CAPSTONE PROJECT -02

JOB MARKET ANALYSIS

ID: PTID-CDA-OCT-25-803

PROJECT ID: PRDA-04 JOB MARKET ANALYSIS

ADITYA HINGRAJIYA

Adityahingrajiya7@gmail.com

INDEX

INTRODUCTION

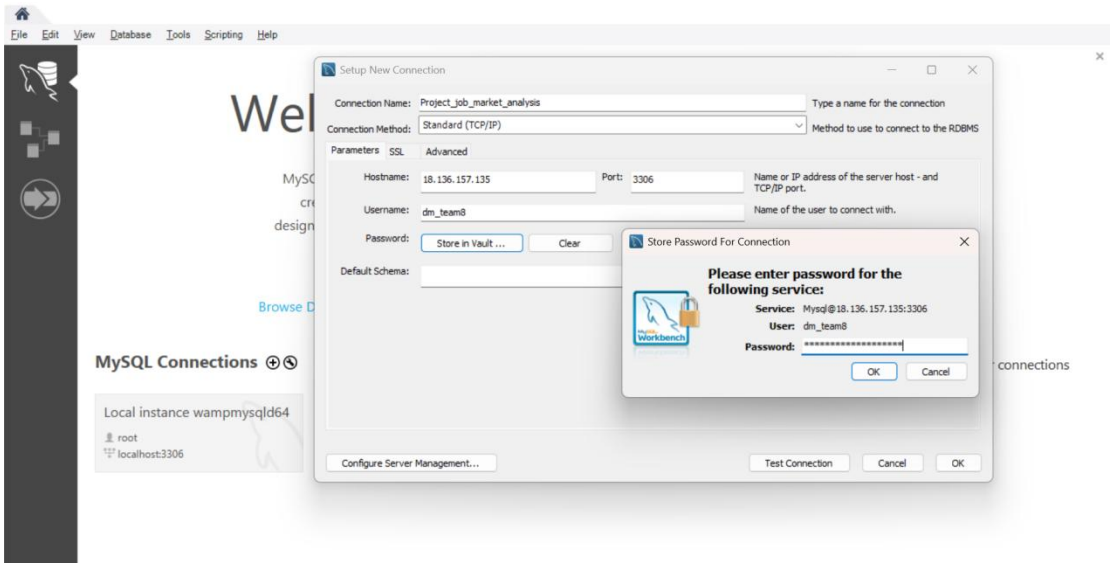
- In the modern data-driven economy, understanding job market trends is crucial for both job seekers and employers. The dataset consists of 742 records and 42 attributes, providing a comprehensive overview of job titles, salary ranges, company details, required skills, and other organizational factors. By examining these features, the study aims to identify which locations, industries, and qualifications are most in demand, helping stakeholders make informed career and business decisions.
- The main factors analyzed are: states with the highest number of job opportunities, average and range of salaries, and industries leading in Data Science job postings. Additionally, we will discuss trends associated with education levels and salary and the most frequent requirements by employers, too.
- Visualization and interpretation of data are effectively done using Tableau and Power BI. These visualizations uncover patterns and correlations that are not immediately apparent from the raw data. The insights drawn from this project will aid job seekers in identifying the most valuable skills and areas that pay the most, while companies benefit by understanding the competition in the market, salary benchmarks, and demand in the industry. The ultimate goal of this analysis is to bridge the gap in supply and demand in the workforce with more informed decisions in hiring, career planning, and skill development.

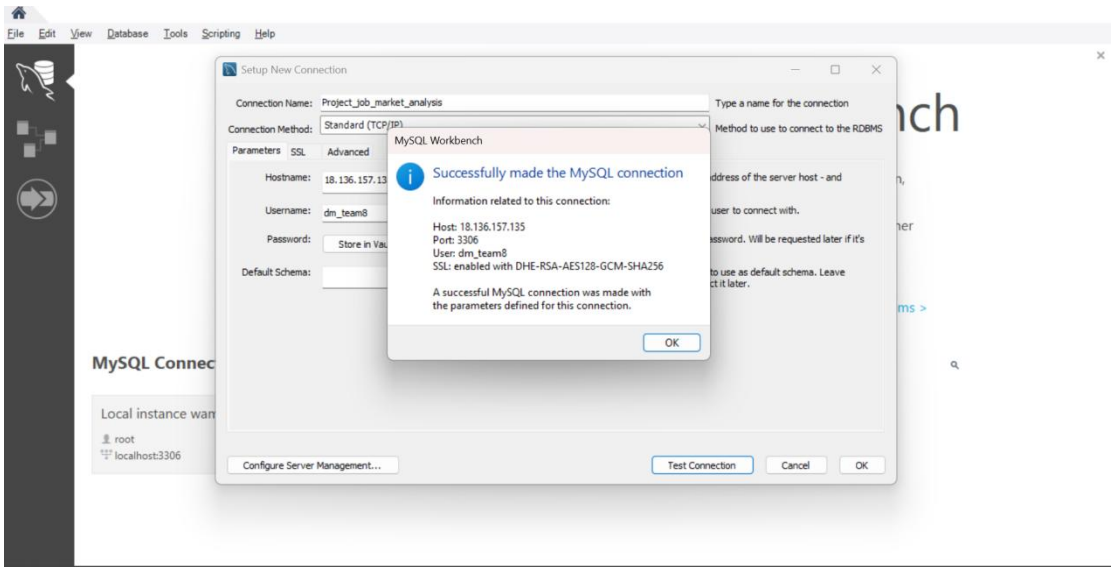
DATA ANALYSIS USING SQL

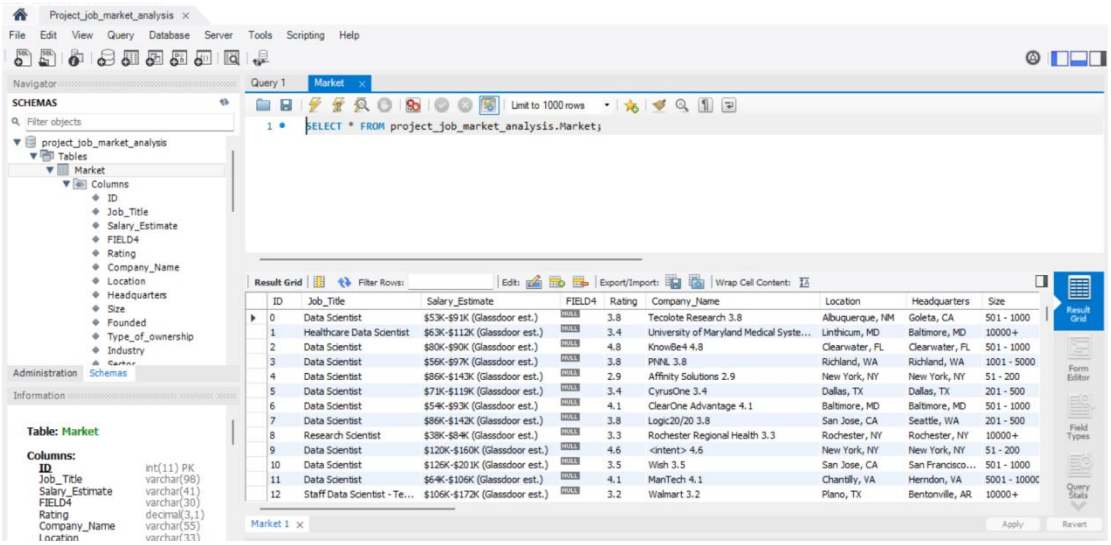
Q.1. Get data from the database with the given credentials.

We connect to the database using valid credentials (username, password, host) through a database driver. Once authenticated, we run SQL queries to retrieve the required data.

Screenshot:



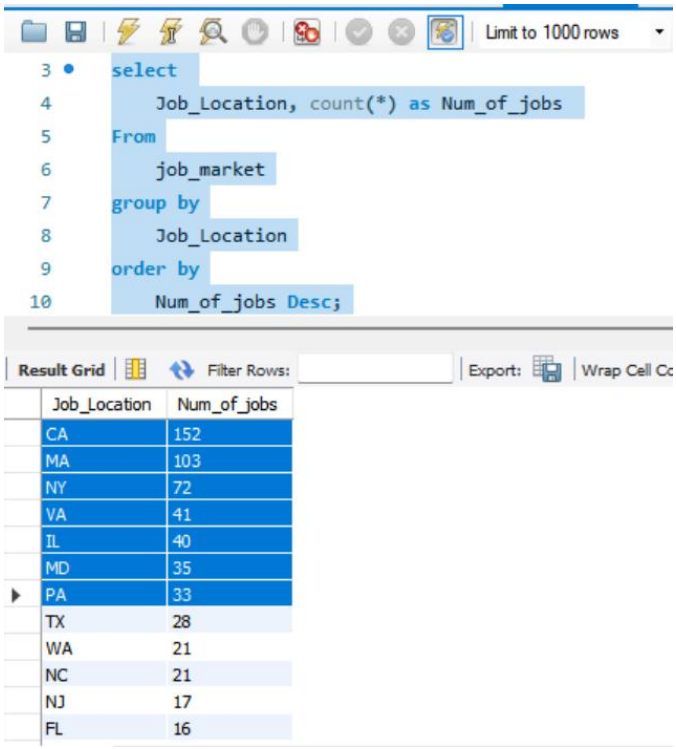



Project_job_market_analysis
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter objects
project_job_market_analysis
Tables
Market
Columns
ID
Job_Title
Salary_Estimate
FIELD4
Rating
Company_Name
Location
Headquarters
Size
Founded
Type_of_ownership
Industry
Administration Schemas
Information
Table: Market
Columns:
ID int(11) PK
Job_Title varchar(98)
Salary_Estimate varchar(41)
FIELD4 varchar(30)
Rating decimal(3,1)
Company_Name varchar(55)
Location varchar(33)
Query 1
Market
Limit to 1000 rows
1 SELECT * FROM project_job_market_analysis.Market;
Result Grid
Filter Rows:
Edit:
Export/Import:
Wrap Cell Content:
ID Job_Title Salary_Estimate FIELD4 Rating Company_Name Location Headquarters Size
0 Data Scientist $53K-$91K (Glassdoor est.) NULL 3.8 Tecolote Research 3.8 Albuquerque, NM Goleta, CA 501 - 1000
1 Healthcare Data Scientist $63K-$112K (Glassdoor est.) NULL 3.4 University of Maryland Medical Syste... Linthicum, MD Baltimore, MD 10000+
2 Data Scientist $80K-$90K (Glassdoor est.) NULL 4.8 KnowBe4 4.8 Clearwater, FL Clearwater, FL 501 - 1000
3 Data Scientist $56K-$97K (Glassdoor est.) NULL 3.8 PNNL 3.8 Richland, WA Richland, WA 1001 - 5000
4 Data Scientist $86K-$143K (Glassdoor est.) NULL 2.9 Affinity Solutions 2.9 New York, NY New York, NY 51 - 200
5 Data Scientist $71K-$119K (Glassdoor est.) NULL 3.4 CyrusOne 3.4 Dallas, TX Dallas, TX 201 - 500
6 Data Scientist $54K-$93K (Glassdoor est.) NULL 4.1 ClearOne Advantage 4.1 Baltimore, MD Baltimore, MD 501 - 1000
7 Data Scientist $86K-$142K (Glassdoor est.) NULL 3.8 Logic20/20 3.8 San Jose, CA Seattle, WA 201 - 500
8 Research Scientist $38K-$84K (Glassdoor est.) NULL 3.3 Rochester Regional Health 3.3 Rochester, NY Rochester, NY 10000+
9 Data Scientist $120K-$160K (Glassdoor est.) NULL 4.6 <intent> 4.6 New York, NY New York, NY 51 - 200
10 Data Scientist $126K-$201K (Glassdoor est.) NULL 3.5 Wish 3.5 San Jose, CA San Francisco... 501 - 1000
11 Data Scientist $64K-$106K (Glassdoor est.) NULL 4.1 ManTech 4.1 Chantilly, VA Herndon, VA 5001 - 10000
12 Staff Data Scientist - Te... $106K-$172K (Glassdoor est.) NULL 3.2 Walmart 3.2 Plano, TX Bentonville, AR 10000+
Result Grid
Form Editor
Field Types
Query Stats
Market 1
Apply
Revert

Q.2.States with most number of Jobs.

Query:

```
select Job_Location, count (*) as Num_of_jobs
From job_market
group by Job_Location
order by Num_of_jobs Desc;
```



```
3  select
4      Job_Location, count(*) as Num_of_jobs
5  From
6      job_market
7  group by
8      Job_Location
9  order by
10     Num_of_jobs Desc;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Co

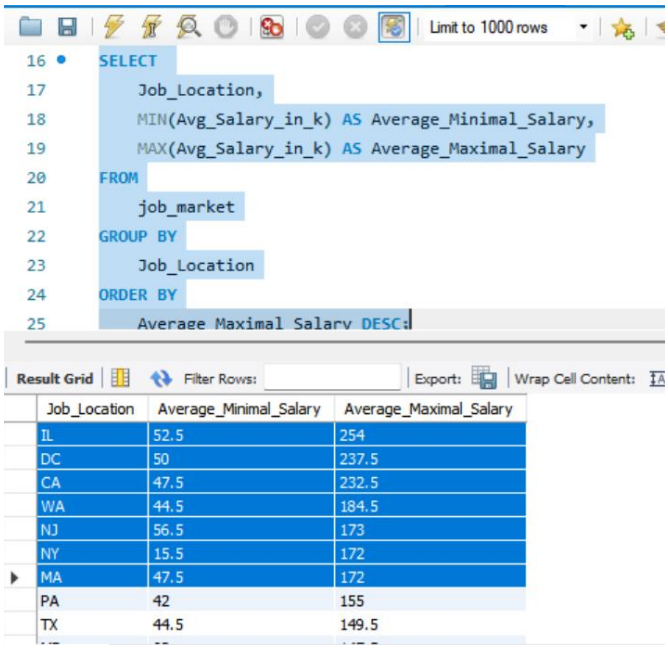
Job_Location	Num_of_jobs
CA	152
MA	103
NY	72
VA	41
IL	40
MD	35
PA	33
TX	28
WA	21
NC	21
NJ	17
FL	16

Q.3. Average Minimal and Maximal Salaries in Different States.

Query:

```
SELECT
   Job_Location,
   MIN(Avg_Salary_in_k) AS Average_Minimal_Salary,
   MAX(Avg_Salary_in_k) AS Average_Maximal_Salary
FROM
```

```
    job_market
GROUP BY
    Job_Location
ORDER BY
    Average_Maximal_Salary DESC;
```



```
SELECT
    Job_Location,
    MIN(Avg_Salary_in_k) AS Average_Minimal_Salary,
    MAX(Avg_Salary_in_k) AS Average_Maximal_Salary
FROM
    job_market
GROUP BY
    Job_Location
ORDER BY
    Average_Maximal_Salary DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

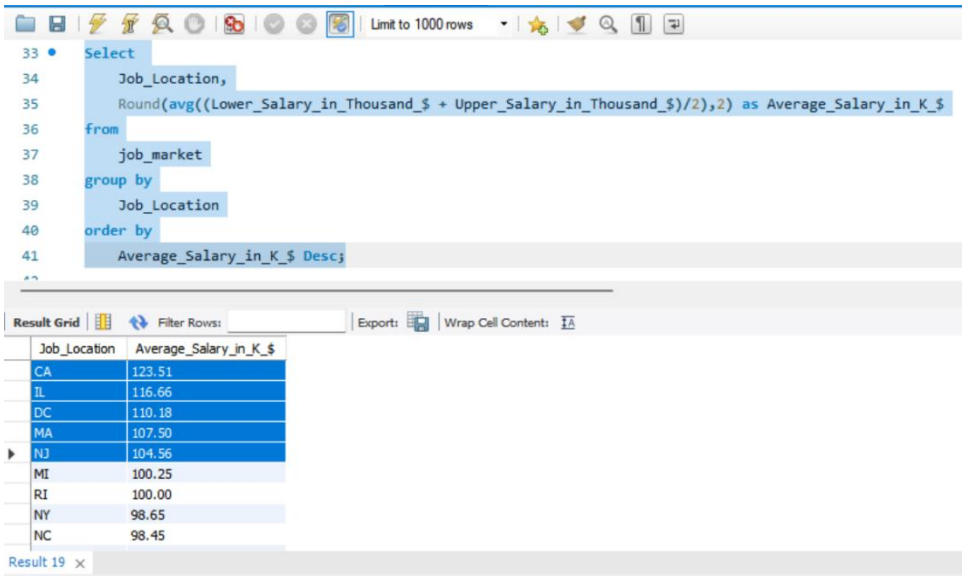
Job_Location	Average_Minimal_Salary	Average_Maximal_Salary
IL	52.5	254
DC	50	237.5
CA	47.5	232.5
WA	44.5	184.5
NJ	56.5	173
NY	15.5	172
MA	47.5	172
PA	42	155
TX	44.5	149.5

Q.4. Average Salary in Different States.

Query:

```
Select Job_Location,
Round (avg ((Lower_Salary_in_Thousand_$ + Upper_Salary_in_Thousand_$)/2),2)
as Average_Salary_in_K_$
from job_market
group by Job_Location
order by Average_Salary_in_K_$ Desc;
```

Screenshot:



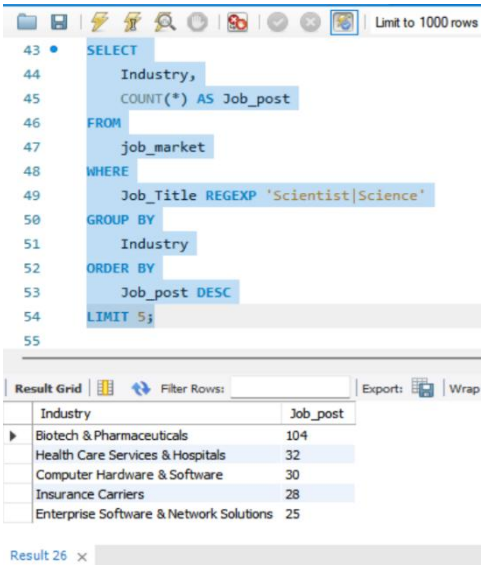
```
Select
    Job_Location,
    Round(avg((Lower_Salary_in_Thousand_$ + Upper_Salary_in_Thousand_$)/2),2) as Average_Salary_in_K_$
from
    job_market
group by
    Job_Location
order by
    Average_Salary_in_K_$ Desc;
```

Job_Location	Average_Salary_in_K_$
CA	123.51
IL	116.66
DC	110.18
MA	107.50
NJ	104.56
MI	100.25
RI	100.00
NY	98.65
NC	98.45

Q.5. Top 5 Industries with maximum number of Data Science related job postings.

Query:

```
SELECT Industry,
COUNT (*) AS Job_post
FROM job_market
WHERE Job_Title REGEXP 'Scientist|Science'
 GROUP BY
 Industry
ORDER BY Job_post DESC LIMIT 5;
```



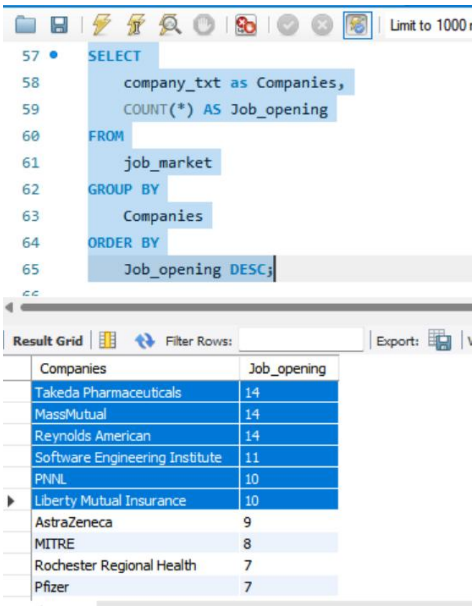
```
SELECT
    Industry,
    COUNT(*) AS Job_post
FROM
    job_market
WHERE
    Job_Title REGEXP 'Scientist|Science'
GROUP BY
    Industry
ORDER BY
    Job_post DESC
LIMIT 5;
```

Industry	Job_post
Biotech & Pharmaceuticals	104
Health Care Services & Hospitals	32
Computer Hardware & Software	30
Insurance Carriers	28
Enterprise Software & Network Solutions	25

Q.6. Companies with maximum Number of job openings.

Query:

```
SELECT company_txt as Companies,
COUNT (*) AS Job_opening
FROM job_market
GROUP BY
Companies ORDER BY Job_opening DESC;
```



```
SELECT
    company_txt as Companies,
    COUNT(*) AS Job_opening
FROM
    job_market
GROUP BY
    Companies
ORDER BY
    Job_opening DESC;
```

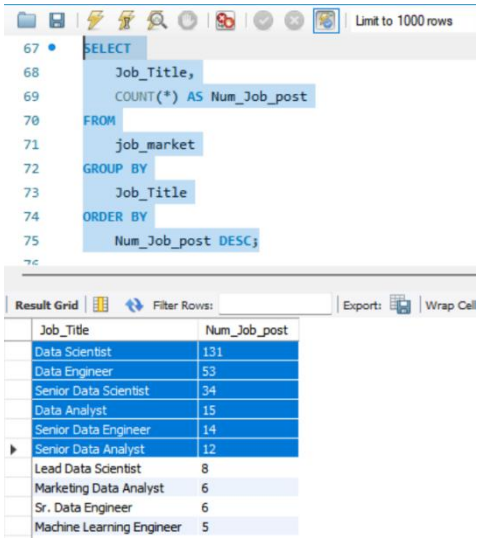
Result Grid | Filter Rows: | Export:

Companies	Job_opening
Takeda Pharmaceuticals	14
MassMutual	14
Reynolds American	14
Software Engineering Institute	11
PNNL	10
Liberty Mutual Insurance	10
AstraZeneca	9
MITRE	8
Rochester Regional Health	7
Pfizer	7

Q.7. Job Titles with most Number of Jobs.

Query:

```
SELECT company_txt as Companies,
COUNT (*) AS Job_opening
FROM job_market
GROUP BY Companies
ORDER BY Job_opening DESC;
```



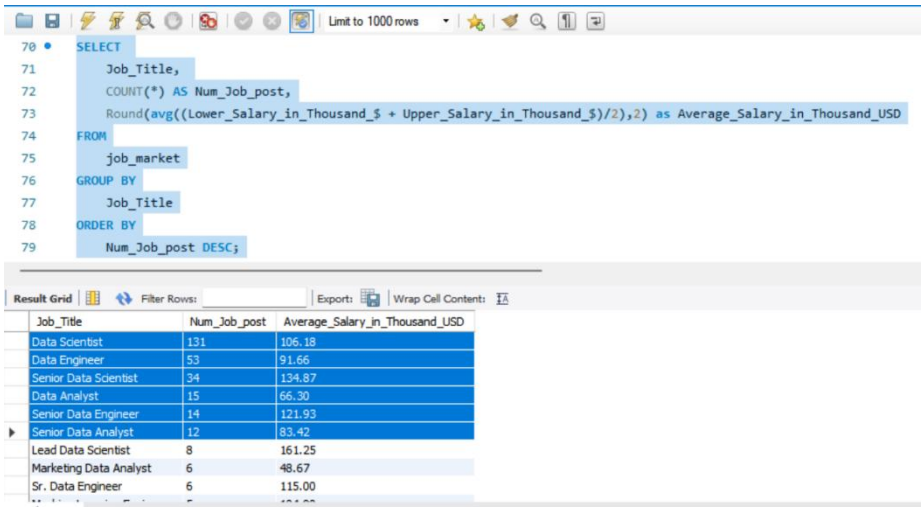
```
67 • SELECT
68       Job_Title,
69       COUNT(*) AS Num_Job_post
70   FROM
71       job_market
72   GROUP BY
73       Job_Title
74   ORDER BY
75       Num_Job_post DESC;
```

Job_Title	Num_Job_post
Data Scientist	131
Data Engineer	53
Senior Data Scientist	34
Data Analyst	15
Senior Data Engineer	14
Senior Data Analyst	12
Lead Data Scientist	8
Marketing Data Analyst	6
Sr. Data Engineer	6
Machine Learning Engineer	5

Q.8. Salary of job titles with most number of jobs.

Query:

```
SELECT Job_Title,
COUNT (*) AS Num_Job_post,
 Round (avg ((Lower_Salary_in_Thousand_$ + Upper_Salary_in_Thousand_$)/2),2)
as Average_Salary_in_Thousand_USD
FROM job_market
GROUP BY Job_Title
ORDER BY Num_Job_post DESC, Average_Salary_in_Thousand_USD DESC;
```



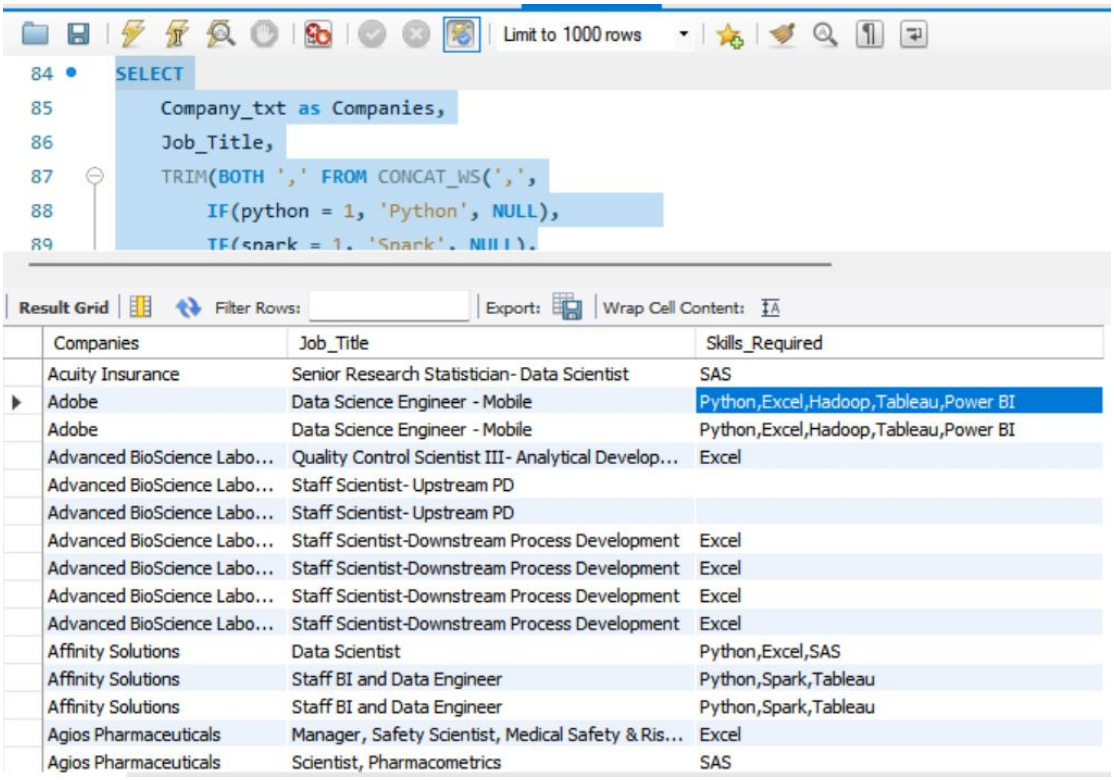
```
70 • SELECT
71       Job_Title,
72       COUNT(*) AS Num_Job_post,
73       Round(avg((Lower_Salary_in_Thousand_$ + Upper_Salary_in_Thousand_$)/2),2) as Average_Salary_in_Thousand_USD
74   FROM
75       job_market
76   GROUP BY
77       Job_Title
78   ORDER BY
79       Num_Job_post DESC;
```

Job_Title	Num_Job_post	Average_Salary_in_Thousand_USD
Data Scientist	131	106.18
Data Engineer	53	91.66
Senior Data Scientist	34	134.87
Data Analyst	15	66.30
Senior Data Engineer	14	121.93
Senior Data Analyst	12	83.42
Lead Data Scientist	8	161.25
Marketing Data Analyst	6	48.67
Sr. Data Engineer	6	115.00

Q.9. Skills required by companies for each job title.

Query:

```
SELECT
   Company_txt as Companies,
   Job_Title,
   TRIM (BOTH ',' FROM CONCAT_WS (',',
      IF (python = 1, 'Python', NULL),
      IF (spark = 1, 'Spark', NULL),
      IF (aws = 1, 'AWS', NULL),
      IF (excel = 1, 'Excel', NULL),
      IF (sas = 1, 'SAS', NULL),
      IF (keras = 1, 'Keras', NULL),
      IF (pytorch = 1, 'PyTorch', NULL),
      IF (scikit = 1, 'Scikit-Learn', NULL),
      IF (tensor = 1, 'TensorFlow', NULL),
      IF (hadoop = 1, 'Hadoop', NULL),
      IF (tableau = 1, 'Tableau', NULL),
      IF (bi = 1, 'Power BI', NULL),
      IF (flink = 1, 'Flink', NULL),
      IF (mongo = 1, 'MongoDB', NULL),
      IF (google_an = 1, 'Google Analytics', NULL)
   )) AS Skills_Required
FROM job_market
ORDER BY Company_txt, Job_Title;
```



```
84 • SELECT
85       Company_txt as Companies,
86       Job_Title,
87       TRIM(BOTH ',' FROM CONCAT_WS(',',
88           IF(python = 1, 'Python', NULL),
89           IF(spark = 1, 'Spark', NULL),
```

Limit to 1000 rows

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

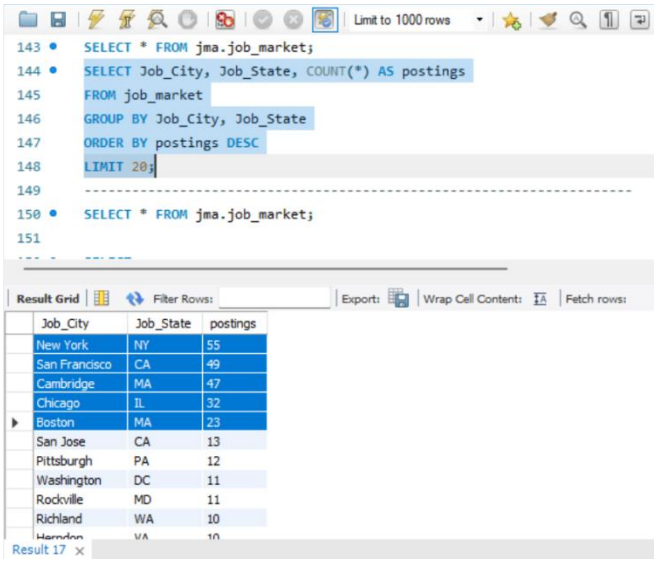
Companies	Job_Title	Skills_Required
Acuity Insurance	Senior Research Statistician- Data Scientist	SAS
Adobe	Data Science Engineer - Mobile	Python,Excel,Hadoop,Tableau,Power BI
Adobe	Data Science Engineer - Mobile	Python,Excel,Hadoop,Tableau,Power BI
Advanced BioScience Labo...	Quality Control Scientist III- Analytical Develop...	Excel
Advanced BioScience Labo...	Staff Scientist- Upstream PD	
Advanced BioScience Labo...	Staff Scientist- Upstream PD	
Advanced BioScience Labo...	Staff Scientist-Downstream Process Development	Excel
Advanced BioScience Labo...	Staff Scientist-Downstream Process Development	Excel
Advanced BioScience Labo...	Staff Scientist-Downstream Process Development	Excel
Advanced BioScience Labo...	Staff Scientist-Downstream Process Development	Excel
Affinity Solutions	Data Scientist	Python,Excel,SAS
Affinity Solutions	Staff BI and Data Engineer	Python,Spark,Tableau
Affinity Solutions	Staff BI and Data Engineer	Python,Spark,Tableau
Agios Pharmaceuticals	Manager, Safety Scientist, Medical Safety & Ris...	Excel
Agios Pharmaceuticals	Scientist, Pharmacometrics	SAS

Q.10. Analyse all the features and derive multiple insights.

1. Cities with most openings.

   Query:

```
SELECT * FROM jma.job_market;
SELECT Job_City, Job_State, COUNT (*) AS postings
FROM job_market
GROUP BY Job_City, Job_State
ORDER BY postings DESC
LIMIT 20;
```



```
143  SELECT * FROM jma.job_market;
144  SELECT Job_City, Job_State, COUNT(*) AS postings
145  FROM job_market
146  GROUP BY Job_City, Job_State
147  ORDER BY postings DESC
148  LIMIT 20;
149  ------------------------------------------------------------------------
150  SELECT * FROM jma.job_market;
151
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

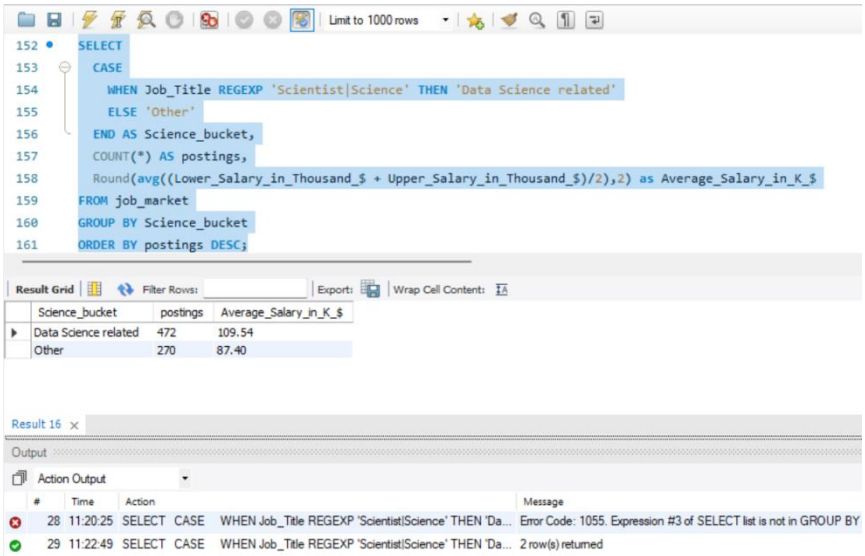
Job_City	Job_State	postings
New York	NY	55
San Francisco	CA	49
Cambridge	MA	47
Chicago	IL	32
Boston	MA	23
San Jose	CA	13
Pittsburgh	PA	12
Washington	DC	11
Rockville	MD	11
Richland	WA	10

Result 17

2. "Data Science" related vs others (volume & salary)

Query:

```
SELECT CASE
   WHEN Job_Title REGEXP 'Scientist|Science' THEN 'Data Science related'
   ELSE 'Other'
 END AS Science_bucket,
 COUNT (*) AS postings,
 Round (avg ((Lower_Salary_in_Thousand_$ +
Upper_Salary_in_Thousand_$)/2),2) as Average_Salary_in_K_$
FROM job_market
GROUP BY Science_bucket
ORDER BY postings DESC;
```



```
152  SELECT
153    CASE
154      WHEN Job_Title REGEXP 'Scientist|Science' THEN 'Data Science related'
155      ELSE 'Other'
156    END AS Science_bucket,
157    COUNT(*) AS postings,
158    Round(avg((Lower_Salary_in_Thousand_$ + Upper_Salary_in_Thousand_$)/2),2) as Average_Salary_in_K_$
159  FROM job_market
160  GROUP BY Science_bucket
161  ORDER BY postings DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Science_bucket	postings	Average_Salary_in_K_$
Data Science related	472	109.54
Other	270	87.40

Result 16

Output

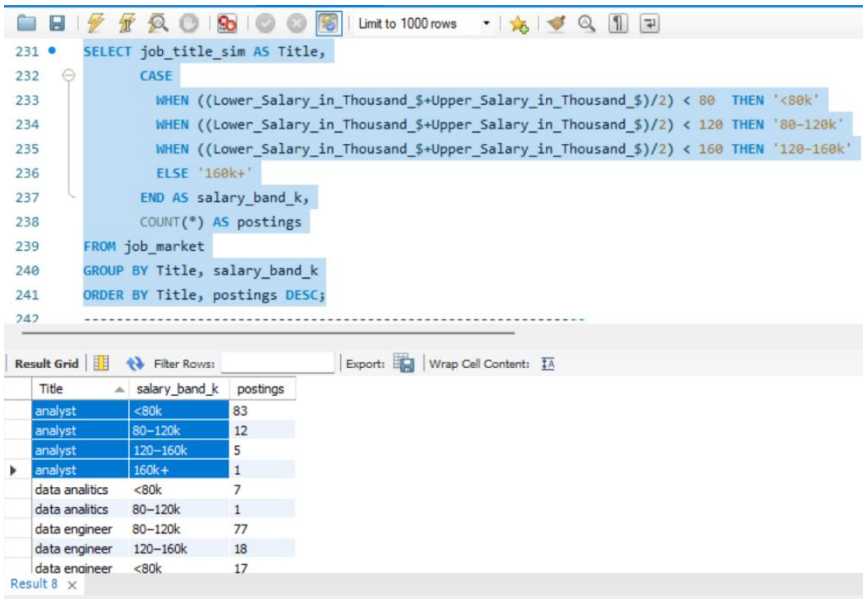
Action Output

#	Time	Action	Message
28	11:20:25	SELECT CASE WHEN Job_Title REGEXP 'Scientist|Science' THEN 'Da...	Error Code: 1055. Expression #3 of SELECT list is not in GROUP BY
29	11:22:49	SELECT CASE WHEN Job_Title REGEXP 'Scientist|Science' THEN 'Da...	2 row(s) returned

3. Salary bands per title (distribution)

Query:

```
SELECT job_title_sim AS Title,
CASE WHEN
((Lower_Salary_in_Thousand_$+Upper_Salary_in_Thousand_$)/2) < 80
THEN '<80k'
 WHEN ((Lower_Salary_in_Thousand_$+Upper_Salary_in_Thousand_$)/2) <
120 THEN '80–120k'
 WHEN ((Lower_Salary_in_Thousand_$+Upper_Salary_in_Thousand_$)/2) <
160 THEN '120–160k'
 ELSE '160k+'
 END AS salary_band_k,
 COUNT (*) AS postings
FROM job_market
GROUP BY Title, salary_band_k
ORDER BY Title, postings DESC;
```



```
SELECT job_title_sim AS Title,
      CASE
        WHEN ((Lower_Salary_in_Thousand_$+Upper_Salary_in_Thousand_$)/2) < 80  THEN '<80k'
        WHEN ((Lower_Salary_in_Thousand_$+Upper_Salary_in_Thousand_$)/2) < 120 THEN '80-120k'
        WHEN ((Lower_Salary_in_Thousand_$+Upper_Salary_in_Thousand_$)/2) < 160 THEN '120-160k'
        ELSE '160k+'
      END AS salary_band_k,
      COUNT(*) AS postings
FROM job_market
GROUP BY Title, salary_band_k
ORDER BY Title, postings DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

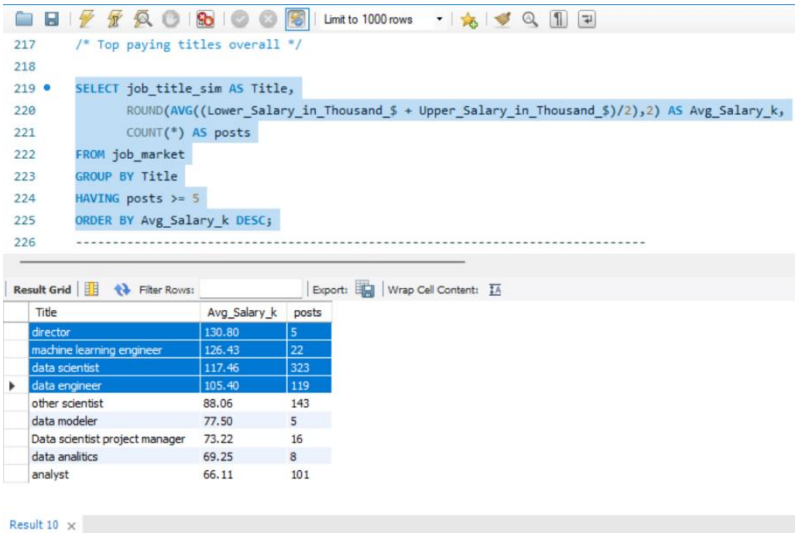
Title	salary_band_k	postings
analyst	<80k	83
analyst	80–120k	12
analyst	120–160k	5
analyst	160k+	1
data analitics	<80k	7
data analitics	80–120k	1
data engineer	80–120k	77
data engineer	120–160k	18
data engineer	<80k	17

Result 8

4. Top paying titles overall

Query:

```
SELECT job_title_sim AS Title,
ROUND (AVG ((Lower_Salary_in_Thousand_$ +
Upper_Salary_in_Thousand_$)/2),2) AS Avg_Salary_k,
COUNT (*) AS posts
FROM job_market
GROUP BY Title
HAVING posts >= 5
ORDER BY Avg_Salary_k DESC;
```



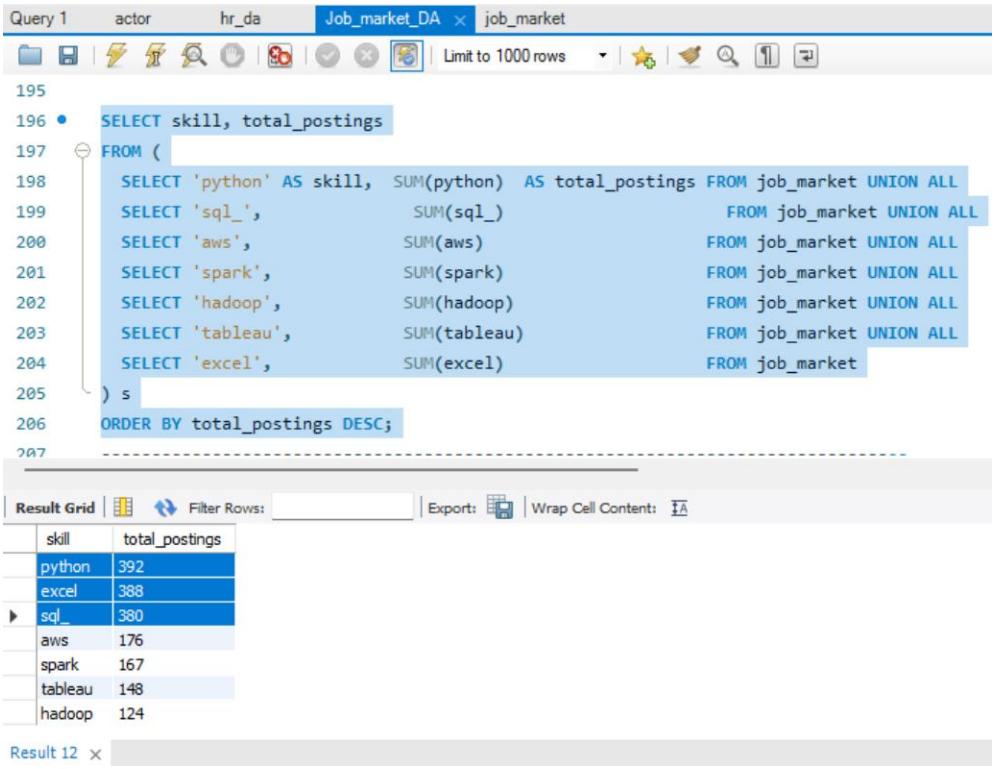
```
/* Top paying titles overall */

SELECT job_title_sim AS Title,
       ROUND(AVG((Lower_Salary_in_Thousand_$ + Upper_Salary_in_Thousand_$)/2),2) AS Avg_Salary_k,
       COUNT(*) AS posts
FROM job_market
GROUP BY Title
HAVING posts >= 5
ORDER BY Avg_Salary_k DESC;
```

Title	Avg_Salary_k	posts
director	130.80	5
machine learning engineer	126.43	22
data scientist	117.46	323
data engineer	105.40	119
other scientist	88.06	143
data modeler	77.50	5
Data scientist project manager	73.22	16
data analitics	69.25	8
analyst	66.11	101

5. Most in-demand skills (overall)

Query:

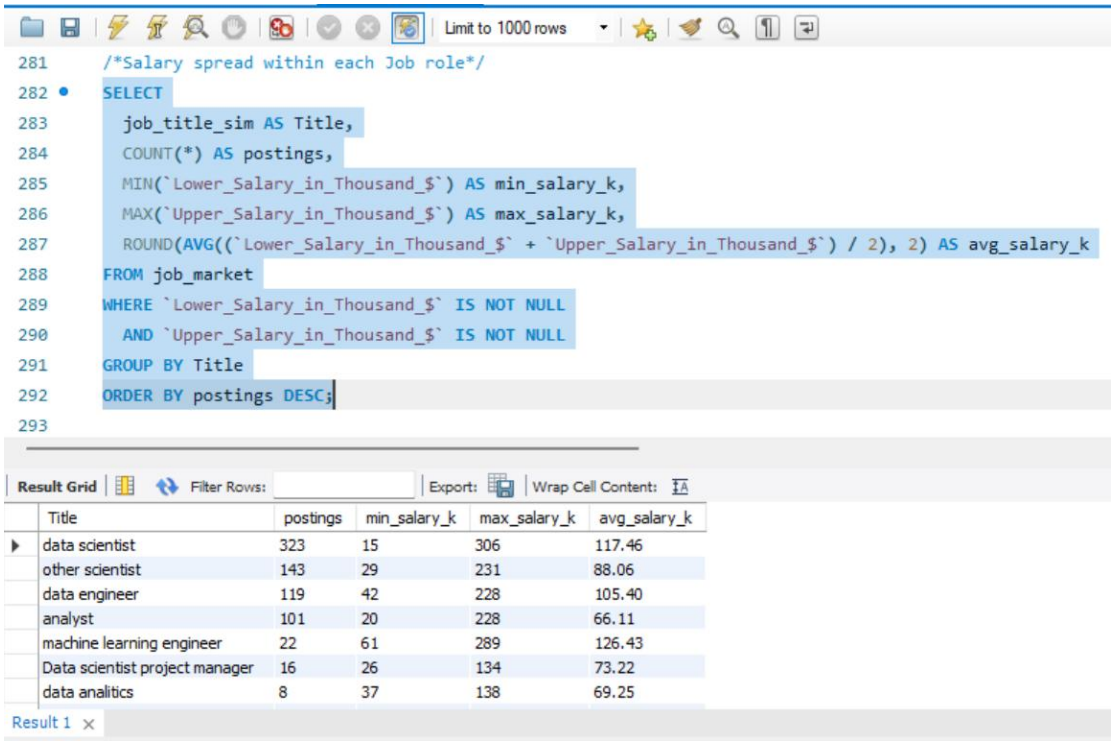
```
SELECT skill, total_postings
FROM (
SELECT 'python' AS skill, SUM (python) AS total_postings FROM job_market
UNION ALL
SELECT 'SQL', SUM (SQL) FROM job_market UNION ALL
SELECT 'aws', SUM (aws) FROM job_market UNION ALL
SELECT 'spark', SUM (spark) FROM job_market UNION ALL
SELECT 'hadoop', SUM (hadoop) FROM job_market UNION ALL
SELECT 'tableau', SUM (tableau) FROM job_market UNION ALL
SELECT 'excel', SUM (excel) FROM job_market) s
ORDER BY total_postings DESC;
```



6.Salary spread within each job role

Query:

```
SELECT job_title_sim AS Title,
COUNT (*) AS postings,
MIN(`Lower_Salary_in_Thousand_$`) AS min_salary_k,
MAX(`Upper_Salary_in_Thousand_$`) AS max_salary_k,
ROUND (AVG ((`Lower_Salary_in_Thousand_$` + `Upper_Salary_in_Thousand_$`) / 2), 2) AS avg_salary_k
FROM job_market
WHERE `Lower_Salary_in_Thousand_$` IS NOT NULL
AND `Upper_Salary_in_Thousand_$` IS NOT NULL
GROUP BY Title
ORDER BY postings DESC;
```



```
/*Salary spread within each Job role*/
SELECT
  job_title_sim AS Title,
  COUNT(*) AS postings,
  MIN(`Lower_Salary_in_Thousand_$`) AS min_salary_k,
  MAX(`Upper_Salary_in_Thousand_$`) AS max_salary_k,
  ROUND(AVG((`Lower_Salary_in_Thousand_$` + `Upper_Salary_in_Thousand_$`) / 2), 2) AS avg_salary_k
FROM job_market
WHERE `Lower_Salary_in_Thousand_$` IS NOT NULL
  AND `Upper_Salary_in_Thousand_$` IS NOT NULL
GROUP BY Title
ORDER BY postings DESC;
```

Result Grid

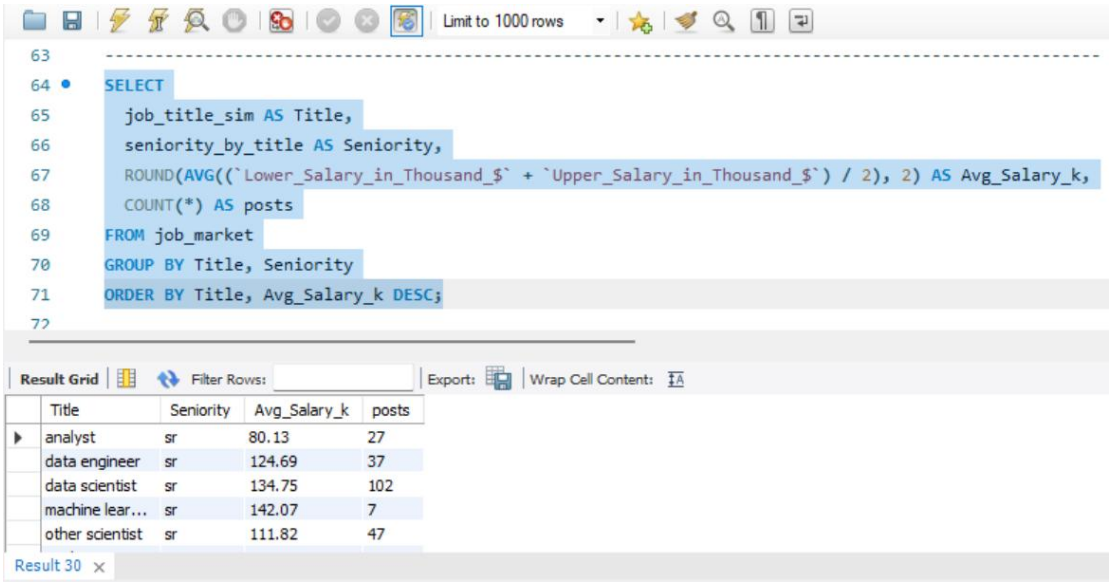
Title	postings	min_salary_k	max_salary_k	avg_salary_k
data scientist	323	15	306	117.46
other scientist	143	29	231	88.06
data engineer	119	42	228	105.40
analyst	101	20	228	66.11
machine learning engineer	22	61	289	126.43
Data scientist project manager	16	26	134	73.22
data analitics	8	37	138	69.25

Result 1

7. Average salary by seniority (within each title).

   Query:

```
SELECT
  job_title_sim AS Title,
  seniority_by_title AS Seniority,
  ROUND (AVG ((`Lower_Salary_in_Thousand_$` +
`Upper_Salary_in_Thousand_$`) / 2), 2) AS Avg_Salary_k,
  COUNT (*) AS posts
FROM job_market
GROUP BY Title, Seniority
ORDER BY Title, Avg_Salary_k DESC;
```



```
63  ------------------------------------------------------------------------------------------------
64  SELECT
65    job_title_sim AS Title,
66    seniority_by_title AS Seniority,
67    ROUND(AVG((`Lower_Salary_in_Thousand_$` + `Upper_Salary_in_Thousand_$`) / 2), 2) AS Avg_Salary_k,
68    COUNT(*) AS posts
69  FROM job_market
70  GROUP BY Title, Seniority
71  ORDER BY Title, Avg_Salary_k DESC;
72
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Title	Seniority	Avg_Salary_k	posts
analyst	sr	80.13	27
data engineer	sr	124.69	37
data scientist	sr	134.75	102
machine lear...	sr	142.07	7
other scientist	sr	111.82	47

Result 30

CHALLENGES DURING EXECUTION

Challenges and Solutions

1. Challenge: Dataset contained inconsistent, missing, and incorrectly formatted values.
Solution: Performed detailed data cleaning using Excel/SQL functions to remove nulls, fix formats, and ensure consistency before querying.

2. Challenge: Large dataset size and many columns resulted in slow query execution and difficulty identifying relevant fields.
Solution: Applied indexing, optimized SQL queries, and selected only necessary columns to improve performance and readability.

3. Challenge: Unclear or inconsistent column names made interpretation difficult.
Solution: Referred to dataset documentation/metadata and used SQL aliasing or renaming to make columns meaningful and easier to understand.

4. Challenge: Incorrect or inconsistent data types (e.g., numbers stored as text) caused errors during analysis.
Solution: Converted data types using SQL functions to ensure accurate filtering, aggregation, and calculations.

5. Challenge: Duplicate records across multiple columns and tables affected accuracy.
Solution: Identified and removed duplicate entries to eliminate redundancy and maintain data reliability.

6. Challenge: Long text fields and too many columns made visualization and chart selection difficult.
Solution: Simplified data by selecting essential fields, summarizing text columns, and using sample data to create clear and effective visualizations.

7. Challenge: Selecting suitable and visually balanced colour combinations for charts was difficult.
Solution: Experimented with different colour palettes and used Power BI's theme customization feature to ensure consistent and professional visuals.

8. Challenge: Crafting meaningful storytelling, clear titles, and impactful insights was challenging.
Solution: Analysed results carefully, identified key trends, and used concise, insight-driven language to improve storytelling quality.

BUSINESS INSIGHTS

- Data driven roles dominate hiring, with Data Scientist and Data Engineer leading, showing companies' strong push toward analytics and automation.
- Python, SQL, and Excel remain the most essential skills, that means the foundational technical capabilities are crucial in most jobs related to data.
- Junior positions are in highest demand, meaning organizations prefer talent that will scale, be cost-efficient, and grow with them.
- Increased average salary for Master's Degree holders. This shows that advanced education provides massive returns related to career and compensation.
- Industries like Biotech, Pharma, Insurance, and IT show the highest job volume, signaling where market growth and digital transformation are strongest.
- States like CA, IL, and DC offer the best opportunities in terms of salary and indicate geographic clusters where companies are willing to pay a premium for skilled talent.

POWERBI DASHBOARDS

INTRODUCTON

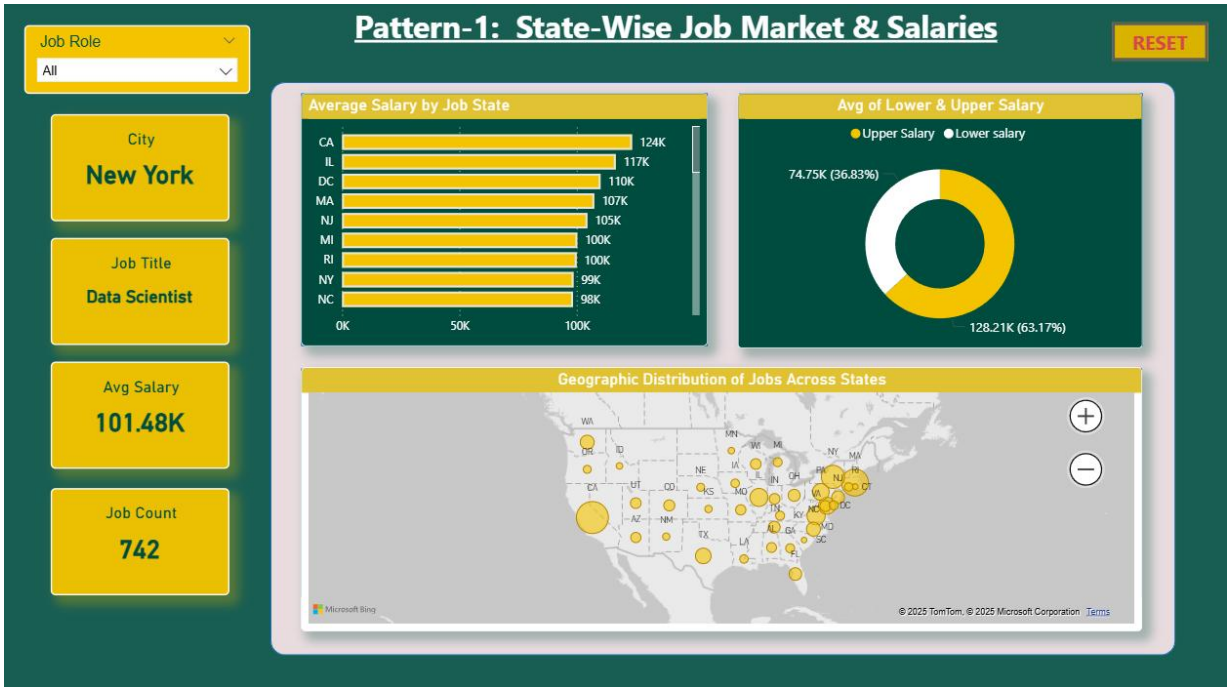
This report provides a comprehensive analysis of the U.S. job market, focusing on state-wise trends, industry demand, and the relationship between skills, education, and salary expectations. The insights are organized into three analytical patterns designed to highlight where opportunities exist, which roles and industries are growing, and what qualifications employers value most. The objective is to support job seekers, workforce planners, and policymakers in making informed, data-driven decisions.

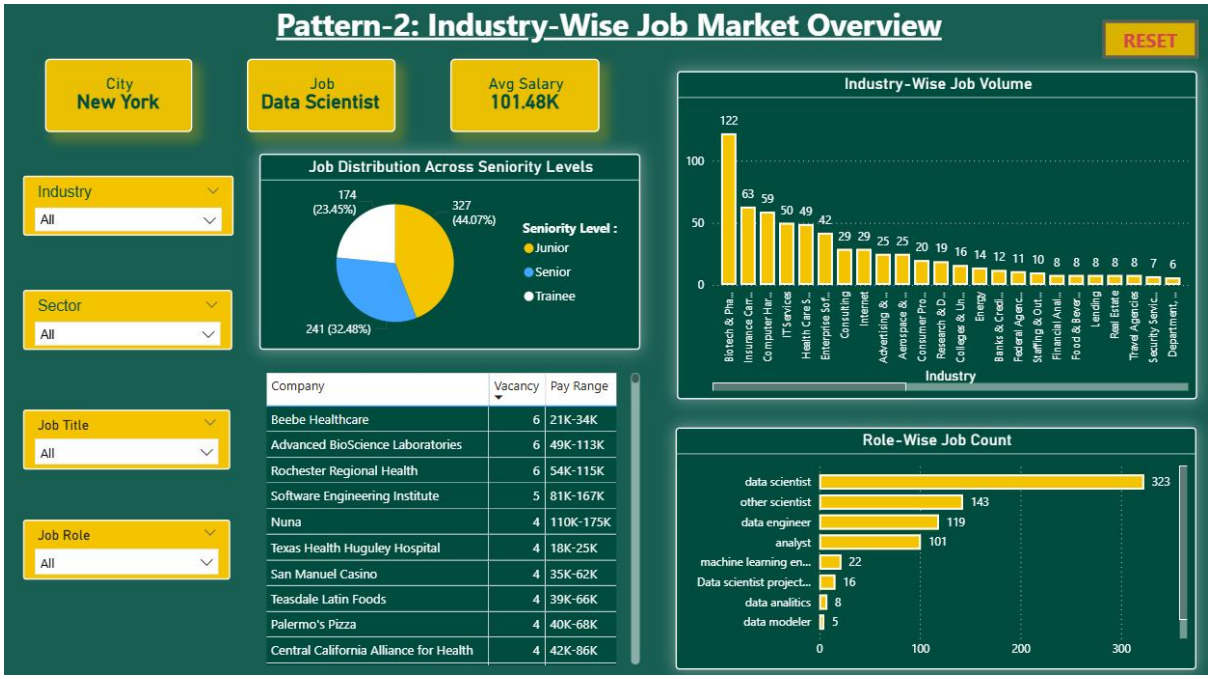
Pattern-1: State Job Trends
Examines how job availability and salary levels vary across states. This pattern reveals high-demand locations, competitive salary regions, and geographic clusters of job opportunities helping users identify where specific roles are most rewarding and accessible.

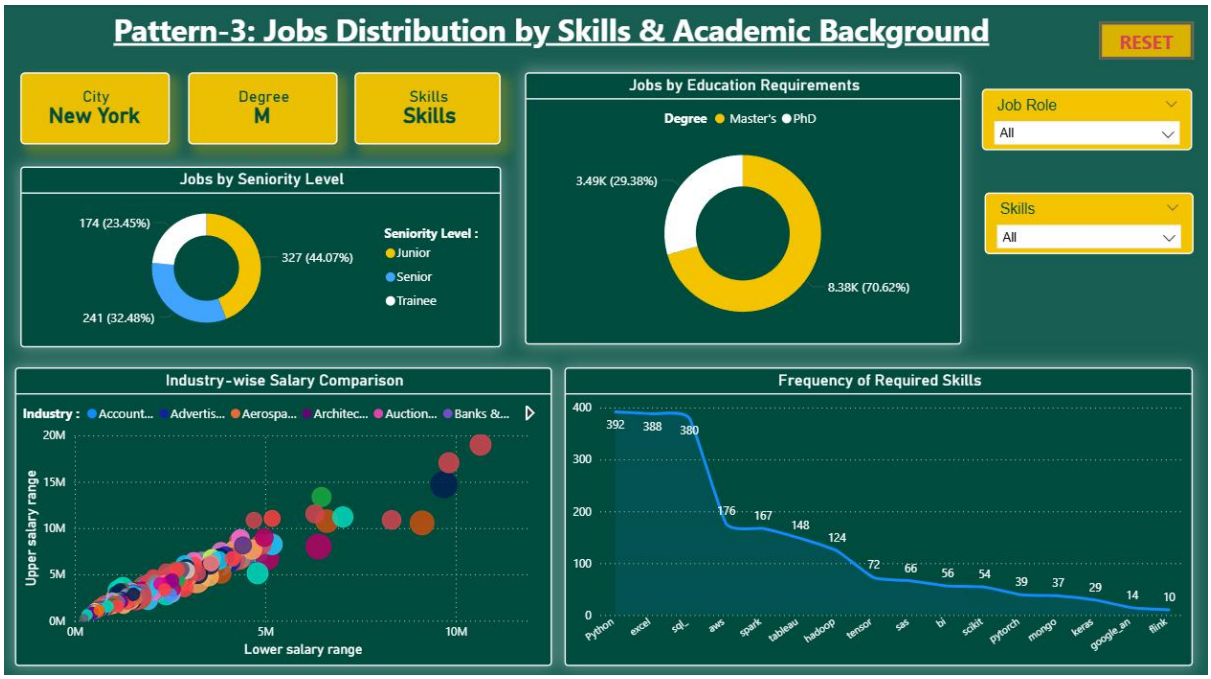
Pattern-2: Industry Job Insights
Analyzes job demand across industries, sectors, and seniority levels. This pattern uncovers which industries are hiring the most, what roles dominate the job market, and how vacancy counts differ across companies supporting strategic career planning and industry selection.

Pattern-3: Skills & Education Impact
Explores how required skills and educational qualifications influence job availability and salary ranges. This pattern highlights the most in-demand technical skills, the value of advanced degrees, and salary variations across industries guiding learners and professionals on which capabilities to develop for career advancement.

Note: All project questions and scenarios are covered within these patterns.




Pattern-2: Industry-Wise Job Market Overview
RESET
City
New York
Job
Data Scientist
Avg Salary
101.48K
Industry
All
Sector
All
Job Title
All
Job Role
All
Job Distribution Across Seniority Levels
174 (23.45%)
327 (44.07%)
241 (32.48%)
Seniority Level :
Junior
Senior
Trainee
Company
Vacancy
Pay Range
Beebe Healthcare 6 21K-34K
Advanced BioScience Laboratories 6 49K-113K
Rochester Regional Health 6 54K-115K
Software Engineering Institute 5 81K-167K
Nuna 4 110K-175K
Texas Health Huguley Hospital 4 18K-25K
San Manuel Casino 4 35K-62K
Teasdale Latin Foods 4 39K-66K
Palermo's Pizza 4 40K-68K
Central California Alliance for Health 4 42K-86K
Industry-Wise Job Volume
122
63
59
50
49
42
29
29
25
25
20
19
16
14
12
11
10
8
8
8
8
7
6
Industry
Role-Wise Job Count
data scientist 323
other scientist 143
data engineer 119
analyst 101
machine learning en... 22
Data scientist project... 16
data analitics 8
data modeler 5


Pattern-3: Jobs Distribution by Skills & Academic Background
RESET
City
New York
Degree
M
Skills
Skills
Jobs by Seniority Level
174 (23.45%)
327 (44.07%)
241 (32.48%)
Seniority Level :
Junior
Senior
Trainee
Jobs by Education Requirements
Degree Master's PhD
3.49K (29.38%)
8.38K (70.62%)
Job Role
All
Skills
All
Industry-wise Salary Comparison
Industry : Account... Advertis... Aerospa... Architec... Auction... Banks &...
Upper salary range
Lower salary range
Frequency of Required Skills
392
388
380
176
167
148
124
72
66
56
54
39
37
29
14
10
Python
excel
sql
aws
spark
tableau
hadoop
tensor
sas
bi
scikit
pytorch
mongo
keras
google_an
flink

CONCLUSION

- Data Scientist, Data Engineer, and Analyst are leading the job market with a high demand for data-oriented talent. This further points out the continued industry movement toward analytics-driven decision-making.
- Python, Excel, and SQL are the most critical skills, these three form the foundation of most high demand roles. The investment in these core skills greatly enhances job readiness and competitiveness.
- Junior level positions dominate the hiring landscape, meaning firms prefer entry-level talent that is scalable. This trend is indicative of strong opportunities for freshers and early-career professionals.
- Master's degree holders have higher average salaries than PhDs, which is indicative of strong ROI from professional degrees. Education still plays a key role in compensation and career advancement.
- High paying states like CA, IL, and DC and growth heavy industries like Biotech, Pharma, and IT are the drivers of the job market. The clusters showcase the areas where job seekers can target to get better salaries and long-term growth.