

---

**SAMPLE PAPER 4**  
**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 a CSV file?
2. Define a factor variable in R.
3. What is a facet in `ggplot2`?
4. Define Python tuple.
5. Mention any one NumPy mathematical function.
6. What is data preprocessing?
7. Define correlation.

**SECTION B**

**Attempt any four questions.**

1. Explain data collection, import, and initial exploration methods in R and Python.
2. Describe different R data structures with appropriate code snippets.
3. Discuss scatter plots, heatmaps, and line plots in depth.
4. Explain Pandas missing value handling methods with examples.
5. Write an essay on Python lists with operations and applications.
6. Discuss features of `apply()`, `lapply()`, `sapply()` in R.
7. Explain grouping and aggregation of data in detail.
8. Describe NumPy's role in machine learning and data preparation.