AI/ML Student Assessment

Total Marks: 50

Instructions:

1. Complete the following tasks using Python and pandas.

2. Write your code in the cells provided below each task.

3. Ensure your code is well-commented and follows best practices.

4. Submit the completed notebook before the deadline.

Question

Task 1: Basic Python Programming (20 Marks)

1. Write a function `fibonacci(n)` that generates the first `n` numbers in the Fibonacci sequence and returns

2. Write a function `is_prime(num)` that checks if a given number is a prime number. Return `True` if the nu

Task 2: Data Manipulation with Pandas (30 Marks)

Using the dataset provided below, perform the following tasks. The dataset contains information about a co

```
import pandas as pd
data = {
    'Date': ['2024-01-01', '2024-01-02', '2024-01-03', '2024-01-04', '2024-01-05'],
    'Sales': [200, 220, 210, 215, 225],
    'Region': ['North', 'South', 'East', 'West', 'North'],
    'Manager': ['John', 'Jane', 'Joe', 'Jake', 'John']
}
df = pd.DataFrame(data)
```

1. Load the dataset into a pandas DataFrame and display its first 5 rows. (5 Marks)

2. Add a new column `Profit` which is 10% of the `Sales` column. (5 Marks)

3. Group the data by `Region` and calculate the total sales for each region. (10 Marks)

4. Find the region with the highest average sales. (10 Marks)

Task 3: Data Loading and Manipulation from GitHub (10 Marks)

1. Load a CSV file from a GitHub repository into a pandas DataFrame. Use the following URL for the datas

   https://raw.githubusercontent.com/datasciencedojo/datasets/master/titanic.csv

   Display the first 5 rows of the DataFrame. (1 Mark)

2. Display summary statistics of the DataFrame. (1 Mark)

3. Check for null values in the DataFrame and display the count of null values in each column. (1 Mark)

4. Fill the null values in the 'Age' column with the mean age. (1 Mark)

5. Drop the rows where 'Embarked' is null. (1 Mark)

6. Create a new column 'FamilySize' which is the sum of 'SibSp' and 'Parch'. (1 Mark)

7. Convert the 'Sex' column to numerical values (male: 0, female: 1). (1 Mark)

8. Group the data by 'Pclass' and calculate the average fare for each class. (1 Mark)

9. Find the most common port of embarkation ('Embarked' column). (1 Mark)

10. Create a pivot table that shows the survival rate ('Survived') for each combination of 'Sex' and 'Pclass'.

End of Assessment

Please ensure your notebook is saved and submitted as per the instructions.