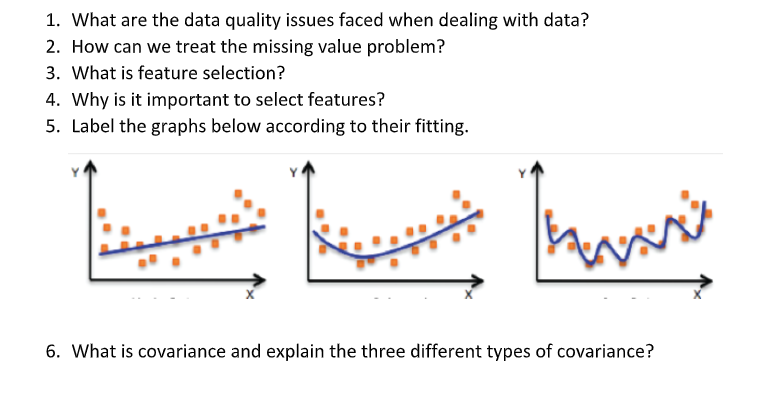
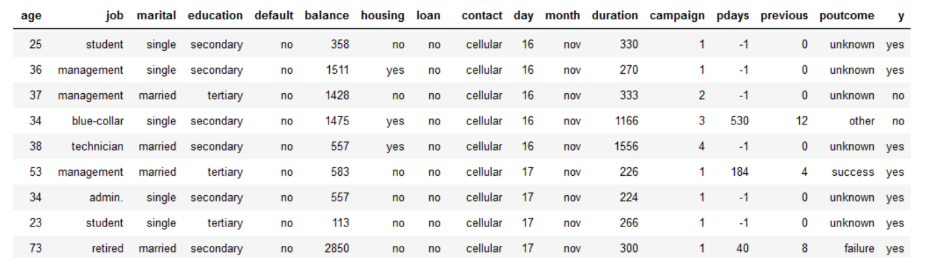
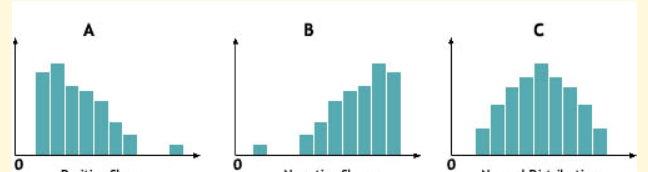
**Quiz 01 – 30 minutes - (12 points)**

Name: Date: 28-02-2021



1. 1) Under fitting 2) Appropriate fitting 3) Over fitting
2. What are the different type of data attributes?
3. Numeric, categorical, Datetime, Boolean, object, int, float
4. In the dataset given below identify each type of the attribute
5. Age=nominal, job=nominal, marital=nominal, education=ordinal, default=Boolean, balance=numeric, housing=Boolean, loan=Boolean, contact=nominal, day=ordinal, month=ordinal, duration=numeric, campaign=nominal, pdays=numeric, previous=numeric, poutcome=nominal, y=boolean
6. What are the three basic measures for the central tendency?
7. 1) Mean 2) Median 3) Mode
8. What does the measure of dispersion tell us about the data?
9. It is a measure of spread of data about the mean.
10. Determine the type of distribution of the given data variables.

Answer. A) left skewed B) right skewed C) uniform distribution

1. Why do we need to cater to missing data?
2. We need to cater missing data because it is part of the sample space.

As it affects the prediction.

1. Why do we need to normalize data?
2. We need to normalize data in order to increase the accuracy of machine learning.
3. What is the purpose of discretization?

A. The purpose of discretization is to reduce the range or variety of values in the dataset and assign them in the form of category or group.

1. What methods can be used for encoding of discretization of features?

Answer. 1) Manual encoding 2) Label encoding 3) One hot encoding