TDS

Section Id : 64065364128

Section Number : 15

Section type : Online

Mandatory or Optional : Mandatory

Number of Questions : 41

Number of Questions to be attempted : 41

Section Marks : 40

Display Number Panel : Yes

Section Negative Marks : 0

Group All Questions : No

Enable Mark as Answered Mark for Review and Clear Response : No

Maximum Instruction Time : 0

Sub-Section Number : 1

Sub-Section Id : 640653134071

Question Shuffling Allowed : No

Question Number : 412 Question Id : 640653903988 Question Type : MCQ Calculator : Yes Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "DIPLOMA LEVEL : TOOLS IN DATA SCIENCE (COMPUTER BASED EXAM)"

ARE YOU SURE YOU HAVE TO WRITE EXAM FOR THIS SUBJECT?
CROSS CHECK YOUR HALL TICKET TO CONFIRM THE SUBJECTS TO BE WRITTEN.

(IF IT IS NOT THE CORRECT SUBJECT, PLS CHECK THE SECTION AT THE TOP FOR THE SUBJECTS REGISTERED BY YOU)

Options :

6406533044048. ✔ YES

6406533044049. ✖ NO

Sub-Section Number : 2

Sub-Section Id : 640653134072

Question Shuffling Allowed : Yes

Question Number : 413 Question Id : 640653903989 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

What does a correlation coefficient close to +1 indicate?

Options :

6406533044050. ✻ A weak negative relationship between variables

6406533044051. ✻ No relationship between variables

6406533044052. ✔ A strong positive relationship between variables

6406533044053. ✻ A strong negative relationship between variables

Question Number : 414 Question Id : 640653903990 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

How can emotions be effectively used in data storytelling?

Options :

6406533044054. ✻ By overwhelming the audience with excessive data

6406533044055. ✔ By highlighting the impact of data insights on real people or situations

6406533044056. ✻ By using only statistical methods

6406533044057. ✻ By avoiding any form of visualization

Question Number : 415 Question Id : 640653903991 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

Which tool is particularly useful for cleaning messy data and transforming it into a structured format?

Options :

6406533044058. ✻ SQL

6406533044059. ✻ Python

6406533044060. ✔ OpenRefine

6406533044061. ✻ Tableau

Question Number : 416 Question Id : 640653903992 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following practices is considered unethical when performing web scraping?

Options :

6406533044062. ✻ Respecting the website's robots.txt file

6406533044063. ✻ Avoiding excessive requests that could overwhelm the server

6406533044064. ✔ Scraping personal data without the user's consent

6406533044065. ✻ Citing the source of data in your analysis

Question Number : 417 Question Id : 640653903993 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

Which API is commonly used for obtaining geolocation data from an address?

Options :

6406533044066. �ע Twitter API

6406533044067. ✔ Nominatim (OpenStreetMap)

6406533044068. �ע Spotify API

6406533044069. �ע OpenWeatherMap API

Question Number : 418 Question Id : 640653903994 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

What is the primary goal of data storytelling?

Options :

6406533044070. �ע To analyze raw data without any context

6406533044071. �ע To collect large datasets

6406533044072. ✔ To present data insights in a way that is engaging and understandable

6406533044073. ✥ To create complex data models

Question Number : 419 Question Id : 640653903995 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

Which tool is commonly used to scrape data directly from Excel?

Options :

6406533044074. ✥ Beautiful Soup

6406533044075. ✥ Selenium

6406533044076. ✔ Power Query

6406533044077. ✥ Scrapy

Question Number : 420 Question Id : 640653903996 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

What does the term "continuous deployment" refer to?

Options :

6406533044078. ✥ Deploying models without any testing

6406533044079. ✔ Automatically deploying updates to production as soon as they pass testing

6406533044080. ✥ Deploying the model only once after completion

6406533044081. ✥ Deploying models to multiple platforms simultaneously

Question Number : 421 Question Id : 640653903997 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is an important security consideration during deployment?

Options :

6406533044082. ✱ Ensuring high model accuracy

6406533044083. ✔ Implementing authentication and access controls

6406533044084. ✱ Minimizing the size of the deployment package

6406533044085. ✱ Using complex data visualization tools

Question Number : 422 Question Id : 640653903998 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is a key consideration when deploying a data science model?

Options :

6406533044086. ✱ The size of the dataset

6406533044087. ✔ The model's interpretability and scalability in a production environment

6406533044088. ✱ The programming language used

6406533044089. ✱ The frequency of data collection

Question Number : 423 Question Id : 640653903999 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Why is version control important in the deployment of data science models?

Options :

6406533044090. ✱ It ensures that only one version of the model exists

6406533044091. ✔ It tracks changes and allows rollback to previous versions if necessary

6406533044092. ✱ It automatically improves model accuracy

6406533044093. ✱ It reduces the need for documentation

Question Number : 424 Question Id : 640653904000 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which factor is most important when choosing a deployment environment?

Options :

6406533044094. ✱ The visual appearance of the platform

6406533044095. ✔ The ease of integration with existing systems and scalability

6406533044096. ✱ The programming language used in the model

6406533044097. ✱ The geographical location of the user

Question Number : 425 Question Id : 640653904001 Question Type : MCQ Calculator : Yes

Question Label : Multiple Choice Question

Which of the following is an example of a data transformation process?

Options :

6406533044098. ✖ Collecting new data

6406533044099. ✔ Changing a categorical variable into numeric values

6406533044100. ✖ Deleting irrelevant datasets

6406533044101. ✖ Visualizing data trends

Question Number : 426 Question Id : 640653904002 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

What is the purpose of the conclusion in a data story?

Options :

6406533044102. ✔ To summarize the key findings and suggest actions or further considerations

6406533044103. ✖ To introduce new data

6406533044104. ✖ To present the raw data in its entirety

6406533044105. ✖ To discuss the data collection process in detail

Question Number : 427 Question Id : 640653904003 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

What is a Geographic Information System (GIS) primarily used for?

Options :

6406533044106. ✖ Creating machine learning models

6406533044107. ✔ Analyzing and visualizing spatial data

6406533044108. ✖ Managing relational databases

6406533044109. ✖ Developing web applications

Question Number : 428 Question Id : 640653904004 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which Excel feature is commonly used to visualize future trends based on historical data?

Options :

6406533044110. ✖ Pivot Table

6406533044111. ✖ Solver

6406533044112. ✔ Forecast Sheet

6406533044113. ✖ VLOOKUP

Question Number : 429 Question Id : 640653904005 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Why is the data discovery process crucial in data science?

Options :

6406533044114. ✳ It ensures the data is free of errors

6406533044115. ✔ It helps in understanding the problem context better

6406533044116. ✳ It allows for immediate model building

6406533044117. ✳ It simplifies the data visualization process

Question Number : 430 Question Id : 640653904006 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following tasks is an example of text extraction using large language models?

Options :

6406533044118. ✳ Summarizing a news article

6406533044119. ✳ Translating text from one language to another

6406533044120. ✔ Identifying the names of entities in a legal document

6406533044121. ✳ Generating creative writing prompts

Question Number : 431 Question Id : 640653904007 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following elements are typically combined in data storytelling?

Options :

6406533044122. ✔ Data analysis, visualizations, and narrative

6406533044123. ✳ Raw data, complex algorithms, and data storage

6406533044124. ✳ Data collection, coding, and debugging

6406533044125. ✳ Hypothesis testing, data cleaning, and model training

Question Number : 432 Question Id : 640653904008 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is a commonly used format when scraping data from websites?

Options :

6406533044126. ✔ JSON

6406533044127. ✳ SQL

6406533044128. ✳ PDF

6406533044129. ✳ CSV

Question Number : 433 Question Id : 640653904009 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which type of data is best represented by a bar chart?

Options :

6406533044130. ✳ Continuous data over time

6406533044131. ✔ Categorical data

6406533044132. ✳ Relationships between two variables

6406533044133. ✳ Proportions of a whole

Question Number : 434 Question Id : 640653904010 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is NOT typically considered a data cleaning task?

Options :

6406533044134. ✳ Removing duplicates

6406533044135. ✳ Normalizing data

6406533044136. ✔ Generating predictive models

6406533044137. ✳ Handling missing values

Question Number : 435 Question Id : 640653904011 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Why is providing context important in data storytelling?

Options :

6406533044138. ✳ It makes the data more complex

6406533044139. ✳ It allows the audience to understand the data in isolation

6406533044140. ✔ It helps the audience grasp the relevance and implications of the data

6406533044141. ✳ It reduces the need for data cleaning

Question Number : 436 Question Id : 640653904012 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is an ethical concern when using large language models?

Options :

6406533044142. ✳ Their inability to perform complex calculations

6406533044143. ✔ Their potential to spread misinformation if not used carefully

6406533044144. ✳ Their high cost of implementation

6406533044145. ✳ Their lack of user-friendly interfaces

Question Number : 437 Question Id : 640653904013 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

What is the purpose of data profiling in data preparation?

Options :

6406533044146. ✻ To visualize data trends

6406533044147. ✔ To assess the quality and structure of data

6406533044148. ✻ To collect new data

6406533044149. ✻ To perform predictive analysis

Question Number : 438 Question Id : 640653904014 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following best describes an ordinal value?

Options :

6406533044150. ✻ Data that is categorized without a specific order

6406533044151. ✔ Data that represents a specific order or ranking

6406533044152. ✻ Data measured on a scale with equal intervals

6406533044153. ✻ Data that can take on any value within a range

Question Number : 439 Question Id : 640653904015 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is true about Heroku?

Options :

6406533044154. ✻ It is a data visualization tool

6406533044155. ✔ It is a cloud platform that allows easy deployment of web applications

6406533044156. ✻ It is used for creating machine learning models

6406533044157. ✻ It is a version control system

Question Number : 440 Question Id : 640653904016 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

What is a primary characteristic of large language models (LLMs)?

Options :

6406533044158. ✻ They are designed to handle small datasets efficiently

6406533044159. ✻ They are based on rule-based systems

6406533044160. ✔ They are trained on vast amounts of text data to generate human-like text

6406533044161. ✻ They are only used for numerical data analysis

Question Number : 441 Question Id : 640653904017 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Why is data normalization important in the data preparation process?

Options :

6406533044162. ✖ To simplify the data collection process

6406533044163. ✖ To ensure that all data follows the same distribution

6406533044164. ✖ To remove outliers from the dataset

6406533044165. ✔ To scale data to a common range for better comparison

Question Number : 442 Question Id : 640653904018 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following best describes linear regression?

Options :

6406533044166. ✖ A method to classify data into categories

6406533044167. ✔ A technique to find the relationship between a dependent variable and one or more independent variables

6406533044168. ✖ A process of clustering similar data points

6406533044169. ✖ A way to visualize data trends over time

Question Number : 443 Question Id : 640653904019 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which feature in Excel can be used to add narrative text that dynamically updates based on data changes?

Options :

6406533044170. ✖ PivotTables

6406533044171. ✖ Conditional Formatting

6406533044172. ✔ Linked text boxes

6406533044173. ✖ Solver

Question Number : 444 Question Id : 640653904020 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

Which Excel tool can help you visually identify outliers in a dataset?

Options :

6406533044174. ✖ Scatter Plot

6406533044175. ✖ Line Chart

6406533044176. ✖ Histogram

6406533044177. ✔ Box Plot

Question Number : 445 Question Id : 640653904021 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is commonly used for time series forecasting?

Options :

6406533044178. ✱ Decision Trees

6406533044179. ✱ K-Means Clustering

6406533044180. ✔ ARIMA (AutoRegressive Integrated Moving Average)

6406533044181. ✱ Principal Component Analysis (PCA)

Question Number : 446 Question Id : 640653904022 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

What is the primary goal of deploying a data science model or application?

Options :

6406533044182. ✱ To visualize data

6406533044183. ✔ To make the model accessible to end-users for real-world application

6406533044184. ✱ To collect more data

6406533044185. ✱ To perform exploratory data analysis

Question Number : 447 Question Id : 640653904023 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

What is a significant requirement for training large language models effectively?

Options :

6406533044186. ✱ High-quality labeled datasets of moderate size

6406533044187. ✱ Limited computational resources

6406533044188. ✔ Large amounts of unstructured text data and powerful computational
resources

6406533044189. ✱ Manual coding of grammatical rules

Question Number : 448 Question Id : 640653904024 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

What is the primary objective of data analysis?

Options :

6406533044190. ✱ To collect as much data as possible

6406533044191. ✱ To transform data into a usable format

6406533044192. ✔ To extract meaningful insights and patterns from data

6406533044193. ✖ To store data efficiently

Question Number : 449 Question Id : 640653904025 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

What is the primary goal of data visualization?

Options :

6406533044194. ✖ To store data efficiently

6406533044195. ✖ To collect data from different sources

6406533044196. ✔ To present data in a way that makes patterns and insights easier to understand

6406533044197. ✖ To build predictive models

Question Number : 450 Question Id : 640653904026 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following is a commonly used format for geospatial data?

Options :

6406533044198. ✖ CSV

6406533044199. ✖ XML

6406533044200. ✔ Shapefile (.shp)

6406533044201. ✖ JSON

Question Number : 451 Question Id : 640653904027 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

In network analysis, which centrality measure indicates the most "connected" node?

Options :

6406533044202. ✖ Betweenness centrality

6406533044203. ✖ Closeness centrality

6406533044204. ✔ Degree centrality

6406533044205. ✖ Eigenvector centrality

Question Number : 452 Question Id : 640653904028 Question Type : MCQ Calculator : Yes
Correct Marks : 1

Question Label : Multiple Choice Question

In which of the following domains is network analysis commonly applied?

Options :

6406533044206. ✖ Social media

6406533044207. ✖ Climate modeling

6406533044208. ✖ Genetic research

6406533044209. ✔ All of these