Installing all the dependencies.¶

In [3]:

```
import pip as pip
!pip install numpy
!pip install scipy
!pip install pandas
!pip install matplotlib
!pip install -U scikit-learn
```

Requirement already satisfied: numpy in /usr/local/lib/python3.7/site-packages
(1.15.4)
You are using pip version 18.1, however version 19.0.3 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
Requirement already satisfied: scipy in /usr/local/lib/python3.7/site-packages
(1.2.0)
Requirement already satisfied: numpy>=1.8.2 in /usr/local/lib/python3.7/site-
packages (from scipy) (1.15.4)
You are using pip version 18.1, however version 19.0.3 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
Requirement already satisfied: pandas in /usr/local/lib/python3.7/site-packages
(0.23.4)
Requirement already satisfied: numpy>=1.9.0 in /usr/local/lib/python3.7/site-
packages (from pandas) (1.15.4)
Requirement already satisfied: pytz>=2011k in /usr/local/lib/python3.7/site-packages
(from pandas) (2018.9)
Requirement already satisfied: python-dateutil>=2.5.0 in
/usr/local/lib/python3.7/site-packages (from pandas) (2.7.5)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/site-packages
(from python-dateutil>=2.5.0->pandas) (1.12.0)
You are using pip version 18.1, however version 19.0.3 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
Requirement already satisfied: matplotlib in /usr/local/lib/python3.7/site-packages
(2.2.4)
Requirement already satisfied: pytz in /usr/local/lib/python3.7/site-packages (from
matplotlib) (2018.9)
Requirement already satisfied: six>=1.10 in /usr/local/lib/python3.7/site-packages
(from matplotlib) (1.12.0)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (2.3.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.7/site-
packages (from matplotlib) (1.0.1)
Requirement already satisfied: numpy>=1.7.1 in /usr/local/lib/python3.7/site-
packages (from matplotlib) (1.15.4)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.7/site-
packages (from matplotlib) (0.10.0)

Importing required libraries:¶

In [4]:

```
import pandas as pd
import numpy as np
```

Loading data¶

In [7]:

```
url = "https://raw.githubusercontent.com/selva86/datasets/master/BostonHousing.csv"

# getting data from online web_file
data = pd.read_csv(
    url,
    delimiter=","
)

data
```

Out[7]:

	crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio	b	lstat	medv
0	0.00632	18.0	2.31	0	0.538	6.575	65.2	4.0900	1	296	15.3	396.90	4.98	24.0
1	0.02731	0.0	7.07	0	0.469	6.421	78.9	4.9671	2	242	17.8	396.90	9.14	21.6
2	0.02729	0.0	7.07	0	0.469	7.185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
3	0.03237	0.0	2.18	0	0.458	6.998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
4	0.06905	0.0	2.18	0	0.458	7.147	54.2	6.0622	3	222	18.7	396.90	5.33	36.2
5	0.02985	0.0	2.18	0	0.458	6.430	58.7	6.0622	3	222	18.7	394.12	5.21	28.7
6	0.08829	12.5	7.87	0	0.524	6.012	66.6	5.5605	5	311	15.2	395.60	12.43	22.9
7	0.14455	12.5	7.87	0	0.524	6.172	96.1	5.9505	5	311	15.2	396.90	19.15	27.1

	crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio	b	lstat	medv
8	0.21124	12.5	7.87	0	0.524	5.631	100.0	6.0821	5	311	15.2	386.63	29.93	16.5
9	0.17004	12.5	7.87	0	0.524	6.004	85.9	6.5921	5	311	15.2	386.71	17.10	18.9
10	0.22489	12.5	7.87	0	0.524	6.377	94.3	6.3467	5	311	15.2	392.52	20.45	15.0
11	0.11747	12.5	7.87	0	0.524	6.009	82.9	6.2267	5	311	15.2	396.90	13.27	18.9
12	0.09378	12.5	7.87	0	0.524	5.889	39.0	5.4509	5	311	15.2	390.50	15.71	21.7
13	0.62976	0.0	8.14	0	0.538	5.949	61.8	4.7075	4	307	21.0	396.90	8.26	20.4
14	0.63796	0.0	8.14	0	0.538	6.096	84.5	4.4619	4	307	21.0	380.02	10.26	18.2
15	0.62739	0.0	8.14	0	0.538	5.834	56.5	4.4986	4	307	21.0	395.62	8.47	19.9
16	1.05393	0.0	8.14	0	0.538	5.935	29.3	4.4986	4	307	21.0	386.85	6.58	23.1
17	0.78420	0.0	8.14	0	0.538	5.990	81.7	4.2579	4	307	21.0	386.75	14.67	17.5
18	0.80271	0.0	8.14	0	0.538	5.456	36.6	3.7965	4	307	21.0	288.99	11.69	20.2
19	0.72580	0.0	8.14	0	0.538	5.727	69.5	3.7965	4	307	21.0	390.95	11.28	18.2
20	1.25179	0.0	8.14	0	0.538	5.570	98.1	3.7979	4	307	21.0	376.57	21.02	13.6
21	0.85204	0.0	8.14	0	0.538	5.965	89.2	4.0123	4	307	21.0	392.53	13.83	19.6
22	1.23247	0.0	8.14	0	0.538	6.142	91.7	3.9769	4	307	21.0	396.90	18.72	15.2
23	0.98843	0.0	8.14	0	0.538	5.813	100.0	4.0952	4	307	21.0	394.54	19.88	14.5
24	0.75026	0.0	8.14	0	0.538	5.924	94.1	4.3996	4	307	21.0	394.33	16.30	15.6
25	0.84054	0.0	8.14	0	0.538	5.599	85.7	4.4546	4	307	21.0	303.42	16.51	13.9
26	0.67191	0.0	8.14	0	0.538	5.813	90.3	4.6820	4	307	21.0	376.88	14.81	16.6
27	0.95577	0.0	8.14	0	0.538	6.047	88.8	4.4534	4	307	21.0	306.38	17.28	14.8
28	0.77299	0.0	8.14	0	0.538	6.495	94.4	4.4547	4	307	21.0	387.94	12.80	18.4
29	1.00245	0.0	8.14	0	0.538	6.674	87.3	4.2390	4	307	21.0	380.23	11.98	21.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
476	4.87141	0.0	18.10	0	0.614	6.484	93.6	2.3053	24	666	20.2	396.21	18.68	16.7
477	15.02340	0.0	18.10	0	0.614	5.304	97.3	2.1007	24	666	20.2	349.48	24.91	12.0
478	10.23300	0.0	18.10	0	0.614	6.185	96.7	2.1705	24	666	20.2	379.70	18.03	14.6
479	14.33370	0.0	18.10	0	0.614	6.229	88.0	1.9512	24	666	20.2	383.32	13.11	21.4
480	5.82401	0.0	18.10	0	0.532	6.242	64.7	3.4242	24	666	20.2	396.90	10.74	23.0
481	5.70818	0.0	18.10	0	0.532	6.750	74.9	3.3317	24	666	20.2	393.07	7.74	23.7
482	5.73116	0.0	18.10	0	0.532	7.061	77.0	3.4106	24	666	20.2	395.28	7.01	25.0
483	2.81838	0.0	18.10	0	0.532	5.762	40.3	4.0983	24	666	20.2	392.92	10.42	21.8
484	2.37857	0.0	18.10	0	0.583	5.871	41.9	3.7240	24	666	20.2	370.73	13.34	20.6

	crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio	b	lstat	medv
485	3.67367	0.0	18.10	0	0.583	6.312	51.9	3.9917	24	666	20.2	388.62	10.58	21.2
486	5.69175	0.0	18.10	0	0.583	6.114	79.8	3.5459	24	666	20.2	392.68	14.98	19.1
487	4.83567	0.0	18.10	0	0.583	5.905	53.2	3.1523	24	666	20.2	388.22	11.45	20.6
488	0.15086	0.0	27.74	0	0.609	5.454	92.7	1.8209	4	711	20.1	395.09	18.06	15.2
489	0.18337	0.0	27.74	0	0.609	5.414	98.3	1.7554	4	711	20.1	344.05	23.97	7.0
490	0.20746	0.0	27.74	0	0.609	5.093	98.0	1.8226	4	711	20.1	318.43	29.68	8.1
491	0.10574	0.0	27.74	0	0.609	5.983	98.8	1.8681	4	711	20.1	390.11	18.07	13.6
492	0.11132	0.0	27.74	0	0.609	5.983	83.5	2.1099	4	711	20.1	396.90	13.35	20.1
493	0.17331	0.0	9.69	0	0.585	5.707	54.0	2.3817	6	391	19.2	396.90	12.01	21.8
494	0.27957	0.0	9.69	0	0.585	5.926	42.6	2.3817	6	391	19.2	396.90	13.59	24.5
495	0.17899	0.0	9.69	0	0.585	5.670	28.8	2.7986	6	391	19.2	393.29	17.60	23.1
496	0.28960	0.0	9.69	0	0.585	5.390	72.9	2.7986	6	391	19.2	396.90	21.14	19.7
497	0.26838	0.0	9.69	0	0.585	5.794	70.6	2.8927	6	391	19.2	396.90	14.10	18.3
498	0.23912	0.0	9.69	0	0.585	6.019	65.3	2.4091	6	391	19.2	396.90	12.92	21.2
499	0.17783	0.0	9.69	0	0.585	5.569	73.5	2.3999	6	391	19.2	395.77	15.10	17.5
500	0.22438	0.0	9.69	0	0.585	6.027	79.7	2.4982	6	391	19.2	396.90	14.33	16.8
501	0.06263	0.0	11.93	0	0.573	6.593	69.1	2.4786	1	273	21.0	391.99	9.67	22.4
502	0.04527	0.0	11.93	0	0.573	6.120	76.7	2.2875	1	273	21.0	396.90	9.08	20.6
503	0.06076	0.0	11.93	0	0.573	6.976	91.0	2.1675	1	273	21.0	396.90	5.64	23.9
504	0.10959	0.0	11.93	0	0.573	6.794	89.3	2.3889	1	273	21.0	393.45	6.48	22.0
505	0.04741	0.0	11.93	0	0.573	6.030	80.8	2.5050	1	273	21.0	396.90	7.88	11.9

506 rows × 14 columns

In [8]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 506 entries, 0 to 505
Data columns (total 14 columns):
crim       506 non-null float64
zn         506 non-null float64
indus      506 non-null float64
chas       506 non-null int64
nox        506 non-null float64
rm         506 non-null float64
age        506 non-null float64
dis        506 non-null float64
rad        506 non-null int64
tax        506 non-null int64
```

```
ptratio    506 non-null float64
b          506 non-null float64
lstat      506 non-null float64
medv       506 non-null float64
dtypes: float64(11), int64(3)
memory usage: 55.4 KB
```

In [9]:

```
data.describe()
```

Out[9]:

	crim	zn	indus	chas	nox	rm	age	dis
count	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000
mean	3.613524	11.363636	11.136779	0.069170	0.554695	6.284634	68.574901	3.795043
std	8.601545	23.322453	6.860353	0.253994	0.115878	0.702617	28.148861	2.105710
min	0.006320	0.000000	0.460000	0.000000	0.385000	3.561000	2.900000	1.129600
25%	0.082045	0.000000	5.190000	0.000000	0.449000	5.885500	45.025000	2.100175
50%	0.256510	0.000000	9.690000	0.000000	0.538000	6.208500	77.500000	3.207450
75%	3.677082	12.500000	18.100000	0.000000	0.624000	6.623500	94.075000	5.188425
max	88.976200	100.000000	27.740000	1.000000	0.871000	8.780000	100.000000	12.126500

In [10]:

```
data.head()
```

Out[10]:

	crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio	b	lstat	medv
0	0.00632	18.0	2.31	0	0.538	6.575	65.2	4.0900	1	296	15.3	396.90	4.98	24.0
1	0.02731	0.0	7.07	0	0.469	6.421	78.9	4.9671	2	242	17.8	396.90	9.14	21.6
2	0.02729	0.0	7.07	0	0.469	7.185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
3	0.03237	0.0	2.18	0	0.458	6.998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
4	0.06905	0.0	2.18	0	0.458	7.147	54.2	6.0622	3	222	18.7	396.90	5.33	36.2

In [ ]: