**EDA with Python**

3 trials. Best of 3. 10 marks per question. Start date: 22 Nov. End date: 26 Nov

Quiz Questions based on the pre-read links:

1. What is the purpose of the following line of code?

**dataframe.isnull()**

1. Display relationships between variables in the dataset
2. Display missing values in the dataset
3. Display only non-missing values in the dataset
4. Display the count of missing values in the dataset

Solution: b. Display missing values in the dataset

1. The **.info()** command helps us to gather information about:

(i) number of rows, columns and missing values in the dataset

(ii) the data type of all columns in the dataset

(iii) the relationship between numeric values in the dataset

1. (i), (ii) and (iii)
2. (i) and (iii)
3. (ii) and (iii)
4. (i) and (ii)

Solution: d. (i) and (ii)

1. A scatter plot gives us insights about the
2. Correlation between variables in a dataset
3. Statistical information about a specific column
4. Row-wise aggregates based on columns
5. Column-wise aggregates based on rows

Solution: a. Correlation between variables in a dataset

Discussion statements:

* Organizations should conduct exploratory data analysis as it enables analysts to examine the data before making any conclusions. It guarantees that the findings are reliable and pertinent to the objectives and consequences of business
* Companies must take a thorough, analytical look at the data collection before they can draw any firm conclusions or make any assumptions.

Data Science, engineering, and other research and development-related professions can all benefit from EDA. With modern analytics technology supported with access to advanced computing power, using EDA can be an exciting and interesting way to discover unanticipated value in a large number of complex data sets.