

GLook

General Data Insights

Univariate Analysis

Bivariate Analysis

Trivariate Analysis

Session State:-> True

Dataframe Overview

	job_title	experience_level	employment_type	company_size	company_location	remote_ratio	salary
0	Data Analyst	Mid	Contract	Medium	Germany	0	INR
1	DevOps Engineer	Mid	Contract	Small	India	100	GBP
2	Research Scientist	Lead	None	Medium	Germany	0	EUR
3	Software Engr	Lead	Full-time	Large	India	50	INR
4	Software Engr	Lead	Intern	Large	Germany	100	INR
5	Data Analyst	Mid	Contract	Large	UK	0	EUR
6	Dt Scientist	Lead	Intern	Small	Canada	0	EUR
7	DevOps Engineer	Lead	Intern	Medium	India	50	INR
8	ML Engineer	None	None	Large	Canada	50	GBP
9	Data Analyst	Senior	Full-time	Large	India	0	CAD

Data dimensions: 100000 rows, 18 columns

Number of rows: 100000

Number of duplicates: 10100

Deep Memory usage: 49.223 MB

Number of features: 18

Number of categorical features: 7

Number of numerical features: 11

Numerical Data Overview

	count	mean	std	min	25%	50%	75%	max
remote_ratio	100000.00	49.90	40.73	0.00	0.00	50.00	100.00	100.00
years_experience	100000.00	10.01	6.06	0.00	5.00	10.00	15.00	20.00
base_salary	100000.00	273915.48	609824.47	-344.34	59139.57	92922.92	126053.91	3121412.46
bonus	100000.00	5000.53	2891.50	0.00	2508.00	5004.00	7504.25	9999.00
stock_options	100000.00	15014.53	8664.14	0.00	7463.75	14995.00	22530.00	29998.00
total_salary	100000.00	105189.40	36335.19	13732.47	74890.96	105372.47	135233.16	196335.84
salary_in_usd	100000.00	103339.98	146128.71	221.01	48338.07	91291.88	133356.75	2354698.24
education	0.00	None	None	None	None	None	None	None
skills	0.00	None	None	None	None	None	None	None
conversion_rate	100000.00	0.83	0.45	0.01	0.75	1.00	1.10	1.30

Categorical Data Overview

	count	unique	top	freq
job_title	100000	12	Data Analyst	16857
experience_level	80000	4	Mid	20079
employment_type	76016	4	Part-time	19146
company_size	100000	3	Small	33538
company_location	100000	6	UK	17087
salary_currency	100000	5	GBP	20197
currency	100000	5	GBP	20202

DataFrame Info:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100000 entries, 0 to 99999
Data columns (total 18 columns):
#   Column                Non-Null Count  Dtype
---  -
0   job_title              100000 non-null object
1   experience_level        80000 non-null object
2   employment_type        76016 non-null object
3   company_size            100000 non-null object
4   company_location        100000 non-null object
5   remote_ratio            100000 non-null int64
6   salary_currency         100000 non-null object
7   years_experience        100000 non-null int64
8   base_salary             100000 non-null float64
9   bonus                  100000 non-null int64
10  stock_options           100000 non-null int64
```

```
11 total_salary      100000 non-null float64
12 salary_in_usd     100000 non-null float64
13 currency          100000 non-null object
14 education         0 non-null float64
15 skills            0 non-null float64
16 conversion_rate    100000 non-null float64
17 adjusted_total_usd 100000 non-null float64
dtypes: float64(7), int64(4), object(7)
memory usage: 13.7+ MB
```

Univariate_Analysis