Assingment 1.

1.What is data science.

Ans: Data science is the study of data to extract meaningful insights for business. It is a multidisciplinary approach that combines principles and practices from the fields of mathematics, statistics, artificial intelligence, and computer engineering to analyze large amounts of data.

2.Importance of statistics in Data Science?

Ans: Statistics, in its broadest sense, refers to a collection of tools and methods for evaluating, interpreting, displaying, and making decisions based on data. Some individuals refer to statistics as the mathematical analysis of technical data.

3. What is Data engineering?

Ans: Data engineering is the process of designing and building systems that let people collect and analyz 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?

Ans: Data visualization is concerned with visually presenting sets of primarily quantitative raw data in a schematic form. The visual formats used in data visualization include tables, charts and graphs (e.g. pie charts, bar charts, line charts, area charts, cone charts, pyramid charts, donut charts, histograms, spectrograms, cohort charts, waterfall charts, funnel charts, bullet graphs, etc.)

5. What is Data cleaning?

Ans: Data cleansing or data cleaning is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data.

6.What is phyton and why we use it.

Ans: Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. 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.

7. What is Data in statistics?

Ans: Data are measurements or observations that are collected as a source of information. There are a variety of different types of data, and different ways to represent data.

8. What is statistcs ?

Ans: Statistics is the discipline that concerns the collection, organization, analysis, interpretation, and presentation of data. In applying statistics to a scientific, industrial, or social problem, it is conventional to begin with a statistical population or a statistical model to be studied.

9.What is mathematics?

Ans: Mathematics is an area of knowledge that includes the topics of numbers, formulas and related structures, shapes and the spaces in which they are contained, and quantities and their changes.

10. What is computer Science?

Ans: Computer Science is the study of computers and computational systems. Unlike electrical and computer engineers, computer scientists deal mostly with software and software systems; this includes their theory, design, development, and application.

11. What is programming language?

Ans: A programming language is a system of notation for writing computer programs. Most programming languages are text-based formal languages, but they may also be graphical. They are a kind of computer language.