

Reg. No.

--	--	--	--	--	--	--	--	--	--



**CSH 202**

**Second Semester M.Sc. Examination, September/October 2022**  
**COMPUTER SCIENCE**  
**Data Science with Python**

Time : 3 Hours

Max. Marks : 70

**Note :** Answer **any five full** questions. **All** question carries **equal** marks.

1. a) What is Data Science and what are the key factors that led to Data Science ? **6**  
b) Explain the descriptive statistics measures that are used to study the characteristics of the dataset. **8**
  2. a) Describe the CRISP-DM framework for the Data Science process. **9**  
b) Explain the different types of real world data with examples. **5**
  3. a) What is data visualization ? Explain any two techniques for univariate and multivariate visualization. **7**  
b) What is the need for feature selection ? Explain the different types of feature selection. **7**
  4. a) What is supervised learning ? Explain the steps involved in supervised learning. **7**  
b) Why kNN method is called as lazy learner ? Explain the kNN method. **7**
  5. a) What is Clustering ? Explain the different types of Clustering. **8**  
b) Explain two applications of Clustering and Classification. **6**
  6. a) Explain the metrics used for evaluating the performance of classifiers. **7**  
b) What is text mining ? Explain the steps involved in processing text data. **7**
  7. a) What is Anomaly detection ? Explain the techniques of Anomaly detection. **8**  
b) What is recommendation Engine ? Explain any two applications of recommendation engine. **6**
  8. a) What is time series data ? Explain the taxonomy of time series forecasting. **7**  
b) What is deep learning ? Explain the functionality of deep learning. **7**
-