

25U2124

(Pages: 2)

Name :

Reg. No :

SECOND SEMESTER B.Voc. DEGREE EXAMINATION, APRIL 2026

(FYUGP)

CC25INT2CJ105 - DATA SCIENCE WITH PYTHON

(B.Voc. IT - Major Course)

(2025 Admission - Regular)

Time: 2.0 Hours

Maximum: 60 Marks

Credit: 4

Section A

Answer *all* questions. Each question carries 2 marks.

1. What is the purpose of the `groupby()` function in Pandas? [Level:1] [CO2]
2. Write one use of Jupyter Notebook. [Level:1] [CO1]
3. What does a scatter plot represent? [Level:1] [CO3]
4. Write the purpose of the `fillna()` method. [Level:1] [CO2]
5. Write the purpose of the `sort_values()` function. [Level:1] [CO2]
6. What is meant by data visualization? [Level:1] [CO3]
7. Which Pandas method is used to plot time series data? [Level:1] [CO3]
8. What is the purpose of a histogram in data visualization? [Level:1] [CO4]
9. Write one use of the `info()` method in Pandas. [Level:1] [CO4]
10. Mention one difference between categorical and ordinal features. [Level:1] [CO4]
11. Define normalization in the context of data preprocessing. [Level:1] [CO4]
12. Which function is used to write data to a CSV file? [Level:1] [CO1]

(Ceiling: 20 Marks)

Section B

Answer *all* questions. Each question carries 6 marks.

13. Compare filtering rows in a DataFrame using conditions with selecting specific columns. [Level:2] [CO1]
14. Explain why the Data Science lifecycle is important for solving real-world problems. [Level:2] [CO1]

15. Discuss how parsing dates makes time-based data easier to analyze. [Level:2] [CO2]
16. Describe the difference between a boxplot and a violinplot in terms of data representation. [Level:2] [CO3]
17. Explain how apply() and map() differ in their usage with Series and DataFrames. [Level:2] [CO3]
18. Explain the difference between Pearson correlation and Spearman rank correlation with examples. [Level:2] [CO4]

(Ceiling: 30 Marks)

Section C

Answer any *one* question. The question carries 10 marks.

19. Explain how Pandas can be used to combine datasets. Show how merge(), join(), and concat() are applied, and describe how handling keys and indexes ensures accurate results. Discuss use-cases where each method is most effective. [Level:3] [CO2]
20. Explain how you can use NumPy to analyze a dataset. Show how arrays can be created, indexed, sliced, reshaped, and how vectorized operations, broadcasting, and statistical functions make the analysis easier. [Level:3] [CO1]

(1 × 10 = 10 Marks)
