Case Study: Data Science Jobs Analysis

Exploring Trends in Data Science Salaries, Remote Work, and Employment Types

Name: Priya Chakradhari
Contact Information: priyachakradhari4@gmail.com
LinkedIn Profile: linkedin.com/in/priya-chakradhari
GitHub: github.com/priya-chakradhari

INTRODUCTION

The "Data Science Jobs 2024" dataset provides a comprehensive overview of data science job salaries and related details for the year 2024. This dataset includes key attributes such as experience level, employment type, job title, salary details, employee residence, remote work ratio, company location, and company size.

The dataset is designed to help analyze trends in data science roles, including salary distributions across different experience levels, employment types, and geographic locations. It also provides insights into the growing trend of remote work and its impact on compensation in the data science field. By exploring this dataset, stakeholders can gain valuable insights into the data science job market, enabling better decision-making for hiring, career planning, and industry benchmarking.

DATASET DESCRIPTION

The provided files contain information about data science job salaries for the year 2024. Here is a summary of the key details:

1. Experience Level: Indicates the level of experience required for the job, with possible values being:
   - EN: Entry-level / Junior
   - M: Mid-level / Intermediate
   - SE: Senior-level / Expert
   - EX: Executive-level / Director

2. Employment Type: Describes the type of employment for the role, including:
   - PT: Part-time
   - FT: Full-time
   - CT: Contract
   - FL: Freelance

3. Job Title: Specifies the role worked in during the year.

4. Salary: The total gross salary amount paid.

5. Salary Currency: The currency of the salary paid, represented as an ISO 4217 currency code.

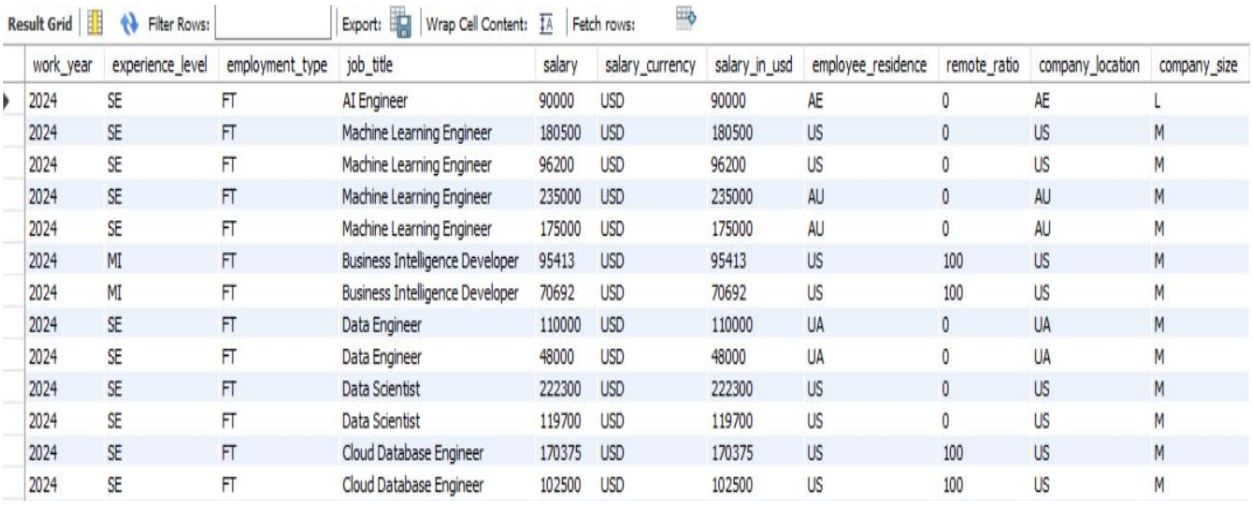
6. Salary in USD: The salary converted to USD using the average USD rate for the respective year.

7. Employee Residence: The primary country of residence of the employee during the work year, represented as an ISO 3166 country code.

8. Remote Ratio: Indicates the amount of work done remotely:
- 0: No remote work (less than 20%)
- 50: Partially remote
- 100: Fully remote (more than 80%)

9. Company Location: The country of the employer's main office or contracting branch, represented as an ISO 3166 country code.

10. Company Size: The average number of people that worked for the company during the year:
- S: Less than 50 employees (small)
- M: 50 to 250 employees (medium)
- L: More than 250 employees (large)


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

work_year	experience_level	employment_type	job_title	salary	salary_currency	salary_in_usd	employee_residence	remote_ratio	company_location	company_size
2024	SE	FT	AI Engineer	90000	USD	90000	AE	0	AE	L
2024	SE	FT	Machine Learning Engineer	180500	USD	180500	US	0	US	M
2024	SE	FT	Machine Learning Engineer	96200	USD	96200	US	0	US	M
2024	SE	FT	Machine Learning Engineer	235000	USD	235000	AU	0	AU	M
2024	SE	FT	Machine Learning Engineer	175000	USD	175000	AU	0	AU	M
2024	MI	FT	Business Intelligence Developer	95413	USD	95413	US	100	US	M
2024	MI	FT	Business Intelligence Developer	70692	USD	70692	US	100	US	M
2024	SE	FT	Data Engineer	110000	USD	110000	UA	0	UA	M
2024	SE	FT	Data Engineer	48000	USD	48000	UA	0	UA	M
2024	SE	FT	Data Scientist	222300	USD	222300	US	0	US	M
2024	SE	FT	Data Scientist	119700	USD	119700	US	0	US	M
2024	SE	FT	Cloud Database Engineer	170375	USD	170375	US	100	US	M
2024	SE	FT	Cloud Database Engineer	102500	USD	102500	US	100	US	M

Questions

1. You're a Compensation analyst employed by a multinational corporation. Your Assignment is to Pinpoint Countries who give work fully remotely, for the title 'managers' Paying salaries Exceeding $90,000 USD

2. AS a remote work advocate Working for a progressive HR tech startup who place their freshers' clients IN large tech firms. You're tasked WITH Identifying top 5 Country Having greatest count of large (company size) number of companies.

3. Picture yourself AS a data scientist Working for a workforce management platform. Your objective is to calculate the percentage of employees. Who enjoy fully remote roles WITH salaries Exceeding $100,000 USD, Shedding light ON the attractiveness of high-paying remote positions IN today's job market.

4. Imagine you're a data analyst Working for a global recruitment agency. Your Task is to identify the Locations where entry-level average salaries exceed the average salary for that job title IN market for entry level, helping your agency guide candidates towards lucrative opportunities.

5. You've been hired by a big HR Consultancy to look at how much people get paid IN different Countries. Your job is to Find out for each job title which. Country pays the maximum average

salary. This helps you to place your candidates IN those countries.

6. AS a data-driven Business consultant, you've been hired by a multinational corporation to analyze salary trends across different company Locations. Your goal is to Pinpoint Locations WHERE the average salary Has consistently Increased over the Past few years (Countries WHERE data is available for 3 years Only(present year and past two years) providing Insights into Locations experiencing Sustained salary growth.

7. Picture yourself AS a workforce strategist employed by a global HR tech startup. Your Mission is to Determine the percentage of fully remote work for each experience level IN 2021 and compare it WITH the corresponding figures for 2024, Highlighting any significant Increases or decreases IN remote work Adoption over the years.

8. AS a Compensation specialist at a Fortune 500 company, you're tasked WITH analyzing salary trends over time. Your objective is to calculate the average salary increase percentage for each experience level and job title between the years 2023 and 2024, helping the company stay competitive IN the talent market.

9. You're a database administrator tasked with role-based access control for a company's employee database. Your goal is to implement a security measure where employees in different

experience level (e.g. Entry Level, Senior level etc.) can only access details relevant to their respective experience level, ensuring data confidentiality and minimizing the risk of unauthorized access.

10. You are working with a consultancy firm, your client comes to you with certain data and preferences such as (their year of experience , their employment type, company location and company size) and want to make an transaction into different domain in data industry (like a person is working as a data analyst and want to move to some other domain such as data science or data engineering etc.) your work is to guide them to which domain they should switch to base on the input they provided, so that they can now update their knowledge as per the suggestion/.. The Suggestion should be based on average salary.

11. As a market researcher, your job is to Investigate the job market for a company that analyzes workforce data. Your Task is to know how many people were employed IN different types of companies AS per their size IN 2021.

12. Imagine you are a talent Acquisition specialist Working for an International recruitment agency. Your Task is to identify the top 3 job titles that command the highest average salary Among part-time Positions IN the year 2023. However, you are Only Interested IN Countries WHERE there are more than 50 employees, Ensuring a robust sample size for your analysis.

13. As a database analyst you have been assigned the task to Select Countries where average mid-level salary is higher than overall mid-level salary for the year 2023.

14. As a database analyst you have been assigned the task to Identify the company locations with the highest and lowest average salary for senior-level (SE) employees in 2023.

15. You're a Financial analyst Working for a leading HR Consultancy, and your Task is to Assess the annual salary growth rate for various job titles. By Calculating the percentage Increase IN salary FROM previous year to this year, you aim to provide valuable Insights Into salary trends WITHIN different job roles.

16. You've been hired by a global HR Consultancy to identify Countries experiencing significant salary growth for entry-level roles. Your task is to list the top three Countries with the highest salary growth rate FROM 2020 to 2023, Considering Only companies with more than 50 employees, helping multinational Corporations identify Emerging talent markets.

17. Picture yourself as a data architect responsible for database management. Companies in US and AU(Australia) decided to create a hybrid model for employees they decided that employees earning salaries exceeding $90000 USD, will be given work from home. You now need to update the remote work ratio for eligible employees, ensuring efficient remote work management while implementing appropriate error handling mechanisms for invalid input parameters.

18. In the year 2024, due to increased demand in the data industry, there was an increase in salaries of data field employees.
a. Entry Level-35% of the salary.
b. Mid junior – 30% of the salary.
c. Immediate senior level- 22% of the salary.
d. Expert level- 20% of the salary.
e. Director – 15% of the salary.
You must update the salaries accordingly and update them back in the original database.

19. You are a researcher and you have been assigned the task to Find the year with the highest average salary for each job title.

20. You have been hired by a market research agency where you been assigned the task to show the percentage of different employment type (full time, part time) in Different job roles, in the format where each row will be job title, each column will be type of employment type and cell value for that row and column will show the % value.

ANALYSIS and INSIGHTS

1. You're a Compensation analyst employed by a multinational corporation. Your Assignment is to Pinpoint Countries who give work fully remotely, for the title 'managers' Paying salaries Exceeding $90,000 USD.

```
select * from salaries
where remote_ratio = 100 and salary_in_usd > 90000 and job_title like '%manager%';
```

OUTPUT:

work_year	experience_level	employment_type	job_title	salary	salary_currency	salary_in_usd	employee_residence	remote_ratio	company_location	company_size
2024	MI	FT	Data Manager	98500	USD	98500	US	100	US	M
2024	MI	FT	Data Product Manager	171000	USD	171000	US	100	US	M
2024	MI	FT	Data Product Manager	136000	USD	136000	US	100	US	M
2024	SE	FT	Data Manager	131200	USD	131200	US	100	US	M
2024	SE	FT	Data Manager	95300	USD	95300	US	100	US	M
2024	MI	FT	Data Manager	127000	USD	127000	US	100	US	M
2024	SE	FT	Data Product Manager	170000	USD	170000	US	100	US	M

2. AS a remote work advocate Working for a progressive HR tech startup who place their freshers' clients IN large tech firms. you're tasked WITH Identifying top 5 Country Having greatest count of large (company size) number of companies.

```
select company_location, count(*) as counts from salaries
where company_size = 'L' and experience_level = 'EN'
group by company_location
order by counts desc
limit 5;
```

OUTPUT:

company_location	counts
US	53
DE	10
CA	10
GB	8
IN	6

3. Picture yourself AS a data scientist Working for a workforce management platform. Your objective is to calculate the percentage of employees. Who enjoy fully remote roles WITH salaries Exceeding $100,000 USD, Shedding light ON the attractiveness of high-paying remote positions IN today's job market.

```
set @total = (select count(*) as remote_emp from salaries where salary_in_usd > 100000);
set @cns = (select count(*) as remote_exd_salaries from salaries where remote_ratio =100 and salary_in_usd > 100000);
select @total;
select @cns;
set @percentage = (select round(((@cns)/(@total))*100,2) as perc) ;
select @percentage ;
```

OUTPUT:

@total
10563

@cns
3405

@percentage
32.24

4. Imagine you're a data analyst Working for a global recruitment agency. Your Task is to identify the Locations where entry-level average salaries exceed the average salary for that job title IN market for entry level, helping your agency guide candidates towards lucrative opportunities.

```
select b.company_location, b.job_title, a.average, b.average_per_country from
(
select job_title, round(avg(salary_in_usd),2) as average from salaries
where experience_level = "EN"
group by  job_title
)a
inner join
(
select company_location,job_title, round(avg(salary_in_usd),2) as average_per_country from salaries
where experience_level = "EN"
group by company_location, job_title
)b
on a.job_title=b.job_title
where average_per_country > average;
```

OUTPUT:

company_location	job_title	average	average_per_country
US	Data Analyst	84808.64	89800.35
US	Analytics Engineer	96722.30	110831.25
US	Data Engineer	92713.47	106791.26
US	Research Analyst	107294.21	110459.54
US	Machine Learning Engineer	110718.38	126188.85
AU	Business Intelligence Analyst	76688.63	91000.00
US	Business Intelligence Analyst	76688.63	79243.95
US	Research Engineer	135467.72	162668.18
CA	Machine Learning Research En...	55653.67	80769.00

5. You've been hired by a big HR Consultancy to look at how much people get paid IN different Countries. Your job is to Find out for each job title which. Country pays the maximum average salary. This helps you to place your candidates IN those countries.

```
select * from
    (
        select *,
        dense_rank() over(partition by job_title order by average desc) num
        from
        (
        select company_location, job_title, avg(salary_in_usd) as average from salaries
        group by job_title, company_location
        ) t
    )k
where num = 1;
```

OUTPUT:

company_location	job_title	average	num
IN	Admin & Data Analyst	60000.0000	1
CA	AI Architect	800000.0000	1
CA	AI Developer	275000.0000	1
QA	AI Engineer	300000.0000	1
US	AI Product Manager	152650.0000	1

6. AS a data-driven Business consultant, you've been hired by a multinational corporation to analyze salary trends across different company Locations. Your goal is to Pinpoint Locations WHERE the average salary Has consistently Increased over the Past few years (Countries WHERE data is available for
3 years Only(present year and past two years) providing Insights into Locations experiencing Sustained salary growth.

```
with t as (
   select * from salaries where company_location in
   (
       select company_location from
       (
           select  company_location,avg(salary_in_usd) as average, count(distinct (work_year)) as cns from salaries
           where work_year >=(year(current_date())-3)
           group by company_location
           having cns =3
       ) a
   )
)
select company_location,
max(case when work_year=2022 then average end) as avg_2022,
max(case when work_year=2023 then average end) as avg_2023,
max(case when work_year=2024 then average end) as avg_2024
from (select company_location, work_year, avg(salary_in_usd) as average from t group by company_location, work_year) q
group by company_location having avg_2024>avg_2023 and avg_2023>avg_2022;
```

OUTPUT:

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

company_location	avg_2022	avg_2023	avg_2024
CA	126009.5526	150724.1414	153611.8077
ES	47997.3415	60327.9857	72184.6667
FI	63040.0000	71259.0000	77777.0000
FR	72684.4667	100411.1905	101370.1667
PT	48921.3750	51521.0000	53054.7500
AR	50000.0000	65000.0000	88500.0000
IN	37328.3333	47777.5217	71538.3333
HU	17684.0000	43000.0000	63333.0000

7. Picture yourself AS a workforce strategist employed by a global HR tech startup. Your Mission is to Determine the percentage of fully remote work for each experience level IN 2021 and compare it WITH the corresponding figures for 2024, Highlighting any significant Increases or decreases
IN remote work Adoption over the years.

```
with t1 as
(
   select a.experience_level, a.total_2021,b.counts_2021,(counts_2021/total_2021)*100 as per_2021 from
   (
       select experience_level, count(experience_level) As total_2021 from salaries where work_year = 2021
       group by experience_level
   ) a
   inner join
   (
       select experience_level, count(experience_level) As counts_2021 from salaries where work_year = 2021 and remote_ratio= 100
       group by experience_level
   ) b
   on a.experience_level=b.experience_level
),
t2 as
(
   select c.experience_level, c.total_2024, d.counts_2024, (counts_2024/total_2024)*100 as per_2024 from
   (
       select experience_level, count(experience_level) As total_2024 from salaries where work_year = 2024
       group by experience_level
   ) c
   inner join
   (
       select experience_level, count(experience_level) As counts_2024 from salaries where work_year = 2024 and remote_ratio= 100
       group by experience_level
   ) d
   on c.experience_level=d.experience_level
)
select * from t1 inner join t2 on t1.experience_level=t2.experience_level;
```

OUTPUT:

experience_level	total_2021	counts_2021	per_2021	experience_level	total_2024	counts_2024	per_2024
SE	75	44	58.6667	SE	1920	483	25.1563
MI	87	45	51.7241	MI	1102	227	20.5989
EN	46	22	47.8261	EN	381	87	22.8346
EX	10	5	50.0000	EX	106	35	33.0189

8. AS a Compensation specialist at a Fortune 500 company, you're tasked WITH analyzing salary trends over time. Your objective is to calculate the average salary increase percentage for each experience level and job title between the years 2023 and 2024, helping the company stay competitive
IN the talent market.

```
select a.experience_level,a.job_title,a.average_2023,b.average_2024,
round((((average_2024-average_2023)/average_2023)*100),2)  AS changes_in_percentage
from
(
select experience_level,job_title,work_year,round(avg(salary_in_usd),2) as average_2023 from salaries where work_year = 2023
group by experience_level, job_title
) a
inner join
(
select experience_level,job_title,work_year, round(avg(salary_in_usd),2) as average_2024 from salaries where work_year = 2024
group by experience_level, job_title
) b
on a.job_title=b.job_title and a.experience_level = b.experience_level
where round((((average_2024-average_2023)/average_2023)*100),2) is not null;
```

OUTPUT:

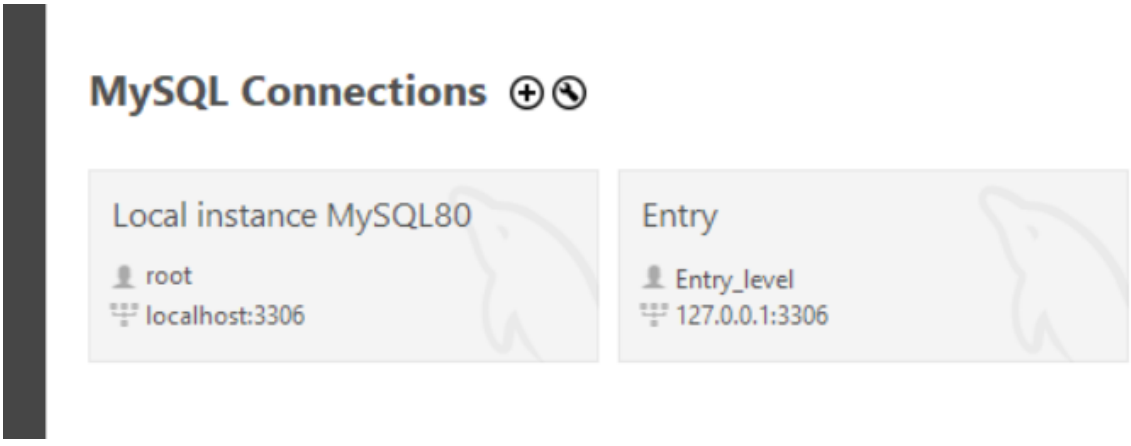
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

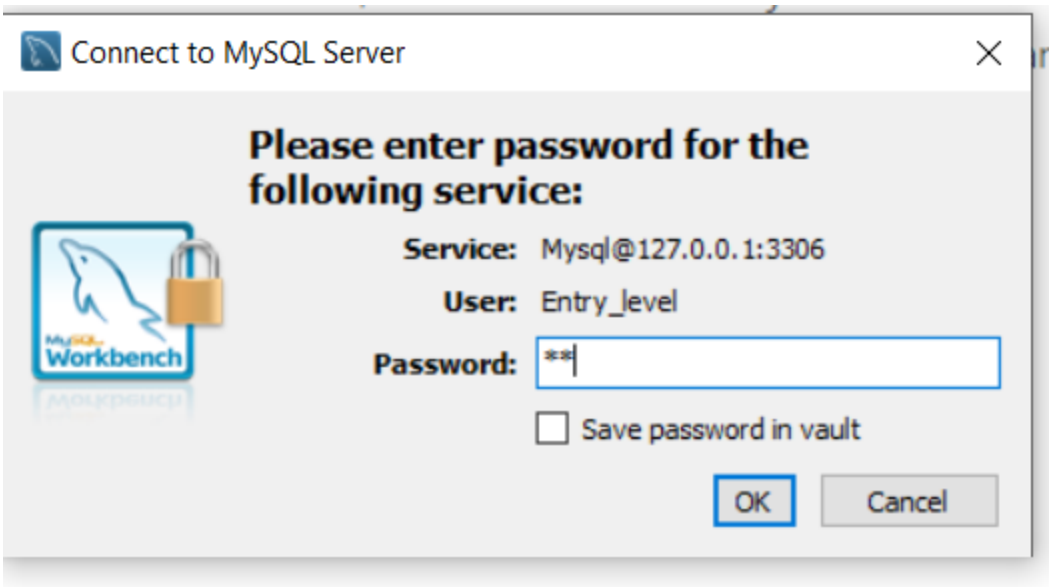
experience_level	job_title	average_2023	average_2024	changes_in_percentage
SE	AI Engineer	172245.94	180068.57	4.54
SE	Machine Learning Engineer	196167.59	206863.44	5.45
MI	Business Intelligence Developer	84032.00	83385.63	-0.77
SE	Data Engineer	158309.32	161949.40	2.30
SE	Data Scientist	173480.98	160234.25	-7.64
SE	Cloud Database Engineer	141666.67	136437.50	-3.69
MI	Data Engineer	124952.02	125574.88	0.50
MI	Machine Learning Engineer	161348.47	162126.76	0.48
MI	Research Engineer	149161.96	233809.50	56.75
EN	Data Analyst	76922.90	94439.16	22.77
SE	Machine Learning Scientist	183504.18	172770.29	-5.85
SE	Research Engineer	193786.28	206686.04	6.66

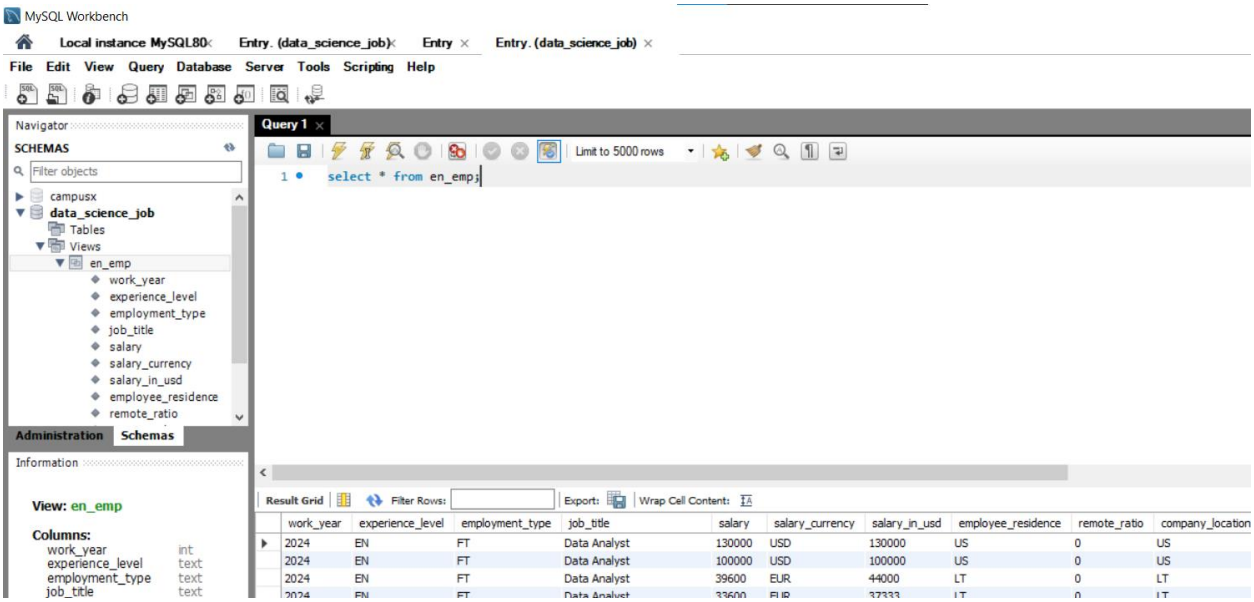
9. You're a database administrator tasked with role-based access control for a company's employee database. Your goal is to implement a security measure where employees in different experience level (e.g. Entry Level, Senior level etc.) can only access details relevant to their respective experience level, ensuring data confidentiality and minimizing the risk of unauthorized access.

```
select * from salaries;
create user 'Entry_level'@'%' identified by 'EN';
create view EN_emp as
(select * from salaries where experience_level = 'EN');

grant select on campusx.EN_emp to 'Entry_level'@'%';

show privileges;
```

OUTPUT:



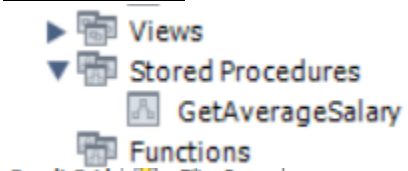

Connect to MySQL Server
Please enter password for the following service:
Service: Mysql@127.0.0.1:3306
User: Entry_level
Password: **
Save password in vault
OK
Cancel
Workbench


MySQL Workbench
Local instance MySQL80
Entry. (data_science_job)
Entry
Entry. (data_science_job)
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter objects
campusx
data_science_job
Tables
Views
en_emp
work_year
experience_level
employment_type
job_title
salary
salary_currency
salary_in_usd
employee_residence
remote_ratio
Administration
Schemas
Information
View: en_emp
Columns:
work_year int
experience_level text
employment_type text
job_title text
Query 1
Limit to 5000 rows
1 select * from en_emp;
Result Grid
Filter Rows:
Export:
Wrap Cell Content:
work_year experience_level employment_type job_title salary salary_currency salary_in_usd employee_residence remote_ratio company_location
2024 EN FT Data Analyst 130000 USD 130000 US 0 US
2024 EN FT Data Analyst 100000 USD 100000 US 0 US
2024 EN FT Data Analyst 39600 EUR 44000 LT 0 LT

10. You are working with a consultancy firm, your client comes to you with certain data and preferences such as (their year of experience, their employment type, company location and company size) and want to make an transaction into different domain in data industry (like a person is working as a data analyst and want to move to some other domain such as data science or data engineering etc.) your work is to guide them to which domain they should switch to base on the input they provided, so that they can now update their knowledge as per the suggestion/. The Suggestion should be based on average salary.

```
DELIMITER //
create procedure GetAverageSalary(in exp_lev varchar(2), IN emp_type varchar(3), IN comp_loc varchar(2), IN comp_size varchar(2))
begin
    select job_title, experience_level, company_location, company_size, employment_type, ROUND(avg(salary), 2) AS avg_salary
    from salaries
    where experience_level = exp_lev and company_location = comp_loc and company_size = comp_size and employment_type = emp_type
    group by experience_level, employment_type, company_location, company_size, job_title order by avg_salary desc ;
end//

call GetAverageSalary('EN','FT','AU','M')
```

OUTPUT:



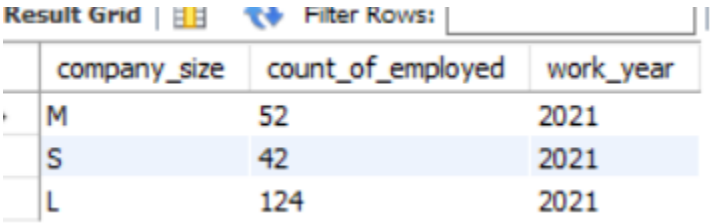
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

job_title	experience_level	company_location	company_size	employment_type	avg_salary
Data Scientist	EN	AU	M	FT	120000.00
Business Intelligence Analyst	EN	AU	M	FT	91000.00
AI Programmer	EN	AU	M	FT	40000.00
Machine Learning Developer	EN	AU	M	FT	40000.00
Data Analyst	EN	AU	M	FT	36276.50

11.As a market researcher, your job is to Investigate the job market for a company that analyzes workforce data. Your Task is to know how many people were employed IN different types of companies AS per their size IN 2021.

```
select  company_size,count(company_size) as count_of_employed, work_year from salaries
where work_year=2021
group by company_size;
```

OUTPUT:



Result Grid | Filter Rows:

company_size	count_of_employed	work_year
M	52	2021
S	42	2021
L	124	2021

12.Imagine you are a talent Acquisition specialist Working for an International recruitment agency. Your Task is to identify the top 3 job titles that command the highest average salary Among part-time Positions IN the year 2023.

```
select job_title, avg(salary_in_usd) as avg_salary from salaries
where employment_type = 'PT'and work_year = 2023
group by job_title
order by avg(salary_in_usd) desc
limit 3;
```

OUTPUT:

	job_title	avg_salary
▶	Data Scientist	95650.0000
	Data Analyst	18160.0000

13. As a database analyst you have been assigned the task to Select Countries where average mid-level salary is higher than overall mid-level salary for the year 2023.

```
select company_location, job_title, salary_in_usd, work_year from salaries
where experience_level='MI' and work_year = 2023 and salary_in_usd >
(
select round(avg(salary_in_usd),2) as avg_salary from salaries
where experience_level = 'MI' and work_year = 2023
)
group by company_location,job_title, salary_in_usd;
```

OUTPUT:

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

company_location	job_title	salary_in_usd	work_year
US	Data Science Manager	190000	2023
US	Data Science Manager	160000	2023
US	Data Engineer	160000	2023
US	Data Engineer	130000	2023
US	Machine Learning Engineer	166000	2023
US	Machine Learning Engineer	165000	2023

14.As a database analyst you have been assigned the task to Identify the company locations with the highest and lowest average salary for senior-level (SE) employees in 2023.

```
select company_location, round(avg(salary_in_usd),2) as highest_avg_salary from salaries
where experience_level = 'SE' and work_year= 2023
group by company_location
order by round(avg(salary_in_usd),2) desc
limit 1;

select company_location, round(avg(salary_in_usd),2) as lowerst_avg_salary from salaries
where experience_level = 'SE' and work_year= 2023
group by company_location
order by round(avg(salary_in_usd),2)
limit 1;
```

OUTPUT:

company_location	highest_avg_salary
IL	266468.50

company_location	lowerst_avg_salary
TR	18381.00

15. You're a Financial analyst Working for a leading HR Consultancy, and your Task is to Assess the annual salary growth rate for various job titles. By Calculating the percentage Increase IN salary FROM previous year to this year, you aim to provide valuable Insights Into salary trends WITHIN different job roles.

```
with t as
(
select a.job_title, avg_2023, avg_2024 from
(select job_title, avg(salary_in_usd) as avg_2023  from salaries
where work_year = 2023
group by job_title) a
inner join
(select job_title, avg(salary_in_usd) as avg_2024  from salaries
where work_year = 2024
group by job_title) b)

select *, round((((avg_2024-avg_2023)/avg_2023)*100),2) as growth_rate_percentage from t;
```

OUTPUT:

job_title	avg_2023	avg_2024	growth_rate_percentage
Azure Data Engineer	100000.0000	164314.7857	64.31
Computer Vision Software Engineer	53984.0000	164314.7857	204.38
Big Data Architect	153799.0000	164314.7857	6.84
Data Analytics Specialist	95000.0000	164314.7857	72.96
Deep Learning Researcher	124163.0000	164314.7857	32.34
Cloud Database Engineer	141666.6667	164314.7857	15.99
BI Data Engineer	60000.0000	164314.7857	173.86
Lead Data Analyst	43120.5000	164314.7857	281.06

16. You've been hired by a global HR Consultancy to identify Countries experiencing significant salary growth for entry-level roles. Your task is to list the top three Countries with the highest salary growth rate FROM 2020 to 2023, helping multinational Corporations identify Emerging talent markets.

```
with t as
(
    select company_location,work_year, avg(salary_in_usd) as average
    from salaries
    where experience_level = 'EN' and work_year in (2021,2023)
    group by company_location,work_year
)
select *, round(((avg_2023-avg_2021)/avg_2021)*100,2) as growth_perc
from
    (
    select company_location,
    max(case when work_year =2021 then average end) as avg_2021,
    max(case when work_year =2023 then average end) as avg_2023
    from t
    group by company_location
    ) a
where round(((avg_2023-avg_2021)/avg_2021)*100,2) is not null
order by round(((avg_2023-avg_2021)/avg_2021)*100,2) desc limit 3 ;
```

OUTPUT:

company_location	avg_2021	avg_2023	growth_perc
AU	42028.0000	53089.3333	26.32
US	88617.6471	101592.8575	14.64
IN	24407.1667	27344.1667	12.03

17.Picture yourself as a data architect responsible for database management. Companies in US and AU(Australia) decided to create a hybrid model for employees they decided that employees earning salaries exceeding $90000 USD, will be given work from home. You now need to update the remote work ratio for eligible employees, ensuring efficient remote work management while implementing appropriate error handling mechanisms for invalid input parameters.

```
-- creating temporary table so that changes are not made in actual table as actual table is being used in other cases also.
create table temp_table as select * from salaries;

 -- by default mysql runs on safe update mode , this mode  is a safeguard against updating
 -- or deleting large portion of  a table.
 -- We will turn off safe update mode using set_sql_safe_updates

SET SQL_SAFE_UPDATES = 0;

update temp_table
set remote_ratio = 100
where (experience_level = 'US' or experience_level = 'AU') and salary_in_usd > 90000;

select * from temp_table where (company_location = 'AU' OR company_location ='US')AND salary_in_usd > 90000;
```

OUTPUT:

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

work_year	experience_level	employment_type	job_title	salary	salary_currency	salary_in_usd	employee_residence	remote_ratio	company_location	company_size
2024	SE	FT	Machine Learning Engineer	180500	USD	180500	US	0	US	M
2024	SE	FT	Machine Learning Engineer	96200	USD	96200	US	0	US	M
2024	SE	FT	Machine Learning Engineer	235000	USD	235000	AU	0	AU	M
2024	SE	FT	Machine Learning Engineer	175000	USD	175000	AU	0	AU	M
2024	MI	FT	Business Intelligence Developer	95413	USD	95413	US	100	US	M
2024	SE	FT	Data Scientist	222300	USD	222300	US	0	US	M

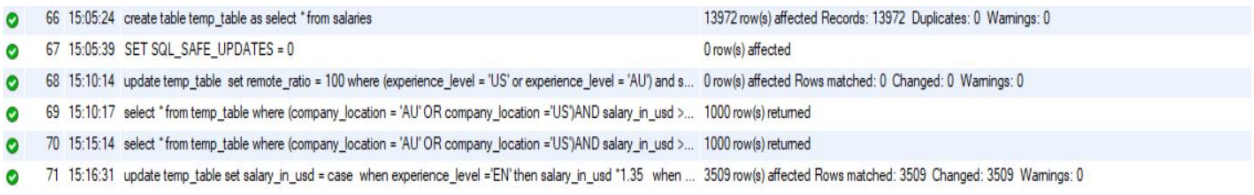
18. In year 2024, due to increase demand in data industry , there was increase in salaries of data field employees.

Entry Level-35% of the salary.
Mid junior – 30% of the salary.
Immediate senior level- 22% of the salary.
Expert level- 20% of the salary.
Director – 15% of the salary.

you have to update the salaries accordingly and update it back in the original database.

```
update temp_table
set salary_in_usd =
case
when experience_level ='EN' then salary_in_usd *1.35
when experience_level ='MI' then salary_in_usd *1.30
when experience_level ='SE' then salary_in_usd *1.22
when experience_level ='EX' then salary_in_usd *1.20
when experience_level ='DX' then salary_in_usd *1.15
else salary_in_usd
end
where work_year = 2024;
```

OUTPUT:



✓	66	15:05:24	create table temp_table as select * from salaries	13972 row(s) affected Records: 13972 Duplicates: 0 Warnings: 0
✓	67	15:05:39	SET SQL_SAFE_UPDATES = 0	0 row(s) affected
✓	68	15:10:14	update temp_table set remote_ratio = 100 where (experience_level = 'US' or experience_level = 'AU') and s...	0 row(s) affected Rows matched: 0 Changed: 0 Warnings: 0
✓	69	15:10:17	select * from temp_table where (company_location = 'AU' OR company_location ='US')AND salary_in_usd >...	1000 row(s) returned
✓	70	15:15:14	select * from temp_table where (company_location = 'AU' OR company_location ='US')AND salary_in_usd >...	1000 row(s) returned
✓	71	15:16:31	update temp_table set salary_in_usd = case when experience_level ='EN' then salary_in_usd *1.35 when ...	3509 row(s) affected Rows matched: 3509 Changed: 3509 Warnings: 0

19. You are a researcher and you have been assigned the task to Find the year with the highest average salary for each job title.

```
with avg_salary_per_year as
(
select job_title, work_year, avg(salary_in_usd) as avg_salary from salaries
group by work_year, job_title
)

select job_title, work_year, avg_salary, rank_by_salary from
(
select job_title, work_year, avg_salary,
rank() over(partition by job_title order by avg_salary desc) as rank_by_salary
from avg_salary_per_year
) as ranked_salary
where rank_by_salary =1;
```

OUTPUT:

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

job_title	work_year	avg_salary	rank_by_salary
Admin & Data Analyst	2022	60000.0000	1
AI Architect	2024	256637.5000	1
AI Developer	2022	275000.0000	1
AI Engineer	2024	164314.7857	1
AI Product Manager	2024	152650.0000	1

20. You have been hired by a market research agency where you been assigned the task to show the percentage of different employment type (full time, part time) in Different job roles, in the format where each row will be job title, each column will be type of employment type and cell value for that row and column will show
the % value.

```
select
    job_title,
    round((SUM(CASE WHEN employment_type = 'PT' THEN 1 ELSE 0 END) / COUNT(*)) * 100, 2) AS PT_percentage, -- Calculate percentage of part-time employment
    round((SUM(CASE WHEN employment_type = 'FT' THEN 1 ELSE 0 END) / COUNT(*)) * 100, 2) AS FT_percentage, -- Calculate percentage of full-time employment
    round((SUM(CASE WHEN employment_type = 'CT' THEN 1 ELSE 0 END) / COUNT(*)) * 100, 2) AS CT_percentage, -- Calculate percentage of contract employment
    round((SUM(CASE WHEN employment_type = 'FL' THEN 1 ELSE 0 END) / COUNT(*)) * 100, 2) AS FL_percentage -- Calculate percentage of freelance employment
from
    salaries
group by
    job_title; -- Group the result by job title
```

OUTPUT:

job_title	PT_percentage	FT_percentage	CT_percentage	FL_percentage
AI Engineer	0.00	98.88	1.12	0.00
Machine Learning Engineer	0.00	99.80	0.14	0.07
Business Intelligence Developer	0.00	100.00	0.00	0.00
Data Engineer	0.13	99.83	0.00	0.03
Data Scientist	0.24	99.65	0.07	0.03
Cloud Database Engineer	0.00	100.00	0.00	0.00
Research Engineer	0.00	100.00	0.00	0.00
Data Analyst	0.29	99.66	0.05	0.00
Machine Learning Scientist	0.00	100.00	0.00	0.00
Applied Scientist	0.00	100.00	0.00	0.00

THANK- YOU