Task: Explore in Google and Make a document:

1. What is Data Science?

Data science is a deep study of the massive amount of data, which involves extracting meaningful insights from raw, structured, and unstructured data that is processed using the scientific method, different technologies, and algorithms. It is a multidisciplinary field that uses tools and techniques to manipulate the data so that you can find something new and meaningful

2. Importance of statistics in Data Science?

In data science, statistics is at the core of sophisticated machine learning algorithms, capturing and translating data patterns into actionable evidence.

* Present key findings revealed by a dataset.
* It summarizes information.
* It calculates the measurement of data by mathematical methods.
* Make future predictions based on previously recorded data.
* It also tests experimental predictions.

3. What is Data engineering?

Data engineering is the process of designing and building systems that let people collect and analyze raw data from multiple sources and formats. These systems empower people to find practical applications of the data, which businesses can use to thrive.

4. What is Data visualization?

Data visualization is the graphical representation of information and data. By using v[isual elements like charts, graphs, and maps](https://www.tableau.com/data-insights/reference-library/visual-analytics), data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. Additionally, it provides an excellent way for employees or business owners to present data to non-technical audiences without confusion. In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

5. What is Data cleaning?

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. When combining multiple data sources, there are many opportunities for data to be duplicated or mislabeled. If data is incorrect, outcomes and algorithms are unreliable, even though they may look correct.

6. What is python and why we use it?

Python is a versatile, general-purpose programming language. It can be used in a variety of fields from data science and machine learning to web development and is a great first language to learn. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems. Python is commonly used for developing websites and software, task automation, data analysis, and data visualization.

7. What is Data in statistics?

Data are the facts and figures that are collected, analyzed, and summarized for presentation and interpretation.

8. What is statistics?

Statistics is the study of the collection, analysis, interpretation, presentation, and organization of data.

9. What is mathematics?

Mathematics is commonly defined as the study of patterns of structure, change, and space

10. What is computer Science?

Computer Science is the study of computers and computational systems. The discipline of computer science includes the study of algorithms and data structures, computer and network design, modeling data and information processes.

11. What is programming language?

A programming language is an artificial language that can be used to control the behavior of a machine, particularly a computer.