KPS Computer Classes

Data Analytics Paper

**MCQ TEST**

**Time: 2 Hours Marks: 30 Marks**

-------------------------------------------------------------------------------------------------------------------------------------

**Instructions**:

- All questions are compulsory.

- Each MCQ carries 1 mark.

- Choose the correct answer from the options provided.

-------------------------------------------------------------------------------------------------------------------------------------

**Questions:**

**1. What is the primary purpose of data analysis?**

a) Data collection

b) Data storage

c) Extracting useful information

d) Data deletion

**2. Which of the following is not a type of data?**

a) Structured

b) Unstructured

c) Semi-structured

d) Hyper-structured

**3. What does ETL stand for in data processing?**

a) Extract, Transform, Load

b) Evaluate, Transfer, Load

c) Extract, Transfer, Load

d) Evaluate, Transform, Load

**4. Which of the following is a common data visualization tool?**

a) Microsoft Word

b) Tableau

c) Notepad

d) Eclipse

**5. What does SQL stand for?**

a) Structured Query Language

b) Simple Query Language

c) Structured Question Language

d) Simple Query Logic

**6. In data analysis, what is a ‘data warehouse’?**

a) A place to store raw data

b) A repository for large volumes of structured data

c) A tool for data visualization

d) A database management system

**7. Which of the following is an example of a NoSQL database?**

a) MySQL

b) PostgreSQL

c) MongoDB

d) SQLite

**8. Which of the following statistical methods is used to predict the value of a dependent variable based on one or more independent variables?**

a) Classification

b) Clustering

c) Regression

d) Association

**9. What is the purpose of a pivot table in Excel?**

a) To create static data

b) To sort and summarize data

c) To create graphs

d) To store data

**10. What is the term for a graphical representation of the distribution of numerical data**?

a) Histogram

b) Pie Chart

c) Bar Chart

d) Line Chart

**11. What is the role of a data analyst in an organization?**

a) To manage the IT infrastructure

b) To develop software applications

c) To collect, process, and perform statistical analyses of data

d) To handle customer queries

**12. Which of the following techniques is used for reducing the dimensionality of data?**

a) Data cleaning

b) Data transformation

c) Principal Component Analysis (PCA)

d) Data integration

**13. What does the term ‘big data’ refer to?**

a) Small datasets

b) Large volumes of data

c) Simple datasets

d) Structured data only

**14. Which programming language is commonly used for statistical analysis and data visualization?**

a) Java

b) Python

c) C++

d) PHP

**15. In the context of data analytics, what is ‘data mining’?**

a) The process of collecting data from multiple sources

b) The process of transforming raw data into meaningful insights

c) The process of cleaning data

d) The process of storing data

**16. What is a ‘data lake’?**

a) A centralized repository for structured and unstructured data

b) A tool for data analysis

c) A type of database

d) A data visualization tool

**17. Which of the following is a measure of central tendency?**

a) Mean

b) Range

c) Variance

d) Standard deviation

**18. What is ‘data normalization’?**

a) The process of organizing data to minimize redundancy

b) The process of collecting data

c) The process of deleting data

d) The process of storing data

**19. Which tool is commonly used for data analysis in Python?**

a) Pandas

b) Flask

c) Django

d) Spring

**20. What is ‘data cleaning’?**

a) The process of collecting data

b) The process of organizing data

c) The process of detecting and correcting errors in data

d) The process of storing data

**21. In data analysis, what is a ‘dataset’?**

a) A collection of related data

b) A type of database

c) A tool for data visualization

d) A method for data processing

**22. What is ‘descriptive statistics’**

a) The use of data to infer the properties of a population

b) The use of data to describe and summarize a dataset

c) The use of data to predict future trends

d) The use of data to create visualizations

**23. Which of the following is not a type of machine learning?**

a) Supervised learning

b) Unsupervised learning

c) Reinforcement learning

d) Data learning

**24. What is the purpose of data visualization?**

a) To collect data

b) To store data

c) To present data in a visual context

d) To delete data

**25. Which of the following is a data visualization library in Python?**

a) NumPy

b) Matplotlib

c) Pandas

d) Scikit-learn

**26. What is ‘inferential statistics’?**

a) The use of data to describe a sample

b) The use of data to infer the properties of a population

c) The use of data to create visualizations

d) The use of data to store information

**27. Which of the following is a measure of variability in data?**

a) Mean

b) Median

c) Mode

d) Standard deviation

**28. What is a ‘time series analysis’?**

a) Analysis of data collected at the same point in time

b) Analysis of data over time

c) Analysis of data from multiple sources

d) Analysis of data from a single source

**29. What is the main purpose of a ‘box plot’?**

a) To show the distribution of data

b) To compare two sets of data

c) To display the relationship between variables

d) To summarize data

**30. Which of the following is an example of a relational database?**

a) MongoDB

b) Cassandra

c) MySQL

d) HBase