pandas

July 10, 2020

[10]:
```
import pandas as pd
days = pd.Series(['Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday',␣
 ↪'Saturday', 'Sunday'])
days[:4]
```

[10]:
```
0       Monday
1      Tuesday
2    Wednesday
3     Thursday
dtype: object
```

[4]:
```
import numpy as np
import pandas as pd
dayss = ['Monday', 'Tuesday', 'Wednesday']
dayslist = np.array(dayss)
numpy_days = pd.Series(dayslist)
numpy_days
```

[4]:
```
0       Monday
1      Tuesday
2    Wednesday
dtype: object
```

[11]:
```
# creating Pandas Series from dictionary
import pandas as pd
my_dict = {'a': 'Monday', 'b': 'Tuesday', 'c': 'Wednesday'}
pdays = pd.Series(my_dict)
pdays
```

[11]:
```
a       Monday
b      Tuesday
c    Wednesday
dtype: object
```

[14]:
```
import pandas as pd
print(pd.DataFrame())
```

```
Empty DataFrame
Columns: []
Index: []
```

[19]:
```
# Creating DataFrame from dictionary
import pandas as pd
df_dict = {'Country': ['Ghana', 'Kenya', 'Nigeria', 'Togo'],
           'Capital': ['Accra', 'Nairobi', 'Abuja', 'Lome'],
           'Population': [10000, 8500, 35000, 12000],
           'Age': [60, 70, 80, 75]
}
df = pd.DataFrame(df_dict, index=[2, 4, 6, 8])
df
```

[19]:

	Country	Capital	Population	Age
2	Ghana	Accra	10000	60
4	Kenya	Nairobi	8500	70
6	Nigeria	Abuja	35000	80
8	Togo	Lome	12000	75

[39]:
```
df.loc[2]
```

[39]:
```
Country       Ghana
Capital       Accra
Population    10000
Age              60
Name: 2, dtype: object
```

[25]:
```
# select the Capital column
df['Capital']
```

[25]:
```
2      Accra
4    Nairobi
6      Abuja
8       Lome
Name: Capital, dtype: object
```

[26]:
```
df.at[6, 'Country']
```

[26]:
```
'Nigeria'
```

[27]:
```
df.iat[2, 0]
```

[27]:
```
'Nigeria'
```

[28]:
```
df['Population'].sum()
```

[28]:
```
65500
```

[29]:
```
df.mean()
```

[29]:
```
Population    16375.00
Age              71.25
dtype: float64
```

[31]:
```
df.describe()
```

[31]:
```
         Population        Age
count      4.000000   4.000000
mean   16375.000000  71.250000
std    12499.166639   8.539126
min     8500.000000  60.000000
25%     9625.000000  67.500000
50%    11000.000000  72.500000
75%    17750.000000  76.250000
max    35000.000000  80.000000
```

[36]:
```
import pandas as pd
new_dict = {'Names': ['Ned', 'Mark', 'Archer'],
            'Profession': ['Engineer', 'Doctor', np.nan]}

new_df = pd.DataFrame(new_dict, index = [7, 8, 9])
new_df
```

[36]:
```
    Names Profession
7     Ned   Engineer
8    Mark     Doctor
9  Archer        NaN
```

[45]:
```
new_df.isnull()
```

[45]:
```
   Names  Profession
7  False       False
8  False       False
9  False        True
```

[46]:
```
new_df.fillna('Writer')
```

[46]:
```
    Names Profession
7     Ned   Engineer
8    Mark     Doctor
9  Archer     Writer
```

[ ]:
```

```