STATISTICS WORKSHEET - 1

Question and Answer from 1 to 9

Ans1 - True
Ans2 - Central Limit Theorem
Ans3 - Modeling event / time data
Ans4 - All of the mentioned
Ans5 - Poisson distribution
Ans6 - False
Ans7 - Hypothesis testing
Ans8 - 0
Ans9 - Outliers can conform to the regression relationship

Question and Answer from 10 to 15

Ans10 - Normal distribution is also known as the Gaussian distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean.

Ans11 - 1- If data missing at random: deletion has no bias effect , but decreases the power of the analysis by decreasing the effective sample size.

2- Recommended : knn imputation , Gaussian mixture imputation.

Ans12 - * Two - sample hypothesis testing
* Randomized experiments with two variants A and B
* A : Control ; B : Variation
* User - experience design : identify changes to web pages that increase clicks on a banner
* Current website : Control ; NULL hypothesis
* New version : variation ; alternative hypothesis

Ans13 - * Bad practice in general
* If just estimating means : mean imputation preserves the mean of the observed dat
* Leads to an underestimate of the standard deviation.
* Distorts relationships between variables by “pulling” estimates of the correlation towards zero.

Ans14 - Linear Regression is an algorithm that provides a linear relationship between an Independent variable and a dependent variable to predict the outcome of future events. It is a statistical method used in data science and machine learning for predictive analysis.

Ans15 - Statistics divided into two branches :-

1- Descriptive Statistics

It deals with collection of data, its presentation in various forms, such as tables, graph and diagrams and findings averages and other measures which would describe the data.

2- Inferential Statistics

It deals with techniques used for analysis of data, making the estimates and drawing conclusions from limited information taken on sample basis and testing the reliability of the estimates.