Kenneth Fields
Batch Code: LISUM11
Week 5: Heroku Deployment
Submission Date: 8/4/22

1. Data from Kaggle for Data Scientist Salaries - ‘https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries’
2. Model developed on Colab - Simple Multiple Linear regression model to predict housing prices based on number of bathrooms, bedrooms, square footage, and year built.
3. Heroku website: https://housing-model-pred.herokuapp.com/
4. Snapshots:
    a. Model Build

```
[ ] import pandas as pd
    import numpy as np
    import matplotlib.pyplot as plt
```

```
[ ] data = pd.read_csv('/content/data.csv')
    data.head()
```

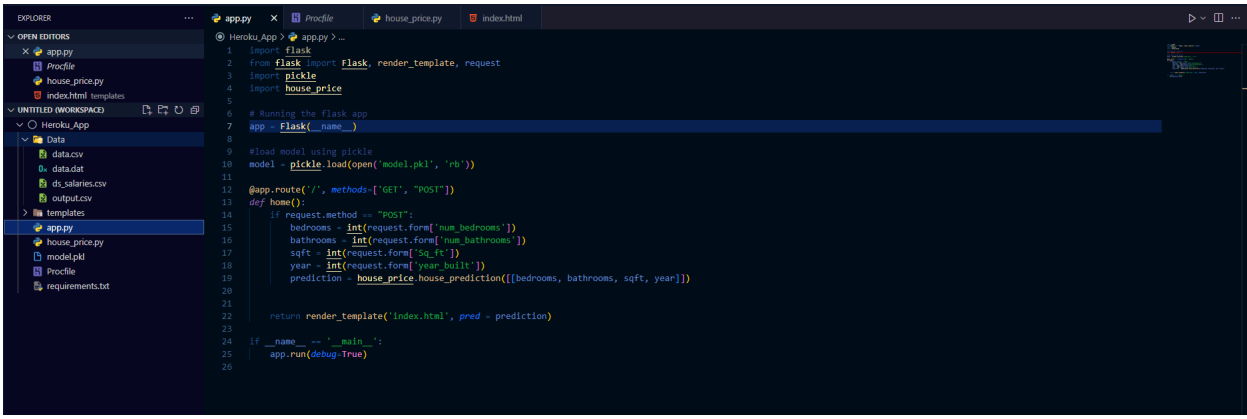
	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	sqft_above	sqft_basement	yr_built	yr_renovated	street	city	statezip	country
0	2014-05-02 00:00:00	313000.0	3.0	1.50	1340	7912	1.5	0	0	3	1340	0	1955	2005	18810 Densmore Ave N	Shoreline	WA 98133	USA
1	2014-05-02 00:00:00	2384000.0	5.0	2.50	3650	9050	2.0	0	4	5	3370	280	1921	0	709 W Blaine St	Seattle	WA 98119	USA
2	2014-05-02 00:00:00	342000.0	3.0	2.00	1930	11947	1.0	0	0	4	1930	0	1966	0	26206-26214 143rd Ave SE	Kent	WA 98042	USA
3	2014-05-02 00:00:00	420000.0	3.0	2.25	2000	8030	1.0	0	0	4	1000	1000	1963	0	857 170th Pl NE	Bellevue	WA 98008	USA
4	2014-05-02 00:00:00	550000.0	4.0	2.50	1940	10500	1.0	0	0	4	1140	800	1976	1992	9105 170th Ave NE	Redmond	WA 98052	USA

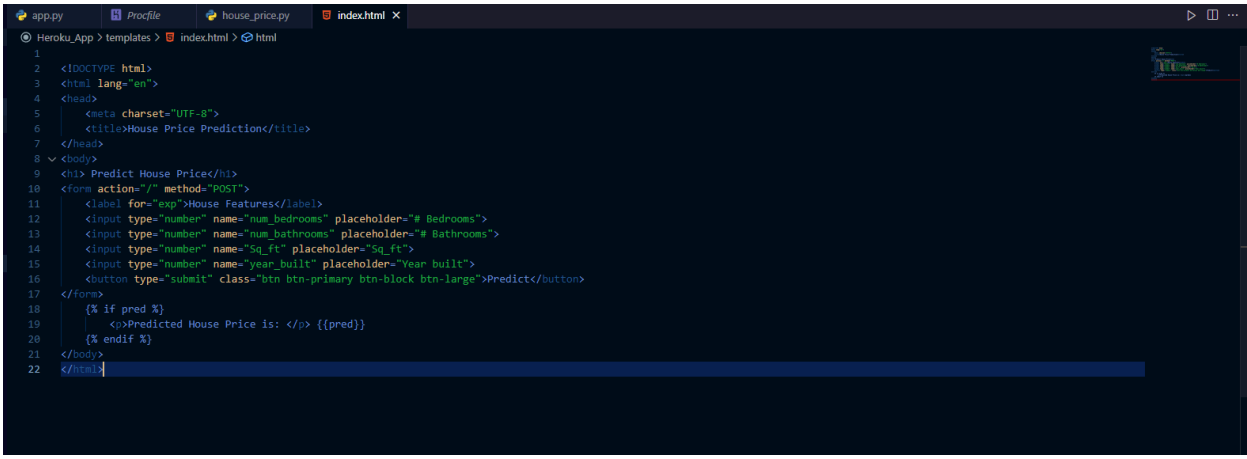
```
[ ] X= data[['bedrooms', 'bathrooms', 'sqft_living', 'yr_built']]
    y = data['price']
```

```
[ ] from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X,y, test_size = .2, random_state = 0)
```

```
[ ] from sklearn.linear_model import LinearRegression
    regressor = LinearRegression()
    regressor.fit(X_train.values, y_train)
```

b. Flask Build:





c. Heroku Deployment: Website - https://housing-model-pred.herokuapp.com/