




Data Science 101

A complete roadmap and guide to learn Data Science 🚩

What we'll learn today?

-  What is Data Science?
-  Why is Data Science important?
-  What are the different job roles in Data Science and Machine Learning fields?
-  What to master to learn Data Science? (Skills)
-  How to learn Data Science? (Learning Plan)

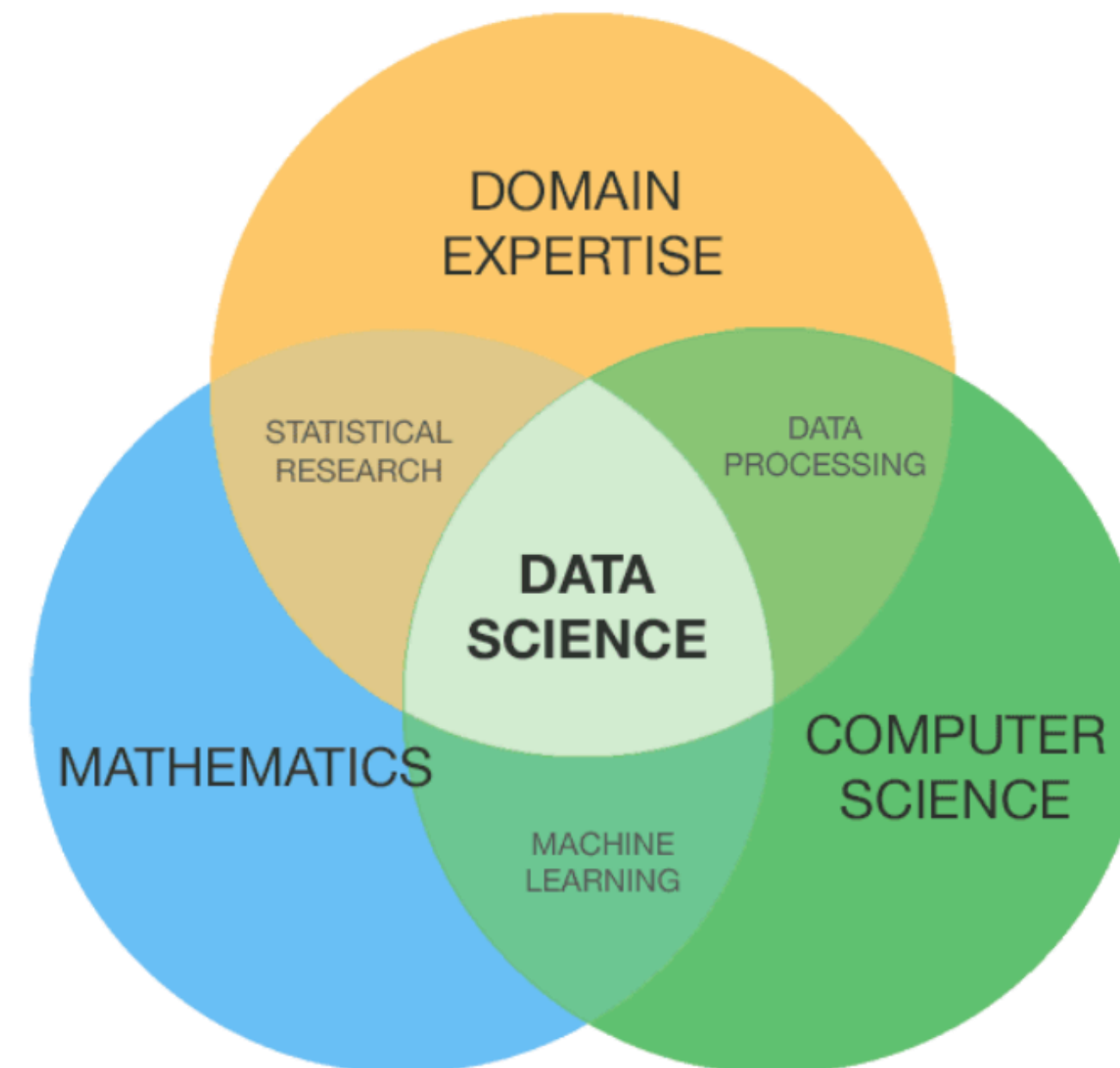


What is Data Science?

Data Science is a field that deals with data processing, analysis, and extraction of insights from the data using various statistical methods and computer algorithms.

It is a multidisciplinary field that combines mathematics, statistics, and computer science.

Data Science churns raw data into meaningful insights.



*Source: Palmer, Shelly. Data Science for the C-Suite.
New York: Digital Living Press, 2015. Print.*

Why is Data Science Important?

Data is the new oil!

- With the emergence of new technologies, there is an immense (exponential increase) amount of data generated by systems, mobile phones, laptops, computers, etc., that are being stored, which holds a lot of valuable knowledge ready to be explored.
- Data Science is the study of data, where it comes from, what it represents, and how you can use it as a valuable resource for creating business strategies.
- Data Scientists deal with the data to assist companies in making proper decisions.
- The majority of businesses and industries nowadays deal with data on a day-to-day basis.
- Data Scientists leverage the power of data and bring out value from it.
- They convert a business problem into a Data Science problem.

Why is Data important and why do we need it?

- Data Science is a recent discipline in Mathematics and Computer Science.
- Computing process + Cloud Storage + Analytical tools → Computer Science merged with Statistics → Data Science is born!
- Before Data scientists, we only had Statisticians existing. They performed a qualitative analysis of data and analyzed companies' overall performance and sales.
- Earlier, performing data analytics involved much manual effort. Then computers could solve more complex statistical problems efficiently.
- With Data Science, companies make the right business decisions to maximize their profits, solve complex problems, and boost customer experiences with the help of Data!

 What are the different job roles in Data Science and Machine Learning fields?

Data Analyst

Responsibilities:

- Tech Focused > Business Focused
- Datasets are examined to draw conclusions about the information they contain
- Performing statistical analysis to discover new patterns and making reports with recommendations
- Extracting data from various sources
- Analysing data and forecasting trends that impact the organisation
- Develop KPIs and visual representation of information contained in the data
- Working to improve data collection and quality

Skills:

SQL, BI Tools - Tableau | PowerBI | Looker, Statistics, Python, R, Data Wrangling and Mining, Database Management, Cloud Technologies - AWS | GCP | Azure

Business Analyst

Responsibilities:

- Business Focused > Tech Focused
- Identifying and prioritizing the organizations functional and technical needs and requirements
- Using SQL and Excel to analyze large data sets
- Compiling charts, tables, and other elements of data visualization
- Creating financial models to support business decisions
- Forecasting, budgeting, and performing both various analyses and financial analysis

Skills:

Data Analysis and Visualisation tools - Excel, Tableau | PowerBI | Looker, Python, R, SQL

Data Engineer

Responsibilities:

- Tech Focused
- Design, build, and maintain data management systems and data pipelines
- Data collection/acquisition and management
- Implement data processing and transformations
- Working to improve data collection and quality processes
- Implementing security measures for database
- Provide ML models with data

Skills:

Scala, Go, Java, Python, Big Data Technologies, Database Warehousing and Management, Cloud Technologies - AWS | GCP | Azure

Data Scientist

Responsibilities:

- Tech + Business Focused
- Developing predictive models
- Using mathematical, statistical, and programming skills to gain insights from data using Machine Learning techniques
- Performing Exploratory Data Analysis to identify trends in data and make predictions
- Gathering, organizing, cleaning, analyzing data, and verifying the integrity of data
- Analyzing large amounts of data to forecast trends and provide reports with recommendations
- Collaborating with business, engineering, and product teams

Skills:

Machine Learning, Statistics, SQL, NoSQL, Python, R, Data Visualisation Tools, Data Wrangling and Mining, Big Data Technologies, Cloud Technologies - AWS | GCP | Azure, BI Tools

Machine Learning Engineer

Responsibilities:

- Tech Focused
- Implementing and deploying Machine Learning models
- Running and analyzing AI/ML system experiments and tests to improve ML algorithms
- Designing and developing machine learning systems
- Automate ML processes and make models work in a production environment
- Create documentation on ML processes
- Performing statistical analysis
- Make business recommendations

Skills:

Machine Learning, Artificial Intelligence, Python, R, Java, Applied Mathematics, Statistics, Cloud Technologies - AWS | GCP | Azure



Data Analyst - DA

Business Analyst - BA

Data Engineer - DE

Data Scientist - DS

Machine Learning Engineer - ML Eng



What to master to learn Data Science? (Skills)

★ Programming and Tools

Python
R
SQL
NoSQL
Big Data Technologies
Data Manipulation
Data Visualization
Exploratory Data Analysis

★ Statistics and Probability

Descriptive Statistics
Inferential Statistics
Normal Distribution
Hypothesis Testing
Probability
Discrete Distributions
Continuous Distributions

★ Mathematics

Linear Algebra
Calculus
Matrices
Vectors and Spaces

★ Machine Learning & AI

Supervised Learning
Unsupervised Learning
Reinforcement Learning
Deep Learning
Computer Vision
Natural Language Processing

★ Domain Knowledge

Business Understanding
Vision and Mission

★ Soft Skills

Communication
Presentation
Storytelling with Data
Social Skills
People Skills

Skipping steps is injurious to learning

Data Science

**Machine Learning &
Deep Learning**

Statistics & Mathematics

Data Visualization & Analysis

Programming







How to Learn Data Science? (Learning Plan)



1. I'm a complete newbie trying to learn from scratch
2. I tried learning on my own earlier but couldn't proceed any further due to the lack of support, motivation, etc...
3. I have good knowledge and can practice Data Science



SQL

- SQL Tutorial - Full Database Course for Beginners by Free Code Camp 
- SQL Masterclass for Data Scientists by Data With Danny 
- Intro to SQL by Kaggle 
- Advanced SQL by Kaggle 

Note: The learning path is an introductory and beginner-friendly plan not limited to the above resources; there is a plethora of information on the internet to browse and learn!



Python

- [Learn Python - Full Course for Beginners by Free Code Camp](#) ➡
- [Python for Data Science by Great Learning](#) ➡
- [Python for Data Science \[Crash Course\] by Frank Andrade](#) ➡
- [Python, Data Science & Machine Learning Mini-Course by Andrew Jones](#) ➡
- [Python for Data Science Cheatsheet by Frank Andrade](#) ➡

Note: The learning path is an introductory and beginner-friendly plan not limited to the above resources; there is a plethora of information on the internet to browse and learn!



Data Analysis and Visualization

- Data Analysis with Python Course by Free Code Camp 
- Data Analysis with Python for Excel Users by Free Code Camp 
- Data Analysis with Python Course - Numpy, Pandas, Data Visualisation by Free Code Camp 
- Replace Excel Pivot Table with Python Pandas by Frank Andrade 
- Python Plotting Tutorial with Matplotlib & Pandas by Keith Galli 

Note: The learning path is an introductory and beginner-friendly plan not limited to the above resources; there is a plethora of information on the internet to browse and learn!



Statistics and Probability

- [Statistics - A Full University Course on Data Science Basics by Free Code Camp](#) ➤
- [Introduction to Statistics by 365 Data Science](#) ➤
- [Introduction to the normal distribution by Khan Academy](#) ➤
- [Confidence Intervals by Khan Academy](#) ➤
- [Hypothesis Testing by Khan Academy](#) ➤
- [Probability by Khan Academy](#) ➤
- [Counting, Permutations, and Combination by Khan Academy](#) ➤
- [Discrete Probability Distributions by Jeremy Balka's Statistics](#) ➤
- [Continuous Probability Distributions by Jeremy Balka's Statistics](#) ➤
- [Probability and Statistics: To p pr not p by University of London](#) ➤
- [Harvard University Lectures](#) ➤
- [MIT OpenCourseWare Lectures](#) ➤

Note: The learning path is an introductory and beginner-friendly plan not limited to the above resources; there is a plethora of information on the internet to browse and learn!



Mathematics

- Linear Algebra - Full College Course by Free Code Camp
- Vectors and Spaces by Khan Academy
- Matrix Transformations by Khan Academy
- Calculus 1 by Khan Academy
- Calculus 2 by Khan Academy
- Introduction to Calculus by the University of Sydney
- 3Blue1Brown
- Mathematics for Machine Learning: Multivariate Calculus by Imperial College London
- Mathematics for Machine Learning and Data Science Specialisation by DeepLearning.AI

Note: The learning path is an introductory and beginner-friendly plan not limited to the above resources; there is a plethora of information on the internet to browse and learn!



Machine Learning and AI

- Machine Learning Course for Beginners by Free Code Camp [↗](#)
- Machine Learning Specialization by Andrew Ng [↗](#)
- Deep Learning Specialization by Andrew Ng [↗](#)
- Deep Learning Lectures by Stanford University [↗](#)
- TensorFlow 2.0 Complete course by Free Code Camp [↗](#)
- Keras with TensorFlow Course by Free Code Camp [↗](#)
- PyTorch for Deep Learning for Free Code Camp [↗](#)
- Natural Language Processing with Python & NLTK by Free Code Camp [↗](#)
- Natural Language Processing with SpaCy & Python by Free Code Camp [↗](#)
- Natural Language Processing with TensorFlow 2 by Free Code Camp [↗](#)
- Natural Language Processing Zero to Hero by TensorFlow [↗](#)
- OpenCV Course by Free Code Camp [↗](#)
- Advanced Computer Vision with Python by Free Code Camp [↗](#)
- TensorFlow for Computer Vision Course by Free Code Camp [↗](#)

Note: The learning path is an introductory and beginner-friendly plan not limited to the above resources; there is a plethora of information on the internet to browse and learn!

Virtual Internships, Competitions, Hackathons, and Practice

- [Awesome YouTubers to follow](#) ➤
- [Forage](#) ➤

[Data Visualisation: Empowering Business with Effective Insights by Tata](#) ➤ | [Data Analytics by Quantum](#) ➤ | [Navigating Numbers by Accenture](#) ➤ | [GE Explorer Series: Digital Technology Data Analytics Program by GE](#) ➤ | [Data Analytics by KPMG](#) ➤ | [Data Science & Analytics by BCG](#) ➤

- [Kaggle](#) ➤
- [Maven Analytics](#) ➤
- [MachineHack](#) ➤
- [DataHack](#) ➤
- [DrivenData](#) ➤
- [AI Planet](#) ➤
- [AI Crowd](#) ➤
- [CodeLabs](#) ➤
- [OpenML](#) ➤
- [UCI Machine Learning Repository](#) ➤
- [Data Flair - Data Science Tutorials](#) ➤
- [Data Flair - Machine Learning Tutorials](#) ➤

Note: The learning path is an introductory and beginner-friendly plan not limited to the above resources; there is a plethora of information on the internet to browse and learn!

Thank you 🙌

[Feedback](#) 🌱

Need any help? 🙋

In case you have any questions on this short course or anything else that I can help you with:

[Book a call](#) 📞

[Message](#) 💬

Support Me ❤️

If you've found value in my work, short courses, or interactions and wanted to say thanks, you can support me or buy me a book to help me create more amazing content for you.

[Support](#) 🤝

Stay in the Loop 🔔

Stay in the Loop about my upcoming webinars and study group sessions without a miss!

[YouTube](#) ✓

[LinkedIn](#) ✓

50% of all proceeds go to feeding stray puppies and dogs!
Be their secret Santa! 🎅

[Woof woof!](#) 🐕❤️

Note: This short course content is based on the learning and personal insights I gained from multiple resources listed in the references section or resources throughout the course.