

PDEA's
BABURAOJI GHOLAP COLLEGE SANGVI, Pune 27
Department of Computer Science
Class: T.Y.B.Sc (Computer Science)

Assignment -1

Submission date: 16-08-2025

Chapter 1: Introduction to Data Science

Course Objective (CO-1): Describe Exploratory and statistical Data Analysis with a variety of statistical methods and models [Marks: 5]

Q.1] Short Answer Questions:

1. Define data science.
2. What is role of data science?
3. Define unstructured Data.
4. Define data source?
5. What is data set?
6. What is CSV format?
7. What are uses of Zip files?
8. What is open data?
9. What is meant by semi-structured data?
10. What is rasterized and vectorized image files?
11. What is SAS?
12. List applications of data science.
13. What is social media data?
14. List tools for data scientist.
15. What is data science life cycle?
16. Give the use of tar files.
17. What is compressed data?
18. What is data source?

Q.2] Long Answer Questions:

1. What is data science? Why the learning of data science is important? Explain in details.
2. Explain 3V's of data science with diagram.
3. What are components of data scientist's toolbox? Explain two of them.
4. What are applications of data science?
5. With the help of diagram describe lifecycle of data science.
6. What are different types of data? Explain in detail with appropriate examples.
7. Distinguish between structured and unstructured data?
8. What are different sources of data in data science? Explain in detail.
9. Explain different data formats in brief.
10. Explain unstructured data with example. What are problems with unstructured data?

Dr. Seema Cowhan
Lecture In-charge

Dr. Seema Cowhan
HOD(CS)
H.O.D.

Department of Computer Science
Baburaoji Gholap College
Sangvi, Pune-27