**Data Science for Beginners: 2023 - 2024 Complete Roadmap.**

**What is data science?**

Data science is a field that involves extracting insights and information from data using various tools and techniques. These insights can be used by organizations in decision-making and strategic planning. Data science combines math and statistics, specialized programming, sophisticated analytics, artificial intelligence (AI), machine learning, and unique subject matter expertise with domain expertise to reveal useful insights hidden in the data of an organization.

**Why data science?**

Data science is a rapidly growing field with a lot of untapped potential. Over the past few years, there has been a high demand for data scientists. To improve productivity, the organization needs to analyze their data. The need for data science has become very crucial in the current due to the enormous amount of data that is generated by businesses and organizations. It provides sophisticated tools to extract meaningful insights from this data and aid in making informed decisions. Data science plays a very crucial role in addressing the world's most pressing issues such as climate change, climate change, finance, etc.

**Data science roadmap**

A data science roadmap can be defined as a representation of a strategic plan to help aspiring data scientists learn and succeed. If you are a beginner in data science, these are the steps you can follow to begin;

**Learn the basics:** There are various fundamental data science ideas that you have to be familiar with. Understand what data is and why they are valuable.

**Learn programming:** Programming is a fundamental skill in data science. R and Python are the two most used programming languages in data science. Python has several data science libraries such as NumPy, Pandas, and Scikit Learn. Anaconda is one of the IDEs that can be used for Python and R studio IDE can be used for R programming. SQL is also a crucial tool used by data analysts to retrieve data and manipulate databases.

**Learn mathematics and statistics:** Maths and statistics are integral skills of data science and machine learning. Some of the mathematics and statistics skills required are descriptive statistics, inferential statistics, Linear algebra, etc.

**Learn Data Visualization:** Visualization is an essential skill for a data scientist. Visualization helps in understanding data and communicating insights. Data visualization libraries such as Matplotlib and Seaborn can be used in Python and ggplot2 can be used in R. Tableau and Power BI are also important tools that can be used in data visualization.

**Data wrangling:** This involves cleaning and pre-processing data. Some of the activities at this stage are handling missing values, and outliers and transforming data into a suitable format for data analysis.

**Machine learning:** This is one of the primary skills that a data scientist needs to have. As a beginner, you should start with simple machine learning techniques such as supervised machine learning (Linear and logistic regression).

**Handle projects:** One of the best ways to learn as a data scientist is handling practice projects. You should gradually start from simple projects as you advance to more complex projects.

**Networking:** Join data science forums, and social media groups, and attend local meetups or conferences. Networking with others in the field can provide valuable insights and opportunities.