IMPORTANT QUESTIONS

DATA VISUALIZATION

Question 1

What is the primary step in acquiring data for visualization?

A) Cleaning the data

B) Collecting the data

C) Analyzing the data

D) Visualizing the data

Answer: B) Collecting the data

Question 2

Which of the following is a common source of data for visualization?

A) Social media

B) Sensor readings

C) Customer reviews

D) All of the above

Answer: D) All of the above

Question 3

What is the purpose of data preprocessing?

A) To visualize the data

B) To analyze the data

C) To clean and transform the data

D) To collect the data

Answer: C) To clean and transform the data

Question 4

Which data type is typically used for categorical data?

A) Numerical

B) Ordinal

C) Nominal

D) Binary

Answer: C) Nominal

Question 5

What is the difference between internal and external data sources?

A) Internal data is from external sources, while external data is from internal sources.

B) Internal data is from within the organization, while external data is from outside.

C) Internal data is more reliable, while external data is less reliable.

D) Internal data is less secure, while external data is more secure.

Answer: B) Internal data is from within the organization, while external data is from outside.

Question 6

Which of the following is an example of a data visualization tool?

A) Excel

B) Tableau

C) Python

D) All of the above

Answer: D) All of the above

Question 7

What is the purpose of data transformation?

A) To clean the data

B) To analyze the data

C) To convert data types

D) To visualize the data

Answer: C) To convert data types

Question 8

Which data type is typically used for numerical data?

A) Nominal

B) Ordinal

C) Interval

D) Ratio

Answer: D) Ratio

Question 9

What is the difference between quantitative and qualitative data?

A) Quantitative data is numerical, while qualitative data is categorical.

B) Quantitative data is categorical, while qualitative data is numerical.

C) Quantitative data is more reliable, while qualitative data is less reliable.

D) Quantitative data is less secure, while qualitative data is more secure.

Answer: A) Quantitative data is numerical, while qualitative data is categorical.

Question 10

Which of the following is an example of a data source?

A) Database

B) API

C) CSV file

D) All of the above

1. **Which of the following libraries is used for data visualization in Python?**

A) pandas  
B) matplotlib  
C) seaborn  
D) All of the above

**Answer**: D) All of the above

2. **What is the purpose of using the** read\_csv **function in Pandas?**

A) To load an Excel file  
B) To load a CSV file  
C) To clean the data  
D) To install necessary libraries

**Answer**: B) To load a CSV file

3. **Which command installs the necessary Python libraries for data visualization in Power BI?**

A) pip install pandas matplotlib seaborn  
B) pip install seaborn matplotlib pandas  
C) pip install pandas seaborn matplotlib  
D) pip install matplotlib seaborn pandas

**Answer**: A) pip install pandas matplotlib seaborn

4. **In which step do you load the data into a Pandas dataframe in Power BI?**

A) Step 1  
B) Step 2  
C) Step 3  
D) Step 4

**Answer**: B) Step 2

5. **What is the first step to acquiring and plotting data in Power BI with Python libraries?**

A) Load the data into Power BI  
B) Install the necessary Python libraries  
C) Clean and pre-process the data  
D) Add custom visual to report

**Answer**: B) Install the necessary Python libraries

6. **Which function in Pandas is used to load an Excel file?**

A) read\_excel  
B) read\_csv  
C) load\_excel  
D) import\_excel

**Answer**: A) read\_excel

7. **What does the** dropna **function in Pandas do?**

A) Fills missing values  
B) Drops rows with missing values  
C) Groups data  
D) Converts data to a DataFrame

**Answer**: B) Drops rows with missing values

8. **Which Python library is used to install pandas, matplotlib, and seaborn?**

A) Python3  
B) pip  
C) Anaconda  
D) conda

**Answer**: B) pip

9. **How can you check the fields you want to visualize in Power BI after loading the data?**

A) In the Python script  
B) By clicking "More" in Power BI  
C) In the dataset window of Power BI  
D) In the report visualizations panel

**Answer**: C) In the dataset window of Power BI

10. **Which function in Pandas is used to fill missing values?**

A) fillna  
B) dropna  
C) groupby  
D) drop

**Answer**: A) fillna

11. **How can you add custom visuals to a Power BI report?**

A) Use Python scripts  
B) Use the "More" option  
C) Import external data  
D) From the dataset panel

**Answer**: A) Use Python scripts

12. **Which method is used to convert the data to a Pandas DataFrame?**

A) pd.read\_csv  
B) pd.DataFrame  
C) df.to\_csv  
D) pd.import\_data

**Answer**: B) pd.DataFrame

13. **Which step involves cleaning and preprocessing the data?**

A) Step 1  
B) Step 2  
C) Step 3  
D) Step 4

**Answer**: C) Step 3

14. **What does the** groupby **function in Pandas do?**

A) Groups data based on a column  
B) Fills missing values  
C) Drops null rows  
D) Creates a DataFrame

**Answer**: A) Groups data based on a column

15. **What is the purpose of the** to\_csv **function in the program?**

A) To load data into Power BI  
B) To install libraries  
C) To save the DataFrame as a CSV file  
D) To display a plot

**Answer**: C) To save the DataFrame as a CSV file

16. **What is the first step in performing time-series analysis using stock market data in Power BI?**

A) Create time-series visualizations  
B) Clean and transform the data  
C) Import the data  
D) Forecast future values

**Answer**: C) Import the data

17. **Where can stock market data for time-series analysis be sourced from?**

A) Only from APIs  
B) From online data providers or APIs  
C) From Excel files only  
D) From Power BI service

**Answer**: B) From online data providers or APIs

18. **Which of the following is NOT part of the data cleaning and transformation process?**

A) Removing null values  
B) Dealing with outliers  
C) Aggregating data to desired time intervals  
D) Creating time-series visualizations

**Answer**: D) Creating time-series visualizations

19. **Which visualization type is recommended for creating time-series visualizations in Power BI?**

A) Line Chart  
B) Pie Chart  
C) Line and Stacked Column Chart  
D) Bar Chart

**Answer**: C) Line and Stacked Column Chart

20. **What should be placed on the x-axis when creating a time-series visualization in Power BI?**

A) Stock price  
B) Date field  
C) Moving averages  
D) Trend lines

**Answer**: B) Date field

21. **Which feature in Power BI helps analyze trends and patterns in stock market data?**

A) Forecasting  
B) Trend lines, moving averages, and regression analysis  
C) Data transformations  
D) Data import from APIs

**Answer**: B) Trend lines, moving averages, and regression analysis

22. **How can you forecast future stock market values in Power BI?**

A) By using the "Forecast" option in the Analytics tab  
B) By using the "Data Transformations" feature  
C) By importing external forecasting data  
D) By applying moving averages

**Answer**: A) By using the "Forecast" option in the Analytics tab

23. **Which of the following features is used to predict future values in Power BI time-series analysis?**

A) Analytics tab  
B) Data cleaning  
C) Trend lines  
D) Regression analysis

**Answer**: A) Analytics tab

24. **What is the purpose of aggregating data in time-series analysis?**

A) To transform the data into another format  
B) To combine data from multiple sources  
C) To group data into desired time intervals (daily, weekly, etc.)  
D) To create a model for forecasting

**Answer**: C) To group data into desired time intervals (daily, weekly, etc.)

25. **After analyzing the stock market data and creating visualizations, what should you do next?**

A) Delete unnecessary data  
B) Publish the report to the Power BI service or share it  
C) Clean the data again  
D) Apply more filters to the data

**Answer**: B) Publish the report to the Power BI service or share it

NATURAL LANGUAGE PROCESSING

Question 1

What is the study of the internal structure of words called?

A) Morphology

B) Syntax

C) Semantics

D) Phonology

Answer: A) Morphology

Question 2

What is the smallest unit of language that carries meaning?

A) Morpheme

B) Phone

C) Syllable

D) Word

Answer: A) Morpheme

Question 3

Identify the root word in the following word: "unhappiness"

A) Happy

B) Unhappy

C) Happiness

D) Unhappiness

Answer: A) Happy

Question 4

What is the process of adding a prefix or suffix to a root word called?

A) Inflection

B) Derivation

C) Compounding

D) Blending

Answer: B) Derivation

Question 5

Identify the prefix in the following word: "unbreakable"

A) Un-

B) Break-

C) -able

D) -break

Answer: A) Un-

Question 6

What is the study of the meaning of words and phrases called?

A) Semantics

B) Syntax

C) Morphology

D) Phonology

Answer: A) Semantics

Question 7

Identify the suffix in the following word: "happiness"

A) -ness

B) -less

C) -ful

D) -ly

Answer: A) -ness

Question 8

What is the process of combining two or more words to form a new word called?

A) Compounding

B) Derivation

C) Inflection

D) Blending

Answer: A) Compounding

Question 9

Identify the root word in the following word: "unemployment"

A) Employ

B) Unemploy

C) Employment

D) Unemployment

Answer: A) Employ

Question 10

What is the study of the sound system of language called?

A) Phonology

B) Morphology

C) Syntax

D) Semantics

Answer: A) Phonology

1. **Which library is used in this exercise for word analysis and morphology?**

A) NumPy  
B) Pandas  
C) NLTK  
D) Matplotlib

**Answer**: C) NLTK

2. **What is the function used to tokenize text into individual words in NLTK?**

A) word\_tokenizer  
B) word\_splitter  
C) word\_tokenize  
D) tokenize\_words

**Answer**: C) word\_tokenize

3. **What does the** word\_tokenize **function do in NLTK?**

A) Splits text into characters  
B) Splits text into sentences  
C) Splits text into words  
D) Tokenizes words into punctuation

**Answer**: C) Splits text into words

4. **Which algorithm is used for applying morphology analysis in this program?**

A) Snowball Stemmer  
B) Lancaster Stemmer  
C) Porter Stemmer  
D) Lemmatization

**Answer**: C) Porter Stemmer

5. **What is the purpose of the PorterStemmer in NLTK?**

A) To lemmatize words  
B) To find the root form of words  
C) To perform part-of-speech tagging  
D) To remove punctuation from words

**Answer**: B) To find the root form of words

6. **Which of the following is the correct syntax for creating a PorterStemmer object in NLTK?**

A) morph = nltk.PorterStemmer()  
B) morph = nltk.Stemmer.Porter()  
C) morph = nltk.Porter()  
D) morph = PorterStemmer()

**Answer**: A) morph = nltk.PorterStemmer()

7. **What type of data structure is used to store the stemmed words in the program?**

A) Tuple  
B) List  
C) Set  
D) Dictionary

**Answer**: B) List

8. **Which function is used to print the tokenized words in the program?**

A) print\_words  
B) tokenize()  
C) word\_tokenize()  
D) print()

**Answer**: D) print()

9. **What is the output of the following code for the input text: "The quick brown fox jumps over the lazy dog"?**

A) List of sentences  
B) List of words  
C) List of characters  
D) List of punctuation

**Answer**: B) List of words

10. **What does the** stem() **function of the PorterStemmer do?**

A) Converts a word to its original form  
B) Removes stopwords from the text  
C) Converts words to their base or root form  
D) Tokenizes the text into sentences

**Answer**: C) Converts words to their base or root form

11. **Which NLTK function is used to import the** word\_tokenize **function?**

A) nltk.download()  
B) from nltk.tokenize import word\_tokenize  
C) import nltk.tokenize  
D) from nltk import word\_tokenize

**Answer**: B) from nltk.tokenize import word\_tokenize

12. **What is the primary aim of performing word analysis and morphology in this exercise?**

A) To convert words to uppercase  
B) To analyze and reduce words to their base form  
C) To translate words into another language  
D) To find the frequency of each word

**Answer**: B) To analyze and reduce words to their base form

13. **Which of the following statements is true about the** PorterStemmer **algorithm?**

A) It is used for lemmatization  
B) It converts words to their base or root form  
C) It removes punctuation from text  
D) It splits words into sentences

**Answer**: B) It converts words to their base or root form

14. **Which of the following steps is NOT performed in this word analysis and morphology exercise?**

A) Tokenizing the input text  
B) Stemmed word list creation  
C) Performing part-of-speech tagging  
D) Printing out tokenized and stemmed words

**Answer**: C) Performing part-of-speech tagging

15. **In the given program, which text is used for word tokenization and stemming?**

A) "The quick brown fox jumps over the lazy dog"  
B) "Hello world"  
C) "Python programming is fun"  
D) "Natural language processing"

**Answer**: A) "The quick brown fox jumps over the lazy dog"

16. **Which module is required to generate n-grams in the NLTK library?**

A) ngram  
B) ngrams  
C) ngram\_util  
D) nltk.util

**Answer**: B) ngrams

17. **What is the purpose of the** ngrams() **function in NLTK?**

A) To generate a list of characters  
B) To generate sequences of words of a specified length  
C) To tokenize sentences into words  
D) To stem words

**Answer**: B) To generate sequences of words of a specified length

18. **In the given program, how is the sentence split into words?**

A) Using split\_words()  
B) Using split()  
C) Using tokenize()  
D) Using word\_tokenize()

**Answer**: B) Using split()

19. **What type of data structure is used to store the bi-gram and tri-gram sequences?**

A) List  
B) Tuple  
C) Dictionary  
D) Set

**Answer**: A) List

20. **What is the length of each sequence in a bi-gram?**

A) 1  
B) 2  
C) 3  
D) 4

**Answer**: B) 2

21. **In the given program, what is the value of "w" after the sentence is split?**

A) A string  
B) A list of words  
C) A tuple of words  
D) A dictionary of words

**Answer**: B) A list of words

22. **What does the** ngrams() **function take as its parameters?**

A) A string and an integer  
B) A list of words and an integer  
C) A tuple and an integer  
D) A list of sentences and an integer

**Answer**: B) A list of words and an integer

23. **Which of the following represents a correct syntax to generate a tri-gram from the word list** w**?**

A) tri\_gram = ngrams(w, 2)  
B) tri\_gram = ngrams(w, 1)  
C) tri\_gram = ngrams(w, 3)  
D) tri\_gram = ngrams(w, 4)

**Answer**: C) tri\_gram = ngrams(w, 3)

24. **What does the** split() **function do to the string "This My laptop it's very comfortable for me"?**

A) Splits the string into characters  
B) Splits the string into individual words  
C) Splits the string into sentences  
D) Converts the string to lowercase

**Answer**: B) Splits the string into individual words

25. **What will be the output for the tri-gram sequence of the sentence "This My laptop it's very comfortable for me"?**

A) [('This', 'My', 'laptop'), ('My', 'laptop', 'it's')]  
B) [('This', 'My', 'laptop'), ('My', 'laptop', 'it's'), ('laptop', 'it's', 'very')]  
C) [('This', 'My', 'laptop')]  
D) [('laptop', 'it's', 'very')]

**Answer**: B) [('This', 'My', 'laptop'), ('My', 'laptop', 'it's'), ('laptop', 'it's', 'very')]