

BCS Level 4 Certificate in Data Analysis Tools QAN 603/0824/2

Specimen Paper A

Record your surname / last / family name and initials on the answer sheet.

Specimen paper only 20 multiple-choice questions – 1 mark awarded to each question. Mark only one answer to each question. There are no trick questions.

A number of possible answers are given for each question, indicated by either **A. B. C. or D.** Your answers should be clearly indicated on the answer sheet.

Pass mark is 13/20.

Copying of this paper is expressly forbidden without the direct approval of BCS, The Chartered Institute for IT.

This qualification is regulated by Ofqual (in England).

1 In a distribution that is positively skewed, which of the following is **TYPICALLY TRUE**?

- A The standard deviation is greater than the range.
- B The median is greater than the mean.
- C The median and mean are equal.
- D The mean is greater than the median.

2 R can easily be extended via what?

- A Models.
- B Packages.
- C Slides.
- D Pockets.

3 What is the median in the following set of data?

1, 2, 2, 3, 5, 6, 9

- A 2.
- B 3.
- C 4.
- D 5.

4 What measure will always change if a single value in a dataset changes?

- A Median.
- B Mode.
- C Mean.
- D All of these.

5 What type of join is used to combine the result-set of two or more SELECT statements?

- A Full.
- B Outer.
- C Union.
- D Inner.

- 6 What will be reduced if multiple disparate datasets are integrated?
- A Data availability.
 - B Data collaboration.
 - C Data complexity.
 - D Data integrity.
- 7 When preparing data for analysis to improve accuracy, quality and usefulness, which of the following are **NOT** suitable ways to respond to an error?
- a) Accept.
 - b) Correct error.
 - c) Ignore.
 - d) Create null value.
- A a and b only.
 - B b and c only.
 - C a and d only.
 - D c and d only.
- 8 Which of the following describes managing an organisation's critical data, providing a single point of reference?
- A Important Data Management.
 - B Master Data Management.
 - C Vital Data Management.
 - D Essential Data Management.
- 9 Which command is used to select only one copy of each set of duplicate rows in SQL?
- A SELECT DIFFERENT
 - B SELECT NONE
 - C SELECT DISTINCT
 - D SELECT ALL

- 10** Which of the following are data quality dimensions?
- a) Uniqueness.
 - b) Timeliness.
 - c) Consistency.
 - d) Cleanliness.
- A** a and b only.
B c and d only.
C a, b and c only.
D b, c and d only.
- 11** Which of the following are user interfaces?
- a) Data dashboard.
 - b) Performance scorecard.
 - c) Data transparency.
 - d) Dynamic charts.
- A** a and b only.
B c and d only.
C b, c and d only.
D a, b and d only.
- 12** Which of the following is a monthly report, which uses dynamic charts and tables to summarise a company's sales turnover?
- A** Dashboard.
B Graph.
C Screen.
D Data.
- 13** Dates in a column of a spreadsheet are an example of which type of data?
- A** Unstructured.
B Semi-structured.
C Structured.
D Multi-structured.

- 14 Which of the following is **NOT** a core data type in Python?
- A Class.
 - B Lists.
 - C Dictionary.
 - D Tuples.
- 15 Which of the following is **NOT** a reason for developing appropriate testing strategies?
- To ensure that unified data sets are...
- A Correct.
 - B Complete.
 - C Legal.
 - D Up to date.
- 16 Which of the following is **NOT** a type of join?
- A Left.
 - B Right.
 - C Select into.
 - D Insert into.
- 17 Which of the following is **NOT** an option to accomplish data migration?
- A Merge two systems into a brand new one.
 - B Merge one system into another.
 - C Remove one of two merged systems from another, without further action.
 - D Create a common view on top of two separate systems.
- 18 Which SQL statement is often used with aggregate functions to assemble the result-set by one or more columns?
- A Count
 - B Order by
 - C Group by
 - D Round

- 19** Which technique can be used to unveil the content and structure of a data set?
- A** Data modelling.
 - B** Data profiling.
 - C** Data processing.
 - D** Data describing.
- 20** Which testing process is used to validate the resource usage, scalability, and reliability of a product?
- A** Owner acceptance testing.
 - B** Technical acceptance testing.
 - C** User acceptance testing.
 - D** Performance stress testing.

-End of Paper-