

**STELLA MARIS COLLEGE (AUTONOMOUS), CHENNAI – 600 086**  
**(For Candidates admitted during the academic year 2023-2024)**

**P.G.D.C.S. EXAMINATION – NOVEMBER 2023**  
**FIRST SEMESTER**

**PAPER : DATA ANALYTICS**

**SUBJECT CODE: 23CS/DC/DA14**

**TIME : 3 HOURS**

**MAX. MARKS : 100**

**SECTION – A**

**ANSWER ALL THE QUESTIONS:**

**(10x2=20)**

1. How to read data from csv file?
2. Write a short note on initial data analysis.
3. What is data transformation?
4. Give an example for data aggregation.
5. What is a time series data? Give an example.
6. Write a short note on converting a date type to a string type in pandas.
7. What is dimensionality reduction?
8. Compare classification and regression.
9. Write a short note on sampling data in a stream.
10. Explain bloom filter.

**SECTION – B**

**ANSWER ANY SIX OF THE FOLLOWING:**

**(6x5=30)**

11. Explain the process of data analysis.
12. Compare Series and DataFrame.
13. Explain handling missing data with example codes.
14. Write short notes on pivot table and cross tabulation.
15. Give any two data analysis examples and explain.
16. Discuss how Machine Learning helps to overcome the challenges in hard coded programs.
17. Compare supervised and unsupervised learning with examples.
18. Explain the three important V's of big data.

**SECTION – C**

**ANSWER ANY FIVE OF THE FOLLOWING:**

**(5x10=50)**

19. Explain the different types of data in detail.
20. What are the different types of analytics? Explain with examples.
21. Explain in detail on combining and merging data sets.
22. What are group operations? Explain in detail with examples.
23. Explain any four different plotting with pandas and seaborn.
24. Explain the following: a) k-nearest Neighbor b) k-means clustering.
25. Discuss on mapreduce and Hadoop in detail with an example.

\*\*\*\*\*

