



SQL PROJECT
MYSKILL.ID

Analysis of Salaries for Data Analysts

ZAHRA ULAYA SIFA



TOOLS



PostgreSQL

A screenshot of a PostgreSQL database browser interface. On the left, there's a tree view of database objects. At the top level, there's a folder icon labeled 'demo_myskill'. Below it are several icons: Casts, Catalogs, Event Triggers, Extensions, Foreign Data Wrappers, Languages, Publications, and Schemas (1). The 'Schemas (1)' node is expanded, showing a red diamond icon followed by 'public'. This 'public' node is also expanded, showing a list of objects: Aggregates, Collations, Domains, FTS Configurations, FTS Dictionaries, FTS Parsers, FTS Templates, Foreign Tables, Functions, Materialized Views, Operators, Procedures, Sequences, and Tables (1). The 'Tables (1)' node is expanded, showing a table icon followed by 'ds_salaries'. The entire interface has a light gray background with white text and icons.



PostgreSQL

Details:

Schema Name: public

Table Name: ds_salaries



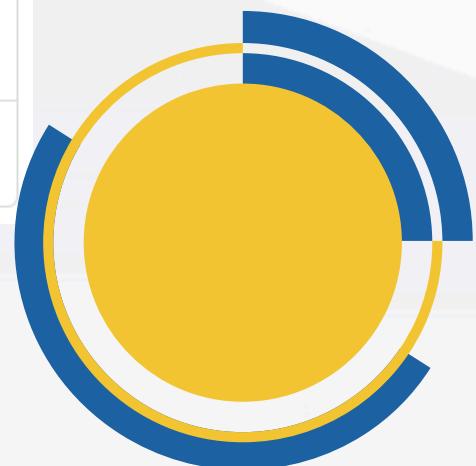
Case Study

Understanding the dynamics of salary structures and employment arrangements is essential for aspiring data analysts seeking to navigate the job market effectively. This case study intends to analyze average compensation patterns based on experience levels, job types, and other relevant factors. By leveraging historical salary data, it aims to provide meaningful insights into the financial viability of a career in data analytics.

About Dataset

Column	Description
work_year	The year the salary was paid.
experience_level	The experience level in the job during the year with the following possible values: EN Entry-level / Junior MI Mid-level / Intermediate SE Senior-level / Expert EX Executive-level / Director
employment_type	The type of employment for the role: PT Part-time FT Full-time CT Contract FL Freelance
job_title	The role worked in during the year.
salary	The total gross salary amount paid.
salary_currency	The currency of the salary paid as an ISO 4217 currency code.
salary_in_usd	The salary in USD (FX rate divided by avg. USD rate for the respective year via fxdata.foorilla.com).
employee_residence	Employee's primary country of residence in during the work year as an ISO 3166 country code.
remote_ratio	The overall amount of work done remotely, possible values are as follows: 0 No remote work (less than 20%) 50 Partially remote 100 Fully remote (more than 80%)
company_location	The country of the employer's main office or contracting branch as an ISO 3166 country code.
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)

www.kaggle.com/datasets/ruchi798/data-science-job-salaries



Check type of dataset

```
-- Viewing data type information in the ds_salaries table
SELECT column_name, data_type
FROM information_schema.columns
WHERE table_schema = 'public'
AND table_name = 'ds_salaries';
```

	column_name	data_type
1	name	character varying
2	no	integer
3	work_year	integer
4	experience_level	text
5	employment_type	text
6	job_title	text
7	salary	integer
8	salary_currency	text
9	salary_in_usd	double precision
10	employee_residence	text
11	remote_ratio	integer
12	company_location	text
13	company_size	text

Check Null values

```
-- Check Null Values
SELECT
    COUNT(*) AS total_rows,
    COUNT(*) FILTER (WHERE no IS NULL) AS null_no,
    COUNT(*) FILTER (WHERE work_year IS NULL) AS null_work_year,
    COUNT(*) FILTER (WHERE experience_level IS NULL) AS null_experience_level,
    COUNT(*) FILTER (WHERE employment_type IS NULL) AS null_employment_type,
    COUNT(*) FILTER (WHERE job_title IS NULL) AS null_job_title,
    COUNT(*) FILTER (WHERE salary IS NULL) AS null_salary,
    COUNT(*) FILTER (WHERE salary_currency IS NULL) AS null_salary_currency,
    COUNT(*) FILTER (WHERE salary_in_usd IS NULL) AS null_salary_in_usd,
    COUNT(*) FILTER (WHERE employee_residence IS NULL) AS null_employee_residence,
    COUNT(*) FILTER (WHERE remote_ratio IS NULL) AS null_remote_ratio,
    COUNT(*) FILTER (WHERE company_location IS NULL) AS null_company_location,
    COUNT(*) FILTER (WHERE company_size IS NULL) AS null_company_size
FROM public.ds_salaries;
```

	total_rows	null_no	null_work_year	null_experience_level	null_employment_type	null_job_title	null_salary	null_salary_currency
1	607	0	0	0	0	0	0	0

null_salary_in_usd	null_employee_residence	null_remote_ratio	null_company_location	null_company_size
0	0	0	0	0

The table contains 607 rows, with **no NULL values** present in any column.

Job titles related to Data Analyst

```
SELECT DISTINCT job_title  
FROM public.ds_salaries  
WHERE job_title  
LIKE '%Data Analyst%'  
ORDER BY job_title;
```

	job_title	text	lock
1	BI Data Analyst		
2	Business Data Analyst		
3	Data Analyst		
4	Finance Data Analyst		
5	Financial Data Analyst		
6	Lead Data Analyst		
7	Marketing Data Analyst		
8	Principal Data Analyst		
9	Product Data Analyst		

Average salary of job title related to Data Analyst

```
SELECT AVG(salary_in_usd) AS avg_salary_annually  
FROM public.ds_salaries  
WHERE job_title LIKE '%Data Analyst%';
```

	avg_salary_annually	double precision	lock
1	93200.3781512605		

Among the 9 job titles related to Data Analyst, **the average annual** salary is approximately **\$93,200.3781512605**

```
SELECT AVG((salary_in_usd/12)) AS avg_salary_monthly  
FROM public.ds_salaries  
WHERE job_title LIKE '%Data Analyst%';
```

	avg_salary_monthly	double precision	lock
1	7766.698179271711		

Among the 9 job titles related to Data Analyst, **the average** salary is approximately **\$7,766.698179271711 per month**

Average salary of job title related to Data Analyst based on experience level

```
SELECT experience_level, AVG((salary_in_usd/12)) avg_salary_monthly  
FROM public.ds_salaries  
WHERE job_title LIKE '%Data Analyst%'  
GROUP BY experience_level  
ORDER BY avg_salary_monthly DESC;
```

	experience_level	avg_salary_monthly
1	EX	10833.33333333334
2	SE	9388.46551724138
3	MI	6495.063008130082
4	EN	4759.323529411764

Among the 9 job titles related to Data Analyst, based on experience level, **the highest average monthly salary** (in USD) is found at the **Executive (EX)** level, followed by Senior (SE), Middle (MI), and Entry (EN)

The company location offering the lowest to highest average salary in Rupiah for positions similar to Data Analyst with full-time employment and experience levels of Medium or Entry has an average salary exceeding Rp1,000,000

```
SELECT company_location, job_title, experience_level,  
AVG(((salary_in_usd*16294)/12)) avg_salary_monthly_Rp  
FROM public.ds_salaries  
WHERE job_title LIKE '%Data Analyst%'  
AND employment_type = 'FT'  
AND experience_level IN ('EN', 'MI')  
GROUP BY 1,2,3  
HAVING AVG(((salary_in_usd*16294)/12)) >= 1000000  
ORDER BY 4;
```

	company_location	job_title	experience_level	avg_salary_monthly_rp
1	IN	Product Data Analyst	MI	8244764
2	IN	Data Analyst	EN	8244764
3	PK	Data Analyst	MI	10862666.666666666
4	KE	BI Data Analyst	EN	12589830.666666666
5	NG	Data Analyst	EN	13578333.33333334
6	IN	Business Data Analy...	MI	25041162.33333332
7	IN	Lead Data Analyst	MI	26625753.83333332
8	HN	Product Data Analyst	MI	27156666.666666668
9	GR	Data Analyst	MI	43876121.11111111
10	ES	Data Analyst	MI	52235848.33333336
11	FR	Data Analyst	MI	63490928.83333336
12	GB	Data Analyst	MI	69080585.53333333
13	US	BI Data Analyst	EN	74680833.33333333



The highest salary increase from MI to EX level for Full-Time positions similar to Data Analyst



	company_location text	job_title text	experience_level text	avg_salary_monthly_rp double precision
14	LU	Business Data Analy...	EN	80250665.66666667
15	FR	Data Analyst	EN	80250665.66666667
16	CA	Data Analyst	EN	80791083.33333333
17	CA	Business Data Analy...	MI	96286677.33333333
18	US	Data Analyst	EN	100253361.11111112
19	CA	Principal Data Analyst	MI	101837500
20	US	BI Data Analyst	MI	106028226.27777778
21	CA	Data Analyst	MI	108626666.66666666
22	US	Lead Data Analyst	MI	118131500
23	US	Data Analyst	MI	130013446.88888888
24	US	Financial Data Analyst	EN	135783333.33333334
25	US	Business Data Analy...	MI	183307500
26	US	Financial Data Analyst	MI	611025000

```
WITH ds_1 AS (
  SELECT work_year, AVG(salary_in_usd) avg_mi
  FROM public.ds_salaries
  WHERE job_title LIKE '%Data Analyst%'
    AND experience_level = 'MI' AND employment_type = 'FT'
  GROUP BY work_year
),
ds_2 AS (
  SELECT work_year, AVG(salary_in_usd) avg_ex
  FROM public.ds_salaries
  WHERE job_title LIKE '%Data Analyst%'
    AND experience_level = 'EX' AND employment_type = 'FT'
  GROUP BY work_year
)
SELECT ds1.work_year, ds1.avg_mi, ds2.avg_ex,
ds2.avg_ex-ds1.avg_mi diff
FROM ds_1 ds1
LEFT JOIN ds_2 ds2
ON ds1.work_year=ds2.work_year ORDER BY diff DESC;
```

The highest salary increase from MI to EX level for Full-Time positions similar to Data Analyst

	work_year integer	avg_mi double precision	avg_ex double precision	diff double precision
1	2020	60728.875	[null]	[null]
2	2022	68970.59090909091	120000	51029.40909090909
3	2021	108398.81818181818	150000	41601.18181818182

The highest salary increase from MI (Mid-level / Intermediate) to EX (Senior-level / Expert) for full-time positions similar to Data Analyst occurred in **2022**, with a difference **\$51,029.40909090909**

The distribution of employment types (full remote, partial remote, or no remote) for employees residing in the same country as their workplace

```
WITH
no_remote AS (
SELECT COUNT(
CASE WHEN job_title LIKE '%Data Analyst%'
AND remote_ratio = 0
AND employee_residence = company_location
THEN 1 END) AS no_remote_count
FROM public.ds_salaries),
partially AS (
SELECT COUNT(
CASE WHEN job_title LIKE '%Data Analyst%'
AND remote_ratio = 50
AND employee_residence = company_location
THEN 1 END) AS partially_count
FROM public.ds_salaries),
full_remote AS (
SELECT COUNT(
CASE WHEN job_title LIKE '%Data Analyst%'
AND remote_ratio = 100
AND employee_residence = company_location
THEN 1 END) AS full_remote_count
FROM public.ds_salaries)
SELECT * FROM no_remote CROSS JOIN partially CROSS JOIN full_remote;
```

	no_remote_count bigint	partially_count bigint	full_remote_count bigint
1	23	10	82

The majority of employees living in the same country as their workplace have **full remote jobs**, with a total of **82 individuals**

Will the average salary for Data Analyst related jobs tend to be higher for employees whose residence is different from the company location?

```
WITH
ga_sama AS (
    SELECT job_title, AVG(salary_in_usd) avg_ga_sama
    FROM public.ds_salaries
    WHERE job_title LIKE '%Data Analyst%'
    AND employee_residence = company_location
    GROUP BY job_title
),
sama AS (
    SELECT job_title, AVG(salary_in_usd) avg_sama
    FROM public.ds_salaries
    WHERE job_title LIKE '%Data Analyst%'
    AND employee_residence != company_location
    GROUP BY job_title
)
SELECT COALESCE(ga_sama.job_title,sama.job_title), avg_ga_sama, avg_sama,
(CASE WHEN avg_ga_sama > avg_sama
    THEN 'ya' WHEN avg_ga_sama <= avg_sama
    THEN 'tidak' END) apakah_avg_ga_sama_lebih_besar
FROM ga_sama
FULL OUTER JOIN sama
ON ga_sama.job_title=sama.job_title;
```

	coalesce text	avg_ga_sama double precision	avg_sama double precision	apakah_avg_ga_sama_lebih_besar text
1	BI Data Analyst	65568	93129.5	tidak
2	Data Analyst	93027.36458333333	80000	ya
3	Marketing Data Anal...	[null]	88654	[null]
4	Product Data Analyst	13036	[null]	[null]
5	Financial Data Analy...	275000	[null]	[null]
6	Principal Data Analyst	122500	[null]	[null]
7	Business Data Analy...	76691.2	[null]	[null]
8	Lead Data Analyst	92203	[null]	[null]
9	Finance Data Analyst	61896	[null]	[null]

There is no consistent pattern indicating that average salaries for Data Analyst-related roles are higher when employees reside in a different country than the company location. In the available data, **only the 'Data Analyst'** role shows a higher average salary when the employee's residence differs from the company's location.

Thank You



Email:
zahraulayyasifa.work@gmail.com



Linkedin
linkedin.com/in/zahraulayyasifa/



Instagram
[@zhrauysf](https://www.instagram.com/@zhrauysf)



Tegal | Available for Remote or On-Site