Name: Husain Rangwala

Date: 23rd February 2021

Title: Module 4 Project

Executive Summary Report 4

Summary:

From the report we can understand what defines data analyst and what were the challenges organization faced with the application/service data and various ways they use data science techniques to solve it.

Once we have a brief understanding on the history of data analyst, we will investigate the various need of the organization by looking into the hierarchy of need as per data science. The image mentioned below shows the various stage and associate job profile a person will work for it. we will also see what the stage does and how the stages are interacting with each other.

Finally, we will recognize what a data scientist is. And analyze the various roles and responsibilities a data scientist deal while using the data science knowledge to predict and provide useful insight to the customer requirement. The report also shows the average salary and the market demand for a data scientist across the globe.

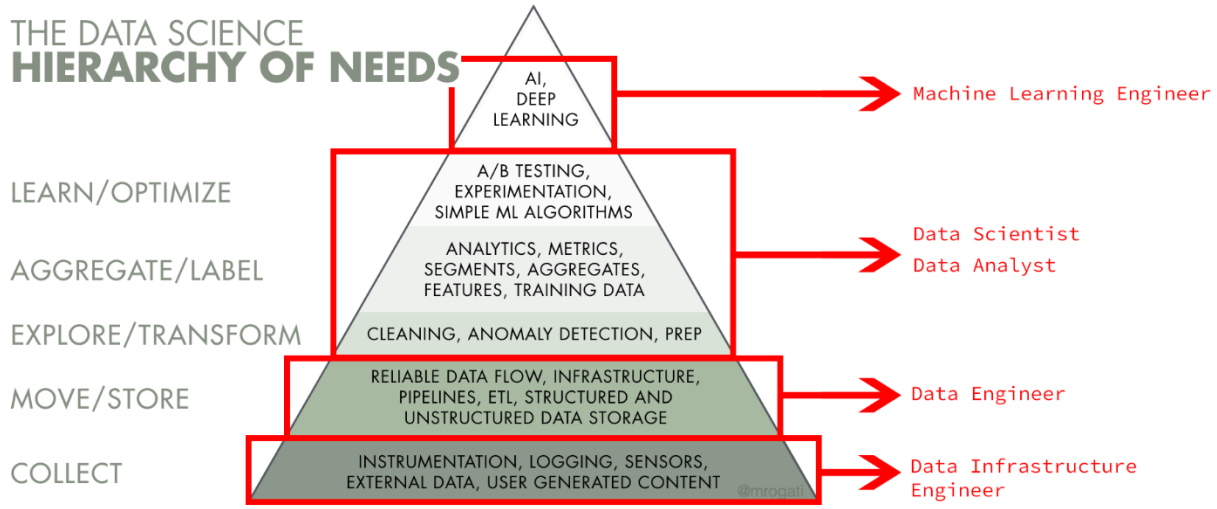
Introduction

Data Analyst is an evolving field with abundant opportunities and requires various skillsets. Almost all well-established organization are adopting and making data analysis one of its cornerstone of their success. This field deals with data and ways on how to normalize and create meaning out of it. Data here refers to the grouped information of people's choice, need, liking etc. It is the work of the data analysts to make that messy data useful and productive.

Since data became the new 21st century currency, work positions in Big Data and Data Science have diversified and branched out at an unparalleled pace. Two of the most promising work positions with an upward career trend are data engineer and data scientist.

Finding

In today's market, data is more valuable than oil. Any and every interaction user makes with the internet or any user made application produces huge and continuous volume of information. This information can be collected through Connected Devices, Data Generation, Ease of Access, Customer Needs and Marketing, Data-Driven Decision-Making. This collected data is only one half of the solution. The second half involved building products that can utilize the data. Till now organizations are trying to analyze and understand as to how could make any sense of the acquired data. Thus, to accomplish this, many organizations are using various machine learning and algorithm predict the information for their user. Thus, there is requirement of skill personal who could analyze the data and compiled smarter data and analytics best practices into a customizable roadmap for the clients.

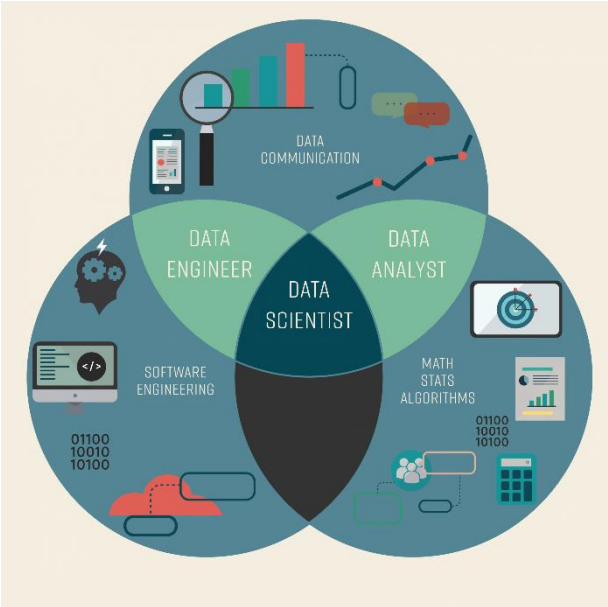


Now let us understand the hierarchy of need from a data science point of view:

1. Collect – This stage is important as continuous and huge volume of data that various application and service that are available on the internet. In this stage there are many ways the data can be stored for further analysis. There are several processes in which the data can be stored one such way is to create a data lake or data warehouse. Through these structure various data such instrumentation, logging, sensors, external data, and user generated content can be stored in the original state.
2. Move/Store – This stage used the raw data stored its previous stage and create pipeline, data flow and select appropriate data set as per business requirements.
3. Explore/ Transform – This stage one explores and transform the data set that was provided from the previous stage and accordingly transform the data once you get the basic requirement from the previous stage. According to the business requirement, you can transform the data and use predictive algorithm to check the dataset.
4. Aggregate/Label – In this stage the transformed data provided are being studied and accordingly the data is aggregated and label for business need. In this stage, the data is modified, and it is made from user understanding. At this stage, the data is presented to the client and accordingly the business can make predictive decision on the data provided.
5. Learn/optimize – Once the data is made as per business demand, many algorithm and artificial learning techniques can be used to predict the user interaction and accordingly provide appropriate output as per their clicks.

What is Data Scientist?

Data scientists are big data wranglers, gathering and analyzing large sets of structured and unstructured data. Data scientists are analytical experts who utilize their skills in both technology and social science to find trends and manage data. They design data modeling processes, create algorithms and predictive models to extract the data the business needs, and help analyze the data and share insights.



According to the above image Data science is the combination of Data scientist, Data analyst and Data engineer. All the three field have their own responsibility and roles they should abide to.

In the data scientist, the responsibilities are as follows:

- Ask the right questions to begin the discovery process: Data scientist work with client and must ask the correct question to their client so that they can use it to create data flow and dashboard accordingly.
- Acquire data: Once they get the business requirement, data scientist can start to collect the raw structured or unstructured data.
- Process and clean the data: Once they raw data is available with data scientist; they start to clean and remove unwanted data and create initial data set for their business requirements.
- Integrate and store data: Once they clean the data, they can integrate the business requirement with the given data and store the new data.
- Initial data investigation and exploratory data analysis: At this step, the start to use data analysis technique and try to test their initial data set and see how well the data shows the output.
- Choose one or more potential models and algorithms: Using the previous step, they use various models and algorithm to predict their output.
- Apply data science techniques, such as machine learning, statistical modeling, and artificial intelligence: Use available techniques and check if the data shows the output as per the business requirement.
- Measure and improve results: Once they get the resultant output, data scientist must measure if they techniques are possible and predict if the resultant can bring more helpful insights.
- Present result to stakeholders: The data is present to the client in a human presentative way.

- Adjust based on feedback: If the business wants to adjust the data models and may want to provide feedback or improve their business requirements.
- Repeat the process to solve a new problem: Once the feedback from the customer is received, data scientist must follow back through the step and come up with better modelling techniques for the clients.

Skills and salary for Data Scientist

Skill that are required for data scientist.

- Fundamentals of Data Science
- Statistics
- Programming knowledge
- Data Manipulation and Analysis
- Data Visualization
- Machine Learning
- Deep Learning
- Big Data
- Software Engineering
- Model Deployment
- Communication Skills
- Storytelling Skills
- Structured Thinking
- Curiosity

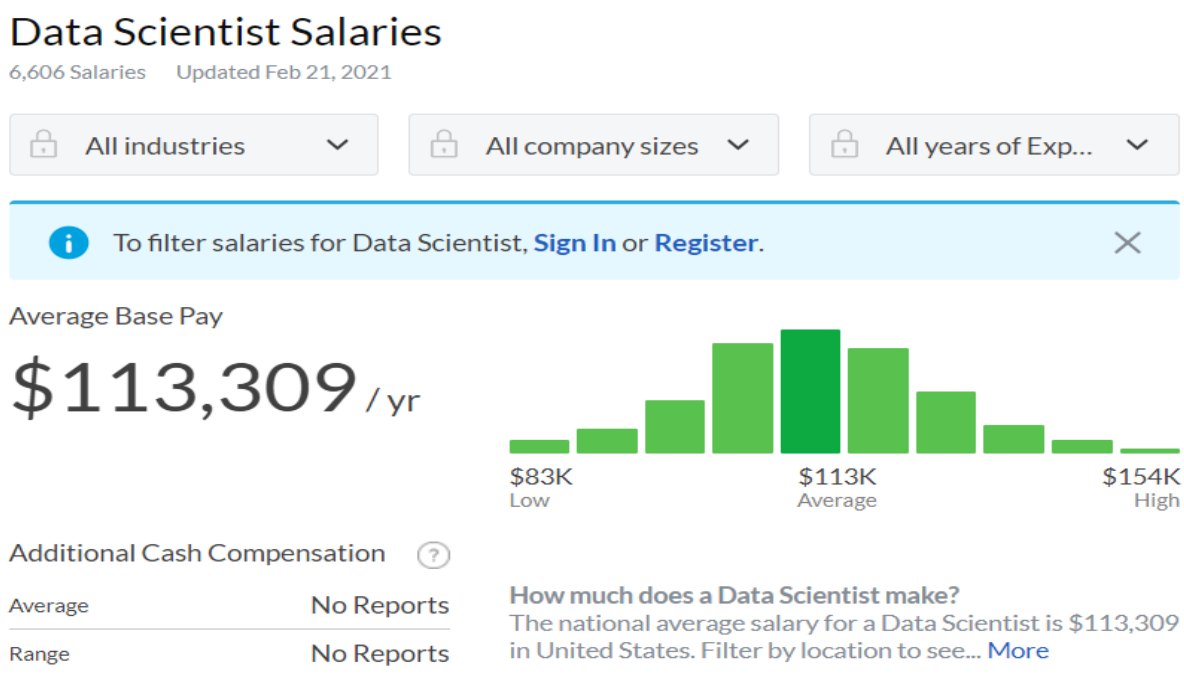
Where Do Data Scientists Work?

Data scientists can work in a variety of settings. Those might include:

- Federal government
- Computer systems design
- Research and development
- Colleges and universities
- Software companies
- Car companies
- Delivery companies
- Tech companies

Technical companies are one of the setting, I would like to be a part of. With my background as a Big data/ Cloud technical support I had a basic understanding on the roles and responsibilities of data infrastructure and data engineer. As my clients were from the technical market, I had insight on what the client requirement are and how to create useful pipeline, workflow to help migrate useful data set to the application team. After I complete my graduation in Data analytics, I would help technical organization with data integration as well as utilize my data science knowledge to offer algorithm and machine learning technique that could help business understand the data.

As per Glassdoor the below image shows the average cost a data scientist earns.



Market demand for Data scientist.

The above image showcases the salary a data scientist can earn across the globe.

References

1. https://knowledge.wharton.upenn.edu/article/whats-driving-demand-data-scientist/
2. https://towardsdatascience.com/why-data-analytics-is-gaining-hype-in-the-21st-century-b7b1ca289f09
3. https://www.purdueglobal.edu/blog/information-technology/rise-of-data-analyst/
4. https://www.pwc.com/us/en/library/data-science-and-analytics.html
5. https://datasciencedegree.wisconsin.edu/data-science/data-scientist-salary/
6. https://www.linkedin.com/salary/data-analyst-salaries-in-san-francisco-bay-area?isQueryParams=true
7. https://www.bioenergyconsult.com/data-science-new-emergence/
8. “https://www.edureka.co/blog/10-reasons-why-big-data-analytics-is-the-best-career-move
9. https://businessoverbroadway.com/2018/04/08/salaries-of-data-scientists-and-machine-learning-engineers-from-around-the-world
10. https://www.cio.com/article/3292983/what-is-a-data-engineer.html
11. https://businessoverbroadway.com/2018/04/08/salaries-of-data-scientists-and-machine-learning-engineers-from-around-the-world
12. “https://towardsdatascience.com/data-engineer-vs-data-scientist-bc8dab5ac124
13. https://intellipaat.com/blog/how-much-does-a-data-scientist-earn/
14. “https://fourpillarfreedom.com/comparing-data-scientist-salaries-and-cost-of-living-in-different-u-s-cities/”
15. https://www.glassdoor.com/Salaries/data-scientist-salary-SRCH_KO0,14.html