JOB MARKET ANALYSIS (PRDA – 04)

Submitted by,

ANNU BABU

mail ID:annubabu97@gmail.com

Individual Project ID : PTID-CDA-JUN-25-535

Project Objective:

1. 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.

2. To identify the most in-demand skills, qualifications, job responsibilities and provide insights that can inform job seekers and employers.

Task 1 – Data Acquiring

- Connected to the Database and exported the data using MySQL Workbench

  Domain Name: projects.datamites.com

  DB Name: project_job_market_analysis

  Table Name : Market

Data cleaning and preprocessing:

- Tool used for data cleaning and preprocessing: Power BI
- Handling Missing Values:
  - Identified and addressed -1 values
  - Strategy: Excluded or replaced with appropriate placeholders
- Removed the ratings from company name by splitting the original column Company_Name
- Extracted and created a new column named state from the column Location

Data analysis and Dashboard creation:

- Tool used: Power BI

Task 2 - States with most Number of Jobs

- Visual used: column chart
- First 5 states with most number of jobs are:
  1. California(CA) – 522 jobs
  2. Massachusetts(MA) – 252 jobs
  3. New York(NY) – 233 jobs
  4. Illinois(IL) – 144 jobs
  5. Virginia(VA) – 139 jobs

Data Scientist is the most in-demand role in every state analyzed among these.

Task 3 - Average Minimal and Maximal Salaries in Different States.

- Visual used: Table

- The states California(CA), Illinois(IL),New Jersey(NJ),Washington(WA) etc. are the states with highest average minimal salaries whereas Nebraska(NE), Idaho(ID), Alabama(AL) ,Louisiana(LA) etc. are the states with lowest average minimal salaries.

- The states California(CA), Illinois(IL),New Jersey(NJ), Massachusetts(MA), North Carolina(NC) etc. are the states with highest average maximal salaries whereas Idaho(ID), Nebraska(NE), Alabama(AL), Louisiana(LA) ,Connecticut(CT) etc. are the states with lowest average maximal salaries among them.

Task 4 - Average Salary in Different States.

- Visual used: Matrix

- California(CA), Illinois(IL), New Jersey(NJ), Massachusetts(MA), Washington(WA) has the highest average salary among all the states.

- Idaho(ID), Nebraska(NE), Alabama(AL) , Louisiana(LA) ,New Mexico(NM) has the lowest average salary among all the states.

Task 5 - Top 5 Industries with Maximum Number of Data Science Related Job Postings

- Visual used: Table
- Top 5 industries are:
    1. Computer Hardware & Software – 119 job postings
    2. Insurance Carriers – 104 job postings
    3. Biotech & Pharmaceuticals – 100 job postings
    4. Consulting – 96 job postings
    5. Enterprise Software & Network Solutions – 92 job postings

Task 6 - Companies with Maximum Number of Job Openings.

- Visual used: Clustered bar chart
- The top 5 companies with maximum number of job openings are:

   1) MassMutual – 46 openings

   2) Takeda Pharmaceuticals – 40 openings

   3) Juniper Networks – 30 openings

   4) MITRE – 29 openings

   5) Esri – 28 openings

Task 7 - Job titles with most number of Jobs

- Visual used: Clustered Bar chart
- Top 5 job titles with most number of jobs are:
    1. Data Scientist   -   529 jobs
    2. Data Engineer   -   189 jobs
    3. Senior Data Scientist   -   138 jobs
    4. Senior Engineer   -   51 jobs
    5. Data Analyst   -   48 jobs

-> Among these top 5 jobs, Senior Data Scientist reports the highest Average salary.

Task 8 - Salary of Job Titles with most Number of Jobs

- Visual used: Clustered Bar Chart
- Top 5 job titles with most number of jobs along with salary

Job title	No. of jobs	Average Salary(K)
Data Scientist	529	105.40
Data Engineer	189	90.57
Senior Data Scientist	138	134.37
Senior Engineer	51	117.70
Data Analyst	48	70.03

- Although the highest average salary marked among the top 5 job titles with the maximum number of job opening is for Senior Data Scientist , the job title Director II, Data Science - GRM Actuarial holds the highest average salary (yearly 254K($)) among all the jobs though the number of jobs for this position is 6.

- The lowest salary marked is for Senior Operations Data Analyst, Call Center Operations among all the jobs whose average salary is27.5K($) yearly.

- Another insight which can be drawn is that the jobs which reports the highest number of job postings has a good average salary on comparing the salary statistics.

- Also, we can observe here that seniority plays less role when it comes to salary drawn and the job field is the one which matters.

Task 9 - Skills Required by Companies for Each Job Title

- Visual used: Table

job_title_sim	Number of Skills	Required Skills
director	2	aws, excel
data analitics	6	bi, excel, Python, sas, sql_, tableau
Data scientist project manager	6	aws, bi, excel, Python, sql_, tableau
data modeler	7	aws, bi, excel, hadoop, Python, spark, sql_
other scientist	7	aws, excel, Python, pytorch, sas, sql_, tensor
machine learning engineer	9	aws, excel, hadoop, Python, pytorch, scikit, spark, sql_, tensor
na	9	aws, bi, excel, hadoop, mongo, Python, spark, sql_, tableau
analyst	11	aws, bi, excel, google_an, hadoop, mongo, Python, sas, spark, sql_, tableau
data engineer	12	aws, bi, excel, flink, hadoop, mongo, Python, sas, spark, sql_, tableau, tensor
data scientist	16	aws, bi, excel, flink, google_an, hadoop, keras, mongo, Python, pytorch, sas, scikit, spark, sql_, tableau, tensor
Total	**85**	

- The insight that can be derived from this observation is that jobs such as Data scientist, Data Engineer, Analyst etc. require more number of skillsets.

Task 10 - Relation between Average Salary and Education

- Visual used: Clustered column chart

Degree	Average of Avg_SalaryK
P	134.82
M	106.46
na	102.49

- People with PhD degree(P) gets more salary when compared to people with masters(M) or others.

- Master's degree (M) is the most in-demand specified qualification.

Insights [Task 11]

- **Top 10 In-Demand Skills**:

  - Python, Excel, and SQL are the leading skills across job listings

  -Tools like AWS, Spark, and Tableau are also frequently requested

- Python, Excel, and SQL are core skills across data-related jobs. Cloud (AWS), big data (Spark, Hadoop), and visualization (Tableau) are also highly valued.

- **Python** dominates in most states, especially in tech hubs.

- **Excel** and **SQL** are consistently strong in almost all top states.

- **AWS** and **Spark** are more concentrated in select states, reflecting demand for cloud and big data skills in those regions.

- A Master's degree is the most common qualification, followed by PhD. Undergraduate degrees might be encoded differently or omitted.

- Best hiring regions for comprehensive skill sets are:

  California**(CA),** Texas**(TX),** New York**(NY),** Massachusetts**(MA),**

  Illinois**(IL),** Pennsylvania**(PA),** Virginia**(VA)**.

PowerBI Dashboard [Task 12]