4_pandas_dataframe

May 16, 2019

In [ ]:
```
from pandas import Series, DataFrame
import pandas as pd
import numpy as np
```

In [ ]:
```
# Example from - https://chrisalbon.com/python/pandas_map_values_to_values.html
raw_data = {'first_name': ['Jason', 'Molly', 'Tina', 'Jake', 'Amy'],
        'last_name': ['Miller', 'Jacobson', 'Ali', 'Milner', 'Cooze'],
        'age': [42, 52, 36, 24, 73],
        'city': ['San Francisco', 'Baltimore', 'Miami', 'Douglas', 'Boston']}
df = pd.DataFrame(raw_data, columns = ['first_name', 'last_name', 'age', 'city'])
df
```

In [ ]:
```
DataFrame(raw_data, columns = ["age", "city"])
```

In [ ]:
```
DataFrame(raw_data,
          columns = ["first_name","last_name","age", "city", "debt"]
         )
```

In [ ]:
```
df = DataFrame(raw_data, columns = ["first_name","last_name","age", "city", "debt"])
df.first_name
```

In [ ]:
```
df["first_name"]
```

In [ ]:
```
df.first_name
```

In [ ]:
```
df
```

In [ ]:
```
df.loc[1]
```

In [ ]:
```
df["age"].iloc[1:]
```

In [ ]:
```
# Example from - https://stackoverflow.com/questions/31593201/pandas-iloc-vs-ix-vs-loc
s = pd.Series(np.nan, index=[49,48,47,46,45, 1, 2, 3, 4, 5])
s.loc[:3]
```

In [ ]:
```
s.iloc[:3]
```

In [ ]:
```
df.age > 40
```

```
In [ ]: df.debt = df.age > 40
        df

In [ ]: df.to_dict()

In [ ]: values = Series(data=["M","F","F"],index=[0,1,3])
        values

In [ ]: df[:1]

In [ ]: df["sex"] = values
        df

In [ ]: df.T

In [ ]: df.values

In [ ]: df.to_csv("test.csv")

In [ ]: df.head()

In [ ]: del df["debt"]

In [ ]: df

In [ ]: # Example from Python for data analyis

        pop = {'Nevada': {2001: 2.4, 2002: 2.9},
         'Ohio': {2000: 1.5, 2001: 1.7, 2002: 3.6}}

        DataFrame(pop)

In [ ]:
```