Project on SQL By Using MYSQL – Company Job Data

1. What is the average salary for all the jobs in the dataset?

```
SELECT AVG(salary) AS AVG_SALARY FROM company
```

AVG_SALARY
190695.5718

2. What is the highest salary in the dataset and which job role does it correspond to?

```
SELECT job_title, MAX(salary) FROM company
GROUP BY job_title
order by MAX(salary) desc
limit 1;
```

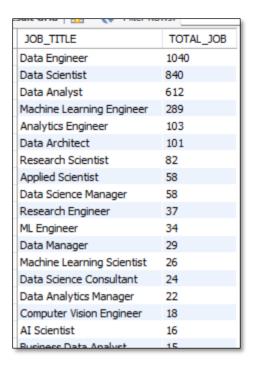| job_title	MAX(salary)
Data Scientist	30400000

3. What is the average salary for data scientists in US?

```
SELECT AVG(salary) FROM company
WHERE job_title = 'Data Scientist' AND company_location = 'US'
```

AVG(salary)
162548.1157

4. What is the number of jobs available for each job title?

```
SELECT JOB_TITLE , COUNT(*) AS TOTAL_JOB FROM company
GROUP BY JOB_TITLE ORDER BY TOTAL_JOB DESC;
```

JOB_TITLE	TOTAL_JOB
Data Engineer	1040
Data Scientist	840
Data Analyst	612
Machine Learning Engineer	289
Analytics Engineer	103
Data Architect	101
Research Scientist	82
Applied Scientist	58
Data Science Manager	58
Research Engineer	37
ML Engineer	34
Data Manager	29
Machine Learning Scientist	26
Data Science Consultant	24
Data Analytics Manager	22
Computer Vision Engineer	18
AI Scientist	16
Business Data Analyst	15

5. What is the total salary paid for all data analyst jobs in DE?

```sql
SELECT job_title,SUM(salary) FROM company
WHERE job_title = 'Data Analyst' AND company_location = 'DE';
```

job_title	TOTAL_SALARY
Data Analyst	54000

6. What are the top 5 highest paying job titles and their corresponding average salaries?

```sql
SELECT JOB_TITLE ,ROUND(AVG(SALARY)) AS AVG_SALARY FROM COMPANY
GROUP BY JOB_TITLE
ORDER BY AVG_SALARY DESC
LIMIT 5;
```

JOB_TITLE	AVG_SALARY
Head of Machine Learning	6000000
Principal Data Architect	3000000
Lead Machine Learning Engineer	2548667
Lead Data Scientist	928485
Data Analytics Lead	922500

7. What is the number of jobs available in each location?

```
select company_location,COUNT(*) AS TOTAL_JOB FROM company
GROUP BY company_location
ORDER BY TOTAL_JOB DESC;
```

company_location	TOTAL_JOB
US	3040
GB	172
CA	87
ES	77
IN	58
DE	56
FR	34
BR	15
GR	14
PT	14
AU	14
NL	13
MX	10
IE	7

8. What are the top 3 job titles that have the most jobs available in the dataset?

```
select company_location,COUNT(*) AS TOTAL_JOB FROM company
GROUP BY company_location
ORDER BY TOTAL_JOB DESC
LIMIT 3;
```

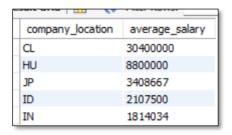| company_location	TOTAL_JOB
US	3040
GB	172
CA	87

9. What is the average salary for data engineers in Boston?

```
SELECT avg(salary) AS average_salary FROM COMPANY
WHERE job_title = 'Data Engineer' AND company_location = 'US';
```

average_salary
152793.6086

10. What are the top 5 cities with the highest average salaries?

```
SELECT company_location, ROUND(AVG(SALARY)) AS average_salary FROM COMPANY
GROUP BY company_location
ORDER BY average_salary DESC
LIMIT 5;
```

company_location	average_salary
CL	30400000
HU	8800000
JP	3408667
ID	2107500
IN	1814034

11. What is the average salary for each job title, and what is the total number of jobs available for each job title?

```
SELECT job_title, ROUND(AVG(SALARY)) AS average_salary,COUNT(*) AS TOTAL_JOB FROM COMPANY
GROUP BY job_title
ORDER BY TOTAL_JOB DESC
```

job_title	average_salary	TOTAL_JOB
Data Engineer	156575	1040
Data Scientist	239073	840
Data Analyst	116374	612
Machine Learning Engineer	182216	289
Analytics Engineer	151353	103
Data Architect	161714	101
Research Scientist	160769	82
Applied Scientist	190264	58
Data Science Manager	379390	58

12. What are the top 5 job titles with the highest total salaries, and what is the total number of jobs available for each job title?

```
SELECT job_title, SUM(SALARY) AS TOTAL_salary,COUNT(*) AS TOTAL_JOB FROM COMPANY
GROUP BY job_title
ORDER BY TOTAL_salary DESC
LIMIT 5;
```

job_title	TOTAL_salary	TOTAL_JOB
Data Scientist	200821720	840
Data Engineer	162837962	1040
Data Analyst	71220916	612
Machine Learning Engineer	52660434	289
Data Science Manager	22004644	58

13. What are the top 5 locations with the highest total salaries, and what is the total number of jobs available for each location?

```
SELECT company_location, SUM(SALARY) AS TOTAL_salary,COUNT(*) AS TOTAL_JOB FROM COMPANY
GROUP BY company_location
ORDER BY TOTAL_salary DESC
LIMIT 5;
```

company_location	TOTAL_salary	TOTAL_JOB
US	479302687	3040
IN	105213997	58
CL	30400000	1
JP	20452000	6
HU	17600000	2

14. What is the average salary for each job title in each location, and what is the total number of jobs available for each job title in each location?

```
SELECT job_title, company_location,ROUND(AVG(SALARY)) AS average_salary, COUNT(*) AS TOTAL_JOB FROM COMPANY
GROUP BY job_title,company_location;
```

job_title	company_location	average_salary	TOTAL_JOB
Data Engineer	US	152794	907
Data Scientist	US	162548	674
Data Analyst	US	116655	525
Machine Learning Engineer	US	174385	219
Data Architect	US	163093	97
Analytics Engineer	US	159024	92
Applied Scientist	US	190264	58
Research Scientist	US	179146	58
Data Engineer	GB	66369	55
Data Science Manager	US	275262	52
Research Engineer	US	168157	31
Data Scientist	GB	72231	31
ML Engineer	US	172712	28
Data Scientist	ES	40236	28
Data Analyst	GB	63782	25
Data Manager	US	125646	24

15. What is the average salary for each job title in each location, and what is the percentage of jobs for each job title in each location?

```sql
SELECT job_title, company_location,ROUND(AVG(SALARY)) AS average_salary,
(COUNT(*)*100/(SELECT COUNT(*) FROM COMPANY)) AS PERCENTAGE_TOTAL_JOB  FROM COMPANY
GROUP BY job_title,company_location
```

job_title	company_location	average_salary	PERCENTAGE_TOTAL_JOB
Data Engineer	US	152794	24.1545
Data Scientist	US	162548	17.9494
Data Analyst	US	116655	13.9814
Machine Learning Engineer	US	174385	5.8322
Data Architect	US	163093	2.5832
Analytics Engineer	US	159024	2.4501
Applied Scientist	US	190264	1.5446
Research Scientist	US	179146	1.5446
Data Engineer	GB	66369	1.4647
Data Science Manager	US	275262	1.3848
Research Engineer	US	168157	0.8256
Data Scientist	GB	72231	0.8256

16. What are the top 5 job titles with the highest average salaries, and what is the total number of jobs available for each job title?

```sql
SELECT job_title, ROUND(AVG(SALARY)) AS average_salary, COUNT(*) AS TOTAL_JOB FROM COMPANY
GROUP BY job_title
ORDER BY  average_salary DESC
LIMIT  5;
```

job_title	average_salary	TOTAL_JOB
Head of Machine Learning	6000000	1
Principal Data Architect	3000000	1
Lead Machine Learning Engineer	2548667	3
Lead Data Scientist	928485	9
Data Analytics Lead	922500	2

17. What is the average salary for each job title, and what is the percentage of jobs for each job title in the dataset?

```sql
SELECT job_title, ROUND(AVG(SALARY)) AS average_salary,
(COUNT(*)*100/(SELECT COUNT(*) FROM COMPANY)) AS PERCENTAGE_TOTAL_JOB FROM COMPANY
GROUP BY job_title;
```

job_title	average_salary	PERCENTAGE_TOTAL_JOB
Principal Data Scientist	190875	0.2130
ML Engineer	609998	0.9055
Data Scientist	239073	22.3702
Applied Scientist	190264	1.5446
Data Analyst	116374	16.2983
Data Modeler	118900	0.0533
Research Engineer	162753	0.9854
Analytics Engineer	151353	2.7430
Business Intelligence Engineer	174150	0.1065
Machine Learning Engineer	182216	7.6964

18. What is the total number of jobs available for each year of experience, and what is the average salary for each year of experience?

```
SELECT experience_level,COUNT(*) NUM_JOBS, AVG(SALARY) FROM COMPANY
GROUP BY experience_level;
```

experience_level	NUM_JOBS	AVG(SALARY)
SE	2516	170048.9654
MI	805	248200.3068
EN	320	188381.1781
EX	114	246802.2018

19. What is the average salary for each experience level, and what is the total number of jobs available for each degree level?

```
select experience_level,AVG(SALARY) AS AVG_SALARY, COUNT(*) AS num_jobs FROM COMPANY
GROUP BY experience_level;
```

experience_level	AVG_SALARY	num_jobs
SE	170048.9654	2516
MI	248200.3068	805
EN	188381.1781	320
EX	246802.2018	114

20. What are the top 5 job titles with the highest salaries, and what is the name of the company that offers the highest salary for each job title?

```
SELECT job_title, MAX(salary) AS HIGHEST_SALARY, companies.company_name
FROM company
INNER JOIN companies ON company.id = companies.id
GROUP BY job_title, company_name
ORDER BY HIGHEST_SALARY DESC
LIMIT 5;
```

job_title	HIGHEST_SALARY	company_name
Data Scientist	30400000	yz
Data Scientist	11000000	jkl
BI Data Analyst	11000000	jkl
ML Engineer	8500000	jkl
Lead Machine Learning Engineer	7500000	ghi

21. What is the total number of jobs available for each job title, and what is the name of the
company that offers the highest salary for each job title?

```
SELECT job_title, COUNT(*) AS num_jobs, company_name
FROM company
INNER JOIN companies ON company.id = companies.id
WHERE salary = (SELECT MAX(salary) FROM company WHERE job_title = company.job_title)
GROUP BY job_title, company_name;
```

job_title	num_jobs	company_name
Data Scientist	1	yz