

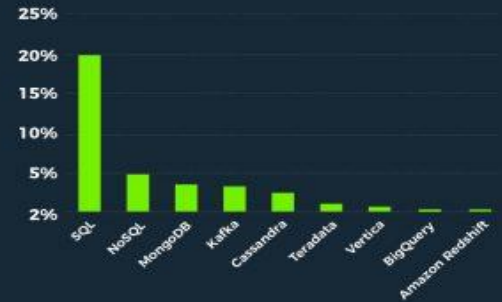
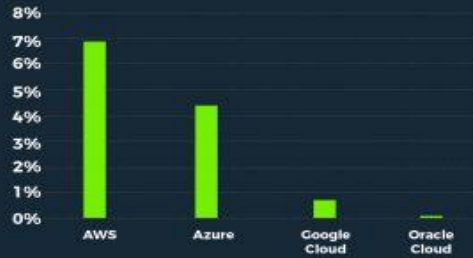
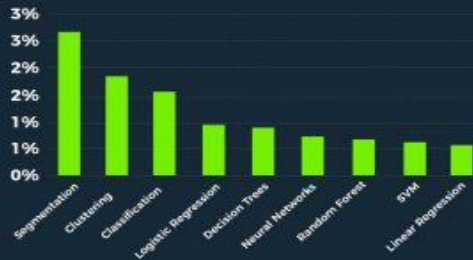
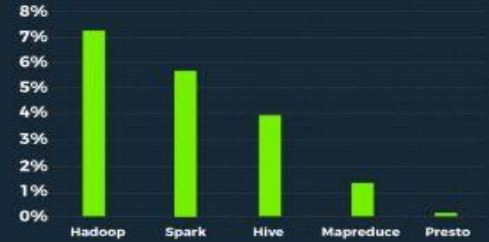
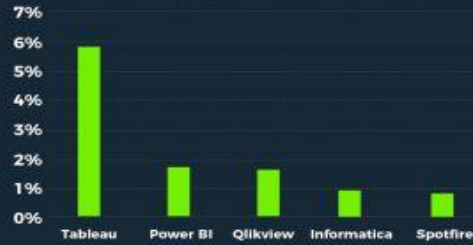
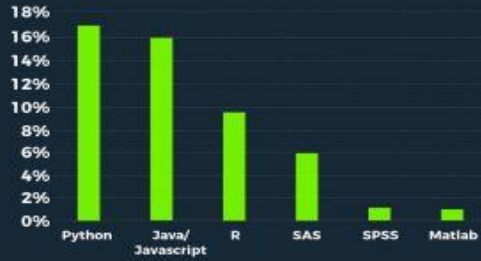
Road-map — Data Scientist



Why Data Scientist?

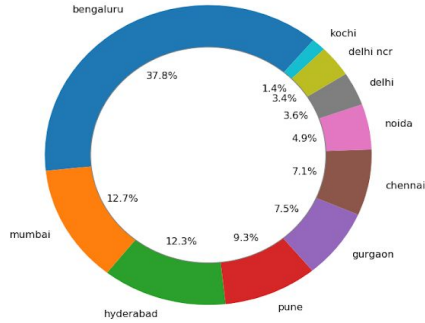
1. Over the past 10 years, there are great requirements for Skilled engineers
2. List of news jobs,
 - a. **Data Scientists**
 - b. **Machine Learning Engineer**
 - c. **Data Analysis**
3. Above declared as a **Sexiest job of the 21st century** with **highly paid** and **future safe** jobs.

Skills in demand in 2019

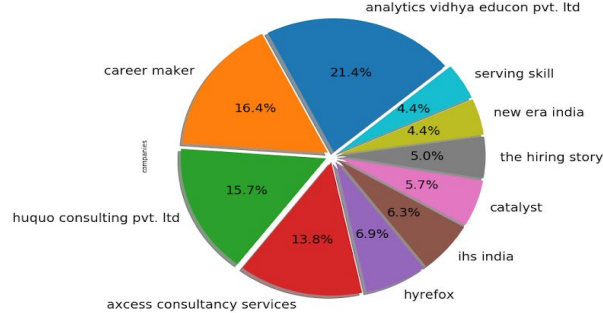


DATA SCIENCE JOB MARKET TREND ANALYSIS - 2020

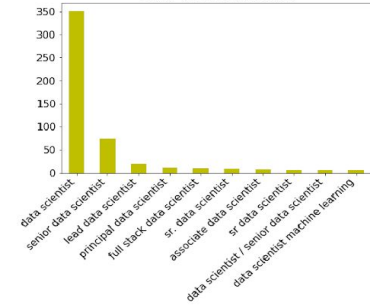
Location Wise Data scientist Jobs



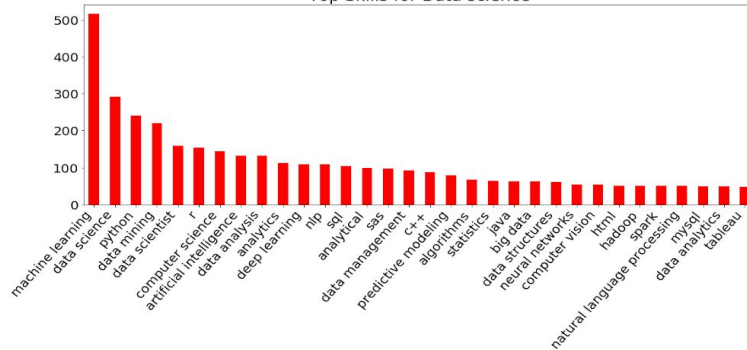
Companies with more job postings



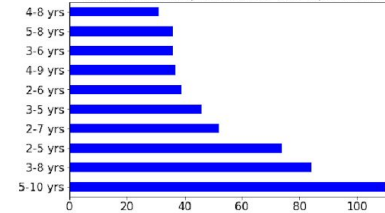
Data Scientist Roles



Top Skills for Data science



Data scientist jobs based on Experience



India / Job / Data Scientist

Average Data Scientist Salary in India

Rs 818,099

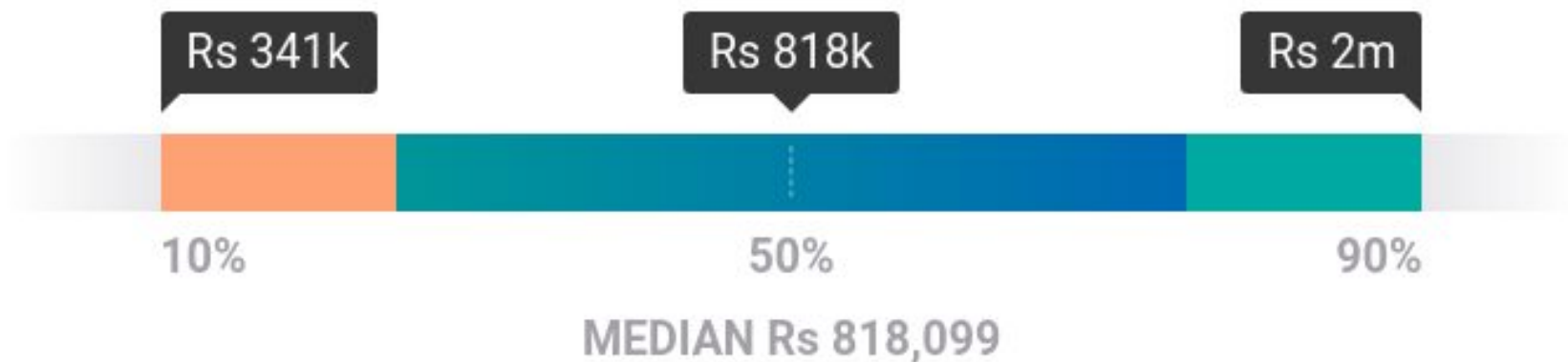
Avg. Salary [Show Hourly Rate](#)

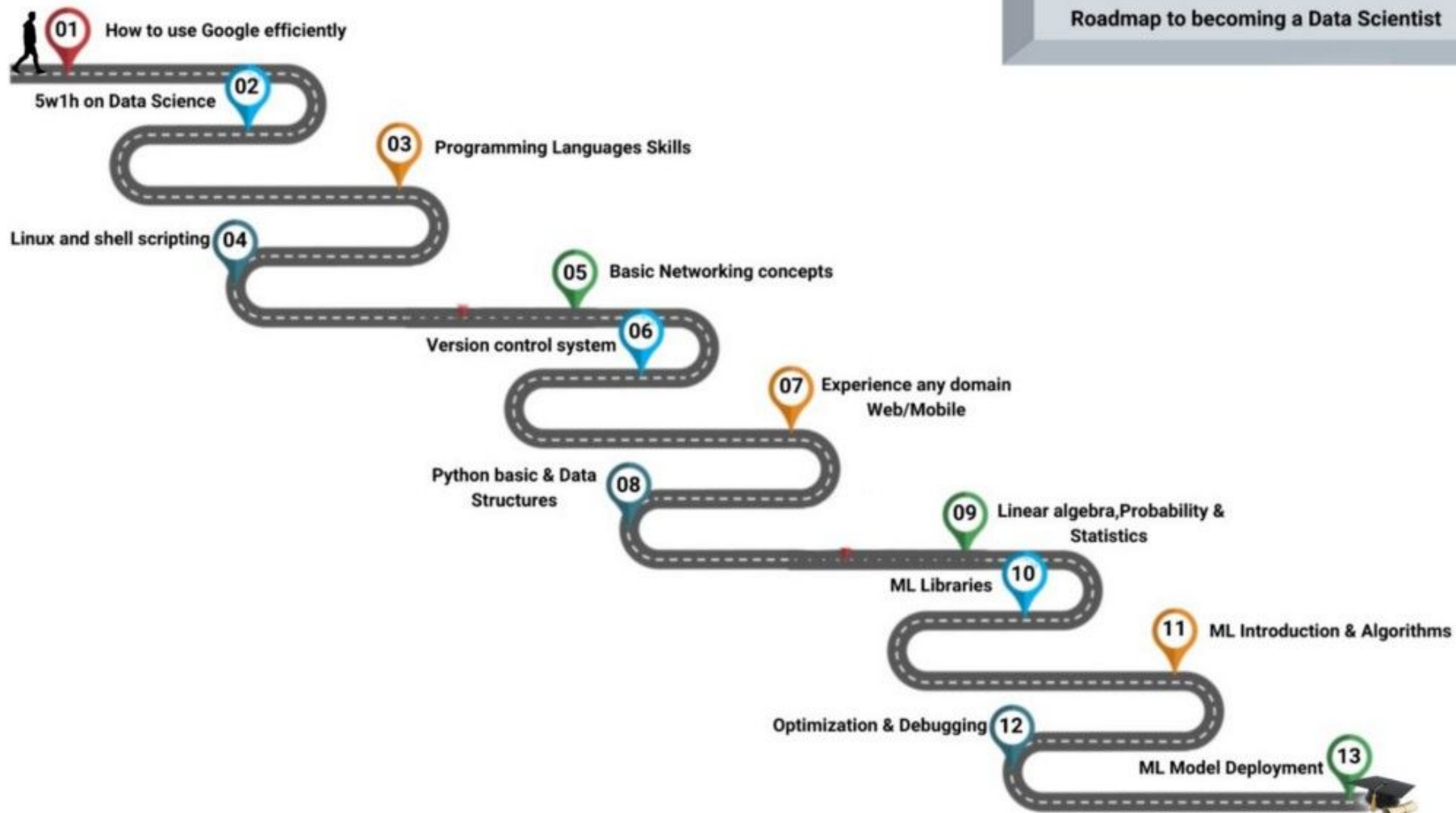
Rs 100,473
BONUS

Rs 32,500
COMMISSION

Rs 36,667
PROFIT SHARING

The average salary for a Data Scientist in India is Rs 818,099.





1. How to use Google efficiently — 2 Days

Before jumping directly into data science first we have to learn how to use the **world's best search engine, "Google"** because nowadays it is placing or involving in all our daily tasks. The more efficiently you learn to use google the best results to get from google

2. 5w1h on Data Science — 2 Days

There is **no use of learning** anything without knowing it's **basics, reasons, applications**, etc, like here also we should know what is data science, why data science, how it works, its application. I **use this rule** whenever I **start learning and exploring a new topic**, it gives me proper information and simple notes.

3. Basic Programming Languages like C, C++, Python, Java, and Data structures — 15 days

Programming is the basic foundation of the technology world and without coding experience, we can not start our career in the Data Science field, so we should learn programming.

4. Linux commands and shell scripting — 7 days

Now you are thinking about why we need this step. Wait I will explain to you see when you are working in any company they mostly preferred to work on open source technologies, libraries, OS. and most of the time we encountered the **basic commands of Linux** or we have to **write some automated scripts** not only in the data science field is required in every designation of the programming world. I **recommended that every programmer** should follow and **learn** the topics till **steps 7**.

5. Basic Networking concepts, commands — 7 days

After completing till four steps, we will proceed to step 5 where you can learn the basic networking concepts and its basic commands, now again you are confused what is the need of networking in the Data Science career path. Hold on all points will be cleared, when we are working on machine learning, we have to deploy models on the cloud, have to do parallel processing training of the model, configuration different things in the cluster, have to work on master and slave architecture, which needed networking knowledge.

6. Version Control System (VCS) —7 days

When we are working on any project. We need a system that can **manage our code**, where VCS comes in the role. With the help of VCS, you can **track your daily coding activities** if you **commit** and **push** your **daily code**. You can use any VCS but most trending and open source and ease of use VCS are Github and I personally use Github. If you **maintain** your **GitHub account** with **proper coding style** and **repositories**, then it helps you in applying for a job also.

7. Need some experience in any domain like Web/Mobile — 15 days

Ok one more, Now why this? This is because machine learning is a concept or in simple terms, we can say that machine learning is all about the algorithms. So we have to integrate it somewhere otherwise, there is no use of doing machine learning if you don't know the integration of the machine learning algorithms for the user.

8. Python and its Data Structures — 10 days

As we all know that Python is one of the most used programming languages in Machine Learning or Data Science world. So you need to have good command in it to do as much as practice on it, with different data structures and even try to learn advanced Python also like **test cases, debugging, optimization** techniques of python code, etc.

9. Linear algebra, Probability, and Statistics — 15 days

Learn the foundation or building blocks of the Machine Learning algorithms. It gives the **matrix calculations, vectors, inverse, and transpose**. Probability gives some machine learning models like **Naive Bayes, probability theorem**, etc. Statistics give us **statistical models like regression, logistics** which helps us in the coming steps a lot

10. Machine Learning Libraries and Tools — 15 days

In this step you need to learn all important machine learning libraries which we are going to use in coming steps like **Pandas for data cleaning**, preprocessing, manipulation, etc, **Numpy for n-dimensional array calculation** and basic arithmetic operations on it, **Matplotlib for graphs** and **Seaborn advanced graphs** for in-depth analysis of the data with the help of advance graphs. You should have to practice on Data Science IDE's like **Jupyter notebook**, Spider etc. In this step, you also have to learn and step up the **Machine Learning Development environment**. Here I **created an automated script** which

11. Machine Learning Introduction, Algorithms, and Building blocks — 20 days

In this step, you have to learn the machine learning introduction and their basic algorithms mathematics like **Regression, Classification, SVM, SVR, KNN, Decision Tree**, etc and its building blocks like **Gradient Descent, Cost Function, Loss Function, Stochastic Gradient Descent, Mini Batch Gradient Descent, Stochastic gradient descent convergence, Online learning, Data Parallelism**, etc. You should have to implement all algorithms in **core python without using any libraries**. If you code all algorithms from scratch in core python then you will get to know how the algorithms work, what are key points and other important and this step is one of the most important, crucial, and time taking steps. When I was learning, **I gave more than 3 months to learn and code every algorithm in core python**. Here is my

12. Optimization, Debugging and Validation Techniques — 10 days

Learn how you can optimize your algorithm, how to test and debug machine learning algorithms or solutions. This is another important step in the roadmap of the data scientist because knowing how to debug or optimize any algorithms or solution is much more important than the implementation of algorithms or solution like **Evaluating a hypothesis, Model selection and training validation test sets, K-Fold Cross Validation, XGBoost, Diagnosis, Regularization and Bias/Variance, Learning curves** etc

13. Project on ML and deploy using flask and AWS — 7 days

After completing all these 12 steps it's time to deploy the machine learning model on any website or app. Here you can use flask or streamlit or fastapi or TensorFlow.js.

Key People to Follow



Mark Zuckerberg, Founder of **Facebook**



Yann Lecun,
Director of AI Research, **Facebook**



Andrew Ng, is VP & **Chief Scientist** of **Baidu**, **Co-Founder** of **Coursera** and an Adjunct Professor at **Stanford University**.

Key People to Follow



Geoffrey Hinton, a researcher at **Google** and professor emeritus at the **University of Toronto**. Also known by **Godfather of Deep Learning**.



Fei Fei Li Associate Professor at *the C.S. Department* at **Stanford University** and Also known by **Expert of Computer Vision**.



Andrej Karpathy
Research Scientist at **OpenAI**.

Like

Comment

Share