|  |  |
| --- | --- |
| **SESSION** | **AUG-SEP’23** |
| **PROGRAM** | **MASTER OF BUSINESS ADMINISTRATION (MBA)** |
| **SEMESTER** | **III** |
| **COURSE CODE & NAME** | **DADS301- PROGRAMMING IN DATA SCIENCE** |
|  |  |
|  |  |

**Assignment Set – 1**

**1. a. What is Data wrangling? Name the package used for Data wrangling in R and describesome of its features.**

**b. What are vectors? Explain the creation of vectors with examples. Also, describe how to identify and handle missing values. 5+5**

**Ans 1.**

**a. Data Wrangling in Data Science**

Data wrangling, often referred to as data munging, is a crucial step in the data science process. It involves transforming and mapping data from one "raw" form into another format with the intent of making it more appropriate and valuable for a variety of downstream purposes such as analytics. This process includes cleaning, structuring, and enriching raw data into a desired format for better decision making in less time. Data wrangling is essential because it ensures the accuracy and consistency of data, which is critical for obtaining meaningful insights.

In the context Its Half solved only

Buy Complete from our online store

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

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

Lowest price guarantee with quality.

Charges**INR 198 only per assignment.**For more information you can get via mail or Whats app also

Mail id is [aapkieducation@gmail.com](mailto: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. a. Describe the steps to initialize a plot, specify aesthetics, create a simple plot, and add titles and labels to the plot.**

**b. Explain the chaining operator with an example. 5+5**

**Ans 2.**

**a. Plot Initialization and Aesthetics in Data Science**

The process of visualizing data in programming, particularly in data science, involves several key steps. Firstly, initializing a plot is the foundational step. It involves setting up a plotting environment or framework within a programming language, such as Python using libraries like Matplotlib or Seaborn. This initialization usually involves commands that create a figure object or a plotting area in the chosen library.

Once the plotting area is initia

**3.a. Describe with example the method to visualize the covariance between two categorical features.**

**b. Explain the ‘Simple If’ construct and the ‘VectorisedIf’ with example. Elaborate the advantages of Vectorised If. 5+5**

**Ans 3.**

**a. Covariance Visualization Between Two Categorical Features**

Covariance is a measure that indicates the extent to which two variables change together. However, when it comes to categorical features, traditional covariance, which is typically used for numerical data, isn't applicable. Instead, we use methods like contingency tables or mosaic plots to understand the relationship between two categorical variables.

Example of

**Assignment Set – 2**

**4. a. Describe the need for Python and applications of Python.**

**b. Explain with examples – Set/List and Tuples. What are the similarities and differences between them? 5+5**

**Ans 4.**

**a. The Need for Python and Its Applications**

Python, a high-level programming language, has become integral in the field of data science due to its simplicity, flexibility, and robust ecosystem. One of the primary reasons for Python's widespread adoption is its readability and ease of learning, making it accessible to both beginners and experienced programmers. The syntax of Python is clean and expressive, which enables developers to write less code for the same task compared to other languages like Java or

**5.a. Explain with examples functions and methods in Python. Differentiate between the two.**

**b. Explain with example the difference between the re.match() and re.search() . 5+5**

**Ans 5.**

**a. Functions and Methods in Python: Definitions and Differences**

In Python, **functions** and **methods** are both blocks of code that perform a specific task, but they have distinct characteristics and uses.

Functions in Python

A **function** in Python is a self-contained block of code that encapsulates a specific task or computation. It can take input, process it, and return an output. Functions are defined using the **def** keyword followed by the function name and parentheses which can enclose parameters. An example is the built-in **len()** function which returns the length of an object. You can also define custom functions, like:

def add\_

**6.a. How are donut charts created in Python. Explain with an example.**

**b. Explain with an example the importance of vectorization and the advantages it offers with an example.**

**Ans 6.**

**a. Creating Donut Charts in Python**

Donut charts, a variant of pie charts, are an effective visual tool for displaying proportional data. In Python, creating donut charts is facilitated primarily by the **matplotlib** library, which offers comprehensive tools for data visualization.

To illustrate, consider the example of visualizing the market share of different smartphone brands. First, import the necessary modules: **matplotlib.pyplot** for plotting and **numpy** for numerical operations. Next, define the data: let's say Brand A holds 40%, Brand B 30%, Brand C 20%,