DWD Final Assessment

.TE MARKS (/PACKAGE-COST-DETAILS/REGENERATE-MARKS?STUDENTID=21950&ASSESSMENTID=227257)

Score Obtained:
49/50 (98%)

RETAKE (/PACKAGE-COST-DETAILS/#) VIEW REPORT

1.
Data Analysis is a process of?

1/1 ATTEMPTED

- inspecting data
- cleaning data
- transforming data
- All of the above

2.
Who is a data scientist?

1/1 ATTEMPTED

- Statistician
- Mathematician
- Software programmer
- All of the above

3.
Point out the correct statement.

1/1 ATTEMPTED

- [x] Raw data is original source of data
- [ ] Preprocessed data is original source of data
- [ ] Raw data is the data obtained after processing steps
- [ ] None of the mentioned

4. Which of the following is a programming language used in Data Science? 1/1 ATTEMPTED

- [ ] Java
- [ ] C++
- [x] Python
- [ ] Ruby

5. Which of the following is a powerful python library for data manipulation and analysis, particularly for working with structured data? 1/1 ATTEMPTED

- [ ] Springboot
- [x] Pandas
- [ ] Matplotlib
- [ ] Seaborn

6.
Which of the tools are Data Science Tools used by Data Scientists to carry out their data operations?

- Apache Spark

1/1 ATTEMPTED

- Excel
- Both A and B
- Ruby

7.

Which of the following is correct skills for a Data Scientist?

1/1 ATTEMPTED

- Probability & Statistics
- Machine Learning / Deep Learning
- Data Wrangling
- All of the above

8.

Point out the wrong statement.

1/1 ATTEMPTED

- Merging concerns combining datasets on the same observations to produce a result with more variables
- Data visualization is the organization of information according to preset specifications
- Subsetting can be used to select and exclude variables and observations
- All of the mentioned

9. Which of the following is organized and formatted data that fits into a predefined model?

1/1 ATTEMPTED

- Structured
- Unstructured
- Semi Structured
- All of the above

10. Why is data visualization important in data analytics? 1/1 ATTEMPTED

- To make the data heavier
- To create complex datasets
- To present findings in a understandable way
- To slow down the analysis process

11.
"Continuous" VS "Discrete" can be written as? 1/1 ATTEMPTED

- Measured Vs Counted
- Counted Vs Measured
- Measured Vs Nominal
- None of these

12.
Using these one can display values they measure in an experiment, sales data, or how your electrical use changes over time? 1/1 ATTEMPTED

Ruby

table or paper

graphs or charts

Fortran

13.

Classify each of the following variables as either nominal or continuous.

age

gender

height

race

The correct combination is:

1/1 ATTEMPTED

a) ordinal b) nominal c) continuous d) nominal

a) continuous b) nominal c) continuous d) nominal

a) continuous b) ratio c) ordinal d) nominal

a) continuous b) interval c)ratio d) nominal

14. What is a frequently used chart to represent the frequency or proportion of different categories in categorical data?

1/1 ATTEMPTED

Scatter plot

Line chart

Bar chart

Box plot

15. In a line graph, what does an upward-sloping line typically indicate?

1/1 ATTEMPTED

- Positive correlation
- Negative correlation
- No correlation
- Categorical data

16.
The goal of _____ is to focus on summarizing and explaining a specific set of data.

1/1 ATTEMPTED

- Inferential statistics
- Descriptive statistics
- Annova statistics
- inference statistics

17.
Approximately what percentage of scores fall within one standard deviation of the mean in a normal distribution?

1/1 ATTEMPTED

- 68%
- 34%
- 95%

- 99%

18.

As a general rule, the _____ is the best measure of central tendency because it is more precise.

1/1 ATTEMPTED

- Mode
- Mean
- Median
- Variance

19.

_____ provides the summary statistics of data.

1/1 ATTEMPTED

- Descriptive Statistics
- Inferential Statistics
- -
- -

20.

Sample is a subset of population.

1/1 ATTEMPTED

- True
- False
- -

-

21.

_____ contains all the elements of a dataset.

1/1 ATTEMPTED

- Sample
- Event
- Population
- None of the options

22.

Descriptive Statistics works on _____ dataset.

1/1 ATTEMPTED

- Sample
- Population
- Both the options
- -

23. Find the mode of the following data: 59, 28, 38, 28, 59, 73, 28, 51, 19

1/1 ATTEMPTED

- 59
- 28
- 38

- 19

24.

If the assumed hypothesis is tested for rejection considering it to be true is called?

1/1 ATTEMPTED

- Null Hypothesis
- Statistical Hypothesis
- Simple Hypothesis
- Composite Hypothesis

25.

If my null hypothesis is 'Dutch people do not differ from English people in height', what is my alternative hypothesis?

1/1 ATTEMPTED

- English people are taller than Dutch people.
- Dutch people are taller than English people.
- Dutch people differ in height from English people.
- All of the statements are plausible alternative hypotheses.

26. Mean is an example of which of the following?

1/1 ATTEMPTED

- Inferential Statistics
- Measures of Central Tendency
- Measures of Variation

- Probability

27. If a distribution is skewed to the right, then it is _____

1/1 ATTEMPTED

- Negatively Skewed
- Positively Skewed
- Symmetrically Skewed
- Symmetric

28. What is the purpose of descriptive statistics?

1/1 ATTEMPTED

- To predict future outcomes
- To summarize and describe the main features of a dataset
- To manipulate data for analysis
- To test hypothesis

29. Which of the following is a measure of central tendency in descriptive statistics? Answer: C) Mean

1/1 ATTEMPTED

- Range
- Variance
- Mean

- Standard Deviation

30. How is the range calculated in descriptive statistics?

1/1 ATTEMPTED

- Subtracting the mean from each data point
- Dividing the dataset into quartiles
- Adding the smallest and largest values in the dataset
- Finding the square root of variance

31. What is the purpose of hypothesis testing in inferential statistics?

1/1 ATTEMPTED

- To summarize and describe data
- To make predictions about future data
- To calculate measures of central tendency
- To draw conclusions about a population based on a sample

32. Which of the following methods is commonly used to estimate population parameters?

1/1 ATTEMPTED

- Descriptive statistics
- Confidence intervals
- Frequency distribution

- Range

33. What does a p-value represent in hypothesis testing? 1/1 ATTEMPTED

- The probability that the null hypothesis is true
- The probability of making a Type I error
- The strength of the evidence against the null hypothesis
- The probability of obtaining the observed results if the null hypothesis is true

34. The p-value quantifies the probability of observing the data, or something more extreme, if the null hypothesis is correct. A low p-value indicates strong evidence against the null hypothesis. 0/1 ATTEMPTED

- Failing to reject the null hypothesis when it is false
- Rejecting the null hypothesis when it is true
- Making a correct decision
- Accepting the alternative hypothesis

35. What does the term "power of a test" refer to? 1/1 ATTEMPTED

- The likelihood of a Type I error
- The likelihood of a Type II error
- The probability of correctly rejecting a false null hypothesis

- The strength of the null hypothesis

36. What does the Central Limit Theorem state?

1/1 ATTEMPTED

- The distribution of a population is always normal.
- The sum of a large number of independent random variables will be normally distributed,
- Sample means will always be equal to the population mean.
- The variance of the sample means will be larger than the population variance.

37. Which of the following is a requirement for the Central Limit Theorem to hold?

1/1 ATTEMPTED

- The population must be normally distributed.
- The sample size must be at least 30.
- The samples must be dependent.
- The population variance must be known.

38. What does a p-value indicate in the context of hypothesis testing?

1/1 ATTEMPTED

- The probability that the null hypothesis is true.
- The probability of making a Type I error.
- The probability of obtaining results at least as extreme as the observed results, assuming the null hypothesis is true.

- The strength of the null hypothesis.

39. Which of the following is a common method for estimating a population mean?

1/1 ATTEMPTED

- Median
- Mode
- Sample mean
- Sample proportion

40. If a 95% confidence interval for a population mean is calculated as (20, 30), what does this mean?

1/1 ATTEMPTED

- The population mean is definitely between 20 and 30.
- There is a 95% probability that the population mean lies between 20 and 30.
- 95% of the sample means will lie between 20 and 30.
- The interval will contain the population mean 95% of the time if the study is repeated many times.

41. What is the main advantage of using surveys for data collection?

1/1 ATTEMPTED

- They are always inexpensive.
- They allow for in-depth qualitative analysis.
- They can reach a large audience quickly.

- They guarantee accurate responses.

42. Which data collection technique involves observing subjects in their natural environment without interference? 1/1 ATTEMPTED

- Controlled experiments
- Case studies
- Naturalistic observation
- Surveys

43. In qualitative research, which of the following is a commonly used data collection technique? 1/1 ATTEMPTED

- Random sampling
- Focus groups
- Statistical analysis
- Structured questionnaires

44. What is a potential disadvantage of using observational methods for data collection? 1/1 ATTEMPTED

- They provide quantitative data.
- They can be time-consuming and expensive.
- They are useful for understanding cause-and-effect relationships.

They eliminate the risk of observer bias.

45. What is the primary goal of data cleaning?

1/1 ATTEMPTED

- To analyze data more quickly.
- To ensure data quality and accuracy.
- To visualize data trends.
- To summarize data findings.

46. Which of the following is a common method for handling missing data?

1/1 ATTEMPTED

- Ignoring missing values.
- Replacing missing values with zeros.
- Imputation using mean, median, or mode.
- Deleting the entire dataset.

47. What is the purpose of deduplication in data preparation?

1/1 ATTEMPTED

- To increase the dataset size.
- To ensure all entries are unique and to eliminate duplicate records.
- To convert data types.
- To visualize the data distribution.

48. What is the first step in the data cleaning process? 1/1 ATTEMPTED

- Removing duplicates
- Identifying data quality issues
- Data transformation
- Data visualization

49. Which of the following is NOT a common issue encountered during data cleaning? 1/1 ATTEMPTED

- Missing values
- Duplicate records
- Data normalization
- Inconsistent formatting

50. Which of the following techniques can be used to standardize data formats? 1/1 ATTEMPTED

- Deduplication
- Data transformation
- Outlier detection
- Data visualization