SAMPLE PAPER 1

End Term Examination

Data Analysis Using R and Python

Time: 3 Hours Maximum Marks: 70

SECTION A (Compulsory)

7 × 2 = 14 Marks

1. What is data normalization?
2. Write any two file formats supported by Pandas.
3. Define a density plot.
4. What is the use of `cbind()` in R?
5. Define a Python dictionary.
6. Give an example of a categorical variable.
7. What is the purpose of `summary()` in R?

SECTION B

Attempt any four questions. Each carries 14 marks.

1. Explain different types of data (categorical, numerical, ordinal) with examples. Discuss their importance in analysis.
2. Describe data import functions in R and Python. Compare their syntax and use cases.
3. Explain the various geoms used in `ggplot2` for data visualization.
4. Discuss data cleaning techniques in Pandas and demonstrate with code.
5. Explain Python control structures (if, loops) with examples.
6. Discuss the role of data frames and matrices in R with detailed examples.
7. Explain the use of groupby operations in Pandas with suitable examples.
8. Describe NumPy's broadcasting rules and show how they help in vectorized computation.