DATA SCIENCE EMPLOYEES' ANALYSIS

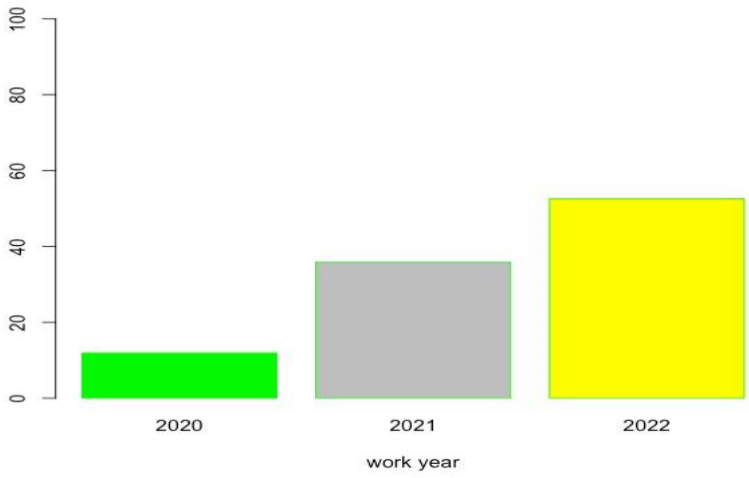
Executive Summary Report

Shivani Vellanki
10/02/2022

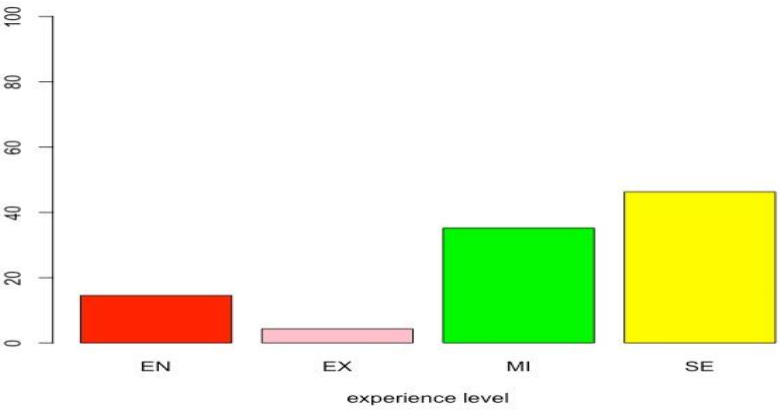
Introduction:

The given dataset represents the data scientist's salary in the years 2020, 2021, and 2022. The data shows various attributes like experience level of the employee i.e., is the employee on an entry level, mid-level, senior or executive level, it also describes the employment type i.e., is it full time or part time? The data consists of 50 unique job titles and their salaries in their currency values and in USD. The data also describes about the employee's residence and the company's location with its size.

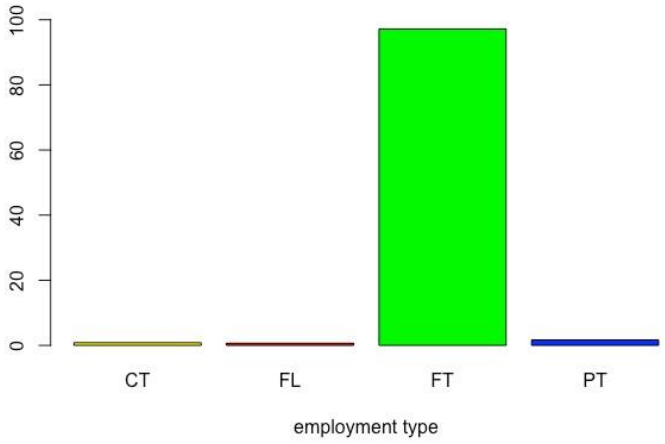
The dataset shows the employee records from 2020, 2021 and 2022. There are a total number of 607 data scientist's employee records, among them, 72 being the least out of 607 belong to 2020, following by 217 employees out of 607 records belong to 2021, and the highest number of employees i.e., 318 out of 607, belong to 2022. This can be graphically represented as follows,



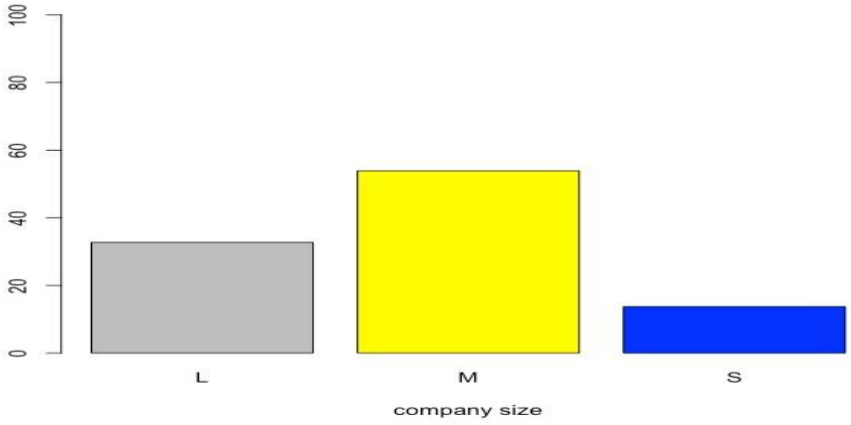
According to the experience level attribute, the employees whose experience levels are EN, EX, MI and SE hold the records of values, 89,27,214 and 281 respectively. It is observed that employees who are of SE level are the highest. This can be graphically represented as,



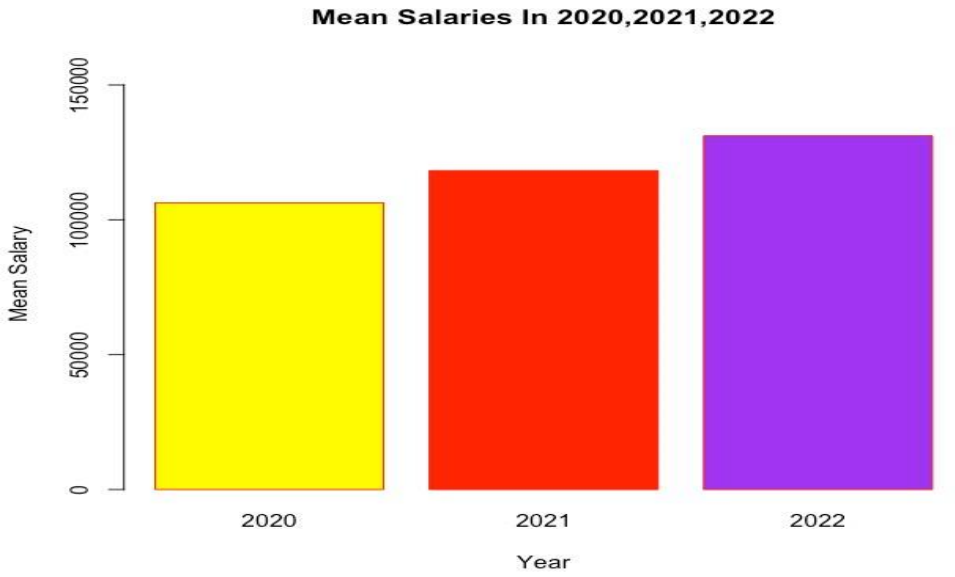
In accordance with the employment type, most number of employees work full time i.e., 588 of the total records in the dataset. The graphs represents, as follows



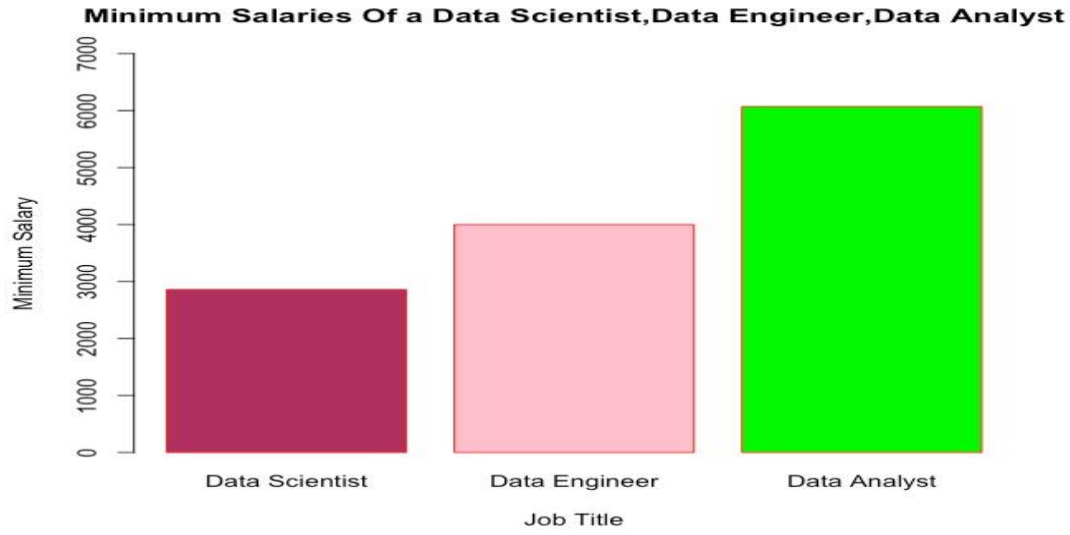
According to the company's size, for the attributes are large(L), medium(M), small(S), the values that hold are 198,326 and 83 respectively.

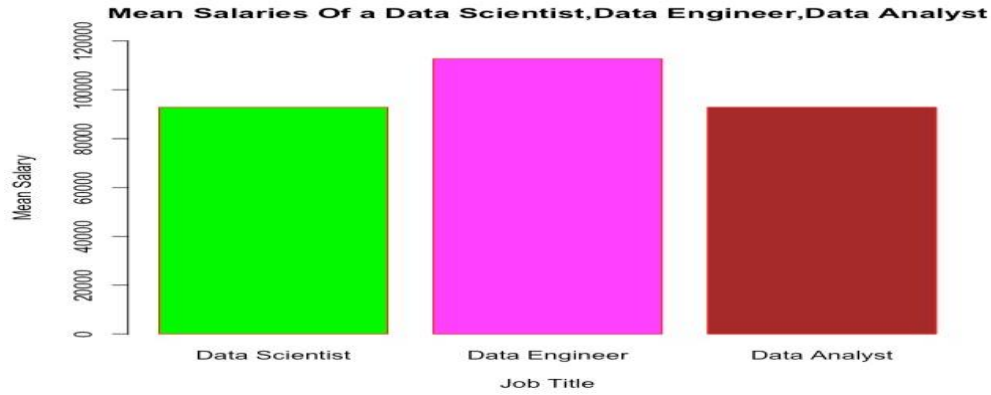


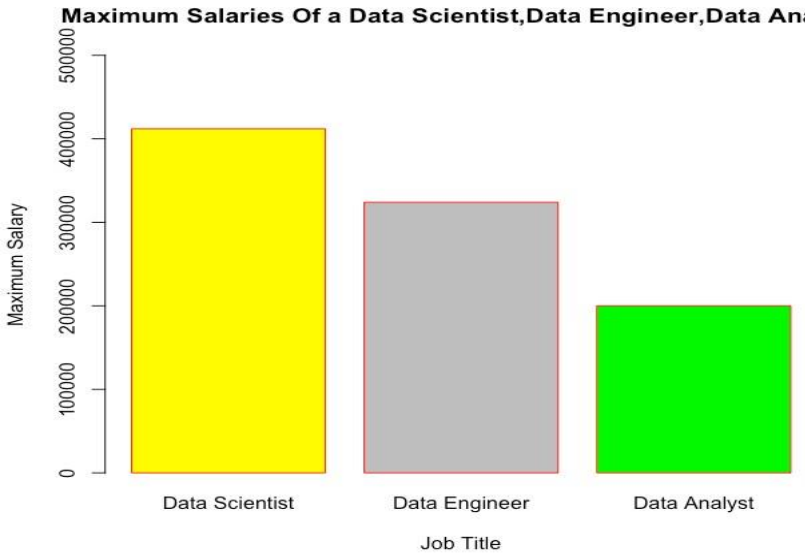
The dataset explains about the mean of the salary in the 3 consecutive years, 2020, 2021 and 2022. The following graph represents the data,



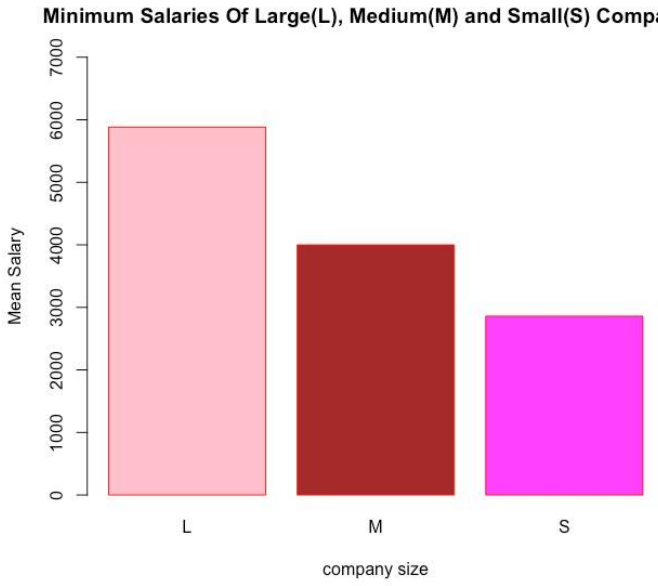
In the given dataset, considering the 3 attributes, data scientist, data engineer and the data analyst; The following graphs represents the data of the minimum, maximum and mean of the salaries for the respective job titles..

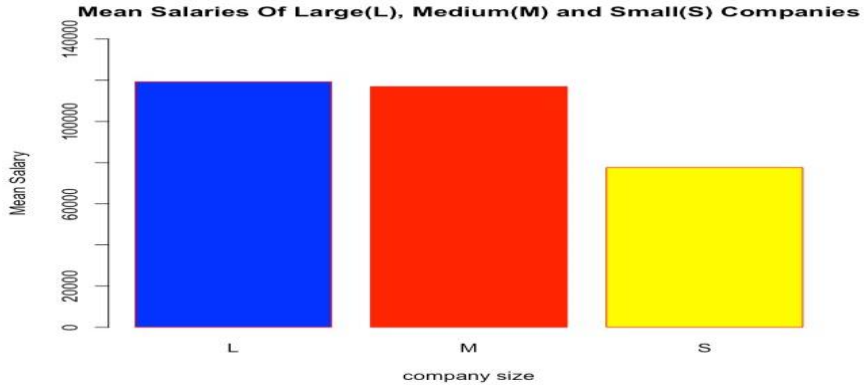


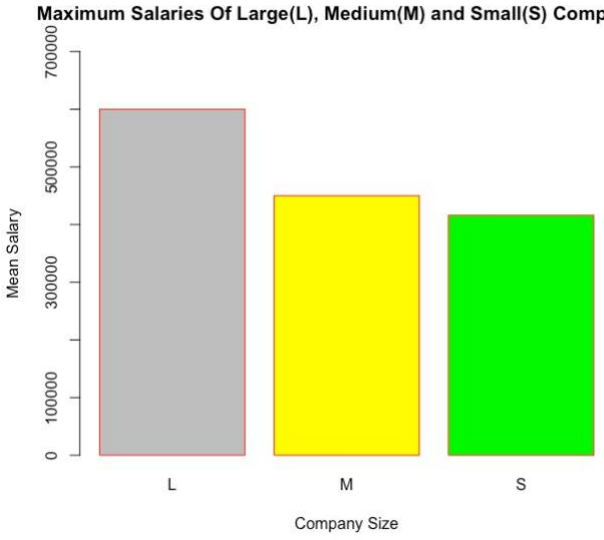




The following graph represents the minimum, mean and maximum salary of large, medium and small companies throughout the dataset.


Minimum Salaries Of Large(L), Medium(M) and Small(S) Companies
Mean Salary
0
1000
2000
3000
4000
5000
6000
7000
L
M
S
company size


Mean Salaries Of Large(L), Medium(M) and Small(S) Companies
Mean Salary
0
20000
60000
100000
140000
L
M
S
company size


Maximum Salaries Of Large(L), Medium(M) and Small(S) Companies
Mean Salary
0
100000
300000
500000
700000
L
M
S
Company Size

The dataset gives an understanding that the employees, who reside and work at different locations are 9 of them out of 607 records, they are

	job_title	salary_in_usd	employee_residence	company_location
40	Data Scientist	45760	PH	US
89	Data Analyst	80000	BG	US
212	Data Engineer	66022	CDC	GB
264	Data Scientist	25532	RS	DE
265	Lead Data Engineer	160000	PR	US
281	Research Scientist	100000	JE	CN
417	Data Science Engineer	60000	AR	MX
489	Applied Machine Learning Scientist	31875	TN	CZ
506	Applied Machine Learning Scientist	75000	BO	US

REFERENCES

https://www.statology.org/xlim-ylim-in-r/

https://www.youtube.com/watch?v=n_ACYLWUmos

https://www.youtube.com/watch?v=Hj1pgap4UOY

https://www.tutorialspoint.com/r/r_histograms.htm