**A z-score of 1 means that \_\_\_\_\_\_.**

(1 mark)

Top of Form

the observation is 1.0 standard deviation to the right of the mean

None of the options.

the observation has no deviation from the mean

the observation is -1.0 standard deviation to the right of the mean

the observation is -1.0 standard deviation to the left of the mean

Bottom of Form

**Which of the following formulas computes variance of a sample?**(1 mark)

Top of Form

Text

Description automatically generated

Text

Description automatically generated

Text

Description automatically generated

None of the options.

Diagram

Description automatically generated with low confidence

Bottom of Form

**What is the variance of the following dataset: 10, 10, 10, 10, 10, 10, 10, 10, 10**

(1 mark)

Top of Form

0.1

100

0

10

10000

Bottom of Form

**4.**

**Which of the following statements about correlation is false?**

(1 mark)

~~Top of Form~~

~~Correlation between variable X and variable Y is equal to the correlation between variable Y and variable X.~~

~~Chart, scatter chart

Description automatically generated~~

~~From this graph, we can infer that X and Y does not have a strong correlation.~~

~~If the covariance is positive, the correlation must be positive as well.~~

A correlation value of -0.5 suggests a negative linear relation between the two variables.

A correlation value of 0.8 between X and Y tells that an increase in X will lead to an increase in Y.

**The \_\_\_\_\_\_\_\_ is the square root of the variance.**(1 mark)

Top of Form

dispersion

Interquartile range (IQR)

average

standard deviation

None of the options.

Bottom of Form

**The mean can be affected by outliers. What are outliers?**(1 mark)

Top of Form

None of the options.

Observations that are radically different from the rest—which pull the value of the mean toward these values.

Observations that should not be included in the data set.

Observations that have the exact value as the mean.

Observations that are radically similar to the rest—which pull the value of the mean toward these values.

Bottom of Form

**The z-score for the *i-th* observation in a sample is calculated as:**

(1 mark)

Top of Form

A picture containing table

Description automatically generatedwhere s is the population standard deviation

A picture containing chart

Description automatically generatedwhere s is the population standard deviation

A picture containing diagram

Description automatically generatedwhere s is the sample standard deviation

Diagram

Description automatically generated with low confidencewhere s is the sample standard deviation

None of the options.

Bottom of Form

**An "outlier" in a data is strictly defined by whether \_\_\_\_\_\_\_\_\_.**

(1 mark)

~~Top of Form~~

~~it lies very far from the mean~~

~~it lies more than 3\*IQR to the left of its first quartile or to the right of its 3rd quartile~~

None of the options.

its z-score is less than -2 or greater than +2

~~its z-score is less than -3 or greater than +3~~

Bottom of FormBottom of Form

**The empirical rule is applicable to data that is \_\_\_\_\_\_\_\_\_\_\_.**

(1 mark)

Top of Form

normally distributed

None of the options.

linearly distributed

uniformly distributed

of any distribution

Bottom of Form

**Which of the following is TRUE of covariance, between two variables, when one of the deviations from the mean is positive and the other is negative?**

(1 mark)

~~Top of Form~~

~~There is no covariance between the two variables.~~

The covariance will be negative.

The covariance will be positive.

~~The covariance can be positive or negative.~~

The degree of linear association is high between the two variables.

Bottom of Form

**\_\_\_\_\_\_\_refers to the peakedness (i.e., high, narrow) or flatness (i.e., short, flat-topped) of a histogram.**

(1 mark)

Top of Form

None of the options.

Coefficient of flatness

Skewness

Coefficient of variation

Kurtosis

Bottom of Form

**The \_\_\_\_\_\_\_\_is the difference between the maximum value and the minimum value in the data set.**

(1 mark)

Top of Form

standard deviation

average

mean

range

None of the options.

Bottom of Form

**The\_\_\_\_\_\_\_\_measures the degree of asymmetry of observations around the mean.**(1 mark)

Top of Form

z-score

coefficient of skewness (CS)

coefficient of variation

average

standardized value

Bottom of Form

**The measure of location that specifies the middle value when the data are arranged from least to greatest is the:**(1 mark)

Top of Form

mean

sum

mode

None of the options.

median

Bottom of Form

**Which of the following values of the coefficients of variation of stocks represents the least risky stock?**

(1 mark)

Top of Form

0.62

1.0

0.005

0.5

0.045

Bottom of Form

**The "describeby" function which can generate summary statistics grouped by some categories, is part of the \_\_\_\_ package in R.**

(1 mark)

Top of Form

tidyr

dplyr

statistical

tidyverse

psych

Bottom of Form

**Process A and B fill up milk cartons with a standard deviation of 19.28ml while Process C fills up milk cartons with a standard deviation of 7.58ml. Which process(es) should a milk packaging company use?**

(1 mark)

Top of Form

C

A and B

A, B and C

B

A

Bottom of Form

**\_\_\_\_\_\_\_is a measure of the linear relationship between two variables, X and Y, which does not depend on the units of measurement.**(1 mark)

Top of Form

Covariance

Standard deviation

Z-score

None of the options.

Correlation

Bottom of Form

**According to the Empirical rule, the proportion of a normally distributed data which falls within 2 standard deviations from its mean is about \_\_\_\_\_\_ .**

(1 mark)

Top of Form

86%

66.7%

99.7%

95%

97%

Bottom of Form

**The \_\_\_\_\_\_\_\_provides a relative measure of the dispersion in data relative to the mean.**(1 mark)

Top of Form

z-score

None of the options.

average

coefficient of variation (CV)

standard value

Bottom of Form

**\_\_\_\_\_\_\_\_ states that for any set of data, the proportion of values that lie within k standard deviations (k > 1) of the mean is at least**

**1 - 1/k2.**

(1 mark)

Top of Form

None of the options

Prime number theorem

Chebyshev's theorem

Bertrand's postulate

Oppermann's conjecture

Bottom of Form

**Which of the following statements is false?**

(1 mark)

Top of Form

None of the above statements are false.

Kurtosis refers to the peakedness or flatness of a histogram.

Data with a Coefficient of Kurtosis (CK) of 3.2 is more dispersed than data with a CK of 2.9.

Data with a Coefficient of Kurtosis (CK) of 1.2 suggests that the data is somewhat flat.

Data with a Coefficient of Kurtosis (CK) of 3.9 is more peaked than data with a CK of 3.1.

Bottom of Form

**The \_\_\_\_\_\_\_\_is the observation that occurs most frequently.**(1 mark)

Top of Form

mean

sum

median

mode

None of the options.

Bottom of Form

**The linear association between two variables, X & Y, can be measured by \_\_\_\_\_\_\_\_.**

(1 mark)

Top of Form

Pearson's correlation coefficient

Z-score

Kurtosis

Standard deviation

None of the options.

Bottom of Form

**\_\_\_\_\_\_\_\_refers to the degree of variation in the data, that is, the numerical spread (or compactness) of the data.**(1 mark)

Top of Form

Z-score

Dispersion

Mean

Average

None of the options

Bottom of Form

Bottom of Form