**Documentation Human Resources Dataset Analysis**

**Week 1: Data Cleaning and Preprocessing (All Teams)**

* **Tasks:**
  + **Data Preprocessing:** Clean and preprocess the data using Power BI.
  + **Tools:** Power BI.
* **Deliverables:**
  + Cleaned dataset ready for analysis.
  + Data preprocessing notebook.

**Week 2: Analysis Questions Phase (All Teams)**

* **Tasks:**
  + **Determine Data Analysis Questions:** Determine all possible analysis questions that can be deducted from the given dataset and would be of interest to the organization’s decision makers, e.g., what is the relation between the employees ages and their satisfaction level?
  + **Tools:** Power BI.
* **Deliverables:**
  + Set of analysis questions that can be answered via the dataset.

**Week 3: Dashboard Phase (All Teams)**

* **Tasks:**
  + **Build Dashboard:** Build a Power BI dashboard that visualize the answers to the asked questions.
  + **Tools:** Power BI.
* **Deliverables:**
  + Power BI dashboard.

**Week 4: Final Presentation (All Teams)**

* **Tasks:**
  + **Final Presentation:** Prepare a report and presentation summarizing the project work, including data analysis, model development, and deployment.
* **Deliverables:**
  + Final report and presentation.

**Details Human Resources Dataset Analysis**

**Task done by Zeinab:**

* Analyze and visualize the company profile section.
* Compile all contributions from team members into a single Power BI file.
* Standardize, organize, and format the final Power BI report.
* Review and ensure the accuracy of all data and visuals provided by the team.

**Query and Transformation steps in “Company Profile” part:**

**Create “All Measures” table that includes all measures done**

Measures created in the Company profile section:

**1-Total Emp** “To calculate the total number of Employees”

**2-Male** “To calculate the total number of males in the company”

**3-Female** “To calculate the total number of females in the company”

**4-% Female** “To Calculate females’ percentage from the total Emp”

**5-% Male** “To Calculate males’ percentage from the total Emp”

**6-Avg Salary** “To calculate the total average salary”

**7-% Due for promotion** “To calculate the percentage of employees due for promotion”

**8-% Not Due**” To calculate the percentage of employees not due for promotion”

**9-Due for Promotion** “To calculate employee due for promotion”

**10-Not due** “To calculate employee due for promotion”

**11-Active Emp** “To calculate current working employees”

**12-Attrition** “To Calculate employees who left the company”

**13-Max of promotion** “To determine the max number of years for promotion=5 years”

**14-Min. of promotion** “To determine the max number of years for promotion=1 year”

**15-Target of promotion** “To target the optimum number of years for promotion=2 Years”

**16-Training Opportunities** “To sum up all the training opportunities offered over the years”

**17-Trainings Taken** “To sum up all the training taken by employees over the years”

**18-Training Utilization** “to calculate the percentage of training utilization by employees”

**Edit In “Employee” Table:**

1. **Create a Conditional column (Promotion status**) to determine which employees are due for promotion and which are not
2. **Create a Conditional column (Distance Status)** to determine distance <= 10KM is very close, distance <=20KM is Close, and distance > 20KM is far
3. **Create a range (Age Range)** to group the employees by age range
4. **Create a calculated column (Year of Departure)** to calculate the departure year of employee using the hiring date and years spent in the company

Analysis Questions:

* Which department has the highest number of employees?
* What is the total number of training opportunities, and how many were taken?
* What percentage of employees are due for promotion?
* Which age range has the most employees in the company?
* How does the number of employees with a bachelor's degree compare to those with a master's degree?
* Which ethnic group has the highest representation in the company?
* How many employees are in the state with the highest workforce concentration?
* What is the total number of employees that are "close" to the workplace compared to those that are "far"?

**Task done by Dalia:**

**Steps for “Data Cleaning” and “Data Transformation” – Attrition Section**

1. **Match Education Degree with Educational Level**:

* Use the **VLOOKUP** formula in Excel to align each employee's education degree with their corresponding educational level.

1. **Calculate Average Performance**:

* Group the performance data by employee and calculate the average performance rating for everyone.

1. **Replace Codes with Descriptive Legends:**

* Replace coded values in the dataset (e.g., satisfaction and rating codes) with their descriptive counterparts from the legends.

1. **Full State Name Replacement**:

* Replace state abbreviations in your dataset with the full state names for clearer reporting and analysis.

**Insights**

1. What percentage of employees left the company over time (e.g., yearly attrition rate)?
2. Compare attrition trends from 2012 to 2022.
3. Break down attrition by gender, age, and department to identify which groups are most affected.
4. Identify the top reasons for attrition, such as low job satisfaction, work-life balance, or salary.

**Task done by Mai:**

**Steps for “Data Cleaning” and “Data Transformation” – Performance Section**

* Performance Section mainly depend on two datasets “Employee Table” and “Performance Table”
* Working on data went through main stages:
* Data Skimming to understand the meanings, context, situation of the company, predict the requirements as well as the sort of requested insights.
* Understanding the measurement of performance level.
* Export the data sets to the power BI
* Transform the different main indicators related to the: education level- rating level and satisfaction level from “numerical view” to “descriptive view”
* Main Steps related to data cleaning and transformation can me summarized in:
* Change data type, merged column, reordered columns, rename columns, insert years- add index column- prompted headers-add conditional columns-custom columns...etc.)
* Prepare group of “lists” to facilitate the visualization and filters (List of education level-list of employee ID- List of satisfaction level…etc)
* Prepare group of “groups” to facilitate the visualization and filters (Group by different dimensions- group education levels.

**Analysis and main forecasted insights- Performance Section**

**The main Steps for analysis:**

* Tracking the aggregate satisfaction level for each measure of performance level (Self rating-manager rating-relationship satisfaction- job satisfaction-work life balance satisfaction-environment satisfaction)

**Forecasted insights:**

* Develop a sort of correlation between every variable and the performance level measurement.
* Apply an assessment to stand on reason behind the dissatisfaction depending on the numerical and descriptive variables, by using the correlation method.
* Align the results with the results conducted from the “Attrition” Section
* Accordingly- Set a group of observation, recommendations and KPIs to enhance the company’s HR Department and Performance

**Task done by Ragab:**

**Steps for “Data Cleaning” and “Data Transformation”**

1. **Data Cleaning**:

* Group the employee data To Five Colum’s.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Personal** | **Job** | **Education** | **Work** | **Place** |
| 1 | First Name | Department | Education | Attrition | Business Travel |
| 2 | Last Name | Job Role | Education Field | Years At Company | Distance From Home (KM) |
| 3 | Gender | Salary |  | Years In Most Recent Role | State |
| 4 | Age | Stock Option Level |  | Years Since Last Promotion | Ethnicity |
| 5 | Marital Status | Over Time |  | Years With Curr Manager |  |
| 6 |  | Hire Date |  |  |  |

1. **Match Satisfy with Satisfied Level:**

* Use the **VLOOKUP** formula in Excel to align each Rating with their corresponding Rating level.

(Unacceptable, Needs Improvement, Meets Expectation, Exceeds Expectation, Above and Beyond)

1. **Documentation**

**•** Prepare Documentation File

1. **Final Presentation:**

* Prepare a report and presentation summarizing the project work, including data analysis, model development, and deployment.