**PRDA – 04 Job Market Analysis**

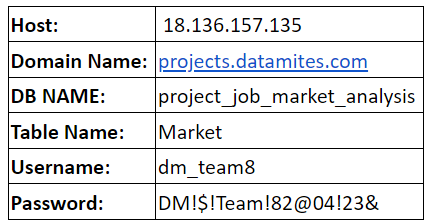
Task:

1. Get data from the database with the given credentials.
2. States with Most Number of Jobs.
3. Average Minimal and Maximal Salaries in Different States.
4. Average Salary in Different States.
5. Top 5 Industries with Maximum Number of Data Science Related Job Postings.
6. Companies with Maximum Number of Job Openings.
7. Job Titles with Most Number of Jobs.
8. Salary of Job Titles with Most Number of Jobs.
9. Skills Required by Companies for Each Job Title.
10. Relation between Average Salary and Education.
11. Analyze all the features and derive multiple insights.
12. Visualize the data using Tableau /PowerBI and derive insights and give your inputs/suggestions to the company.

About Dataset:

Your aim of this project is to analyze the job market trends for positions by analyzing job data which includes 742 rows and 42 features like Job title, Salary Estimate, Job Description, Rating, Company, Location, Company, Headquarters and many more acquired from various sources. Your goal is to identify the most in-demand skills, qualifications, job responsibilities and provide insights that can inform job seekers and employers.

Database Credentials:



Attribute Information:

**Job Title:** The title of job, eg. Data scientist, junior data scientist, senior data scientist etc.

**Salary Estimate:** Range of salary.

**Job Description:** Tells us what is expected out of the job title.

**Rating:** It gives the rating of the company

**Company Name:** Name of the company

**Location:** Location of the job

**Headquarters:** location of headquarter of the company

**Size:** Range of number of employee working in the company

**Founded:** Company founded in Year

**Type of ownership:** Tells us if the company is private, public or government owned.

**Industry:** Industry of the company

**Sector:** Sector in which company works

**Revenue:** Total revenue of the company per year

**Competitors:** Current competitor of the company in the same sector

**Hourly:** Tells us if the salary reported was hourly or yearly.

1: Hourly, 0: not hourly.

**Employer provided:** 1: If the salary was provided by the employee of the company, 0: otherwise.

**Lower Salary:** Lower salary reported for the job.

**Uppr Salary:** Upper salary reported for the job.

**Avg Salary(K):** Average of Lower and Upper salary yearly. K is the unit of the column, it means 1000. Also, the salary is in ($) U.S. dollars.

**company\_txt:** It contains the name of the company.

**Multiple skill columns** (python, spark, aws, excel etc): 1: Skill is required by the company, 0: It is not required.

**Jobtitle\_sim:** It contains the title of the job like Data scientist, ML engineer etc.

**seniority\_by\_title:** Seniority of the position, it is extracted from the Job Title.

**Degree:** If the job description mention that the company gives experience credit for a master(M) or Ph.D degree(P).

**Note*:*** -1 value in any column means the information was not available. Take the necessary steps to impute those values as per your choice.

You can provide your inputs/solution as a PPT presentation and you can explain your project, record it and send it with the PPT file.