LAB-2
Housing Dataset
End-End Machine leaning Project

1) We frame the problem, use median & RMSE regression model.

2) Getting the data:-
Download-Root holds URL of having housing data
Housing-KML Path is used to join "data" & "01" directories
Housiong-URL has it combined.
OS mordus for creating a directory extract the files using tarfile.open() .extractall().

3) imports required-
os for joining paths
tarfile for opening tarfiles
urllib for downloading data
matplot.pyput for plotting data

4) load dataframe -
pd.read_csv(data-path) to create a dataframe

5) Create Test-data:
np.random.permutation() ⇒ generates random number
test-size ⇒ stov the % of data in test-sige
indices [: test_set_size] ⇒ first set
indices [test_set_size:] ⇒ second set.

use a seed to get the same data always
on store in a file