THEORY QUESTIONS ASSIGNMENT

Data Science Stream

Maximum score: 100

KEY NOTES

- This assignment to be completed at student's own pace and submitted before given deadline.

- There are 10 questions in total and each question is marked on a scale 1 to 10. The maximum possible grade for this assignment is 100 points.

- Students are welcome to use any online or written resources to answer these questions.

- The answers need to be explained clearly and illustrated with relevant examples where necessary. Your examples can include code snippets, diagrams or any other evidence-based representation of your answer.

Theory questions	10 point each

1. What does "Data Cleansing" mean? What are the best ways to practice this?

2. What is the difference between data profiling and data mining?

3. Define Outlier with an example.

4. What is "Collaborative Filtering"?

5. What is "Time Series Analysis"?

6. Explain the core steps of a Data Analysis project?

7. What are the characteristics of a good data model?

8. Explain and provide examples of univariate, bivariate, and multivariate analysis?

9. What is a Linear Regression?

10. In terms of modelling data, what do we mean by Over-fitting and Under-fitting?