **Dashboard** may include:
    - **Card**: Overall attrition rate (58% in the sample).
    - **Stacked bar chart**: Attrition by *BusinessTravel* or *OverTime* (e.g., *Travel_Frequently* linked to higher attrition).
    - **Map**: If hypothetical geographic coordinates are added for *DistanceFromHome*.
    - **Scatter plot**: Relationship between *MonthlyIncome* and *YearsAtCompany*, colored by attrition (e.g., salaries below 5k are more often linked to attrition).
- **Interactivity**: Add slicers to filter by gender, education field, etc., allowing dynamic exploration of results.
- **DAX calculations**: Create measures like "Average Salary Hike" = AVERAGE([PercentSalaryHike]) or *What-If Parameters* for scenario analysis (e.g., impact of salary increases).
- **Forecasting**: Use Power BI's forecasting feature to predict future attrition trends based on *YearsAtCompany*.
