1. **What is Data Science?**

Ans: **Data science** is an [interdisciplinary](https://en.wikipedia.org/wiki/Interdisciplinary) academic field [[1]](https://en.wikipedia.org/wiki/Data_science#cite_note-1) that uses [statistics](https://en.wikipedia.org/wiki/Statistics), [scientific computing](https://en.wikipedia.org/wiki/Scientific_computing), [scientific methods](https://en.wikipedia.org/wiki/Scientific_method), processes, [algorithms](https://en.wikipedia.org/wiki/Algorithm) and systems to extract or extrapolate [knowledge](https://en.wikipedia.org/wiki/Knowledge) and insights from noisy, structured, and [unstructured data](https://en.wikipedia.org/wiki/Unstructured_data).[[2]](https://en.wikipedia.org/wiki/Data_science#cite_note-2)

Data science also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine).[[3]](https://en.wikipedia.org/wiki/Data_science#cite_note-3) Data science is multifaceted and can be described as a science, a researchS paradigm, a research method, a discipline, a workflow, and a profession.[[4]](https://en.wikipedia.org/wiki/Data_science#cite_note-4)

Data science is a "concept to unify [statistics](https://en.wikipedia.org/wiki/Statistics), [data analysis](https://en.wikipedia.org/wiki/Data_analysis), [informatics](https://en.wikipedia.org/wiki/Informatics), and their related [methods](https://en.wikipedia.org/wiki/Scientific_method)" in order to "understand and analyse actual [phenomena](https://en.wikipedia.org/wiki/Phenomena)" with [data](https://en.wikipedia.org/wiki/Data).[[5]](https://en.wikipedia.org/wiki/Data_science#cite_note-5) It uses techniques and theories drawn from many fields within the context of [mathematics](https://en.wikipedia.org/wiki/Mathematics), statistics, [computer science](https://en.wikipedia.org/wiki/Computer_science), [information science](https://en.wikipedia.org/wiki/Information_science), and [domain knowledge](https://en.wikipedia.org/wiki/Domain_knowledge).[[6]](https://en.wikipedia.org/wiki/Data_science#cite_note-:2-6) However, data science is different from [computer science](https://en.wikipedia.org/wiki/Computer_science) and information science. [Turing Award](https://en.wikipedia.org/wiki/Turing_Award) winner [Jim Gray](https://en.wikipedia.org/wiki/Jim_Gray_(computer_scientist)) imagined data science as a "fourth paradigm" of science ([empirical](https://en.wikipedia.org/wiki/Empirical_research), [theoretical](https://en.wikipedia.org/wiki/Basic_research), [computational](https://en.wikipedia.org/wiki/Computational_science), and now data-driven) and asserted that "everything about science is changing because of the impact of [information technology](https://en.wikipedia.org/wiki/Information_technology)" and the [data deluge](https://en.wikipedia.org/wiki/Information_explosion).[[7]](https://en.wikipedia.org/wiki/Data_science#cite_note-TansleyTolle2009-7)[[8]](https://en.wikipedia.org/wiki/Data_science#cite_note-BellHey2009-8)

**2) Importance of statistics in Data Science?**

+ Data Visualization:

Data visualization helps to tell stories by curating data into a form easier to understand, highlighting the trends and outliers.

+ It enables you to Analyze Data efficiently

+it enables you to reach better and more accurate conclusions.

**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 helps to tell stories by curating data into a form easier to understand, highlighting the trends and outliers.

**5) what is data cleaning?**

**Ans)** Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset

**6) what is python 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 computer science?**

Ans) Computer Science is the study of computers and computational systems