|  |  |
| --- | --- |
| **SESSION** | **March 2023** |
| **PROGRAM** | **MASTER OF BUSINESS ADMINISTRATION (mba)** |
| **SEMESTER** | **III** |
| **course CODE & NAME** | **DADS302 EXPLORATORY DATA ANALYSIS** |
| **CREDITS** | **4** |

**Assignment Set – 1**

**1. What is Data Science? Discuss the role of Data Science in various domains.**

**ANS: Data Science** is an interdisciplinary field that combines scientific methods, algorithms, and systems to extract knowledge and insights from structured and unstructured data. It involves utilizing techniques from mathematics, statistics, computer science, and domain knowledge to uncover patterns, make predictions, and drive informed decision-making. The role of Data Science has become increasingly significant in various domains due to the

Its Half solved only

Buy Complete from our online store

<https://smuassignment.in/online-store/>

MUJ Fully solved assignment available for**session March 2023.**

Lowest price guarantee with quality.

Charges**INR 200 only per assignment.**For more information you can get via mail or Whats app also

Mail id is aapkieducation@gmail.com

Our website www.smuassignment.in

After mail, we will reply you instant or maximum

1 hour.

Otherwise you can also contact on our

whatsapp no 8791490301.

**2. Explain various measures of dispersion in detail using specific examples.**

**ANS:** Measures of dispersion, also known as measures of variability or spread, quantify the extent to which data points deviate from the central tendency and provide insights into the spread or distribution of values within a dataset.

**Here are some commonly used measures of dispersion along with specific examples:**

**3. Discuss various techniques used for Data Visualization.**

**ANS: Data visualization** is an essential component of data analysis and communication. It involves the creation of visual representations, such as charts, graphs, and maps, to effectively convey insights and patterns hidden within data.

**Here are various techniques commonly used for data visualization:**

**Bar Charts:** Bar charts display categorical data as rectangular bars with lengths proportional

**Assignment Set – 2**

**1. What is feature selection? Discuss any two feature selection techniques used to get optimal feature combinations.**

**ANS: Feature selection** is the process of selecting a subset of relevant features (variables, attributes) from a larger set of available features. The goal of feature selection is to identify the most informative and discriminative features that contribute the most to the predictive performance of a machine learning model. It helps in reducing the dimensionality of the data, improving model interpretability, reducing computational complexity, and avoiding over

**2. Discuss in detail the concept of Factor Analysis.**

**ANS: Factor analysis** is a statistical technique used to explore and uncover underlying latent variables, known as factors, from a set of observed variables. It aims to identify the common sources of variation among observed variables and reduce them to a smaller number of unobservable factors.

**3. Differentiate between Principal Component Analysis and and Linear Discriminant Analysis.**

**ANS: Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA)** are both dimensionality reduction techniques commonly used in machine learning and data analysis. While they serve a similar purpose of reducing the dimensionality of data, they have different objectives and are applied in different contexts.

**Here are the key differences between PCA and LDA:**