MYSKILL.ID

SALARY TRENDS ANALYSIS USING SQL

Suci Wulan Dari

MYSKILL.ID

TABLE OF CONTENTS

PROJECT OVERVIEW

Salary analysis helps companies determine competitive salary structures and helps job seekers understand the salary market. With the right data, companies can attract the best talent and retain existing employees.

Project Objectives:

- Analyze salary trends by position, experience, location, and job type.
- Provides insights for better salary decision making in the company.

MYSKILL.ID

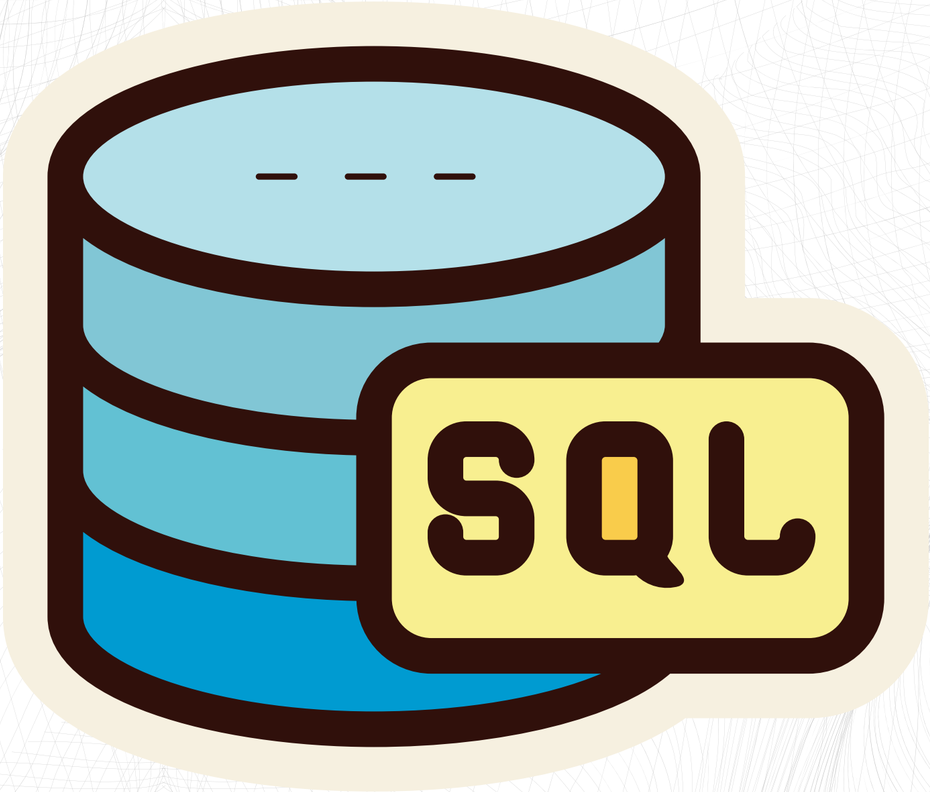
DATASET INFORMATION

- work_year: Work year (2020, 2021, 2022)
- experience_level: Experience level (Entry, Mid, Senior, Lead)
- employment_type: Employment type (Full-time, Contract, Part-Time, Freelance)
- job_title: Job position
- salary: Salary in local currency
- salary_currency: Salary currency
- salary_in_usd: Salary in USD (for cross-country comparison)
- employee_residence: Employee residence location
- remote_ratio: Percentage of remote work
- company_location: Company location
- company_size: Company size (e.g., Small, Medium, Large)

Column Name	Data Type
id	int
work_year	int
experience_level	nvarchar(50)
employment_type	nvarchar(50)
job_title	nvarchar(50)
salary	int
salary_currency	nvarchar(50)
salary_in_usd	int
employee_residence	nvarchar(50)
remote_ratio	tinyint
company_location	nvarchar(50)
company_size	nvarchar(50)

TOOLS

SQL SERVER



DATA PREPROCESSING

1. MISSING VALUE CHECKING

```
SELECT * FROM ds_salaries
    WHERE work_year is NULL
        OR experience_level is NULL
        OR employment_type is NULL
        OR job_title is NULL
        OR salary is NULL
        OR salary_currency is NULL
        OR salary_in_usd is NULL
        OR employee_residence is NULL
        OR remote_ratio is NULL
        OR company_location is NULL
        OR company_size is NULL;
```

Results | Messages

id	work_year	experience_level	employment_type	job_title	salary	salary_currency

2. DELETE DUPLICATE DATA

```
WITH CTE AS (
    SELECT *, ROW_NUMBER () OVER
    (PARTITION BY work_year,
                  experience_level,
                  employment_type,
                  job_title,
                  salary,
                  salary_currency,
                  salary_in_usd,
                  employee_residence,
                  remote_ratio,
                  company_location,
                  company_size ORDER BY id) AS RowNum
    FROM ds_salaries)

DELETE FROM CTE WHERE RowNum > 1;
```

%

Messages

```
(42 rows affected)

Completion time: 2025-02-02T02:40:27.8087459+07:00
```

MYSKILL.ID

BUSINESS QUESTIONS

01 How have average salaries changed over the years?

02 Which job positions have the highest average salaries?

03 How does salary vary by experience level?

04 How do salaries differ by company location?

05 How do remote jobs compare to on-site jobs in terms of salary?

HOW HAVE AVERAGE SALARIES CHANGED OVER THE YEARS?

```sql
SELECT work_year,
       ROUND(
            AVG(salary_in_usd),
                2)
            AS avg_salary
FROM ds_salaries
    GROUP BY work_year
    ORDER BY work_year;
```

Results | Messages

	work_year	avg_salary
1	2020	95813
2	2021	99430
3	2022	123089

- Salary Growth Trend:

From 2020 to 2021, the average salary increased by 3,617 (approximately 3.8%).
From 2021 to 2022, the average salary rose by 23,659 (approximately 23.8%).

- No Salary Decline:

Based on the available data, there was no decline in salary over the years.

WHICH JOB POSITIONS HAVE THE HIGHEST AVERAGE SALARIES?

```sql
SELECT TOP 10 job_title,
        ROUND(
          AVG(salary_in_usd),
              2)
          AS avg_salary,
          COUNT(*) AS num_employees
FROM ds_salaries
    GROUP BY job_title
    ORDER BY avg_salary DESC;
```

	job_title	avg_salary	num_employees
1	Data Analytics Lead	405000	1
2	Principal Data Engineer	328333	3
3	Financial Data Analyst	275000	2
4	Principal Data Scientist	215242	7
5	Director of Data Science	195074	7
6	Data Architect	177873	11
7	Applied Data Scientist	175655	5
8	Analytics Engineer	175000	4
9	Data Specialist	165000	1
10	Head of Data	160162	5

The highest-paying jobs based on the average salary include **Data Analytics Lead** (405,000), **Principal Data Engineer** (328,333), a**nd Financial Data Analyst** (275,000).

High-paying jobs not only have competitive salaries but also have smaller workforces, perhaps due to the higher skills and specialization required for these positions.

HOW DOES SALARY VARY BY EXPERIENCE LEVEL?

```sql
SELECT experience_level,
       ROUND(
         AVG(salary_in_usd),
             2)
         AS avg_salary
FROM ds_salaries
  GROUP BY experience_level
  ORDER BY avg_salary DESC;
```

	experience_level	avg_salary
1	EX	199392
2	SE	138374
3	MI	87792
4	EN	61643

- From Entry (EN) to Mid (MI): Increase of $26,149 (+42.4%)
- From Mid (MI) to Senior (SE): Increase of $50,582 (+57.6%)
- From Senior (SE) to Executive (EX): Increase of $61,018 (+44.1%)

There is a significant gap between each experience level, especially from Mid to Senior (+$50,582 or 57.6%).

HOW DO SALARIES DIFFER BY COMPANY LOCATION?

```sql
SELECT company_location,
       ROUND(
         AVG(salary_in_usd),
             2)
         AS avg_salary
FROM ds_salaries
  GROUP BY company_location
  ORDER BY avg_salary DESC;
```

	company_location	avg_salary
1	RU	157500
2	US	144292
3	NZ	125000
4	IL	119059
5	JP	114127
46	UA	13400
47	PK	13333
48	KE	9272
49	IR	4000
50	VN	4000

- **Russia** offers the highest average salary ($157,500), followed by the **US** ($144,292) and **New Zealand** ($125,000).
- There is a massive salary gap between top-paying and lower-paying countries, with some earning 40 times more than others.
- Economic factors, cost of living, and demand for skilled professionals contribute to these differences.

MYSKILL.ID

HOW DO REMOTE JOBS COMPARE TO ON-SITE JOBS IN TERMS OF SALARY?

- **Remote workers** earn an average salary of $120,763, which is $14,978 higher than **on-site workers (**$105,785) and **hybrid workers**($80,721).
- This trend is likely to continue as **remote work** becomes **more attractive** and **widely adopted**.

```sql
SELECT
    CASE
        WHEN remote_ratio = 0 THEN 'On-Site'
        WHEN remote_ratio = 50 THEN 'Hybrid'
        WHEN remote_ratio = 100 THEN 'Remote'
    END AS work_type,
    ROUND(AVG(salary_in_usd), 2) AS avg_salary
FROM ds_salaries
GROUP BY
    CASE
        WHEN remote_ratio = 0 THEN 'On-Site'
        WHEN remote_ratio = 50 THEN 'Hybrid'
        WHEN remote_ratio = 100 THEN 'Remote'
    END
ORDER BY avg_salary DESC;
```

	work_type	avg_salary
1	Remote	120763
2	On-Site	105785
3	Hybrid	80721

THANK YOU