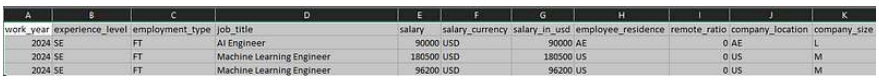
SQL-Case Study

data-science-jobs-sql-case-study

Dataset: https://github.com/dk1coding1zone/data-science-jobs-sql-case-study/blob/main/salaries.csv

overview:



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

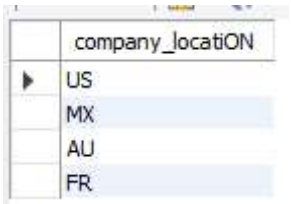
demo-data-screenshot

remote_ratio 0 means offline and 100 means online

Task-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 distINct(company_locatiON) FROM salaries
```



company_locatiON
US
MX
AU
FR

OUTPUT:

Task-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(company_size) AS
FROM (
    SELECT * FROM salaries WHERE experience_leve
) AS t
GROUP BY company_locatiOn
ORDER BY cnt DESC
LIMIT 5;
```

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

OUTPUT:

Task-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 @COUNT= (SELECT COUNT(*) FROM salaries  WHER
set @total = (SELECT COUNT(*) FROM salaries wher
set @percentage= round((((SELECT @COUNT)/(SELECT
SELECT @percentage AS '%  of people workINg remo
```

	% of people workINg remotly and havINg salary >100,000 USD
▶	32.24

OUTPUT:

Task-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 countries.

```
ge_per_country, average FROM

ry_IN_usd) AS average_per_country FROM  salaries WH



erage FROM  salaries  WHERE experience_level = 'EN'

_per_country> averages
```

	company_locatiON	job_title	average_per_country	average
▶	US	Data Analyst	89800.3519	84808.6361
	US	Analytics Engineer	110831.2500	96722.3000
	US	Data Engineer	106791.2584	92713.4701
	US	Research Analyst	110459.5385	107294.2143
	US	Machine Learning Engineer	126188.8529	110718.3778
	AU	Business Intelligence Analyst	91000.0000	76688.6333
	US	Business Intelligence Analyst	79243.9500	76688.6333
	US	Research Engineer	162668.1818	135467.7241
	CA	Machine Learning Research Engineer	80769.0000	55653.6667
	US	Data Scientist	104573.4333	87028.3731
	MX	Data Analyst	429950.0000	84808.6361
	US	Data Science	100450.0000	85511.7857
	IT	AI Engineer	44444.0000	33679.0000
	US	Research Scientist	168206.9231	158548.5517
	CA	Machine Learning Engineer	131600.0000	110718.3778

OUTPUT:

Task-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 company_locatiON , job_title , average FR
(
SELECT *, dense_rank() over (partitiON by job_ti
(
SELECT company_locatiON , job_title , AVG(salary
)k
)t  WHERE num=1
```

company_locatiON	job_title	average
IN	Admin & Data Analyst	60000.0000
CA	AI Architect	800000.0000
CA	AI Developer	275000.0000
QA	AI Engineer	300000.0000
US	AI Product Manager	152650.0000
FR	AI Programmer	120000.0000
US	AI Research Engineer	175000.0000
DE	AI Research Scientist	88888.0000
IL	AI Scientist	417937.0000
EG	AI Software Engineer	174100.0000
FR	Analytics Engineer	188000.0000
GB	Analytics Engineerin...	399880.0000
US	Applied Data Scientist	238000.0000
US	Applied Machine Lear...	177500.0000
US	Applied Machine Lear...	141550.0000

OUTPUT:

Task-6

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, total_remote ,total
  (
     SELECT experience_level, COUNT(experience_l
  )a
  INNER JOIN
  (
    SELECT  experience_level, COUNT(experience_l
  )b ON a.experience_level= b.experience_level
  ),
  t2 AS
     (
  SELECT a.experience_level, total_remote ,total
  (
  SELECT experience_level, COUNT(experience_leve
  )a
  INNER JOIN
  (
  SELECT  experience_level, COUNT(experience_lev
  )b ON a.experience_level= b.experience_level
  )

 SELECT * FROM t1 INNER JOIN t2 ON t1.experience
```

experience_level	total_remote	total_2021	2021 remote %	experience_level	total_remote	total_2024	2024 remote %
SE	44	75	58.67	SE	483	1920	25.16
MI	45	87	51.72	MI	227	1102	20.60
EN	22	46	47.83	EN	87	381	22.83
EX	5	10	50.00	EX	35	106	33.02

OUTPUT:

Task-7

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.

```
WITH t AS
(
SELECT experience_level, job_title ,work_year, r
)  -- step 1



SELECT *,round((((AVG_salary_2024-AVG_salary_202
FROM
(
 SELECT
  experience_level, job_title,
  MAX(CASE WHEN work_year = 2023 THEN average EN
  MAX(CASE WHEN work_year = 2024 THEN average EN
 FROM  t GROUP BY experience_level , job_title -
)a WHERE (((AVG_salary_2024-AVG_salary_2023)/AVG
```

experience_level	job_title	AVG_salary_2023	AVG_salary_2024	changes
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
MI	Data Scientist	124881.49	135296.68	8.34
MI	Applied Scientist	173830.71	181031.25	4.14
SE	Data Analyst	120853.50	127195.41	5.25

OUTPUT: