Data Analysis and Python Fundamentals Quiz

Part I: Fundamentals of Data Analysis

1. Which of the following best describes the principle of **modularization** in data analysis?
   (a) Breaking down a program into smaller, reusable parts.
   (b) Combining multiple data sources into a single dataset.
   (c) Visualizing data using modules.
   (d) Encrypting data for security purposes.
2. What task is better suited for a **GPU**?
   (a) Complicated calculation with many parameters.
   (b) Storing data for long-term use.
   (c) Optimizing user input for a web application.
   (d) Matrix multiplication in machine learning.
3. Where would you store a lot of experimental data which you need to access once every 5 years.
   (a) USB Drive
   (b) SSD
   (c) RAM
   (d) Tape Storage
4. The **FAIR** principles stand for:
   (a) Findable, Accessible, Interoperable, Reusable.
   (b) Fast, Accurate, Intelligent, Reliable.
   (c) Flexible, Adaptable, Integrative, Robust.
   (d) Free, Available, Interactive, Real-time.
5. In data analysis, **data cleaning** refers to:
   (a) Organizing data into tables.
   (b) Removing or correcting errors in the dataset.
   (c) Deleting unused datasets.
   (d) Encrypting data for privacy.
6. The term **reproducibility** in data analysis means:
   (a) Achieving the same results using the same data and methods.
   (b) Sharing data publicly.
   (c) Visualizing data in multiple formats.
   (d) Using proprietary software.
7. **Modularity** in programming helps in:
   (a) Increasing code redundancy.
   (b) Making code less readable.
   (c) Reusing code and simplifying debugging.
   (d) Slowing down program execution.
8. **Data interoperability** refers to:
   (a) The ability to use data across different systems.
   (b) The speed at which data is processed.
   (c) The size of the dataset.
   (d) The security level of data storage.
9. The **open data** movement encourages:
   (a) Keeping data proprietary.

(b) Charging fees for data access.

(c) Sharing data freely for transparency and collaboration.

(d) Encrypting all data.

10. The purpose of **data visualization** is to:

(a) Encrypt data for security.

(b) Represent data graphically to identify patterns and insights.

(c) Store data more efficiently.

(d) Increase data redundancy.

Part II: Python Fundamentals

11. What is the correct way to create a list in Python?

(a) `my_list = (1, 2, 3)`

(b) `my_list = [1, 2, 3]`

(c) `my_list = {1, 2, 3}`

(d) `my_list = <1, 2, 3>`

12. Which of the following is a **dictionary** in Python?

(a) `my_dict = { 'a':1, 'b':2 }`

(b) `my_dict = ['a', 'b', 'c']`

(c) `my_dict = (1, 2, 3)`

(d) `my_dict = {1, 2, 3}`

13. What is the output of the following code?

```
print(2 ** 3)
```

(a) `6`

(b) `8`

(c) `9`

(d) `23`

14. Which of the following is a **tuple** in Python?

(a) `my_tuple = [1, 2, 3]`

(b) `my_tuple = (1, 2, 3)`

(c) `my_tuple = {1, 2, 3}`

(d) `my_tuple = <1, 2, 3>`

15. How do you start a **function definition** in Python?

(a) `function my_func():`

(b) `def my_func():`

(c) `define my_func():`

(d) `func my_func():`

16. Which of the following is the correct way to write an **if** statement in Python?

(a) `if x > 0 then:`

(b) `if x > 0:`

(c) `if (x > 0):`

(d) `if x > 0;`

17. What is the output of the following code?

```
print('Hello' + 'World')
```

(a) `Hello World`

(b) `HelloWorld`

(c) `Hello+World`

(d) `Error`

18. Which of the following is not a valid **Python data type**?

(a) `List`

(b) `Tuple`

(c) `Integer`

(d) `Character`

19. How do you insert comments in Python code?

(a) `/* This is a comment */`

(b) `// This is a comment`

(c) `# This is a comment`

(d) `<!-- This is a comment -->`

20. What does **IDE** stand for?

    (a) Integrated Development Environment

    (b) Interactive Data Exploration

    (c) Intelligent Data Extraction

    (d) Integrated Data Encryption

Part III: Advanced Questions

21. In Python, what is a **list comprehension** used for?

    (a) Iterating over lists in a compact form.

    (b) Compressing lists to save memory.

    (c) Sorting lists in place.

    (d) Generating random lists.

22. For a dataset with limited statistics which visualization method best serves the purpose of understanding the data?

    (a) Pie Chart

    (b) Line Graph

    (c) Scatter Plot

    (d) Histogram with error bars

Open Question:

Shortly describe how would you structure data analysis for one of your projects (Lab Reports, BSc Thesis, etc.).