Pandas Assignments

December 18, 2024

[2]:
```
import pandas as pd
import numpy as np
```

[3]:
```
# 1: Create and display a one-dimensional array-like object

data = pd.Series([1, 2, 3, 4, 5])
print(data)
```

```
0    1
1    2
2    3
3    4
4    5
dtype: int64
```

[4]:
```
# 2: Convert a Pandas Series to Python list and its type

data_list = data.tolist()
print(data_list)
print("Type of the list:", type(data_list))
```

```
[1, 2, 3, 4, 5]
Type of the list: <class 'list'>
```

[5]:
```
# 3: Add, subtract, multiply and divide two Pandas Series

series1 = pd.Series([2, 4, 6, 8, 10])
series2 = pd.Series([1, 3, 5, 7, 9])
print("Addition:\n", series1 + series2)
print("Subtraction:\n", series1 - series2)
print("Multiplication:\n", series1 * series2)
print("Division:\n", series1 / series2)
```

```
Addition:
 0     3
1     7
2    11
3    15
```

```
4    19
dtype: int64
Subtraction:
 0    1
1    1
2    1
3    1
4    1
dtype: int64
Multiplication:
 0     2
1    12
2    30
3    56
4    90
dtype: int64
Division:
 0    2.000000
1    1.333333
2    1.200000
3    1.142857
4    1.111111
dtype: float64
```

[6]:
```python
# 4: Convert a dictionary to a Pandas Series

data_dict = {'a': 1, 'b': 2, 'c': 3}
series_from_dict = pd.Series(data_dict)
print(series_from_dict)
```

```
a    1
b    2
c    3
dtype: int64
```

[7]:
```python
# 5: Convert a NumPy array to a Pandas Series

numpy_array = np.array([10, 20, 30, 40])
series_from_numpy = pd.Series(numpy_array)
print(series_from_numpy)
```

```
0    10
1    20
2    30
3    40
dtype: int64
```

[8]:
```
# 6: Change the data type of a given column or Series

series_float = pd.Series([1.0, 2.0, 3.0])
series_int = series_float.astype(int)
print(series_int)
```

```
0    1
1    2
2    3
dtype: int64
```

[9]:
```
# 7: Convert a given Series to an array

array_from_series = series_from_numpy.to_numpy()
print("\nTask 7: Convert Series to array:")
print(array_from_series)
```

```
Task 7: Convert Series to array:
[10 20 30 40]
```

[10]:
```
# 8: Create a DataFrame from a dictionary and display it

data_dict_df = {'Name': ['Alice', 'Bob', 'Charlie'], 'Age': [25, 30, 35]}
df = pd.DataFrame(data_dict_df)
print(df)
```

```
      Name  Age
0    Alice   25
1      Bob   30
2  Charlie   35
```

[11]:
```
# 9: Employee records
emp_ids = [101, 102, 103, 104, 105, 106, 107, 108, 109, 110]
emp_names = ['John', 'Jane', 'Doe', 'Smith', 'Emily', 'Michael', 'Sarah',␣
 ↪'David', 'Chris', 'Anna']
emp_series = pd.Series(emp_names, index=emp_ids)
print(emp_series)

# Convert Series to DataFrame
emp_df = emp_series.reset_index()
emp_df.columns = ['EmpID', 'Name']
print(emp_df)

# Save to CSV
emp_df.to_csv('MyRecord.csv', index=False)
```

```
101       John
```

```
102       Jane
103        Doe
104      Smith
105      Emily
106    Michael
107      Sarah
108      David
109      Chris
110       Anna
dtype: object
   EmpID     Name
0    101     John
1    102     Jane
2    103      Doe
3    104    Smith
4    105    Emily
5    106  Michael
6    107    Sarah
7    108    David
8    109    Chris
9    110     Anna
```

[12]:
```
# 10: Read MyRecord.csv and create a new column Salary

my_record_df = pd.read_csv('MyRecord.csv')
my_record_df['Salary'] = [50000, 60000, 55000, 70000, 65000, 80000, 75000,␣
 ↪90000, 85000, 95000]
print(my_record_df)
```

```
   EmpID     Name  Salary
0    101     John   50000
1    102     Jane   60000
2    103      Doe   55000
3    104    Smith   70000
4    105    Emily   65000
5    106  Michael   80000
6    107    Sarah   75000
7    108    David   90000
8    109    Chris   85000
9    110     Anna   95000
```

[13]:
```
# 11: Work with Advertising.csv

advertising_df = pd.read_csv('./50_Startups.csv')
print(advertising_df.head())
```

```
         RnD  Administration  Marketing       State     Profit
0  165349.20       136897.80  471784.10    New York  192261.83
```

```
1  162597.70       151377.59  443898.53  California  191792.06
2  153441.51       101145.55  407934.54     Florida  191050.39
3  144372.41       118671.85  383199.62    New York  182901.99
4  142107.34        91391.77  366168.42     Florida  166187.94
```

[14]:
```
# Task 12: Work with Salaries.csv

salaries_df = pd.read_csv('./Salary_Data.csv')
print(salaries_df.head())
```

```
   YearsExperience  Salary
0              1.1   39343
1              1.3   46205
2              1.5   37731
3              2.0   43525
4              2.2   39891
```