Job Posting Report (Data Analysis using Power BI)

1. Project Introduction

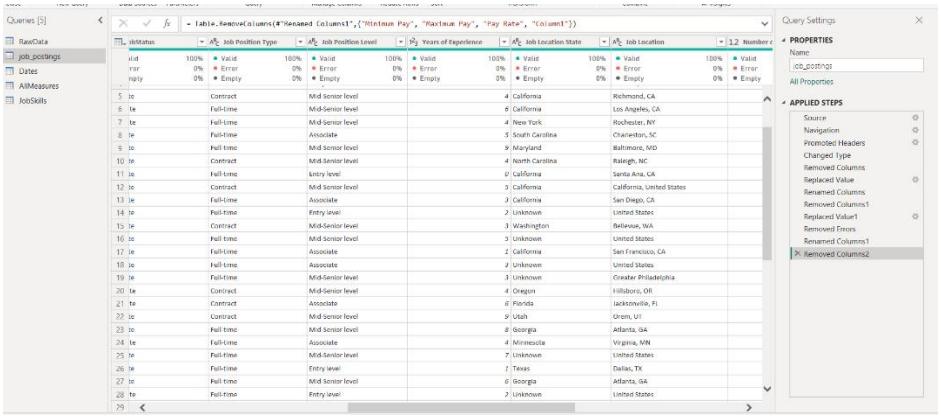
- **Project Overview:** This project focuses on analyzing job postings from Kaggle's dataset to understand job trends, types, skills required, and company insights in the U.S. tech sector.
- **Objectives:**
    1. Clean and preprocess the job posting data.
    2. Gain insights into job positions, seniority levels, and popular skills.
    3. Provide actionable insights for job seekers and businesses.
- **Target Audience:** Job seekers, companies, recruiters, and individuals interested in U.S. tech job trends and right final project to DEBI Power Bi track

2. Scope of Work

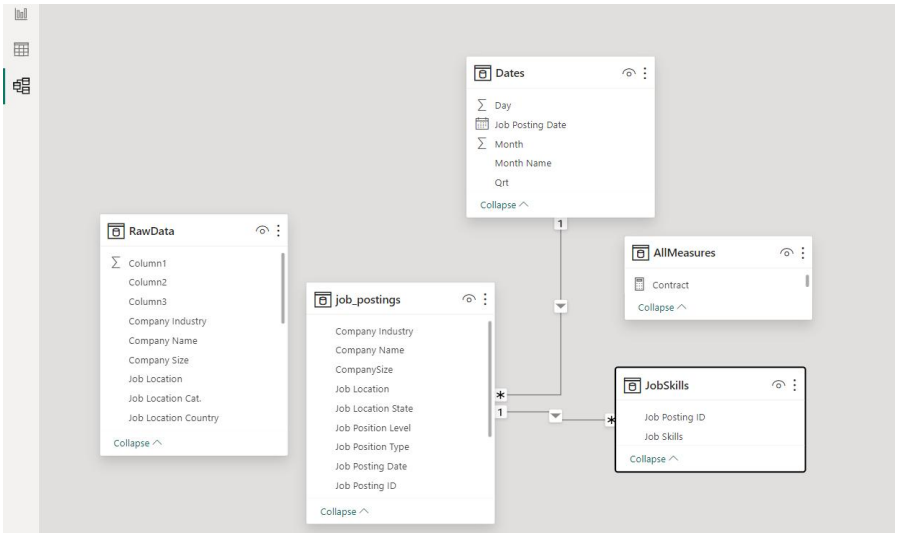
- **Data Sources:** Kaggle job posting dataset.
- **Data Domain:** Job titles, companies, countries, number of applicants, job location, skills, and job types.

3. Analysis Methodology

- **Analysis Tools:** Power BI, including Power Query for data transformation and DAX for calculation.
- **Data Cleaning Process:**
    1. **Change data types** where necessary for consistency.
    2. **Remove irrelevant columns** like "Job Title Additional Info" and "Job Skills" to focus on key data.
    3. **Handle missing values** in the "Number of Applicants" column by replacing null values with 0.
    4. **Create a new column** "Company Type" based on "Company Size," categorizing companies into 4 distinct types, and removing the original "Company Size" column.
    5. **Create a "Job Location State" column** by deriving it from "Job Location," and replace 'US' with 'Unknown' for more granularity.
    6. **Check and remove duplicates** to ensure data integrity.
    7. **Correct data errors** found during the preprocessing phase.
    8. **Remove rows with missing values** from the "Company Name" column (61 rows removed).
    9. **Add Job Status Column** based on Remote or On-Site workplace extracted from "Job Title Additional Info" column.
    10. **Extract three categories** "Startup, Medium-Sized and Large-Sized" out of Company Size field instead of categorizing them by the number of employees, "51-200 employees" for example.



- **Date Table Creation:**
    1. A separate date table was created, with columns for quarters, months, and days, improving the accuracy of time-based analysis.
- **Job Skills Table:**
    1. **Create a new "Job Skills" table** to identify the most in-demand individual skills.
    2. **Clean the table** by retaining only the relevant "Job Skills" column.
    3. **Use the TRIM and CLEAN functions** to improve the text format.
    4. **Perform additional cleaning steps** to remove unwanted characters or text from the "Job Skills" column.
    5. **Split the "Job Skills" column by delimiter** to distinguish multiple skills listed together.
    6. **Remove blank rows** (2,210 rows) to ensure data quality.
- **Data Model:**
    1. Establish a **many-to-one relationship** between "Posting ID" and the main dataset.
    2. Implement relevant **measures** throughout the project to track key metrics.

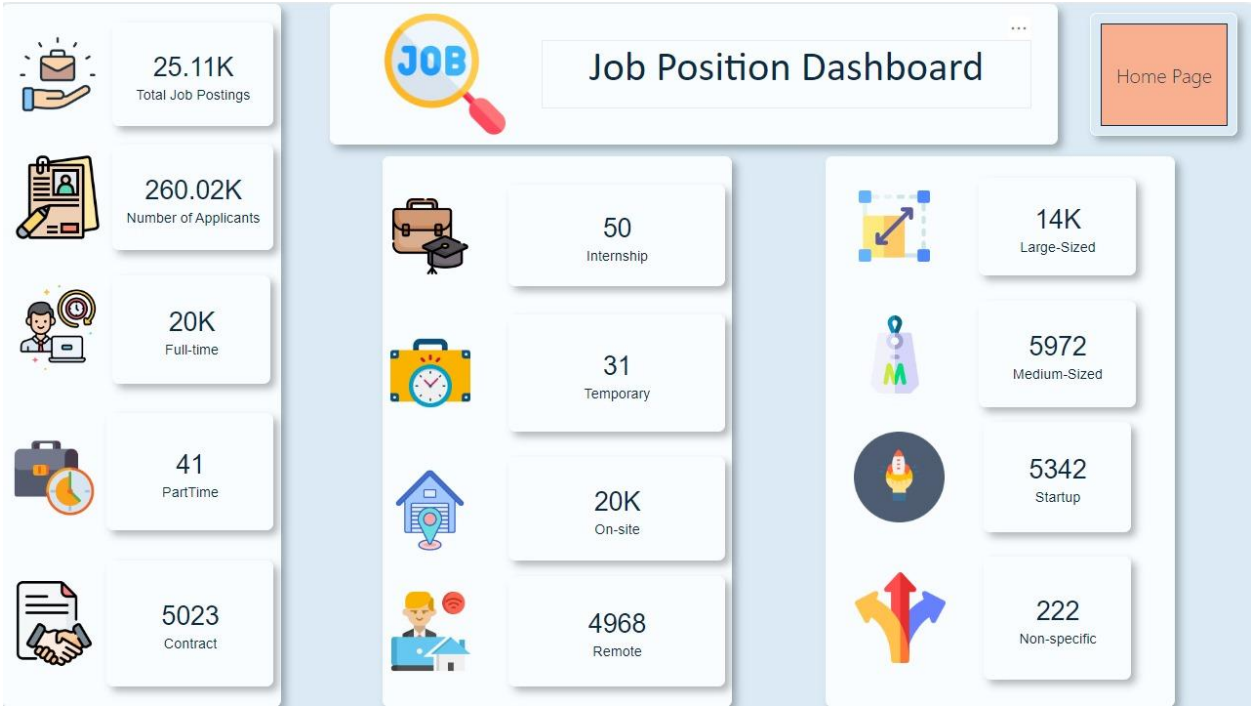


4. Analysis Results

- **Graphs and Visualizations:** Multiple Power BI visualizations were created, focusing on:

1. **Total number of job postings** in the dataset.
2. **Number of countries** represented by job postings.
3. **Variety of job titles** in different sectors.
4. **Total number of applications** across all job postings.
5. **Breakdown of job positions** by type and seniority level.
6. **Types of work** (e.g., remote, on-site).
7. **Methods of salary payment** across jobs.
8. **Breakdown of company types.**
9. **Top 10 most in-demand skills** identified from the skills table.

- **Conclusions:** The analysis highlights key job trends in the U.S. tech field, with remote work being a growing trend and certain skills in high demand. The dataset also shows the variety of job types and the importance of company size in determining job postings.
- **Recommendations:** Job seekers should focus on developing the top 10 most required skills and pay attention to senior roles, while companies should cater to the growing demand for flexible work options.



5. Challenges and Solutions

- **Challenges:** Handling missing data in key columns, such as "Number of Applicants," and managing large blank rows in the "Job Skills" table.
- **Solutions:** Null values were replaced appropriately, irrelevant columns removed, and text cleaning was performed rigorously on the "Job Skills" data.

6. Project Future

- **Next Steps:** Future analysis could expand to other industries beyond the U.S. tech sector, and job postings can be tracked over time for emerging trends.
- **Future Recommendations:** Continuously refine the data model and measures, incorporating more nuanced analysis of job levels, salary data, and specific technical skills.

7. Appendices and References

- **Data References:** Kaggle job posting dataset.

- **Scientific References:** N/A for this analysis.
- **Appendices:** Additional documentation of Power BI steps or scripts can be included for reproducibility.

8. Conclusion

- **Summary:** This project provided comprehensive insights into job postings in the U.S. tech industry, focusing on trends in job types, skills, and company characteristics.
- **Acknowledgments:** Special thanks to Kaggle for the dataset and contributors who assisted in data preprocessing and model creation.

Mission and Deeper Insights:

1. **Types of Work in U.S. Tech Field:** The dataset offers insights into the balance between remote and on-site work, highlighting a significant shift towards remote jobs in tech.
2. **Senior and Expert-Level Roles:** Senior roles are in high demand, reflecting the industry's need for experience and specialized knowledge.
3. **Criteria for Senior Professionals:** To be considered a "Senior," one typically needs several years of experience and mastery in key technical areas.
4. **What's Next for Seniors:** Senior professionals are encouraged to seek opportunities in large companies to gain experience before potentially starting their own businesses or taking on more leadership responsibilities.

What can I do if I'm an owner company?

From this data can help to understand the standard for other companies and the skills require globally

If I'm a job seeker?

U can utilize this to understand what the company they are looking for and to gain this skills from courses or scholarships etc..

if I'm a normal person looking for a business to launch

u can provide courses regarding these skills