

Working with Dataset

1. Import the load_iris dataset

Analyze the dataset and print below values. Learn what is the mean values.

```
In [1]: import numpy as np
import pandas as pd
from sklearn.datasets import load_iris
df=load_iris()
```

```
In [2]: df #It's Load data frame (iris _flower)
```

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	class
0	5.1	3.5	1.4	0.2	Iris-Setosa
1	4.9	3.0	1.2	0.2	Iris-Setosa
2	4.7	3.2	1.3	0.2	Iris-Setosa
3	4.6	3.1	1.5	0.2	Iris-Setosa
4	5.0	3.6	1.4	0.2	Iris-Setosa
5	5.4	4.4	1.5	0.4	Iris-Setosa
6	4.8	3.9	1.4	0.3	Iris-Setosa
7	5.2	4.7	1.6	0.3	Iris-Setosa
8	5.2	3.4	1.4	0.2	Iris-Setosa
9	5.2	3.6	1.4	0.2	Iris-Setosa
10	5.2	3.5	1.5	0.2	Iris-Setosa
11	5.1	3.8	1.6	0.3	Iris-Setosa
12	4.4	3.0	1.4	0.2	Iris-Setosa
13	4.9	3.1	1.5	0.2	Iris-Setosa
14	5.0	3.5	1.4	0.2	Iris-Setosa
15	4.4	2.9	1.4	0.2	Iris-Setosa
16	4.5	2.2	1.4	0.2	Iris-Setosa
17	4.7	3.2	1.3	0.2	Iris-Setosa
18	4.4	3.0	1.3	0.2	Iris-Setosa
19	4.8	3.1	1.6	0.3	Iris-Setosa
20	5.0	3.5	1.4	0.2	Iris-Setosa
21	4.4	2.9	1.4	0.2	Iris-Setosa
22	4.9	3.1	1.5	0.2	Iris-Setosa
23	5.4	4.4	1.5	0.4	Iris-Setosa
24	4.7	3.2	1.3	0.2	Iris-Setosa
25	4.4	3.0	1.4	0.2	Iris-Setosa
26	4.9	3.1	1.5	0.2	Iris-Setosa
27	5.0	3.5	1.4	0.2	Iris-Setosa
28	4.4	2.9	1.4	0.2	Iris-Setosa
29	4.8	3.1	1.6	0.3	Iris-Setosa
30	5.0	3.5	1.4	0.2	Iris-Setosa
31	4.4	2.9	1.4	0.2	Iris-Setosa
32	4.9	3.1	1.5	0.2	Iris-Setosa
33	5.4	4.4	1.5	0.4	Iris-Setosa
34	4.7	3.2	1.3	0.2	Iris-Setosa
35	4.4	3.0	1.4	0.2	Iris-Setosa
36	4.9	3.1	1.5	0.2	Iris-Setosa
37	5.0	3.5	1.4	0.2	Iris-Setosa
38	4.4	2.9	1.4	0.2	Iris-Setosa
39	4.8	3.1	1.6	0.3	Iris-Setosa
40	5.0	3.5	1.4	0.2	Iris-Setosa
41	4.4	2.9	1.4	0.2	Iris-Setosa
42	4.9	3.1	1.5	0.2	Iris-Setosa
43	5.4	4.4	1.5	0.4	Iris-Setosa
44	4.7	3.2	1.3	0.2	Iris-Setosa
45	4.4	3.0	1.4	0.2	Iris-Setosa
46	4.9	3.1	1.5	0.2	Iris-Setosa
47	5.0	3.5	1.4	0.2	Iris-Setosa
48	4.4	2.9	1.4	0.2	Iris-Setosa
49	4.8	3.1	1.6	0.3	Iris-Setosa
50	5.0	3.5	1.4	0.2	Iris-Setosa
51	4.4	2.9	1.4	0.2	Iris-Setosa
52	4.9	3.1	1.5	0.2	Iris-Setosa
53	5.4	4.4	1.5	0.4	Iris-Setosa
54	4.7	3.2	1.3	0.2	Iris-Setosa
55	4.4	3.0	1.4	0.2	Iris-Setosa
56	4.9	3.1	1.5	0.2	Iris-Setosa
57	5.0	3.5	1.4	0.2	Iris-Setosa
58	4.4	2.9	1.4	0.2	Iris-Setosa
59	4.8	3.1	1.6	0.3	Iris-Setosa
60	5.0	3.5	1.4	0.2	Iris-Setosa
61	4.4	2.9	1.4	0.2	Iris-Setosa
62	4.9	3.1	1.5	0.2	Iris-Setosa
63	5.4	4.4	1.5	0.4	Iris-Setosa
64	4.7	3.2	1.3	0.2	Iris-Setosa
65	4.4	3.0	1.4	0.2	Iris-Setosa
66	4.9	3.1	1.5	0.2	Iris-Setosa
67	5.0	3.5	1.4	0.2	Iris-Setosa
68	4.4	2.9	1.4	0.2	Iris-Setosa
69	4.8	3.1	1.6	0.3	Iris-Setosa
70	5.0	3.5	1.4	0.2	Iris-Setosa
71	4.4	2.9	1.4	0.2	Iris-Setosa
72	4.9	3.1	1.5	0.2	Iris-Setosa
73	5.4	4.4	1.5	0.4	Iris-Setosa
74	4.7	3.2	1.3	0.2	Iris-Setosa
75	4.4	3.0	1.4	0.2	Iris-Setosa
76	4.9	3.1	1.5	0.2	Iris-Setosa
77	5.0	3.5	1.4	0.2	Iris-Setosa
78	4.4	2.9	1.4	0.2	Iris-Setosa
79	4.8	3.1	1.6	0.3	Iris-Setosa
80	5.0	3.5	1.4	0.2	Iris-Setosa
81	4.4	2.9	1.4	0.2	Iris-Setosa
82	4.9	3.1	1.5	0.2	Iris-Setosa
83	5.4	4.4	1.5	0.4	Iris-Setosa
84	4.7	3.2	1.3	0.2	Iris-Setosa
85	4.4	3.0	1.4	0.2	Iris-Setosa
86	4.9	3.1	1.5	0.2	Iris-Setosa
87	5.0	3.5	1.4	0.2	Iris-Setosa
88	4.4	2.9	1.4	0.2	Iris-Setosa
89	4.8	3.1	1.6	0.3	Iris-Setosa
90	5.0	3.5	1.4	0.2	Iris-Setosa
91	4.4	2.9	1.4	0.2	Iris-Setosa
92	4.9	3.1	1.5	0.2	Iris-Setosa
93	5.4	4.4	1.5	0.4	Iris-Setosa
94	4.7	3.2	1.3	0.2	Iris-Setosa
95	4.4	3.0	1.4	0.2	Iris-Setosa
96	4.9	3.1	1.5	0.2	Iris-Setosa
97	5.0	3.5	1.4	0.2	Iris-Setosa
98	4.4	2.9	1.4	0.2	Iris-Setosa
99	4.8	3.1	1.6	0.3	Iris-Setosa
100	5.0	3.5	1.4	0.2	Iris-Setosa
101	4.4	2.9	1.4	0.2	Iris-Setosa
102	4.9	3.1	1.5	0.2	Iris-Setosa
103	5.4	4.4	1.5	0.4	Iris-Setosa
104	4.7	3.2	1.3	0.2	Iris-Setosa
105	4.4	3.0	1.4	0.2	Iris-Setosa
106	4.9	3.1	1.5	0.2	Iris-Setosa
107	5.0	3.5	1.4	0.2	Iris-Setosa
108	4.4	2.9	1.4	0.2	Iris-Setosa
109	4.8	3.1	1.6	0.3	Iris-Setosa
110	5.0	3.5	1.4	0.2	Iris-Setosa
111	4.4	2.9	1.4	0.2	Iris-Setosa
112	4.9	3.1	1.5	0.2	Iris-Setosa
113	5.4	4.4	1.5	0.4	Iris-Setosa
114	4.7	3.2	1.3	0.2	Iris-Setosa
115	4.4	3.0	1.4	0.2	Iris-Setosa
116	4.9	3.1	1.5	0.2	Iris-Setosa
117	5.0	3.5	1.4	0.2	Iris-Setosa
118	4.4	2.9	1.4	0.2	Iris-Setosa
119	4.8	3.1	1.6	0.3	Iris-Setosa
120	5.0	3.5	1.4	0.2	Iris-Setosa
121	4.4	2.9	1.4	0.2	Iris-Setosa
122	4.9	3.1	1.5	0.2	Iris-Setosa
123	5.4	4.4	1.5	0.4	Iris-Setosa
124	4.7	3.2	1.3	0.2	Iris-Setosa
125	4.4	3.0	1.4	0.2	Iris-Setosa
126	4.9	3.1	1.5	0.2	Iris-Setosa
127	5.0	3.5	1.4	0.2	Iris-Setosa
128	4.4	2.9	1.4	0.2	Iris-Setosa
129	4.8	3.1	1.6	0.3	Iris-Setosa
130	5.0	3.5	1.4	0.2	Iris-Setosa
131	4.4	2.9	1.4	0.2	Iris-Setosa
132	4.9	3.1	1.5	0.2	Iris-Setosa
133	5.4	4.4	1.5	0.4	Iris-Setosa
134	4.7	3.2	1.3	0.2	Iris-Setosa
135	4.4	3.0	1.4	0.2	Iris-Setosa
136	4.9	3.1	1.5	0.2	Iris-Setosa
137	5.0	3.5	1.4	0.2	Iris-Setosa
138	4.4	2.9	1.4	0.2	Iris-Setosa
139	4.8	3.1	1.6	0.3	Iris-Setosa
140	5.0	3.5	1.4	0.2	Iris-Setosa
141	4.4	2.9	1.4	0.2	Iris-Setosa
142	4.9	3.1	1.5	0.2	Iris-Setosa
143	5.4	4.4	1.5	0.4	Iris-Setosa
144	4.7	3.2	1.3	0.2	Iris-Setosa
145	4.4	3.0	1.4	0.2	Iris-Setosa
146	4.9	3.1	1.5	0.2	Iris-Setosa
147	5.0	3.5	1.4	0.2	Iris-Setosa
148	4.4	2.9	1.4	0.2	Iris-Setosa
149	4.8	3.1	1.6	0.3	Iris-Setosa
150	5.0	3.5	1.4	0.2	Iris-Setosa
151	4.4	2.9	1.4	0.2	Iris-Setosa
152	4.9	3.1	1.5	0.2	Iris-Setosa
153	5.4	4.4	1.5	0.4	Iris-Setosa
154	4.7	3.2	1.3	0.2	Iris-Setosa
155	4.4	3.0	1.4	0.2	Iris-Setosa
156	4.9	3.1	1.5	0.2	Iris-Setosa
157	5.0	3.5	1.4	0.2	Iris-Setosa
158	4.4	2.9	1.4	0.2	Iris-Setosa
159	4.8	3.1	1.6	0.3	Iris-Setosa
160	5.0	3.5	1.4	0.2	Iris-Setosa
161	4.4	2.9	1.4	0.2	Iris-Setosa
162	4.9	3.1	1.5	0.2	Iris-Setosa
163	5.4	4.4	1.5	0.4	Iris-Setosa
164	4.7	3.2	1.3	0.2	Iris-Setosa
165	4.4	3.0	1.4	0.2	Iris-Setosa
166	4.9	3.1	1.5	0.2	Iris-Setosa
167	5.0	3.5	1.4	0.2	Iris-Setosa
168	4.4	2.9	1.4	0.2	Iris-Setosa
169	4.8	3.1	1.6	0.3	Iris-Setosa
170	5.0	3.5	1.4	0.2	Iris-Setosa
171	4.4	2.9	1.4	0.2	Iris-Setosa
172	4.9	3.1	1.5	0.2	Iris-Setosa
173	5.4	4.4	1.5	0.4	Iris-Setosa
174	4.7	3.2	1.3	0.2	Iris-Setosa
175	4.4	3.0	1.4	0.2	Iris-Setosa
176	4.9	3.1	1.5	0.2	Iris-Setosa
177	5.0	3.5	1.4	0.2	Iris-Setosa
178	4.4	2.9	1.4	0.2	Iris-Setosa
179	4.8	3.1	1.6	0.3	Iris-Setosa
180	5.0	3.5	1.4	0.2	Iris-Setosa
181	4.4	2.9	1.4	0.2	Iris-Setosa
182	4.9	3.1	1.5	0.2	Iris-Setosa
183	5.4	4.4	1.5	0.4	Iris-Setosa
184	4.7	3.2	1.3	0.2	Iris-Setosa
185	4.4	3.0	1.4	0.2	Iris-Setosa
186	4.9	3.1	1.5	0.2	Iris-Setosa
187	5.0	3.5	1.4	0.2	Iris-Setosa
188	4.4	2.9	1.4	0.2	Iris-Setosa
189	4.8	3.1	1.6	0.3	Iris-Setosa
190	5.0	3.5	1.4	0.2	Iris-Setosa
191	4.4	2.9	1.4	0.2	Iris-Setosa
192	4.9	3.1	1.5	0.2	Iris-Setosa
193	5.4	4.4	1.5	0.4	Iris-Setosa
194	4.7	3.2	1.3	0.2	Iris-

```
In [3]: print(df.data)      #Only print it's data value
```

```
[5.1 3.5 1.4 0.2]
[4.9 3.  1.4 0.2]
[4.7 3.2 1.3 0.2]
[4.6 3.1 1.5 0.2]
[5.  3.6 1.4 0.2]
[5.4 3.9 1.7 0.4]
[4.6 3.4 1.4 0.3]
[5.  3.4 1.5 0.2]
[4.4 2.9 1.4 0.2]
[4.9 3.1 1.5 0.1]
[5.4 3.7 1.5 0.2]
[4.8 3.4 1.6 0.2]
[4.8 3.  1.4 0.1]
[4.3 3.  1.1 0.1]
[5.8 4.  1.2 0.2]
[5.7 4.4 1.5 0.4]
[5.4 3.9 1.3 0.4]
[5.1 3.5 1.4 0.3]
[5.7 3.8 1.7 0.3]
```

```
In [4]: df.data.shape      #That is shape or dimention of data model
```

```
Out[4]: (150, 4)
```

```
In [5]: df.feature_names      # Here by , Total 4 column which
```

```
Out[5]: ['sepal length (cm)',
         'sepal width (cm)',
         'petal length (cm)',
         'petal width (cm)']
```

```
In [6]: df.target_names      # this respresent the Three diffrent
```

```
Out[6]: array(['setosa', 'versicolor', 'virginica'], dtype='<U16')
```

[6.1, 3. , 4.9, 1.8],
[6.4, 2.8, 5.6, 2.1],
[7.2, 3. , 5.8, 1.6],
[7.4, 2.8, 6.1, 1.9],
[7.9, 3.8, 6.4, 2.],
[6.4, 2.8, 5.6, 2.2],
[6.3, 2.8, 5.1, 1.5],
[6.1, 2.6, 5.6, 1.4],
[7.7, 3. , 6.1, 2.3],
[6.3, 3.4, 5.6, 2.4],
[6.4, 3.1, 5.5, 1.8],
[6. , 3. , 4.8, 1.8],
[6.9, 3.1, 5.4, 2.1],
[6.7, 3.1, 5.6, 2.4],
[6.9, 3.1, 5.1, 2.3],
[5.8, 2.7, 5.1, 1.9],
[6.8, 3.2, 5.9, 2.3],
[6.7, 3.3, 5.7, 2.5],
[6.7, 3. , 5.2, 2.3],
[6.3, 2.5, 5. , 1.9],

[illegible]

In [9]: df.DESCR *# Description of DATA*

```
Out[9]: '.. _iris_dataset:\n\nIris plants dataset\n-----\n
haracteristics:**\n\n      :Number of Instances: 150 (50 i
s)\n      :Number of Attributes: 4 numeric, predictive att
:Attribute Information:\n          - sepal length in cm\n
cm\n          - petal length in cm\n          - petal width
s:\n          - Iris-Setosa\n          - Iri
- Iris-Virginica\n          \n      :Summary Statist
=== =====\n
Mean    SD    Class Correlation\n      =====
===== \n      sepal length:    4.3  7.9   5.84
al width:    2.0  4.4   3.05   0.43   -0.4194\n      petal
3.76  1.76    0.9490 (high!)\n      petal width:    0.1
0.9565 (high!)\n      =====
=\n\n      :Missing Attribute Values: None\n      :Class Dis
ch of 3 classes.\n      :Creator: R.A. Fisher\n      :Donor:
HALL%PLU@io.arc.nasa.gov)\n      :Date: July, 1988\n\nThe
first used by Sir R.A. Fisher. The dataset is taken\nfrom
that it\'s the same as in R, but not as in the UCI\nMach
y, which has two wrong data points.\n\nThis is perhaps t
to be found in the\npattern recognition literature. Fis
ic in the field and\nis referenced frequently to this da
or example.) The\ndata set contains 3 classes of 50 ins
class refers to a\ntype of iris plant. One class is lir
e other 2; the\nlatter are NOT linearly separable from e
c:: References\n\n      - Fisher, R.A. "The use of multiple
mic problems"\n      Annual Eugenics, 7, Part II, 179-188
ibutions to\n      Mathematical Statistics" (John Wiley,
R.O., & Hart, P.E. (1973) Pattern Classification and Sce
7.D83) John Wiley & Sons. ISBN 0-471-22361-1. See page
B.V. (1980) "Nosing Around the Neighborhood: A New Syste
lassification Rule for Recognition in Partially Exposed\
EEE Transactions on Pattern Analysis and Machine\n      I
2, No. 1, 67-71.\n      - Gates, G.W. (1972) "The Reduced M
IEEE Transactions\n      on Information Theory, May 1972,
o: 1988 MLC Proceedings, 54-64. Cheeseman et al\'s AUTOC
l clustering system finds 3 classes in the data.\n      - M
```

In []:

2. Write a logic to Filter the rows of iris data that have petal length > 4.5 and sepal length < 5.0

```
In [10]: for i in df.data:
          for j in i:
              if i[0] < 5 and i[2] > 1.5:
                  print(i)
          print("\n")
```

```
[4.8 3.4 1.6 0.2]
[4.8 3.4 1.6 0.2]
[4.8 3.4 1.6 0.2]
[4.8 3.4 1.6 0.2]
[4.8 3.4 1.9 0.2]
[4.8 3.4 1.9 0.2]
[4.8 3.4 1.9 0.2]
[4.8 3.4 1.9 0.2]
[4.8 3.4 1.9 0.2]
[4.7 3.2 1.6 0.2]
[4.7 3.2 1.6 0.2]
[4.7 3.2 1.6 0.2]
[4.7 3.2 1.6 0.2]
[4.8 3.1 1.6 0.2]
[4.8 3.1 1.6 0.2]
[4.8 3.1 1.6 0.2]
[4.8 3.1 1.6 0.2]
[4.9 2.4 3.3 1. ]
[4.9 2.4 3.3 1. ]
[4.9 2.4 3.3 1. ]
[4.9 2.4 3.3 1. ]
[4.9 2.5 4.5 1.7]
[4.9 2.5 4.5 1.7]
[4.9 2.5 4.5 1.7]
[4.9 2.5 4.5 1.7]
```

3. Calculate the mean, median and standard deviation

```
In [11]: Mean = np.mean(df.data)
          Mean                                     #Mean of data
```

```
Out[11]: 3.4644999999999997
```

```
In [12]: Median = np.median(df.data)
          Median                                   #Median of data
```

```
Out[12]: 3.2
```

```
In [13]: STd = np.std(df.data)
          STd                                     #Standard deviation
```

```
Out[13]: 1.9738430577598278
```

4. Use the petal length (3rd) column of iris data to categorical data, such that if petal length is:

- a. Less than 3 --> 'Small'
- b. Between 3to 5 --> 'Medium'
- c. Greater than 5 --> 'Large'

```
In [14]: for i in df.data:
          for j in i:
              if i[2] < 3 :
                  print("Smaller : {}".format(i[2]))
              elif i[2] >= 3 and i[2] <= 5 :
                  print("Medium : {}".format(i[2]))
              else :
                  print("Larger : {}".format(i[2]))
```

```
Larger : 5.1
Larger : 5.1
Larger : 5.1
Larger : 5.1
Larger : 5.9
Larger : 5.9
Larger : 5.9
Larger : 5.9
Larger : 5.7
Larger : 5.7
Larger : 5.7
Larger : 5.7
Larger : 5.2
Larger : 5.2
Larger : 5.2
Larger : 5.2
Medium : 5.0
Medium : 5.0
Medium : 5.0
.. ..
```

In []: