# K. K. Wagh Institute of Engineering Education & Research, Nashik

(An Autonomous Institute From A.Y. 2022-23)

| WINTER-2025 | |
|---|---|
| Exam Seat No.: | |
| Academic Year:2025-2026 | Semester:III |
| Class:SY | Program:B.Tech |
| Branch Code:ETC | Pattern:2023 |
| Name of Course:Data Analytics with Python | Course Code:2302206(B) |
| Max. Marks:60 | Duration:2.30 Hrs. |

**Instructions:** Candidates should read carefully the instructions printed on the Question Paper and on the cover page of the Answer Book, which is provided for their use.

1. This question paper contains _02____page(s).
2. Answer to each new question is to be started on a new page.
3. Assume suitable data wherever required, but justify it.
4. Draw the neat labelled diagrams, wherever necessary.
5. The last columns indicates the Course Outcome and level of Blooms Taxonomy of the Question/sub-question.

| | | Marks | CO |
|---|---|---|---|
| **Question No. 1** | | | |
| 1a) | What is skewness? Differentiate between positive and negative skewness with diagrams. | (6) | 1 |
| **Question No. 2** | | | |
| 2a) | Define variance and standard deviation. Given the data: 5, 7, 9, 10, 9, 6, calculate the variance and standard deviation. Interpret what the standard deviation indicates about the data spread. | (6) | 2 |
| **Question No. 3** | | | |
| 3a) | Explain the concept of one-way and two-way ANOVA with suitable examples. | (8) | 3 |
| | **OR** | | |
| 3b) | Define null and alternative hypotheses. Explain the significance of the p-value and confidence level | (8) | 3 |
| 3c) | Write short notes on the following hypothesis tests and also give examples where they can be used<br>1. t-test (Two-sample test),<br>2. ANOVA (Analysis of Variance) | (8) | 3 |
| | **OR** | | |
| 3d) | What are the possible decision errors in hypothesis testing? Explain with examples. | (8) | 3 |
| **Question No. 4** | | | |
| 4a) | Discuss the construction and interpretation of the Confusion Matrix and ROC curve. | (8) | 4 |
| | **OR** | | |

| | | | |
|---|---|---|---|
| 4b) | A hospital's AI-based diagnostic tool was tested for detecting a specific disease. The following results were obtained:<br><br>True Positives (TP) = 150, True Negatives (TN) = 820 , False Positives (FP) = 50, False Negatives (FN) = 30<br><br>Calculate (a) Accuracy, (b) Precision, (c) Recall (Sensitivity),(d) F1-Score | (8) | 4 |
| 4c) | Describe the performance evaluation metrics : Accuracy, Precision, Recall, F1-Score along with its formulas. | (8) | 4 |
| | **OR** | | |
| 4d) | Explain the concept of Logistic Regression. How is it different from Linear Regression? | (8) | 4 |
| **Question No. 5** | | | |
| 5a) | Explain the Chi-Square Test of Independence with an example. How is it used to determine the association between variables? | (8) | 5 |
| | **OR** | | |
| 5b) | Differentiate between K-Means and Hierarchical Clustering methods. | (8) | 5 |
| 5c) | Write a short note on Classification and Regression Trees. Explain its advantages and limitations. Also Mention which python instruction is used to implement CART? | (8) | 5 |
| | **OR** | | |
| 5d) | Elaborate concept of regression analysis. Explain in detail linear regression and multiple linear regression. Which python instruction is used to implement Linear regression model. | (8) | 5 |

**····· End of question paper·····**