



INDIAN INSTITUTE OF MATERIALS MANAGEMENT
PGDMM/PGDSCM &L– Two Years Course

COMPUTER MARKED ASSIGNMENT (CMA) For Jan-June 2020 semester.

Semester 4
PAPER – 21
BUSINESS ANALYTICS

Instructions:

1. Answer all 50 questions (compulsory). Each question carries 2 marks **Total : 100 Marks**
2. Read instructions given on the OMR answer sheet and answer the questions.

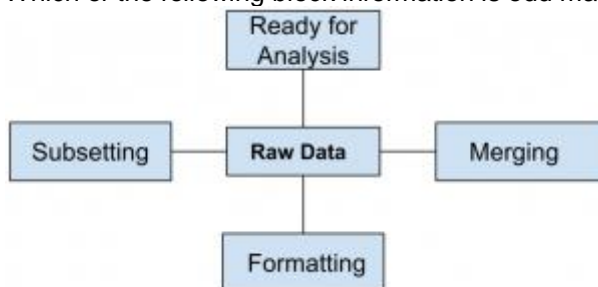
1. Which of the following would be more appropriate to be replaced with question mark in the \ following figure?



- a) Data Analysis
 - b) Data Science
 - c) Descriptive Analytics
 - d) None of the mentioned
2. Point out the correct statement.
- a) Raw data is original source of data
 - b) Preprocessed data is original source of data
 - c) Raw data is the data obtained after processing steps
 - d) None of the mentioned
3. Which of the following is performed by Data Scientist?
- a) Define the question
 - b) Create reproducible code
 - c) Challenge results
 - d) All of the mentioned
4. Which of the following is the most important language for Data Science?
- a) Java
 - b) Ruby
 - c) R
 - d) None of the mentioned
5. Which of the following approach should be used to ask Data Analysis question?
- a) Find only one solution for particular problem
 - b) Find out the question which is to be answered
 - c) Find out answer from dataset without asking question
 - d) None of the mentioned

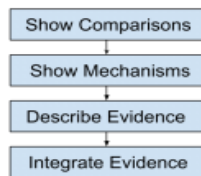
6. Which of the following is one of the key data science skills?
- Statistics
 - Machine Learning
 - Data Visualization
 - All of the mentioned
7. Which of the following is a key characteristic of a hacker?
- Afraid to say they don't know the answer
 - Willing to find answers on their own
 - Not Willing to find answers on their own
 - All of the mentioned
8. Which of the following is characteristic of Processed Data?
- Data is not ready for analysis
 - All steps should be noted
 - Hard to use for data analysis
 - None of the mentioned

9. Which of the following block information is odd man out?



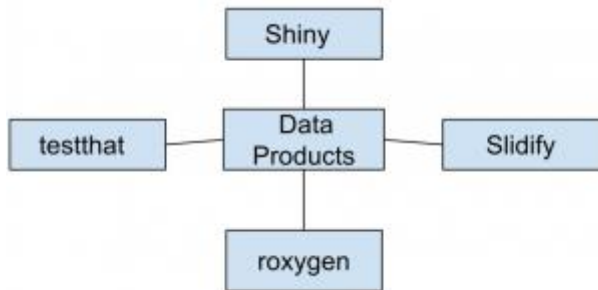
- Subsetting
 - Raw data
 - Ready for analysis
 - None of the mentioned
10. Point out the correct statement.
- Data has only qualitative value
 - Data has only quantitative value
 - Data has both qualitative and quantitative value
 - None of the mentioned
11. Data that summarize all observations in a category are called _____ data.
- frequency
 - summarized
 - raw
 - none of the mentioned
12. Which of the following is an example of raw data?
- original swath files generated from a sonar system
 - initial time-series file of temperature values
 - a real-time GPS-encoded navigation file
 - all of the mentioned
13. Which of the following data is put into a formula to produce commonly accepted results?
- Raw
 - Processed
 - Synchronized
 - All of the Mentioned

14. Processing data includes subsetting, formatting and merging only.
- True
 - False
 - Collating
 - Coding
15. Which of the following is another name for raw data?
- Destination data
 - eggy data
 - secondary
 - machine learning
16. Which type of data is generated by POS terminal in a busy supermarket each day?
- Source
 - Processed
 - Synchronized
 - All of the mentioned
17. Which of the following principle is incorrectly represented in the below figure?



- Show Comparisons
 - Integrate Evidence
 - Describe Evidence
 - None of the mentioned
18. Point out the correct statement.
- Least square is an estimation tool
 - Least square problems falls in to three categories
 - Compound least square is one of the category of least square
 - None of the mentioned
19. How many principles of analytical graphs exist?
- 3
 - 4
 - 6
 - None of the mentioned
20. Which of the following is not a step in data analysis?
- Obtain the data
 - Clean the data
 - EDA
 - None of the mentioned
21. Which of the following technique comes under practical machine learning?
- Bagging
 - Boosting
 - Forecasting
 - None of the mentioned

22. Data Products shown in the below figure is built using which programming language?



- a) S
- b) Python
- c) R
- d) Java

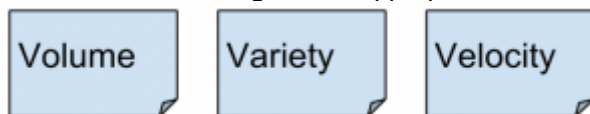
23. Which of the following technique is also referred to as Bagging?

- a) Bootstrap aggregating
- b) Bootstrap subsetting
- c) Bootstrap predicting
- d) All of the mentioned

24. Which of the following is characteristic of Raw Data?

- a) Data is ready for analysis
- b) Original version of data
- c) Easy to use for data analysis
- d) None of the mentioned

25. Which of the following term is appropriate to the below figure?



- a) Large Data
- b) Big Data
- c) Dark Data
- d) None of the mentioned

26. Point out the correct statement.

- a) Machine learning focuses on prediction, based on known properties learned from the training data
- b) Data Cleaning focuses on prediction, based on known properties learned from the training data
- c) Representing data in a form which both mere mortals can understand and get valuable insights is as much a science as much as it is art
- d) None of the mentioned

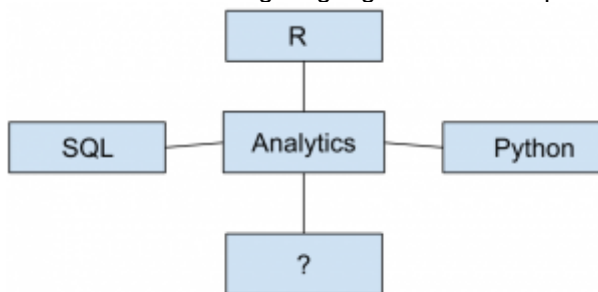
27. Which of the following characteristic of big data is relatively more concerned to data science?

- a) Velocity
- b) Variety
- c) Volume
- d) None of the mentioned

28. Which of the following analytical capabilities are provided by information management company?

- a) Stream Computing
- b) Content Management
- c) Information Integration
- d) All of the mentioned

29. Which of the following step is performed by data scientist after acquiring the data?
- Data Cleansing
 - Data Integration
 - Data Replication
 - All of the mentioned
30. 3V's are not sufficient to describe big data.
- True
 - False
 - Cannot say
 - Incomplete data
31. Which of the following focuses on the discovery of (previously) unknown properties on the data?
- Data mining
 - Big Data
 - Data wrangling
 - Machine Learning
32. Which of the following language should be replaced with the question mark in the below figure?



- Java
 - PHP
 - COBOL
 - None of the mentioned
33. Which of the following is the probability calculus of beliefs, given that beliefs follow certain rules?
- Bayesian probability
 - Frequency probability
 - Frequency inference
 - Bayesian inference
34. Point out the correct statement.
- Bayesian inference is the use of Bayesian probability representation of beliefs to perform inference
 - NULL is the standard missing data marker used in S
 - Frequency inference is the use of Bayesian probability representation of beliefs to perform inference
 - None of the mentioned
35. Which of the following can be considered as random variable?
- The outcome from the roll of a die
 - The outcome of flip of a coin
 - The outcome of exam
 - All of the mentioned
36. Which of the following random variable that take on only a countable number of possibilities?
- Discrete
 - Non Discrete
 - Continuous
 - All of the mentioned

37. Which of the following is also referred to as random variable?
- a) stochast
 - b) aleatory
 - c) eliette
 - d) all of the mentioned
38. Which of the following testing is concerned with making decisions using data?
- a) Probability
 - b) Hypothesis
 - c) Causal
 - d) None of the mentioned
39. Point out the correct statement.
- a) Power of a one sided test is lower than the power of the associated two sided test
 - b) Power of a two sided test is greater than the power of the associated one sided test
 - c) Hypothesis testing is less commonly used
 - d) None of the mentioned
40. Which of the following value is the most common measure of “statistical significance”?
- a) P
 - b) A
 - c) L
 - d) All of the mentioned
41. Point out the wrong statement with respect to FDR.
- a) FDR is difficult to calculate
 - b) FDR is relatively less conservative
 - c) FDR allows for more false positives
 - d) None of the mentioned
42. Which of the following is the oldest multiple testing correction?
- a) Bonferroni correction
 - b) Bernoulli correction
 - c) Likelihood correction
 - d) All of the mentioned
43. The pooled estimator is a mixture of the group variances, placing greater weight on whichever has a larger sample size.
- a) True
 - b) False
 - c) Continuous
 - d) All of the mentioned
44. Which of the following tool is used for constructing confidence intervals and calculating standard errors for difficult statistics?
- a) baggyer
 - b) bootstrap
 - c) jackknife
 - d) none of the mentioned
45. Which of the following tool is used for estimating standard errors and the bias of estimators?
- a) knitr
 - b) jackknife
 - c) ggplot2
 - d) all of the mentioned

46. Power is the probability of rejecting the null hypothesis when it is true.
- a) True
 - b) False
 - c) Alternative hypothesis
 - d) All of the above
47. Which of the following gave rise to need of graphs in data analysis?
- a) Data visualization
 - b) Communicating results
 - c) Decision making
 - d) All of the mentioned
48. Which of the following is characteristic of exploratory graph?
- a) Made slowly
 - b) Axes are not cleaned up
 - c) Color is used for personal information
 - d) All of the mentioned
49. Point out the correct statement.
- a) coplots are one dimensional data graph
 - b) Exploratory graphs are made quickly
 - c) Exploratory graphs are made relatively less in number
 - d) All of the mentioned
50. Which of the following graph can be used for simple summarization of data?
- a) Scatterplot
 - b) Overlaying
 - c) Barplot
 - d) All of the mentioned
