HR Data Analytics Project

The purpose of this project is to showcase my data analytical skills using an HR dataset. The tools I will use to perform my analysis will be Excel and Tableau to clean, validate, analyze, and visualize my findings.

Resource: My data set was downloaded from <https://www.kaggle.com/datasets/rhuebner/human-resources-data-set>

Problem Statement: What is causing employment turnover and why?

What is the overall diversity profile of the organization?

Which recruitment sources provide the organization with diverse candidates?

ASK

1. **Question formulation**

Research Questions

Voluntary turnover

* Does engagement predict voluntary turnover?
* How does engagement relate to voluntary turnover?
* Does the employee satisfaction score relate to voluntary turnover?
* What work characteristics predict voluntary turnover?
* Which departments have the highest voluntary turnover?
* Does salary correlate with voluntary turnover?
* Which manager had the most voluntary turnover?

Diversity profile

* Does the organization have a diverse portfolio based on gender?
* Which gender is represented the least in the company?
* Does the organization have a diverse portfolio based on ethnicity?
* Which ethnic group is represented the least?
* Does the organization have a diverse portfolio based on age?
* Which age is represented the least?

Hiring/Termination

* Which recruitment resource returns the most diverse group of candidates based on race hired?
* Which month had the most employees hired?
* Which month had the most employees terminated?
* What were the top three termination reasons?

DATA CLEANING AND PROCESSING

Resource: My data set was downloaded from <https://www.kaggle.com/datasets/rhuebner/human-resources-data-set>

1. The data set was downloaded from Kaggle, imported validated, and cleaned in an Excel spreadsheet.

* I replaced the missing values in the date\_of\_termination with “0000-00-00” for the employees who are still employed with the company. I then reformatted all the cells with dates to “yyyy-dd-mm”. I used the Find and Replace function to execute my tasks.

1. Create an EER through MYSQL workbench

* I imported my CSV file from Excel to MYSQL, created the EER, and lastly created the tables based on the EER.

The tables I created were “Employee”, “Department, Emp\_Feedback, Recruitment, Performance, Position, and Status.

* I used the INSERT INTO function to insert columns from HR\_Analysis to the 7 other tables. I successfully transferred the columns to the 7 tables except for the Performance table, which was returning NULLS. To fix this I dropped the table, created a new performance table, Inserted the columns from HR\_Analysis to the Performance table, and the columns were successfully inserted with no NULLS.

ANALYZE

1. I performed an exploratory analysis of the data corresponding to the questions in the ASK section. The code was saved and uploaded to GitHub as well as the corresponding screenshots of the queries returned.
2. After performing exploratory analysis using SQL, I imported the CSV file to Tableau to further my analysis and create visual representations of my findings.
3. I created 3 dashboards corresponding to the topics of Voluntary Turnover, Diversity Profile, and Hiring/Termination in Tableau.

SHARE

After creating the dashboards I shared and presented my findings.