**Data Science Roadmap 2023-2024 for beginners.**

Before we get to the roadmap, let’s answer the question, What’s data science?

Data science is the science of combining maths, statistics, data analysis, artificial intelligence(AI) & machine learning(ML) with specific subject matter expertise to uncover actionable insights hidden in an organization’s data.

If you would like to learn data science in 6 or so months, here is a roadmap one can follow to learn and master Data Science fundamentals and advanced techniques.

1. Week 1 to 3: Learning a programming language:

You can learn Python or R, the most used programming languages in data science.

If you are interested in learning Python, here is a roadmap you can follow:

* Learn Python's basic syntax, e.g., data structures like variables, integers, and strings.
* Lists, Dictionaries, Sets, and Tuples.
* Functions.
* Advanced features like Object Oriented Programming(OOP), Handling exceptions(try…except…), and Regular Expressions(RegEx).
* Handling files, read and write files in JSON or XML.

1. Week 3 to 4: Learn Pandas, numpy, matplotlib/seaborn.

Pandas library in Python is mainly used for data analysis.

Numpy provides several techniques for data visualization.

Matplotlib is a popular Python library for displaying data and creating static, animated, and interactive plots. One can also learn Seaborn but matplotlib is mostly used.

1. Week 5 to 8: Learn statistics and math.

In Maths and statistics, one should learn:

* Linear algebra: Scalar, vector matrices, and their operations.
* Calculus: Differentiation and its rules, partial differentiation, and integration and its rules.
* Probability: Rules, Dependent and independent events, Mutually exclusive events, etc.
* Statistics(stats): Sampling techniques, testing data, regression modeling, etc.

1. Week 9 to 12: Big Data and Tools:

Big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just can't manage them.

Big data tools and techniques include:

* Apache Hadoop and Apache Spark
* Apache Spark framework.
* Spark RDD’s
* Dataframes & Spark SQL
* Machine learning using Spark MLLib
* Understanding Apache Kafka & Apache Flume
* Apache Spark streaming

1. Machine Learning (ML) and Artificial Intelligence(AI:

Machine learning is a branch of artificial intelligence (AI) and computer science that focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Artificial intelligence (AI) is the ability of a computer or a robot controlled by a computer to do tasks that are usually done by humans because they require human intelligence and discernment.

In Machine Learning(ML), you can learn how ML works and build ML models like

* Supervised learning models
* Unsupervised learning models
* Reinforcement learning e.g. dimensionality reduction, time series analysis, model selection & boosting.

1. Week 13 to 15: Machine Learning Model projects with Deployment:

After learning ML and AI, you can finish 2 end to end ML model projects from Kaggle on (i) Regression and (ii) Classification and deploy them using web frameworks e.g. Django, Flask, or FastAPI.

1. Week 16 to 17: Learn SQL or any other database management system.

SQL stands for Structured Query Language and is used to communicate with a database. It is the standard language for relational database management systems.

1. Week 18 to 20: Learn and interact with BI tools.

Business Intelligence (BI) tools e.g. PowerBI and Tableau are all about helping you understand trends and derive insights from your data so that you can make tactical and strategic business decisions which also help you identify patterns.

1. Week 21 to 25: Learn Deep Learning:

In Deep learning, on should learn and understand the fundamentals of Deep Learning for example:

* Tensorflow 2.0
* Single Layer Perceptron
* Neural Networks
* Conventional Neural Networks(CNN)
* Regional CNN
* Emotion and Gender detection among others.

1. Week 26 onwards:

(i). Focus on building more projects and uploading them to your Github account

(ii). Build an online brand via Kaggle, X(formerly Twitter) and LinkedIn.

(iii). Join and participate in online forums and communities on X or LinkedIn to form meaningful connections in the Data Science field and learn more from other data scientists, engineers, and analysts.

(iv). Have fun in the learning process too!

I wish you success in the journey of learning and getting to being a world-class data scientist, engineer, or analyst!