07RealEstatePractice

November 4, 2018

```python
In [1]: import pandas as pd
        import os
        import urllib.request
```

```python
In [2]: if 'df_realestate.csv' not in os.listdir('data'):
            url = 'https://s3.amazonaws.com/datasets-jeremy/df_realestate.csv'
            urllib.request.urlretrieve(url, os.path.join('data', 'df_realestate.csv'))

        if 'df_realestate_processed.csv' not in os.listdir('data'):
            url = 'https://s3.amazonaws.com/datasets-jeremy/df_realestate_processed.csv'
            urllib.request.urlretrieve(url, os.path.join('data', 'df_realestate_processed.csv')
```

```python
In [4]: # before preprocessing
        file = os.path.join('data', 'df_realestate.csv')
        df_realestate = pd.read_csv(file, encoding='big5')
        df_realestate

        # processed
        path = "data//df_realestate_processed.csv"
        df_realestate_processed = pd.read_csv(path)
        X = df_realestate_processed.drop(["price_per_meter", "total_price"], axis=1)
        Y = df_realestate_processed['total_price']
```

```python
In [ ]:
```