

AN ANALYSIS AND VISUALIZATION OF DATA PROFESSIONALS' SALARY IN INDIA – 2022 USING R

Dr. Vijay Gupta, Associate Professor, International School of Informatics & Management, Jaipur

Abstract

Data Science is an interdisciplinary field that seeks to retrieve information from huge amounts of data. Data scientists use algorithms, mathematics and science to uncover hidden patterns in data and increase efficiency, productivity, profitability for business. Analytics is used for interpretation, systematic computational analysis and communication of meaningful patterns in data. Effective decision making can be made due to data analytics. To improve business performance organizations apply analytics. This paper visualizes and analyzes the data professionals' salary in India-2022 by using R programming.

Keywords: R, R-Studio, Data Visualization, Data Analytics

Introduction

Job of Data Analyst is in very high demand around the world. As a result, other software professionals have less salary as compared to Data Analyst in India. They have very high salaries. If a person updated itself regularly and have the required skill set, Data Analyst career keeps growing onwards and upwards. Salary of Data analyst in India is totally dependent on skilled and updated knowledge. If a person has good and updated knowledge, any company can hire him and give good salary.

Data analysts use figures, charts, numbers, statistics and convert into meaningful, understandable English language. Data professionals have higher salary due to their good skills in mathematics, statistics and Computer Science. As a Data Analyst the person should have knowledge of data mining and good knowledge of languages like SQL, Python, and R etc. Data Analyst extracts significant information from different data sets and visualizes that information using reports and graphs (Sharma, 2021).

A Data Analyst gathers, stores, and deciphers information to change it into important business bits of knowledge that can be utilized to further develop business tasks and cultivate information driven direction. This occupation is absolutely insightful, since it includes information parsing, deciphering and analyzing. The expanding interest for data analysts is increasing the data analyst salary in India day by day. Data Analysts decipher the stowed away data from inside enormous datasets and convey the data actually to every one of the partners in question - associates, business and supervisory groups, clients, organization accomplices, and so forth In view of the interest, Data Analyst salary in India is one of the highest

Skills required for a Data Analyst

Domain Understanding

To find the solution of any problem domain understanding plays very important role for data analyst. Without understanding the domain, data analysts do not understand the parameters require in a business for a significant change. Factors that affect the growth of a business must be known to data

analyst. Data analyst salary in India directly related to business understanding. The growth in business can be achieved by good understanding and deep insights about business. This skill set must be in data analyst.

Mathematics & Statistics

By good decision making any problem can be solved easily. Mathematics and Statistics makes a person confident and reliable to take decisions objectively. (Sharma, 2021). Data Analyst can make predictions from the historical information. He has the ability to generate patterns, segments, conclusions by using Mathematics and Statistics. (Kandel et.al., 2011)(Sharma, 2021).

Technical Skills

Data Analyst must have technical skills. By using domain understanding and mathematical skills data analyst identify and solve a problem easily. A data analyst needs to go through data sets and replicate the algorithms and provide with better business solutions. Tools such as Python, SQL and R play a crucial role in providing business solutions.

Soft Skills

Data Analyst must communicate the solution in the most simple and understandable format to the stakeholders. Soft skills plays a very important role for communication, whether a person might not know anything about Statistics, or root mean square error, or clustering algorithm or classification but where soft skills come in, everything becomes very easy to understand.

Data Analyst Job Roles

Data Science means Statistics can be implemented through different programming languages. Languages like R, Python is used in data science by which data can be retrieved and sorted out as per the requirements. Now a days, in India data science programmers or software developers earn more as compared to other software developers. Each company requires information of market and customers. After collection of data, data cleaning and processing is done to retrieve meaningful information. In major metros like Delhi-NCR, Mumbai, Pune, Bangalore and Hyderabad data scientists are getting very good salaries.

Experiments and Observations

Data Set

The Dataset consists of salaries for Data Scientists, Machine Learning Engineers, Data Analysts; Data Engineers in various cities across India 2022. It contains 4344 records and five columns Company Name, Job Title, Salaries Reported, Location and Salary.(Mooney, 2018).

Titles present in this Dataset

Data Scientist, Data Science Associate, Data Science Consultant, Data Science, Senior Data Scientist, Junior Data Scientist, Lead Data Scientist, Data Science Manager, Data Scientist - Trainee, Data Science Lead, Data Analyst, Data Engineer, Machine Learning Engineer, Machine Learning Software Engineer, Software Engineer - Machine Learning, Machine Learning Engineer/Data Scientist, Machine Learning Consultant, Machine Learning Data Associate, Machine Learning Data Associate I, Machine Learning Associate, Machine Learning Data Associate II, Associate Machine Learning Engineer. **(Refer Figure 1)**

Methodology

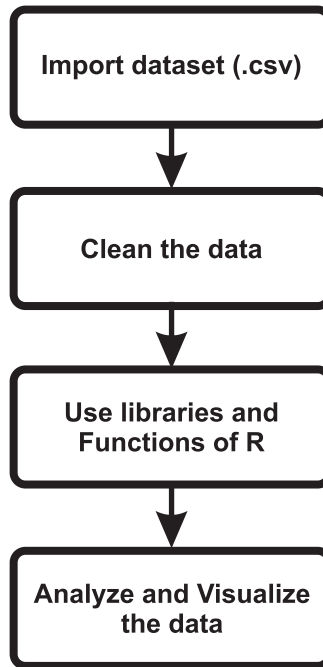


Figure 1 : Methodology

Parameters that affect the Salaries

Multiple factors that might influence the data scientists' salary-location, industry, years of experience, roles, and tools used so that professionals can maximize their data science salaries and work to their benefit as great enterprise data scientists.

Results and Observations

Table 1 gives the complete description of the data. It shows number of rows, number of columns, number of discrete columns, missing columns, complete rows and total observations.

Table 1 : Description of Data

Rows	4344
Columns	10
Discrete_columns	6
Continuous columns	4
All_missing_column	0
Total_missing_value	12974
Complete_rows	20
Total_observations	43440

By applying ggplot (Datacarpentry 2021), library and function plot of R to create this graph. It gives complete visualization of data. (Wilke 2019)

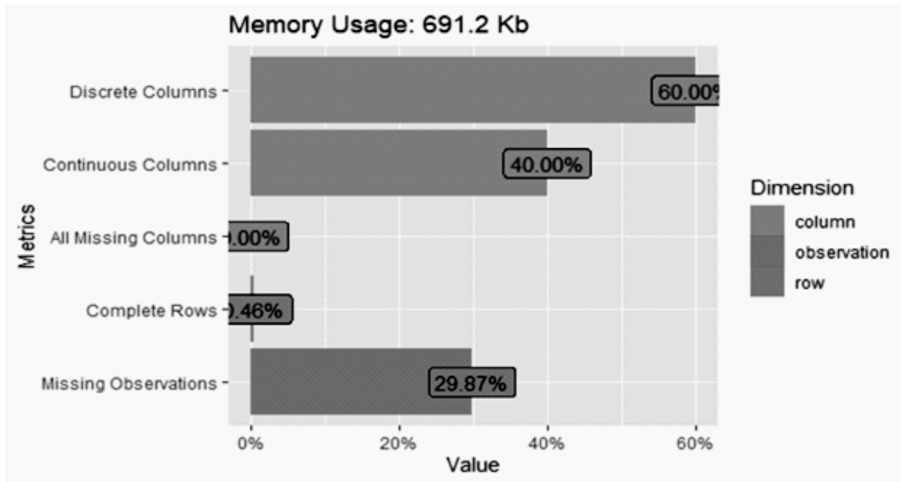


Figure 2 : Memory usage of Data Set

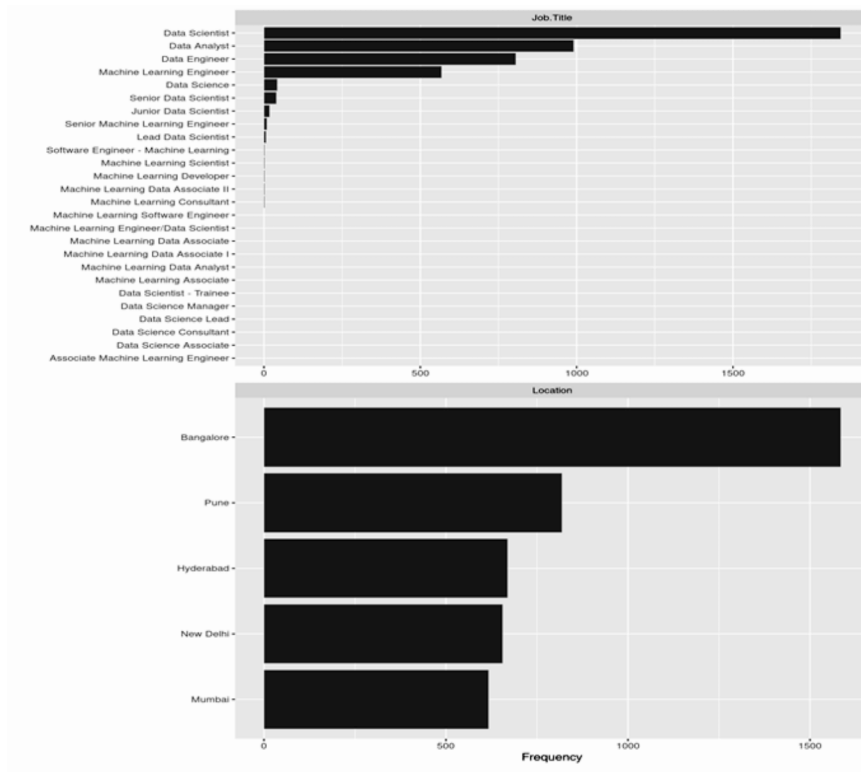


Figure 3: Job title, Salary and Location about the Data

Figure -3 describes the job title, salary and location about the data. It describes mainly 5 locations Bangalore, Pune, Hyderabad, New Delhi and Mumbai. (Wickham 2018)

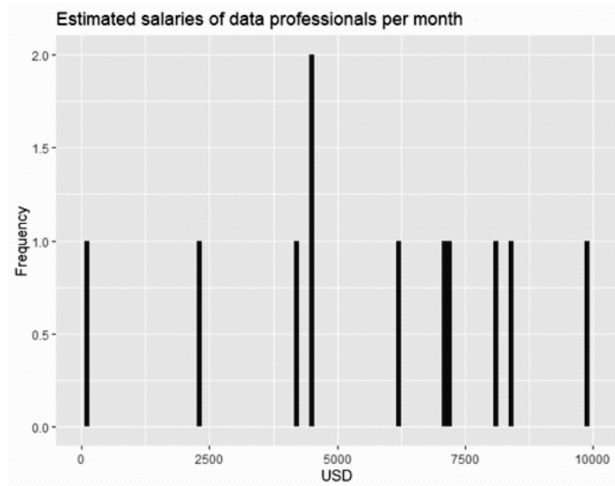


Figure 4 : Estimated Salaries of Data Professionals per Month

Figure-4 describes estimated salaries of data professionals per month. It shows that maximum data professionals have salary between 3500 to 4000 USD. Data Scientists have salaries near about 6000USD per month.

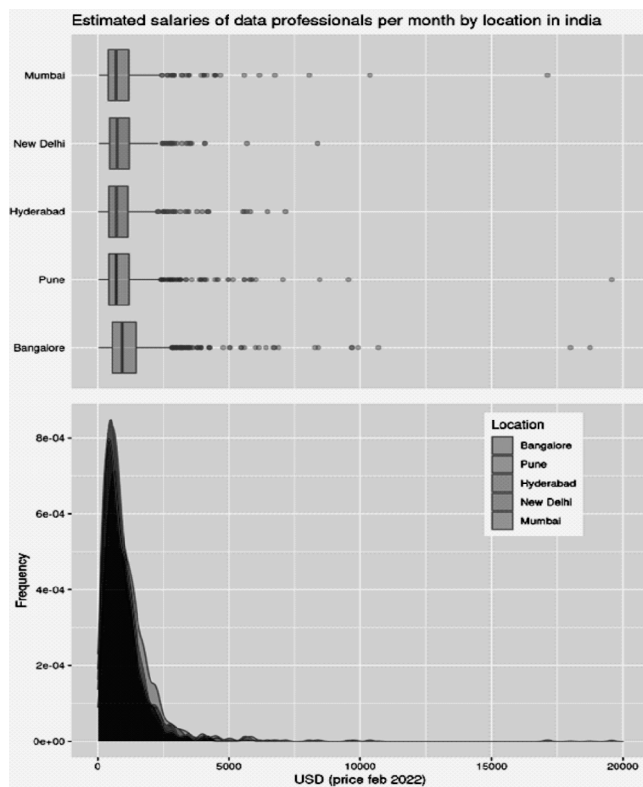


Figure 5: Estimated Salaries of Data Professionals Per Month by Location in India

Table 2 : Mean Salaries of Cities

Location	mean_usd	dev_usd
Bangalore	1318.5120	4305.3384
Pune	1325.1763	5238.1048
Hyderabad	1319.0484	5587.4395
New Delhi	935.9838	732.8731
Mumbai	1142.0104	3034.9252

Table-2 describes that location does not matter. Distributions are almost similar. New Delhi and Mumbai lag behind with the mean.

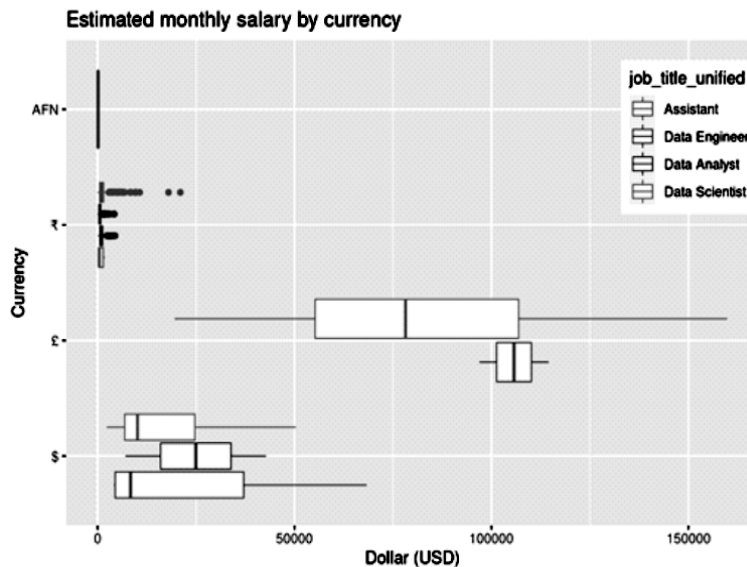


Figure 6 :Estimated Monthly Salary by Currency

Figure-6 shows monthly salary by currency. Salaries paid in Western currencies are higher than most of those paid in the local currency, reaching very high points, and establishing much higher averages. The euro even outperforms the U.S. dollar. It is clear that the jobs paid in these currencies is because they are foreign companies, which implies that within the country with a weaker currency, the profits are supremely higher.

Conclusion

This paper gives analysis and visualization of data professionals salary in India. It concludes that location does not affect the salary very much. Salary will be effected by designations and skill sets. It also describes monthly salaries paid in western currencies are higher then those paid in local currency.

References:

- Kandel, S., Heer, J., Plaisant, C., Kennedy, J., Van Ham, F., Riche, N. H., & Buono, P. (2011). Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization*, 10(4), 271-288.
- Sharma, R. (2021). Data Scientist Salary in India in 2022 [For Freshers & Experienced]<https://www.upgrad.com/blog/data-scientist-salary-india-freshers-experienced/>
- Mooney, P. (2018). Kaggle Machine Learning & Data Science Survey.
- Wickham, H., Francois, R., Henry, L., & Müller, K. (2018). dplyr: A Grammar of Data Manipulation. R package version 0.7. 8. R package version 0.7. 8.
- Wilke, C. O. (2019). Fundamentals of data visualization: A Primer on Making Informative and Compelling Figures. O'Reilly Media.