

Project Proposal

Course: INST 733 – Database Design and Modeling

Date: 03/07/2023

Team Members: Aditya Aswin Kumar, Pranav Tejasvi Adiraju, Srikanth Parvathala

DEI Data Hub

Building a Comprehensive Database System for a Workplace

Introduction:

Diversity, equity, and inclusion (DEI) are crucial to the success of any organization in the modern world. The success of an organization's DEI efforts may be tracked and measured in large part because of HR analytics. A well-designed database is necessary for tracking and measuring DEI parameters effectively. In this project, we will create a database for HR metric analytics based on a DEI Measurements Dataset. This database will give HR professionals the resources they need to monitor and assess the DEI initiatives of their organization and base decisions on the data. We will concentrate on creating a database that is simple to use, effective, and scalable. By doing this, we intend to give businesses a useful tool for enhancing their DEI metrics and fostering a more welcoming and diverse workplace.

Target Audience :

The Human Resource professionals, managers, and executives in organizations of any size and industry who are interested in improving their DEI metrics and promoting a more inclusive and diverse workplace may be the target audience for the database created based on the Diversity, Equity and Inclusion Measures Dataset for HR metric analytics. This database can be used to monitor and examine many HR indicators connected to DEI, such as employee engagement, diversity in the workforce, pay equity, and retention rates. Making data-driven decisions to

enhance HR policies and practices linked to DEI can be done using the insights obtained from the database.

Here are a few questions which the database can be able to answer:

1. What is the satisfactory rate for employees with low salaries?
2. What is the average salary of employees in different demographic groups?
3. Is there a significant pay gap based on gender?
4. Is there a significant pay gap based on ethnicity?
5. Which departments have the most ethnicity groups?
6. Which departments have the least nationality groups?
7. What department exhibits the highest level of sexual orientation diversity?
8. Is the job satisfaction rate higher among employees with disabilities compared to non-disabled employees?
9. What are the top hobbies/interests of employees?
10. How many employees have disclosed their preferred pronouns, and what are the most used pronouns in the organization?
11. How is the diversity of employees in higher management positions?

Potential entities/tables with potential attributes/columns:

Entities:

1. **Employees** - The Employees table contains data about the individuals who have enrolled or registered to work with a particular organization. This table typically includes various employee-related information such as their names, contact details, employment status, and other relevant details necessary for managing their employment within the organization.
2. **Departments** - In a Departments table, information about all the different departments that are present within an organization is stored. This table typically includes details such as department names, IDs, and descriptions, and is used to manage and organize the various departments and their related information within the organization.
3. **Nationality** - A Nationality table contains information regarding the different nationalities of individuals who are a part of a specific organization. This table is used to track and

manage data related to the nationalities of employees or members, and can be used to promote diversity, inclusion, and equal opportunity within the organization.

4. ***Employee_Satisfaction*** - An Employee_Satisfaction table provides information about the different levels of satisfaction that are used to measure the overall job satisfaction of employees within an organization. This table is used to track and analyze data related to employee satisfaction and can help organizations to identify areas where improvements can be made to create a more fulfilling and rewarding work environment for their employees.
5. ***Ethnicity*** - The Ethnicity table provides information on various ethnicity groups associated with individuals within the organization. This table is useful in promoting diversity and inclusion within the workplace and can help the organization to identify any disparities or imbalances that may exist among its workforce.
6. ***Sexual_Orientation*** - A Sexual_Orientation table contains information about the different sexual orientations of individuals who are part of a particular organization. This table is used to track and manage data related to the sexual orientations of employees or members and can be used to promote diversity and inclusivity within the organization, as well as to ensure that appropriate policies and practices are in place to support individuals of different sexual orientations.
7. ***Disabilities*** - The Disabilities table is a table that contains information about any disabilities that may be associated with individuals who are part of an organization. This table is used to keep track of the disability-related information and to ensure that appropriate accommodations are made for these individuals.
8. ***Pronouns*** - The Pronouns table contains information about the pronouns that employees within the organization use to identify themselves. This table helps to ensure that employees are referred to using their preferred pronouns, and that the organization is able to create an inclusive and respectful environment for all individuals regardless of their gender identity.
9. ***Hobbies*** - A Hobbies table is a table that contains information about the hobbies that are listed by employees within a specific organization. This table can be used to encourage employee engagement, team building, and to identify potential shared interests among employees.

10. ***Employee_Ethnicity*** - The Employee_Ethnicity table contains information regarding the ethnic background of individual employees within a particular organization.
11. ***Employees_Pronouns*** - The Employee_Pronouns table provides information about the specific pronouns that are used by individual employees within the organization. This table is used to ensure that employees are referred to in a way that aligns with their gender identity and promotes a culture of inclusivity and respect within the workplace.
12. ***Employees_Hobbies*** - In an Employee_Hobbies table, information about the hobbies of individual employees within an organization is stored.

Sample Attributes:

Table 1 - Employees
Employee_ID (PK)
Employee_fname
Employee_lname
Prefered_name
Department_ID (FK)
Employee_age
Salary
mobile_number
Email_address
Gender
Nationality_ID (FK)
Satisfaction_ID (FK)
Sexual_orientation_ID (FK)
Disability_ID (FK)

Table 2 - Departments
Department_ID (PK)
Department_name

Table 3 - Nationality
Nationality_ID (PK)
Country_name

Table 4 -Employees_Satisfaction
Satisfaction_ID (PK)
Satisfaction_description

Table - 5 Ethnicities
Ethnicity_ID (PK)
Ethnicity

Table 6 - Sexual_Orientation
Sexual_orientation_ID (PK)
Sexual_orientation

Table 7 - Disabilities

Disability_ID (PK)
Disability

Table 8 - Pronouns
Pronoun_ID (PK)
Pronouns

Table 9 - Hobbies
Hobby_ID (PK)
Hobby

Table 10 - Employees_Ethnicity
Employee_ID (FK/CPK)
Ethnicity_ID (FK/CPK)

Table 11 - Employees_Pronouns
Employee_ID (FK/CPK)
Pronoun_ID (FK/CPK)

Table 12 - Employees_Hobbies
Employee_ID (FK/CPK)
Hobby_ID (FK/CPK)

Entities/tables you will not include in the database:

- ***Retention table***: Data regarding Employee retention
- ***Engagement table***: How actively engaged the employees are.
- ***Survey table***: Contains the survey questions and the responses.
We have included only salary but avoided other things like ***bonuses*** and ***insurance*** etc.
- ***Training table*** that contains DEI training initiatives and attendance records for employees

Sample data:

Our project involves merging two datasets, the Human Resources Data Set and the Diversity, Equity and Inclusion Measures Dataset from Kaggle, to create a unified database that will provide insights into DEI metrics and enable workplaces to foster organizational success and employee satisfaction.

To ensure that our database is manageable and focused, we have limited the rows to 100 from the original 311 rows in the Human Resources dataset and 10,000 rows in the DEI dataset. We will be using columns such as ***Nationality***, ***Ethnicity***, ***Employee_Satisfaction*** (which has levels from 1 to 5), and ***Hobbies*** to address diversity and employee engagement within the organization. Additionally, ***Sexual_Orientation***, ***Disabilities***, and ***Pronouns*** columns will be used to address inclusiveness.

For the scope of this project, we have discarded columns such as Employee Retention, Employee Engagement, Employee Survey, and DEI training facilitation from the dataset. We have also normalized our tables using the ***third normal form (3NF)*** to ensure that the database is free from anomalies and unwanted dependencies.

These datasets are open source and provide licenses to anyone who wishes to use this dataset for learning or teaching.

References:

Dr. Carla Patalano and Dr. Rich Huebner. Updated in 2023, January. Kaggle Human Resources Data Set, Version 14. Retrieved March 3, 2023, from [Kaggle dataset link](#)

Kerem Kurt. Updated in December 2022, December. Diversity, Equity and Inclusion Measures Dataset, Version 3. Retrieved March 3, 2023 from [Kaggle dataset link](#)

THE END