CPSC 250 - Programming for Data Manipulation
Final Exam Review Version

Instructions

You have 150 minutes. This is a closed-book, in-class review exam. You may not use ChatGPT or any programming tools. Answer clearly and legibly. Show all work.

Part I: Multiple Choice (30 points, 2 points each)

Circle the best answer.

1. What is the output of the expression type((5,)) in Python?

   A. int
   B. list
   C. tuple
   D. set

2. What does the __lt__ method define for a class?

   A. Equality
   B. String conversion
   C. Less-than comparison
   D. Logical negation

3. Given x = [1, 2, 3] and y = x, which statement is true?

   A. x and y refer to different lists
   B. y is a shallow copy of x
   C. x == y is False
   D. x is y is True

4. What does plt.hist(data) do?

   A. Plots a histogram of data
   B. Computes the mean of data
   C. Sorts data
   D. Removes outliers from data

5. What is encapsulation in OOP?

   A. Storing data in lists
   B. Restricting access to internal object details
   C. Using recursion inside a class
   D. Replacing operators like + and ==

6. Which of the following is true about default arguments in Python functions?

   A. Mutable defaults like lists are always safe to use
   B. Default values are evaluated each time the function is called
   C. Default arguments must appear after non-default ones
   D. You cannot use keyword arguments with default values

7. What does df.loc[0] return for a DataFrame df?

   A. First row
   B. First column
   C. First value
   D. Column names

8. In binary search trees, what is the time complexity of search in a balanced tree?

   A. O(1)
   B. O(log n)
   C. O(n)
   D. O(n log n)

9. Which method from statsmodels.api is used to perform OLS regression?

   A. lm()
   B. OLS()
   C. linreg()

D. curve_fit()

10. What happens if a base class method is overridden in a derived class?

    A. The base class method runs first
    B. The derived method replaces the base version
    C. The method becomes private
    D. Python raises an error

11. Which of the following best describes inheritance in Python?

    A. Classes cannot inherit from multiple parents
    B. Inheritance allows one class to reuse code from another
    C. Derived classes must override all methods
    D. Only abstract classes can be inherited

12. Which of the following best describes method overloading in Python?

    A. Not supported directly, but can be simulated with default arguments
    B. Multiple methods with the same name are allowed
    C. Python only allows one constructor
    D. You must use decorators for it to work

13. What does the following expression return? {x for x in [1, 2, 2, 3, 3, 3]}

    A. [1, 2, 3]
    B. {1, 2, 2, 3, 3, 3}
    C. {1, 2, 3}
    D. A list of tuples

14. Which of the following is true about tuples in Python?

    A. They are mutable
    B. They are slower than lists
    C. They are immutable
    D. They cannot be nested

15. What is the primary use of np.linspace()?

    A. Count elements in a list
    B. Sort an array
    C. Create a sequence of evenly spaced values
    D. Generate random numbers

Part II: Error Identification (15 points)

The following code is intended to create a dictionary mapping each unique word in a file to the number of times it appears. Find and correct all errors.

```
def count_words(filename):
    with open(filename) as f:
        words = f.read().split()

    counts = {}
    for word in words:
        if word in counts:
            counts[word] = 1
        else:
            counts[word] += 1

    return counts
```

Part III: Code Writing (30 points)

Q1. (6 points) Write a function factorial(n) that uses recursion to compute n!.

Q2. (8 points) Write a class Book with private fields for title and author, and a ___str___ method that returns "<title>" by <author>.

Q3. (8 points) Write a function read_and_plot(filename) that reads a CSV file with columns x and y and produces a scatterplot using matplotlib.

Q4. (8 points) Write a base class Shape with a method area() that returns 0. Then

write derived classes Square and Triangle, each with their own constructor and overridden area() method. Demonstrate polymorphism using a list of Shape objects.

Part IV: Code Comprehension + Commenting (25 points)

The following function is meant to perform simple multi-variable regression using statsmodels. Add comments explaining each line and describe what this code is doing.

```
import statsmodels.api as sm
import pandas as pd

def regress(df, yname, xnames):
    X = df[xnames]
    X = sm.add_constant(X)
    y = df[yname]
    model = sm.OLS(y, X)
    results = model.fit()
    print(results.summary())
```

What does the function output represent?