

PROJECT TITLE: TECHCORP EMPLOYEE DATA ANALYSIS

INTRODUCTION:

In the dynamic landscape of modern businesses, data-driven decision-making has emerged as a crucial strategy for growth and sustainability. TechCorp, a pioneering technology company, recognizes the power of data in shaping its workforce dynamics. To gain deeper insights into its employee landscape, TechCorp embarked on an in-depth analysis of its employee data. This endeavor aimed to uncover trends, patterns, and key metrics that could guide strategic decisions and foster a more engaging and productive work environment. By delving into the nuances of its workforce, TechCorp sought to answer critical questions about retention, diversity, and performance.

PROBLEM STATEMENT:

As TechCorp continues to navigate the competitive tech industry, optimizing its workforce composition and retention strategies has become paramount. The company is confronted with questions regarding employee turnover, demographic representation, and the overall impact of these factors on its growth trajectory. To address these challenges, TechCorp undertook a comprehensive analysis of its employee data. The goal was to uncover underlying patterns that could shed light on the optimal strategies for enhancing employee satisfaction, reducing attrition, and harnessing the potential of its diverse talent pool. By leveraging data-driven insights, TechCorp aims to proactively shape its workforce dynamics and create an environment that nurtures longevity, diversity, and performance. The findings derived from this analysis serve as a compass for TechCorp's strategic decisions, enabling the company to chart a course towards sustained success.

DATASET:

The dataset consists of the following columns:

id: Employee ID.

first_name: First name of the employee.

last_name: Last name of the employee.

birthdate: Date of birth of the employee.

gender: Gender of the employee (e.g., male or female).

race: Race/ethnicity of the employee.

department: The department in which the employee works.

jobtitle: Job title of the employee.

location: General location of the employee (e.g., office branch).

hire_date: Date when the employee was hired.

termdate: Date when the employee's employment was terminated (if applicable).

location_city: City of the employee's location.

location_state: State of the employee's location.

DATA CLEANING PROCESS

Data cleaning, also known as data cleansing or data scrubbing, is a crucial process in data preparation and analysis. Its main objective is to identify and correct errors, inconsistencies, and inaccuracies in datasets to improve their quality and reliability. The ultimate goal of data cleaning is to ensure that the data used for analysis is accurate, complete, and consistent, leading to more reliable insights and conclusions. In this project, we utilized Microsoft Excel to clean the data, making it more useful before visualizing it and drawing conclusions from the insights. Below is a detailed description of the data cleaning process we performed:

Data Cleaning in Excel:

1. **Checked for Duplicates:** We used the "remove duplicates" feature to identify and remove any duplicate records in the dataset. Fortunately, no duplicates were found.
2. **Checked for Blanks or Null Values:** We thoroughly reviewed all columns to identify any blank or null values that needed attention.
3. **Inserted New Column for Age:** We created a new column, "Age," by extracting the year of birthdate for each employee and calculating their age by subtracting it from the current year (2023)
4. **Standardized Date Format:** We utilized the "text to columns" feature to standardize all date formats, ensuring consistency throughout the dataset.
5. **Standardized Job Titles:** We used the "Find and Replace" function to standardize job titles, consolidating variations of the same role into one uniform category. For example, "Accountant I," "Accountant II," etc., were standardized as "Accountant."
6. **Created Age Brackets:** We created an "AgeBracket" column to categorize employees based on age. The formula assigned employees to one of the following age brackets:

Young Adults: 21 - 30 years old

Adults: 31 - 40 years old

Middle-aged Adults: 41 - 50 years old

Senior Adults: > 50 years old

7. **Deleted Birthdate Column:** Since we had already calculated the age of employees, the birthdate column was no longer necessary and was removed from the dataset.

Data Cleaning in Power BI:

After importing the data into Power BI, we entered Power Query mode to ensure data accuracy. Here are some of the actions we performed in Power Query:

1. **Verified Data Types:** We carefully reviewed and verified all data types to ensure they were correctly assigned before proceeding to visualization.
2. **Enhanced Age Bracket Labels:** To add clarity to the age bracket labels, we used Power Query to find and replace values in the original table. We included the age range alongside each age bracket, as follows:

Young Adults (21-30)

Adults (31-40)

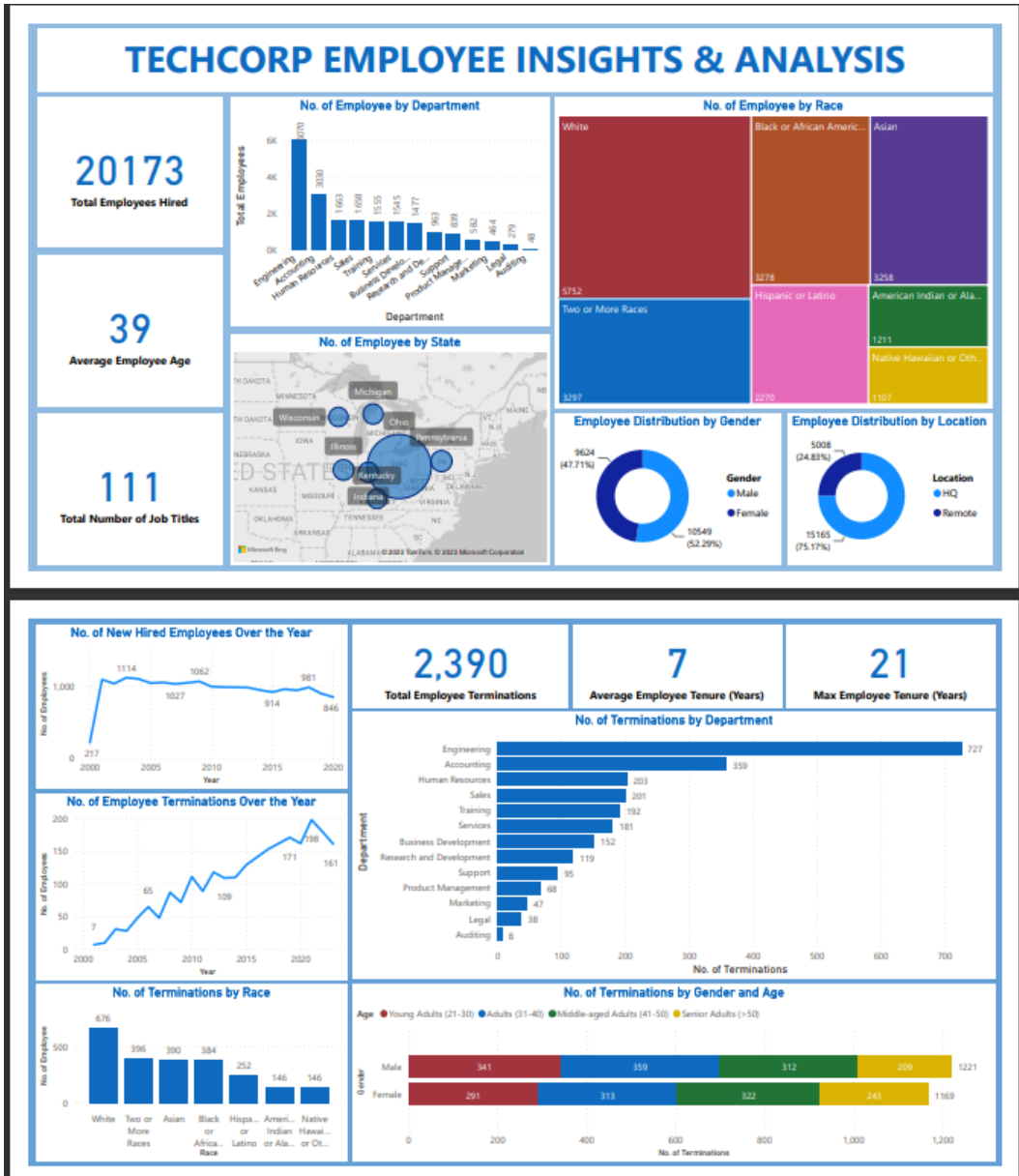
Middle-aged Adults (41-50)

Senior Adults (>50)

3. **Created a Sorting Table:** For sorting purposes in a stacked bar chart, we created another table in data modeling. This table had a unique key for the original table and was used to sort the age brackets.
4. **Standardized "Headquarters" to "HQ" in Location Column:** To enhance consistency and clarity, a uniform approach was applied to the "Location" column. We replaced instances of "Headquarters" with the abbreviation "HQ," ensuring uniformity across the dataset.
5. **Added Tenure Columns:** To provide an additional layer of context, we introduced new columns to capture "Tenure" information. This valuable data was skillfully incorporated to further enrich the dataset.

The data cleaning process ensures that the data used for analysis is accurate, consistent, and ready for visualization in Power BI. By addressing data issues and preparing the dataset diligently, we can draw meaningful insights from the data and make informed decisions based on reliable information.

DATA VISUALIZATION IN POWER BI: FINAL DASHBOARD



KEY FINDINGS & INSIGHTS:

1. **Overview of the Company's Employees:** From 2000 to 2020, Techcorp hired a total of 20,173 employees, with an average employee age of 39 years. The company has a diverse workforce with 111 different job titles.
2. **Number of Employees by Departments:** The Engineering Department stands out as the largest department in Techcorp, employing 6,070 individuals across 42 job titles. In contrast, the Auditing Department is the smallest, with only 48 employees and a single job title.
3. **Employee's Race:** The majority of Techcorp's workforce is composed of employees from the White racial group, totaling 5,752 individuals. On the other hand, the Native Hawaiian or Other Pacific Islander group represents the smallest segment with 1,107 employees.
4. **Gender Distribution:** Male employees comprise more than half of Techcorp's workforce, accounting for 10,549 employees (52.29%). The number of female employees is slightly lower, totaling 9,624 (47.71%).
5. **Working Employee Location:** A significant majority of employees (more than 70%) work at the company's headquarters rather than remotely.
6. **Employee's State:** The bubble map highlights that Ohio has the highest number of employees, represented by the largest bubble, with over 16,000 employees hailing from this state.
7. **Total Employee Terminations:** Between 2001 and 2023, 2,390 employees said goodbye to Techcorp. On average, they spent around 7 years here, with the longest stay being an impressive 21 years.
8. **Employee's Terminations by Department and Race:** The Engineering Department takes the lead with the highest number of employee terminations by department, accounting for 727 departures. Similarly, among racial groups, the White racial category holds the record for the highest number of terminations, totaling 676.
9. **Terminations in Gender and Age:** Overall, males constitute the majority among employee terminations, and individuals aged 31-40 (referred to as 'Adults') hold the highest departure rate.
10. **Termination Patterns :** Between 2001 and 2023, we observe a growing trend in terminations, with variations along the way. The highest number of terminations in a year occurred in 2021, with 198 employees leaving.
11. **New Hires Trend:** The number of newly hired employees witnessed a significant increase, jumping from 217 to 1,032 during the sharp rise between 2000 and 2002.

CONCLUSION:

Roughly 11.8% of the total employees have terminated. While the absolute number of terminations might not seem extremely high, it's essential to analyze this figure within the context of the company's goals, industry standards, and historical data. A deeper examination of the reasons behind these terminations and a comparison to industry benchmarks could provide a more accurate assessment of whether this number is significant for TechCorp. This can involve targeted initiatives to address the higher departure rate among employees aged 31-40 (referred to as 'Adults'). Implementing measures to support this age group, such as professional growth opportunities and work-life balance initiatives, could contribute to reducing departures. Furthermore, exploring ways to tap into the potential of the Engineering Department, which experiences the highest termination rate, could involve conducting exit interviews to identify areas for improvement. By leveraging these insights, TechCorp can create a workplace culture that not only attracts top talent but also nurtures long-term commitment and growth.