|  |  |  |
| --- | --- | --- |
| Quiz 1 data warehouse and CRISP-DM | | |
| **1** | **Which of the following features usually applies to data in a data warehouse?** | **C** |
| A | Data are often deleted | |
| B | Most application consist of transaction | |
| C | Data are rarely deleted | |
| D | Relatively few records are processed by application | |
|  |  | |
| **2** | **The time horizon in data warehouse is usually:** | **C** |
| A | 3-4 years | |
| B | 5-6 years | |
| C | 5-10 years | |
| D | None of the above | |
|  |  | |
| **3** | **What is true about the multidimensional model?** | **B** |
| A | Require less disk storage | |
| B | Require more disk storage | |
| C | Increasing the size of a dimension is difficult | |
| D | None of the above | |
|  |  | |
| **4** | **A null value indicates** | **B** |
| A | A numerical value with value 0 | |
| B | The absence of a value | |
| C | A very small value | |
| D | An erroneous value | |
|  |  | |
| **5** | **The purpose of data warehouse is** | **A** |
| A | Data analysis | |
| B | Data retrieval | |
| C | Data updating | |
| D | Data management | |
|  |  | |
| **6** | **Which of the following is true about data warehousing?** | **C** |
| A | Can be updated by user | |
| B | Contain only current data | |
| C | Organized around important subject-oriented data | |
| D | Use for retrieving data | |
|  |  | |
| **7** | **Data warehouse is volatile, because obsolete data are discarded.** | **B** |
| A | True | |
| B | False | |
|  |  | |
| **8** | **ETL is an abbreviation for elevation, transformation and loading.** | **B** |
| A | True | |
| B | False | |
|  |  | |
| **9** | **The full for of OLAP is** | **A** |
| A | Online analytical processing | |
| B | Online advanced processing | |
| C | Online analytical performance | |
| D | Online advanced preparation | |
|  |  | |
| **10** | **One can perform Query operation in the data present in Data warehouse.** | **A** |
| A | True | |
| B | False | |
|  |  | |
| **11** | **Data granularity is \_\_\_\_ of detail of data.** | **C** |
| A | Transformation | |
| B | Summarization | |
| C | Level | |
|  |  | |
| **12** | **When the level of details of data is reducing the data granularity goes higher** | **B** |
| A | True | |
| B | False | |
|  |  | |
| **13** | **Data warehouses are having summarized and reconciled data that can be used by decision makers.** | **A** |
| A | True | |
| B | False | |
|  |  | |
| **14** | **The data is stored, retrieved & updated in \_\_.** | **A** |
| A | OLTP | |
| B | OLAP | |
| C | SMTP | |
| D | FTP | |
|  |  | |
| **15** | **\_\_ is the heart of the warehouse.** | **B** |
| A | Data mining database servers. | |
| B | Data warehouse database servers. | |
| C | Data mart database servers. | |
| D | Relational database servers. | |
|  |  | |
| **16** | **Which of the methodology to commonly used for data analytics / data mining?** | **B** |
| A | SEMMA | |
| B | CRISP-DM | |
| C | KDD Process | |
| D | None | |
|  |  | |
| **17** | **What is CRISP-DM?** | **D** |
| A | A decision tree developed in the 1980’s. | |
| B | A six-phase method for predicting e-commerce buying habits. | |
| C | Microsoft’s linear regression algorithm. | |
| D | A cross-industry standard process for data mining. | |
|  |  | |
| **18** | **What is the first stage of the CRISP-DM model?** | **A** |
| A | Business understanding | |
| B | Data understanding | |
| C | Data preparation | |
| D | Modelling | |
|  |  | |
| **19** | **What are the tasks carried out in Business Understanding in CRISP-DM?** | **D** |
| A | Determine business objectives | |
| B | Assess situation | |
| C | Determine data mining goals | |
| D | All of the above | |
|  |  | |
| **20** | **What is the stage after the Business understanding phase in the CRISP-DM model?** | **A** |
| A | Data understanding | |
| B | Data preparation | |
| C | Modelling | |
| D | Deployment | |
|  |  | |
| **21** | **Which of these is not part of the CRISP-DM data understanding phase?** | **C** |
| A | Collecting relevant data. | |
| B | Finding and identifying any problems within the data sets. | |
| C | Cleaning and addressing any problems with the data sets. | |
| D | All of the above | |
|  |  | |
| **22** | **What is the stage after the data understanding phase in the CRISP-DM model?** | **D** |
| A | Business understanding | |
| B | Modelling | |
| C | Deployment | |
| D | Data preparation | |
|  |  | |
| **23** | **Which of these is part of the CRISP-DM data preparation phase?** | **D** |
| A | Data selection | |
| B | Data cleaning | |
| C | Data integration | |
| D | All of the above | |
|  |  | |
| **24** | **Which of these is NOT part of the CRISP-DM data preparation phase?** | **D** |
| A | Data collection | |
| B | Data exploration | |
| C | Data verification | |
| D | All of the above | |
|  |  | |
| **25** | **What is the stage after the data preparation phase in the CRISP-DM model?** | **B** |
| A | Business understanding | |
| B | Modelling | |
| C | Deployment | |
| D | Evaluation | |
|  |  | |
| **26** | **Which of these is part of the CRISP-DM modelling phase?** | **D** |
| A | Select modelling technique | |
| B | Build the model | |
| C | Discuss results with a domain expert in the business context | |
| D | All the above | |
|  |  | |
| **27** | **Which of these is NOT part of the CRISP-DM modelling phase?** | **C** |
| A | Interpret the model based on domain knowledge | |
| B | Judge the success of the application of the modelling | |
| C | Determine how to proceed at this stage | |
| D | Generate a procedure to test the model’s quality | |
|  |  | |
| **28** | **Which of the following is part of the deployment phase in the CRISP-DM Model?** | **D** |
| A | Plan deployment | |
| B | Plan monitoring and maintenance | |
| C | Assess what went right and wrong, what else to improve? | |
| D | All of the above | |
|  |  | |
| **29** | **Which of the following is the correct sequence of CRISP-DM?** | **B** |
| A | Data u-> data p->modelling -> evaluation -> deployment | |
| B | Business u -> data u->data p ->modelling -> evaluation-> deployment | |
| C | Business u -> data p -> data u-> modelling -> evaluation -> deployment | |
| D | Business u -> data u-> data p-> modelling -> evaluation | |
|  |  | |
| **30** | **Which task of the evaluation phase in CRISP-DM do you assess the degree to which the model meets the business objective?** | **C** |
| A | Describe the data | |
| B | Review the process | |
| C | Evaluate the results | |
| D | Cleaning the data | |
|  |  | |
| **31** | **What is predictive modelling?** | **A** |
| A | The process of using decision trees to predict certain outcomes. | |
| B | Is a process of developing clusters to discover relevant data categories | |
| C | The process of discovering association rules between variables | |
| D | None of the above | |

|  |  |  |
| --- | --- | --- |
| 1 | **Point out the correct statement** | B |
| A | None of the statements | |
| B | Raw data is the original source of data | |
| C | Raw data is the data obtained after processing steps | |
| D | Pre-processed data is the original source of data | |
| 2 | **Which part of CRISP-DM is this: Collect Initial Data - Initial Data Collection Report?** | D |
| A | Data preparation | |
| B | Business understanding | |
| C | Modelling | |
| D | Data Understanding | |
| 3 | **Variety of methods, technologies and associated tools for creating new knowledge/insight to solve complex problems and make a better and faster decision is referred to as\_\_\_\_** | D |
| A | Predictive | |
| B | Prescriptive | |
| C | Analysis | |
| D | Analytics | |
| 4 | **Which of the following referred to the type of data analytics which could answer the question “what should I do about it”?** | D |
| A | Descriptive analytics | |
| B | Predictive analytics | |
| C | Diagnostic analytics | |
| D | Prescriptive analytics | |
| 5 | **Data mart is subsets of data warehouses that is highly focused and isolated for a specific population of Users** | A |
| A | True | |
| B | False | |
| 1 |  |  |
| A |  | |
| B |  | |
| C |  | |
| D |  | |