

Chapter 1: First Steps

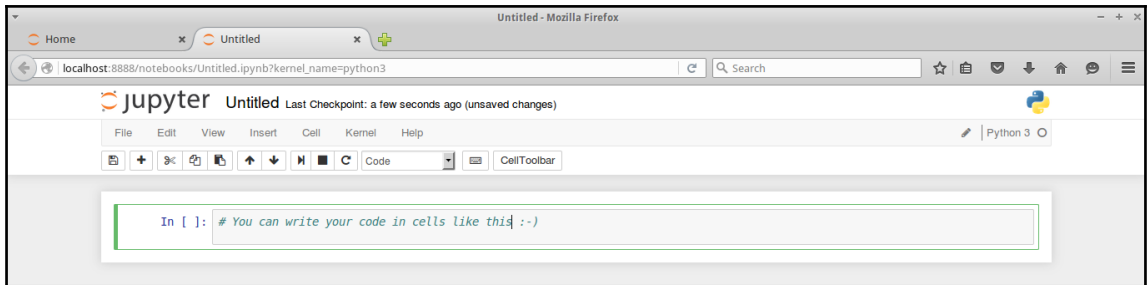
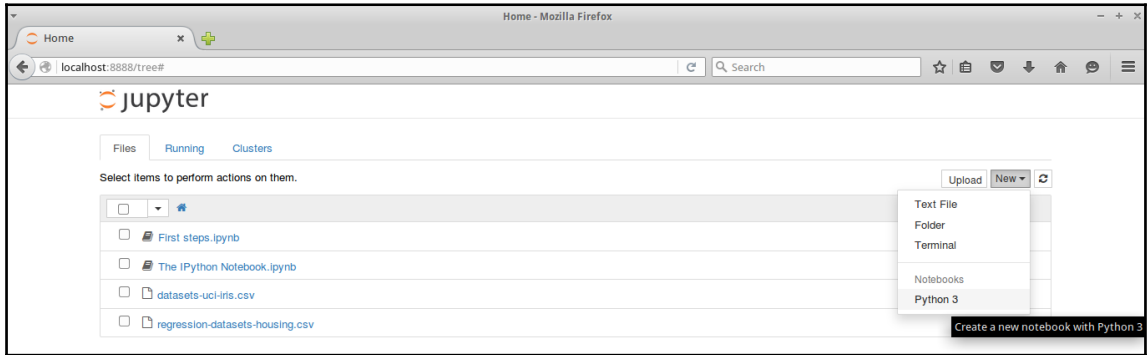
This is a \LaTeX inline equation: $x = Ax + b$

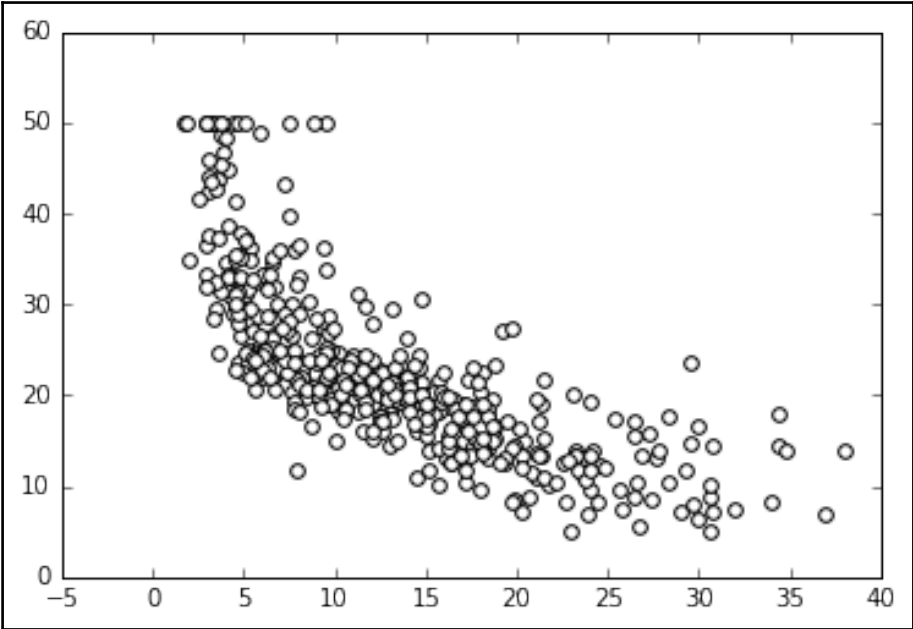
And this is a one-liner:

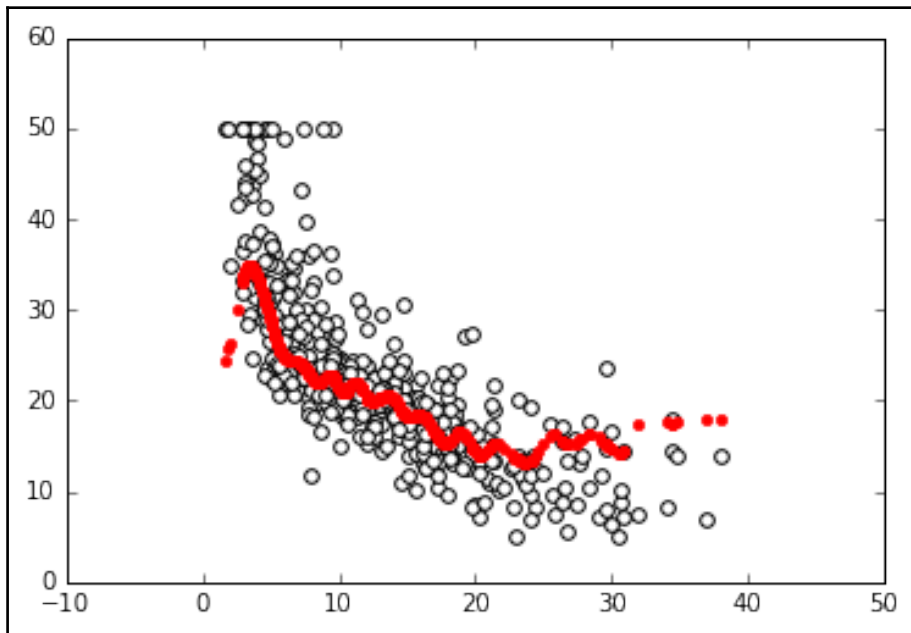
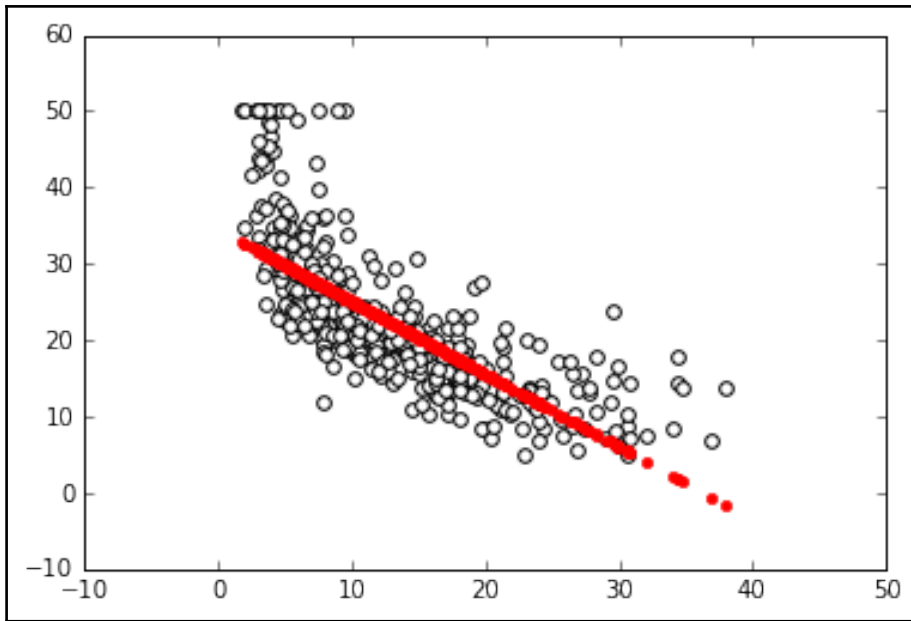
$$x = Ax + b$$

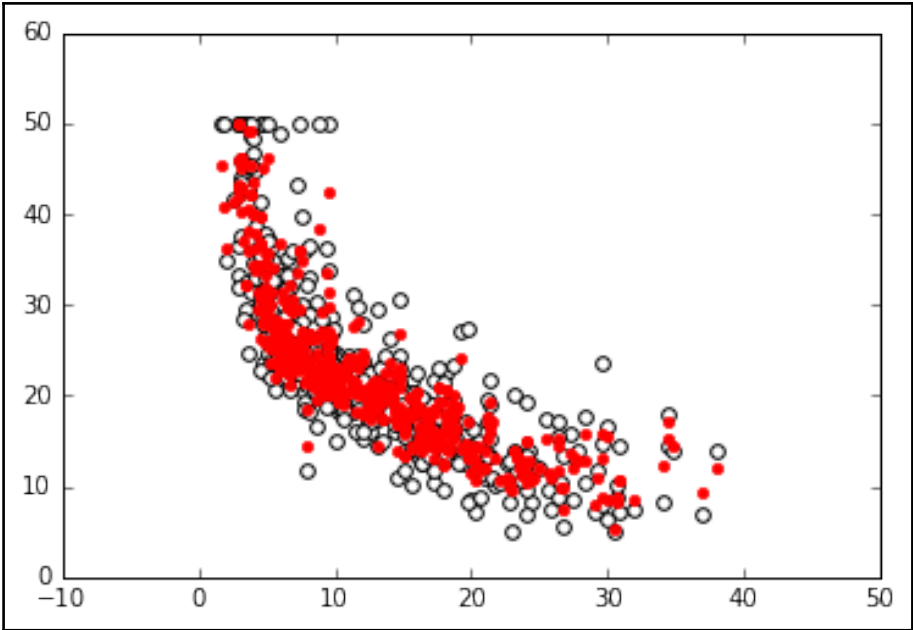
$$|u(t)| = \begin{cases} u(t) & \text{if } t \geq 0 \\ -u(t) & \text{otherwise} \end{cases}$$

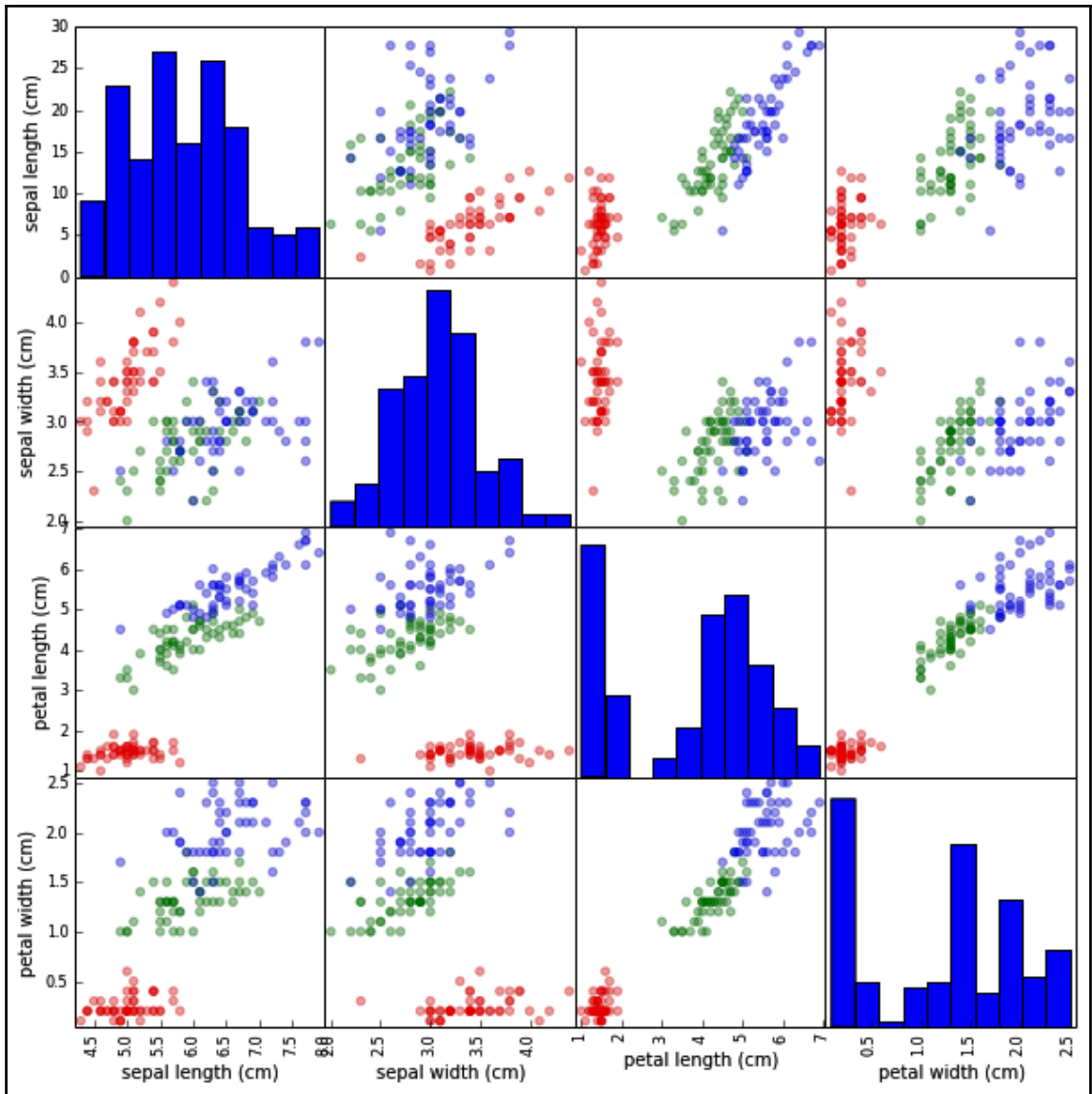
$$\begin{aligned} f(x) &= (a+b)^2 \\ &= a^2 + (a+b) + (a+b) + b^2 \\ &= a^2 + 2 \cdot (a+b) + b^2 \end{aligned}$$



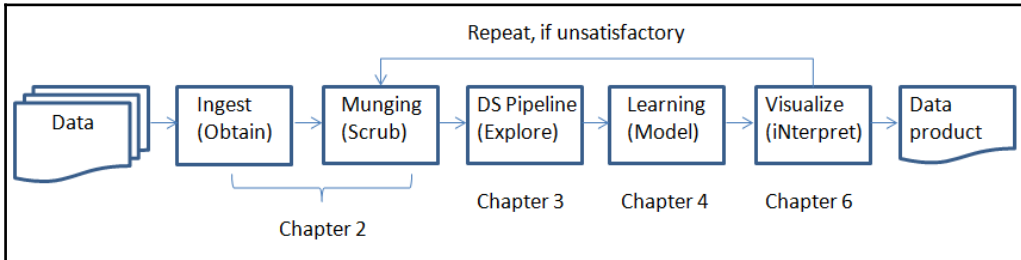








Chapter 2: Data Munging



	sepal_length	sepal_width	petal_length	petal_width	target
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	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

	sepal_length	sepal_width
0	5.1	3.5
1	4.9	3.0
2	4.7	3.2
3	4.6	3.1

146	6.3	2.5
147	6.5	3.0
148	6.2	3.4
149	5.9	3.0

	Date	Temperature_city_1	Temperature_city_2	Temperature_city_3	Which_destination
0	20140910	80.0	32.0	40	1
1	20140911	100.0	50.0	36	2
2	20140912	102.0	55.0	46	1
3	20140913	60.0	20.0	35	3
4	20140914	60.0	NaN	32	3
5	20140915	NaN	57.0	42	2

	Date	Temperature_city_1	Temperature_city_2	Temperature_city_3	Which_destination
0	2014-09-10	80.0	32.0	40	1
1	2014-09-11	100.0	50.0	36	2
2	2014-09-12	102.0	55.0	46	1
3	2014-09-13	60.0	20.0	35	3
4	2014-09-14	60.0	NaN	32	3
5	2014-09-15	NaN	57.0	42	2

	Date	Temperature_city_1	Temperature_city_2	Temperature_city_3	Which_destination
0	2014-09-10	80.0	32.0	40	1
1	2014-09-11	100.0	50.0	36	2
2	2014-09-12	102.0	55.0	46	1
3	2014-09-13	60.0	20.0	35	3
4	2014-09-14	60.0	50.0	32	3
5	2014-09-15	50.0	57.0	42	2

	Val1	Val2	Val3
0	0	0	0
1	1	1	1
2	3	3	3

	C1	C2	C3	C4	C5
0	4.8	3.1	1.6	0.2	Iris-setosa
1	5.4	3.4	1.5	0.4	Iris-setosa

	Col1	Col2	Col3	Col4
0	0	1.0	1.0	Hello World!
1	1	1.0	1.0	Hello World!
2	2	1.0	1.0	Hello World!
3	3	1.0	1.0	Hello World!
4	4	1.0	1.0	Hello World!

	sepal_length	sepal_width	petal_length	petal_width
target				
Iris-setosa	5.006	3.418	1.464	0.244
Iris-versicolor	5.936	2.770	4.260	1.326
New label	6.588	2.974	5.552	2.026

	sepal_length	sepal_width	petal_length	petal_width
target				
Iris-setosa	0.124249	0.145180	0.030106	0.011494
Iris-versicolor	0.266433	0.098469	0.220816	0.039106
New label	0.404343	0.104004	0.304588	0.075433

	sepal_length	sepal_width	petal_length	petal_width	target
13	4.3	3.0	1.1	0.1	Iris-setosa
42	4.4	3.2	1.3	0.2	Iris-setosa
38	4.4	3.0	1.3	0.2	Iris-setosa
8	4.4	2.9	1.4	0.2	Iris-setosa
41	4.5	2.3	1.3	0.3	Iris-setosa

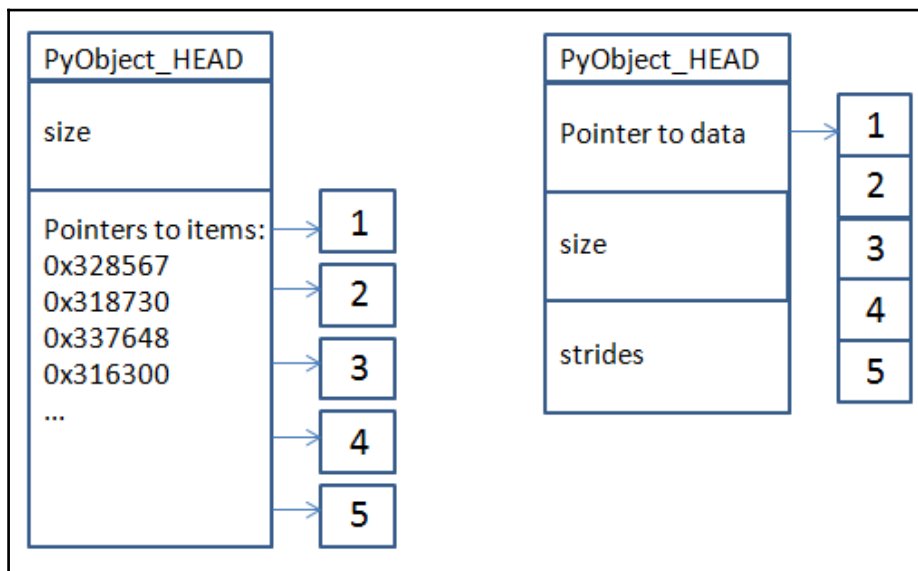
	sepal_length	sepal_width	petal_length	petal_width	target
0	3	3	3	3	11
1	3	3	3	3	11
2	3	3	3	3	11
3	3	3	3	3	11
4	3	3	3	3	11

	n	val1	val2	val3
0	100	10	10	C
1	101	10	20	C
2	102	10	30	B
3	103	10	40	B
4	104	10	50	A

	val1	val2	val3
n			
100	10	10	C
101	10	20	C
102	10	30	B
103	10	40	B
104	10	50	A

	val3	val2
n		
100	C	10
101	C	20

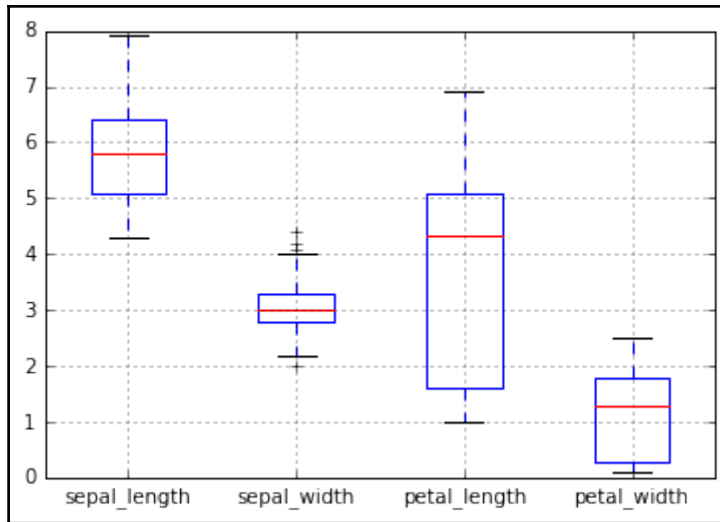
	cloudy	foggy	rainy	snowy	sunny
0	0.0	0.0	0.0	0.0	1.0
1	1.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	1.0	0.0
3	0.0	0.0	1.0	0.0	0.0
4	0.0	1.0	0.0	0.0	0.0



Chapter 3: The Data Pipeline

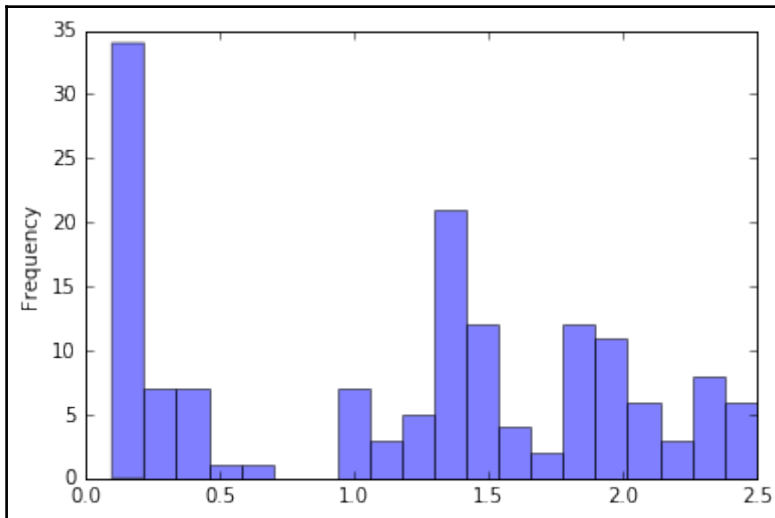
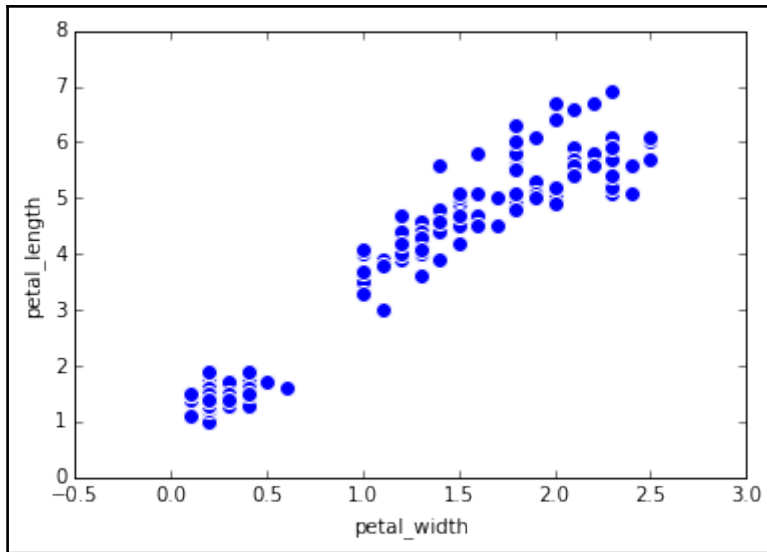
	sepal_length	sepal_width	petal_length	petal_width	target
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

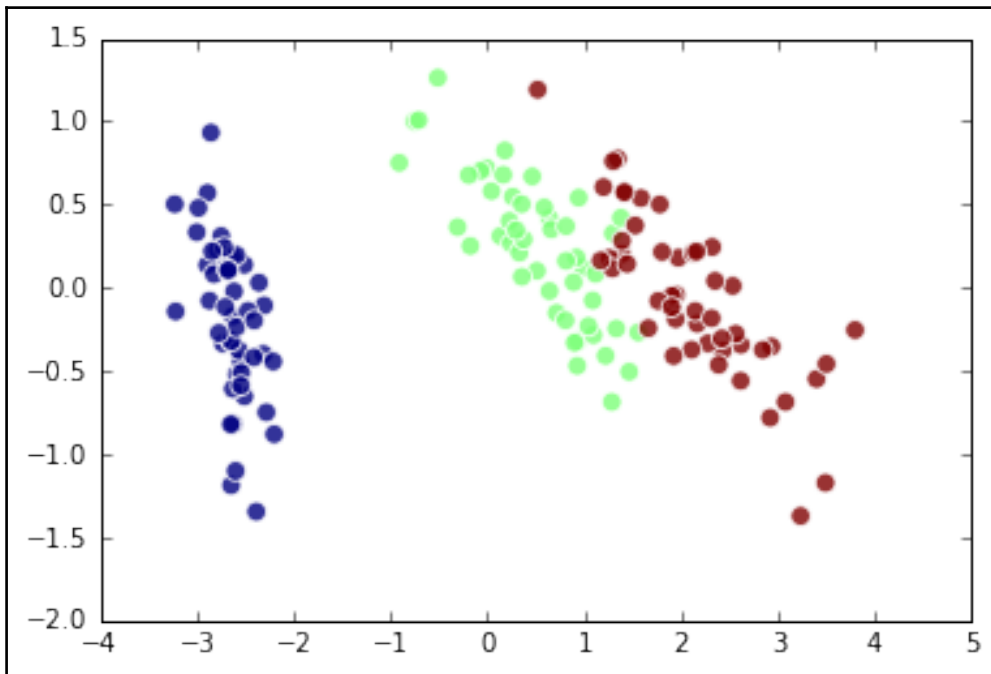
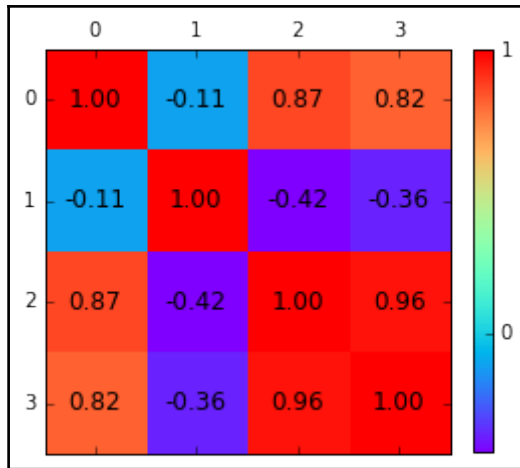
	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

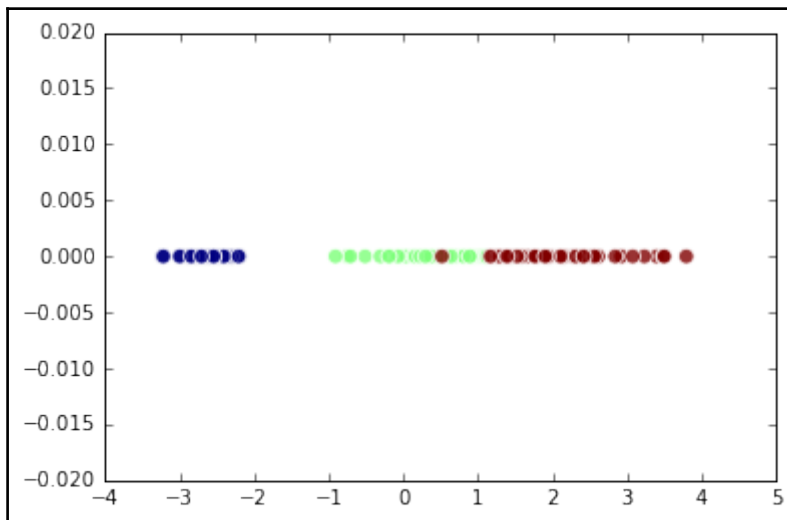
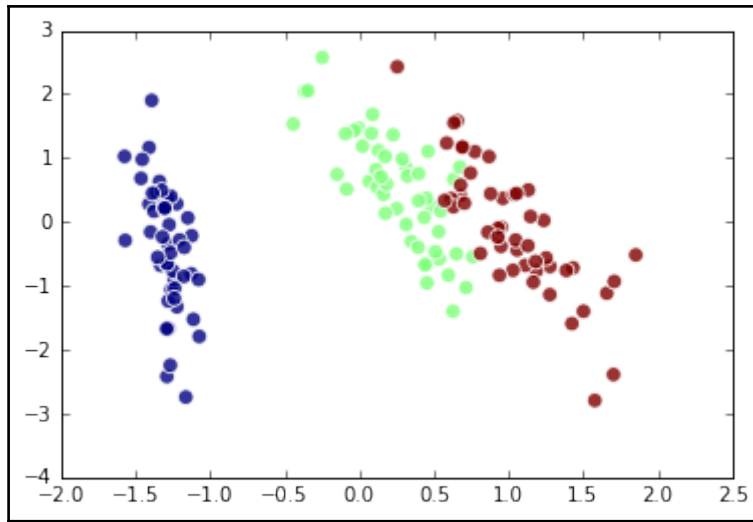


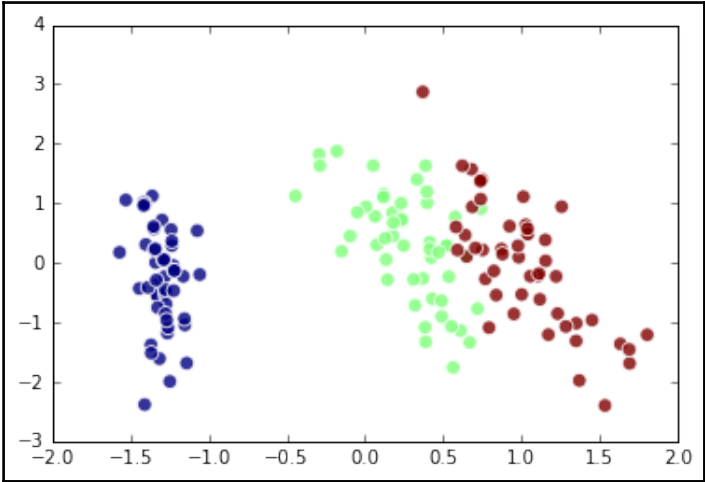
	sepal_length	sepal_width	petal_length	petal_width
0.1	4.8	2.50	1.4	0.2
0.9	6.9	3.61	5.8	2.2

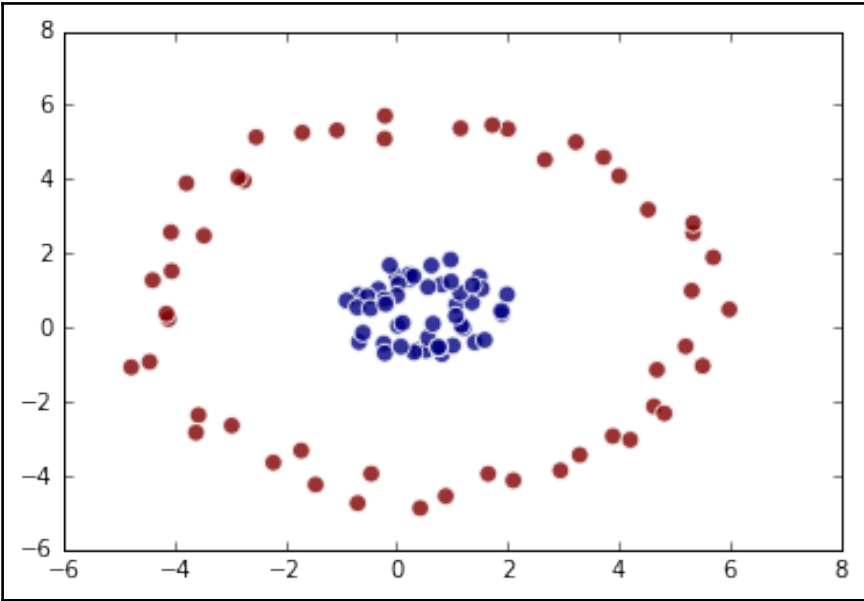
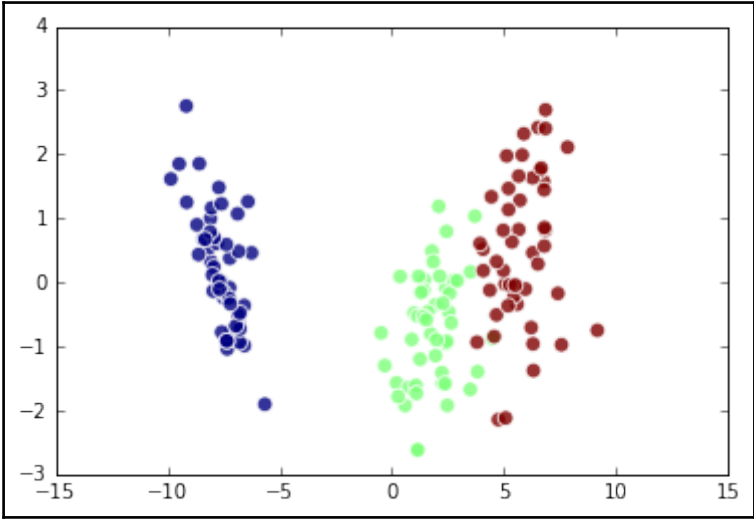
petal_width	False	True
petal_length		
False	56	1
True	4	89

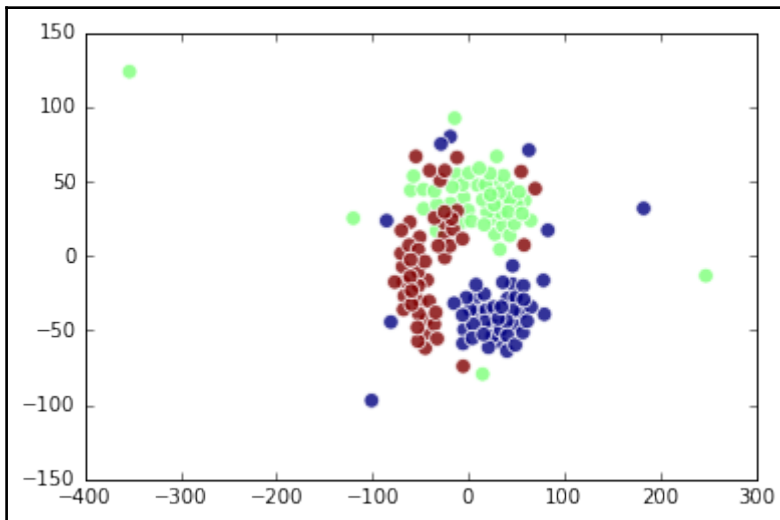
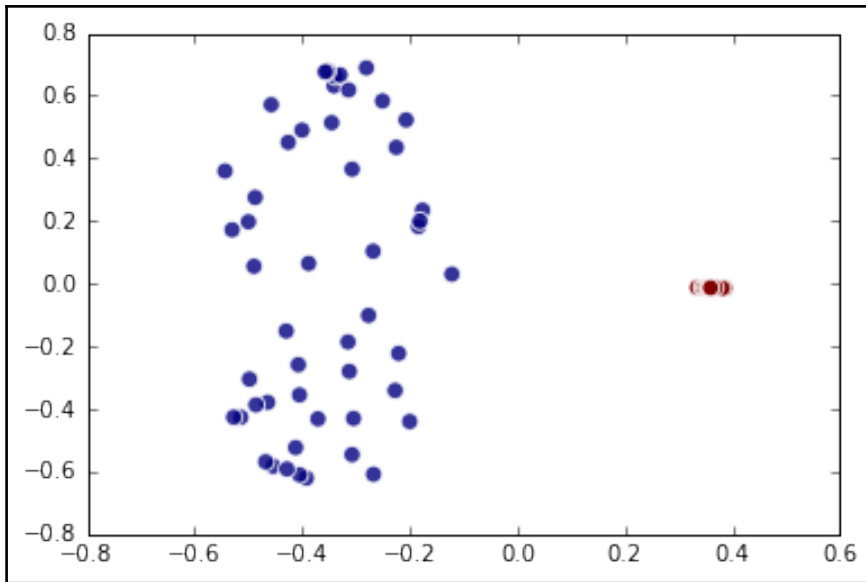


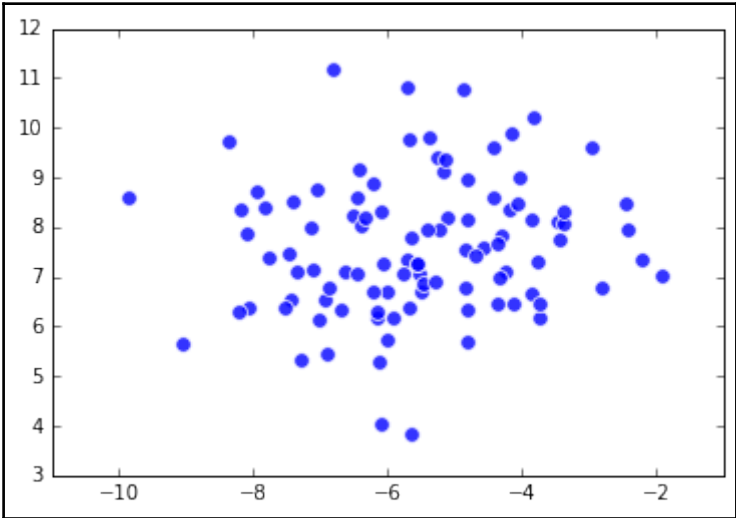
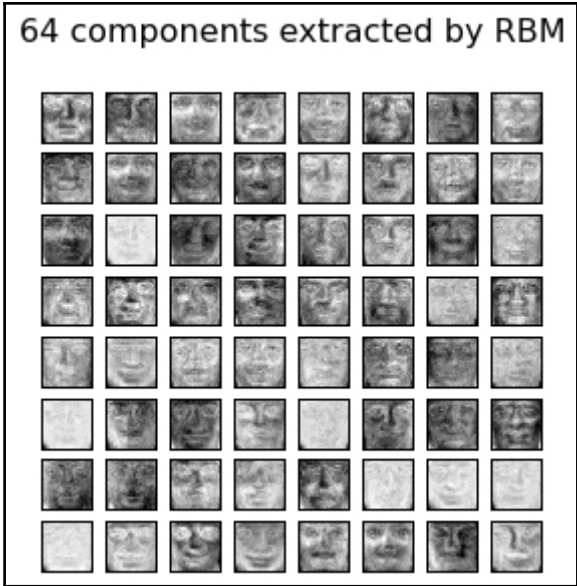


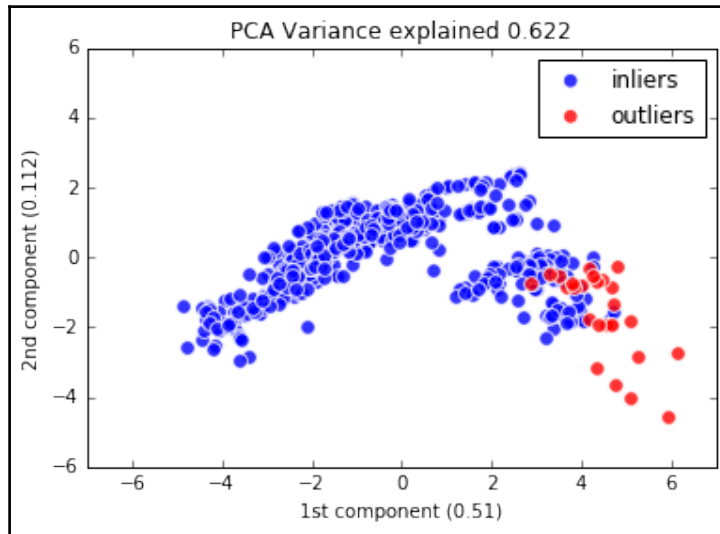
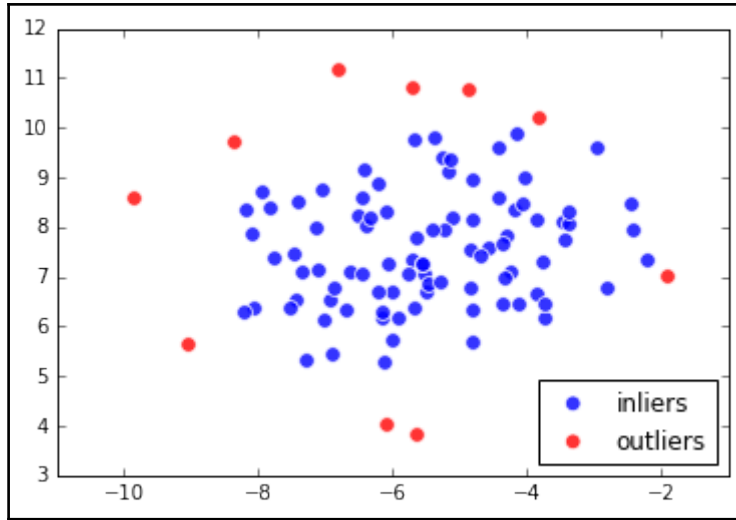


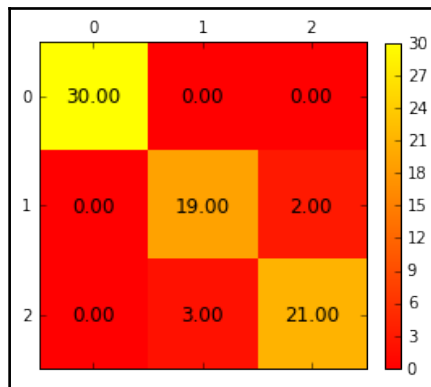
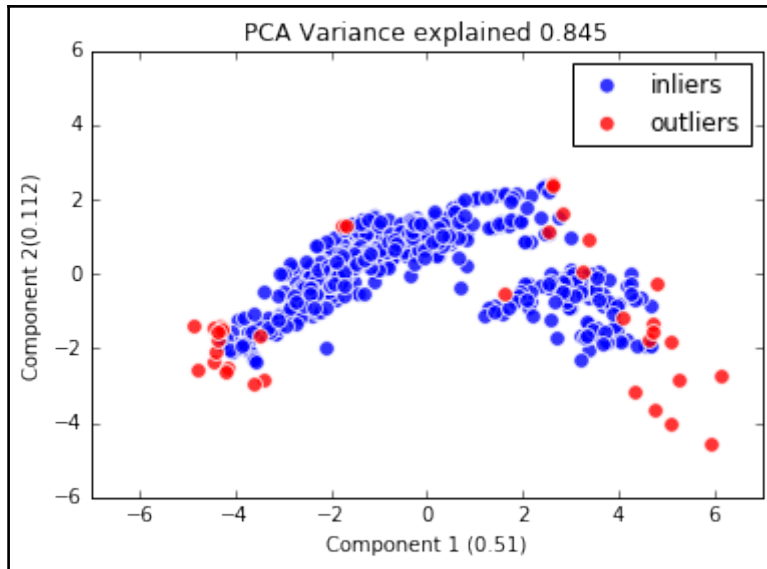




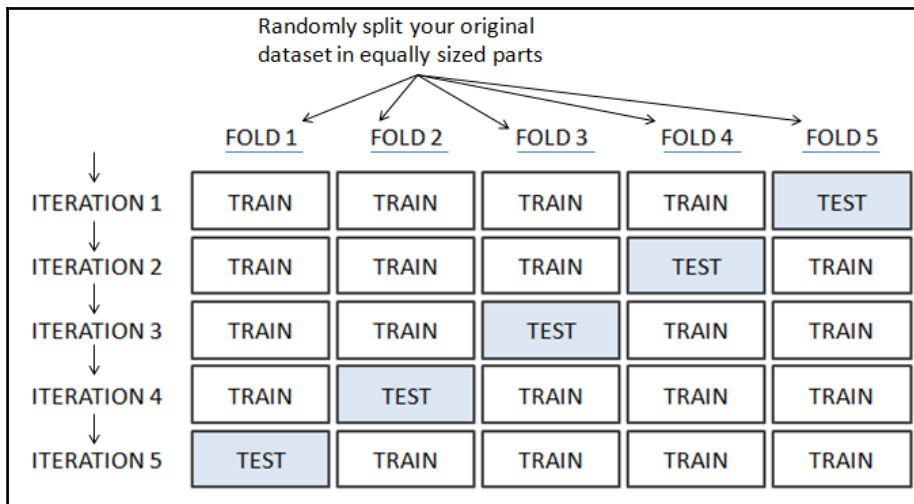
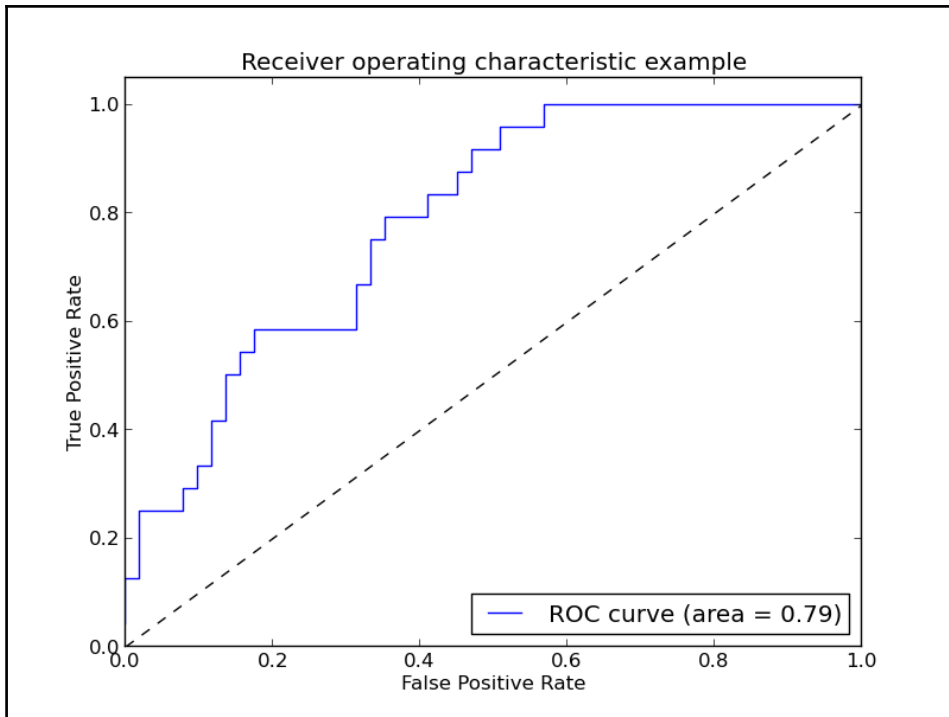








	precision	recall	f1-score	support
setosa	1.00	1.00	1.00	30
versicolor	0.90	0.86	0.88	22
virginica	0.88	0.91	0.89	23
avg / total	0.93	0.93	0.93	75



Chapter 4: Machine Learning

$$\begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(n) \end{bmatrix} * W = \begin{bmatrix} Y(0) \\ Y(1) \\ \vdots \\ Y(n) \end{bmatrix}$$

$$y_pred_i = sign(X_i * W)$$

$$Prob(y_i = +1 | X_i) = logistic(X_i \cdot W)$$

$$logistic(\alpha) = e^\alpha / (1 + e^\alpha)$$

$$Prob(Ck | X(i))$$

$$Prob(Ck | X(i)) = \frac{Prob(Ck) Prob(X(i) | Ck)}{Prob(X(i))}$$

$$Prob(Ck, X(i,0), \dots, X(i,n)) = Prob(X(i,0), \dots, X(i,n) | Ck)$$

$$Prob(X(i,0) | Ck) Prob(X(x,1), \dots, X(i,n) | Ck, X(i,0))$$

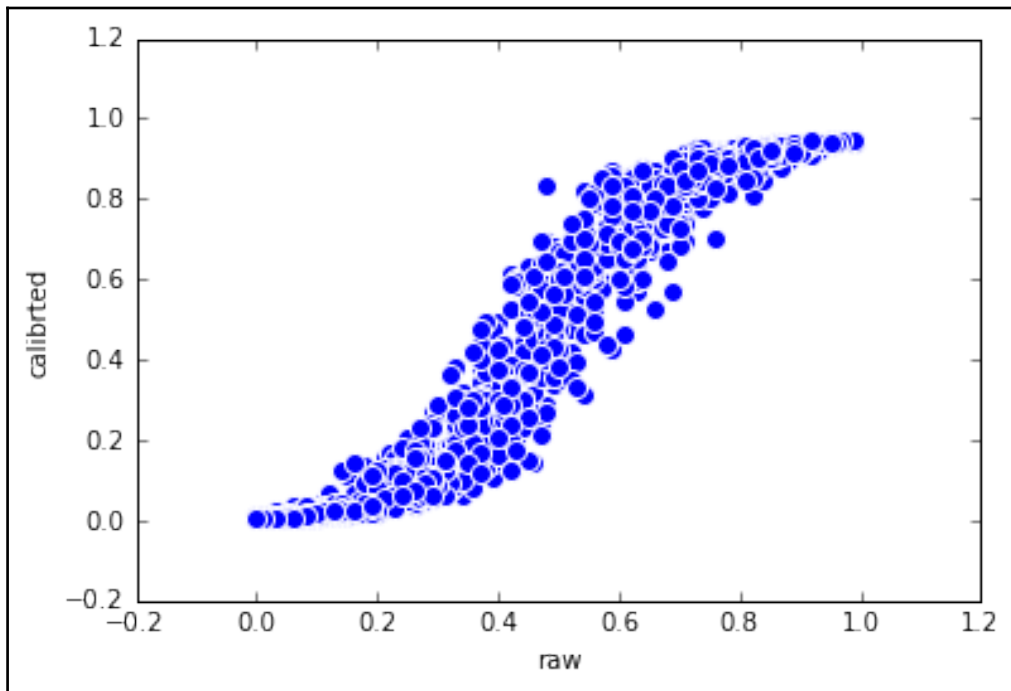
$$Prob(Ck, X(i,0), \dots, X(i,n)) =$$

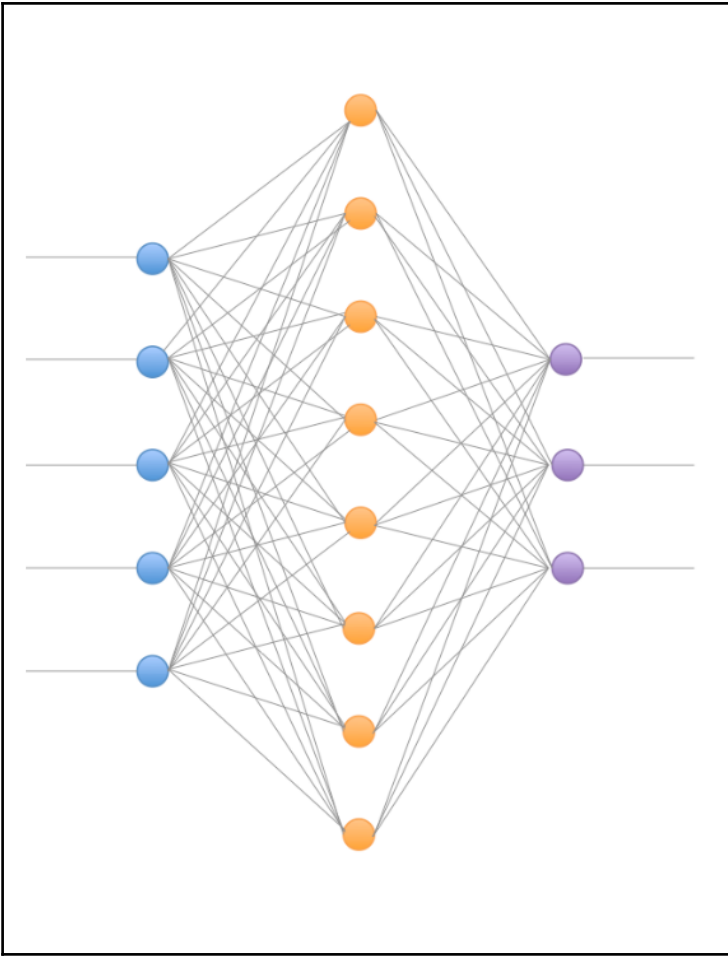
...

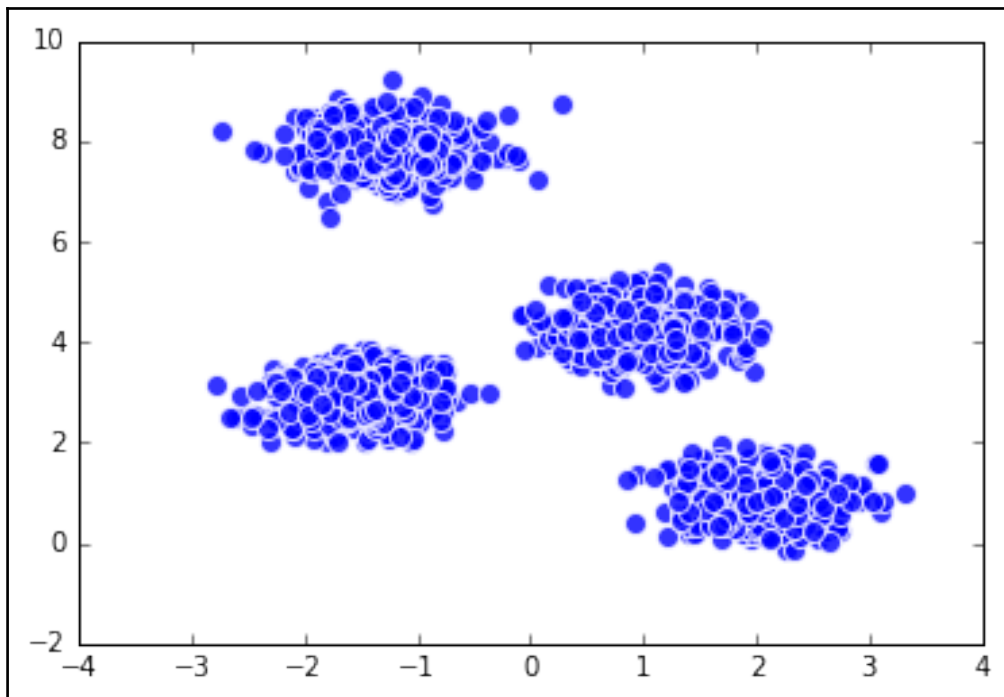
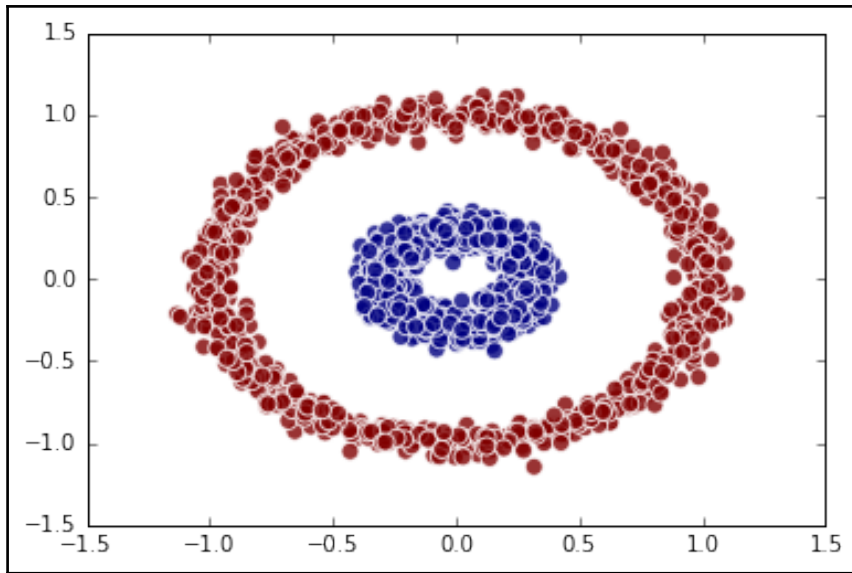
$$Prob(Ck) Prob(X(i,0) | Ck) Prob(X(i,1) | Ck, X(i,0)) \dots$$

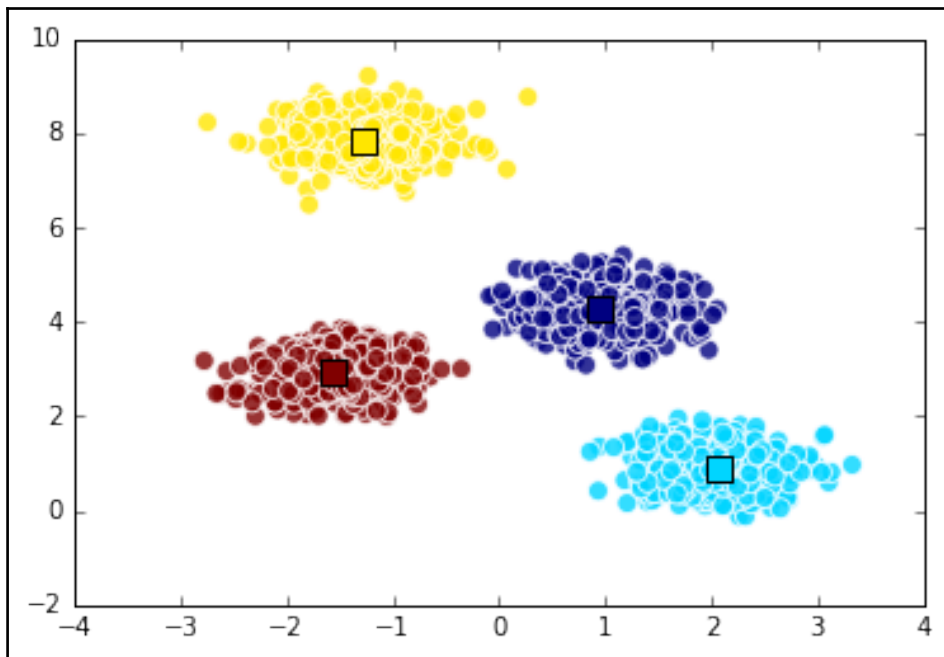
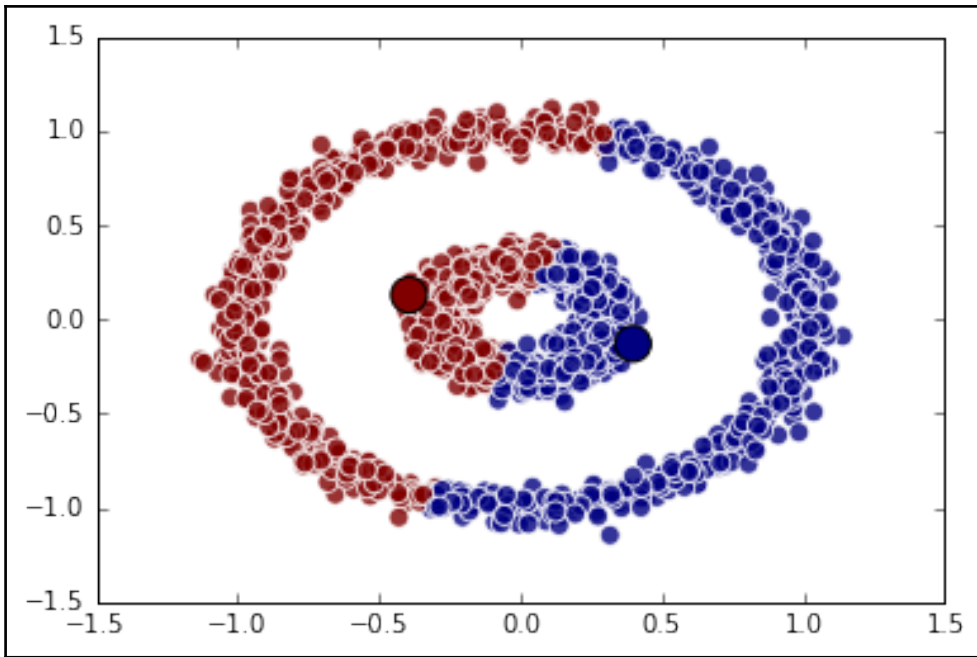
$$Prob(X(i,0) | Ck, X(:)) = Prob(X(i,0) | Ck)$$

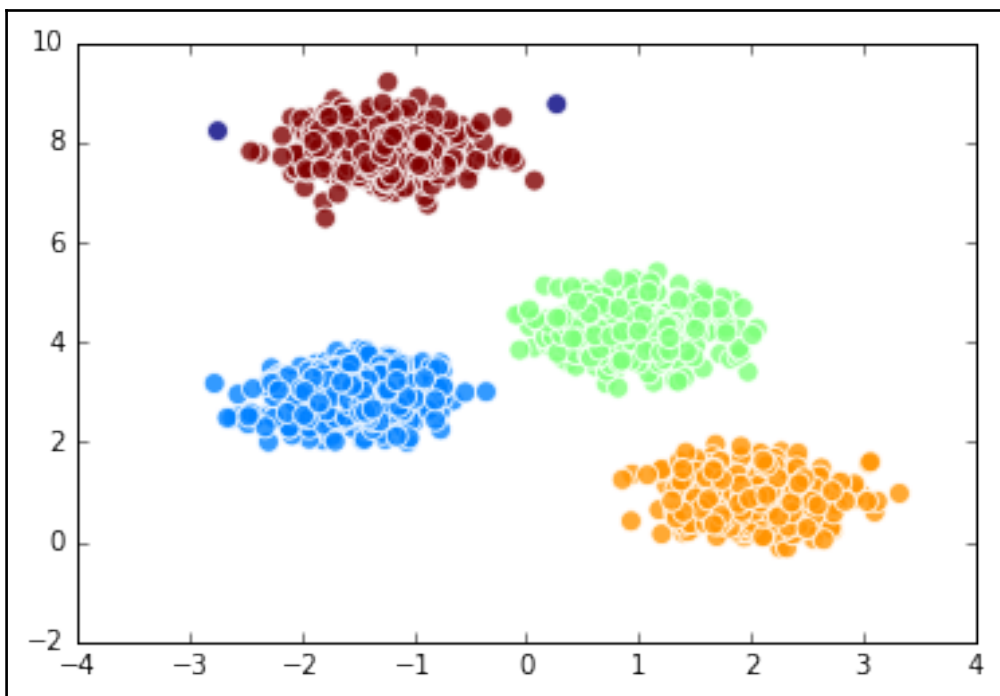
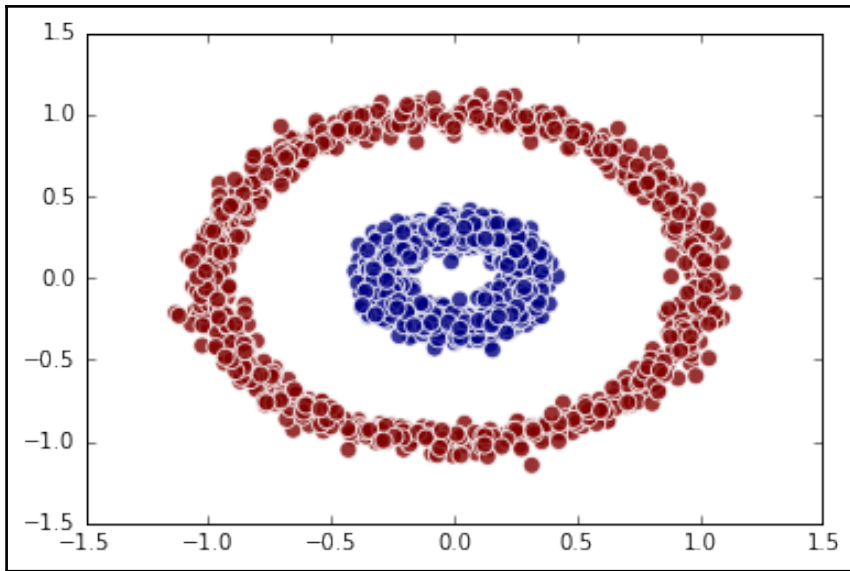
$$Y_{pred}(i) = \underset{k=0,1,2,3}{\operatorname{argmax}} \operatorname{Prob}(C_k) \prod_{k=0}^{n-1} \operatorname{Prob}(X(i,k) | C_k)$$



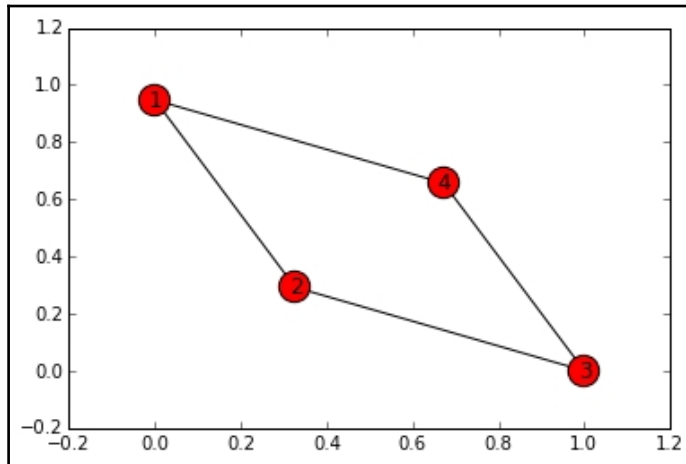
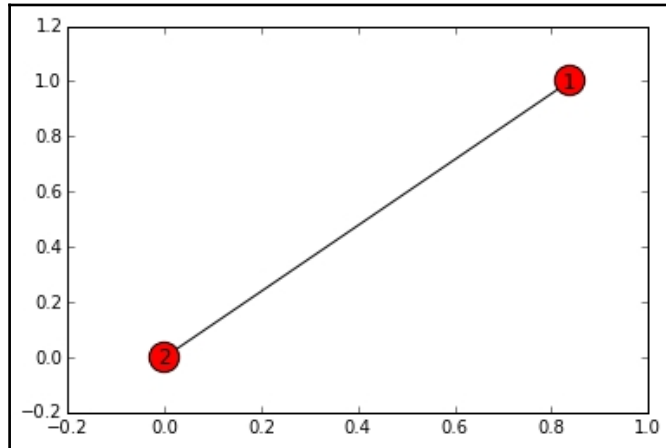




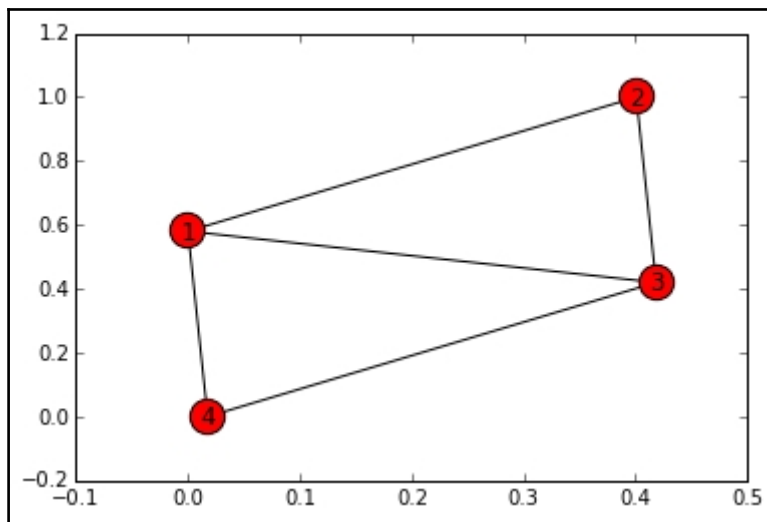


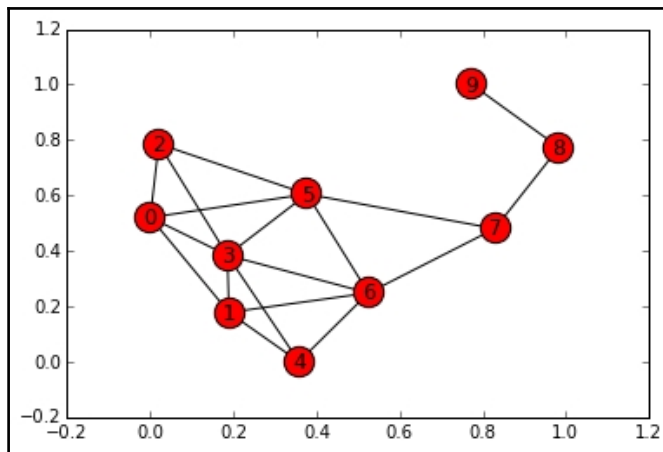
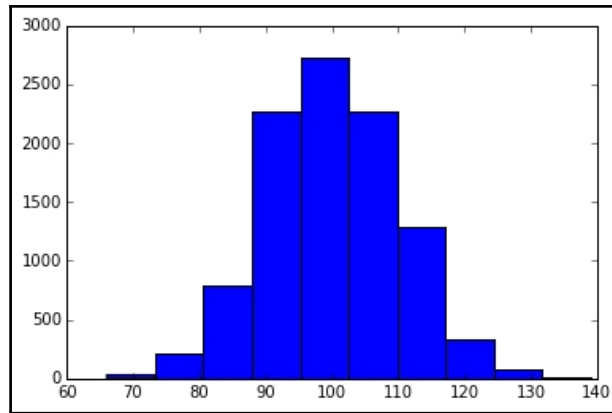


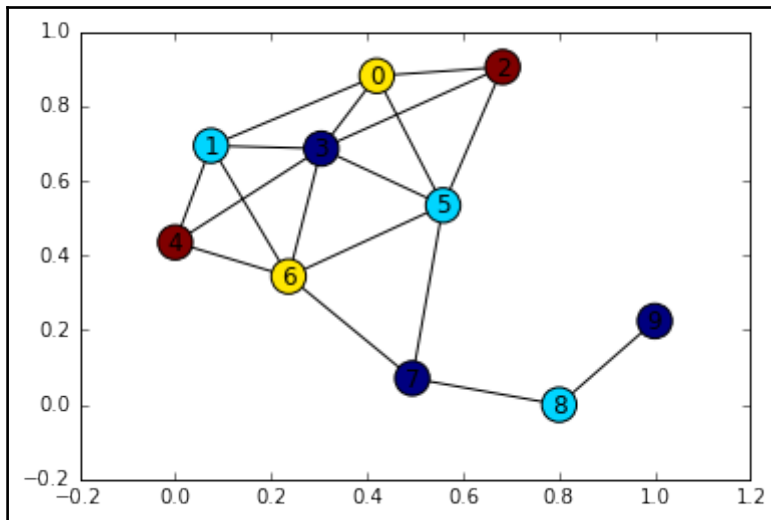
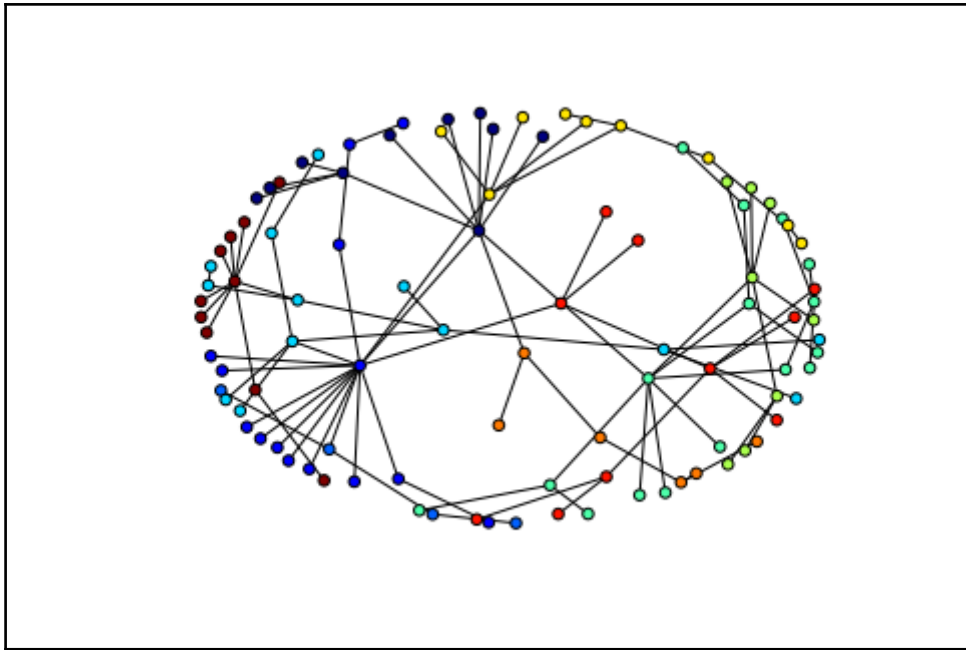
Chapter 5: Social Network Analysis

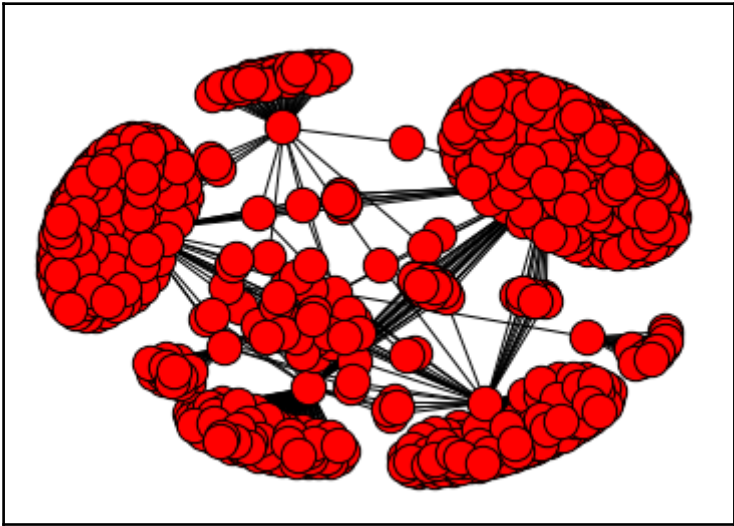
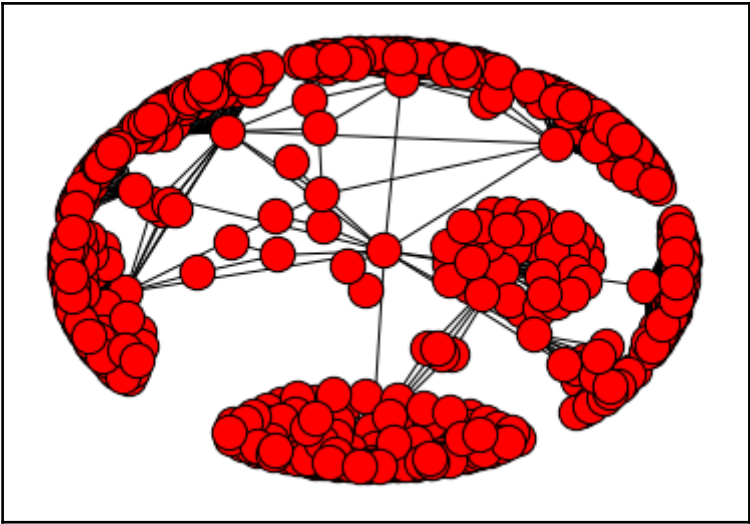


	0	1	2	3	4	5	6	7	8	9
0	0.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0
1	1.0	0.0	0.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0
2	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0
3	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0
4	0.0	1.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0
5	1.0	0.0	1.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0
6	0.0	1.0	0.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	1.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0

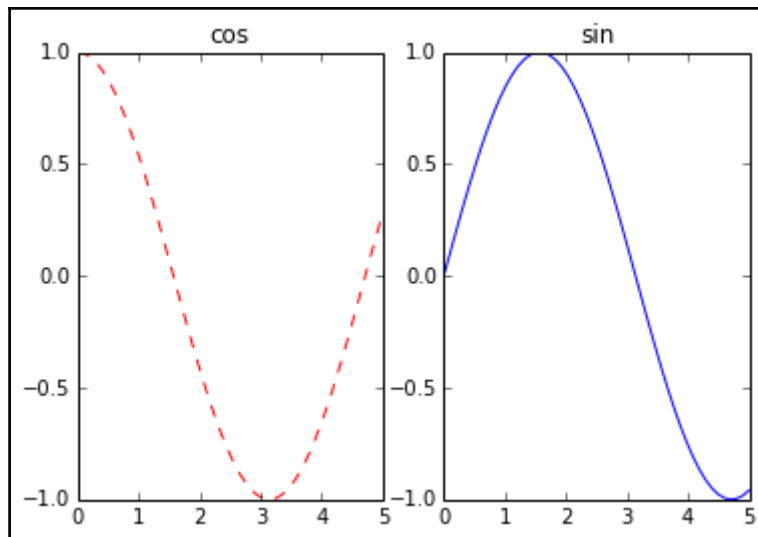
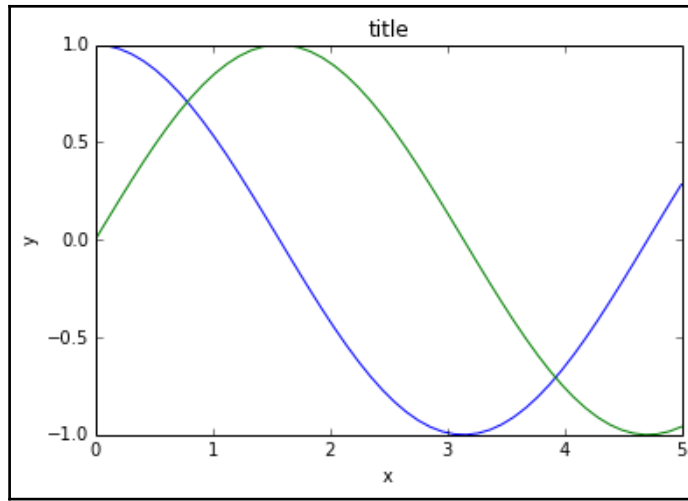


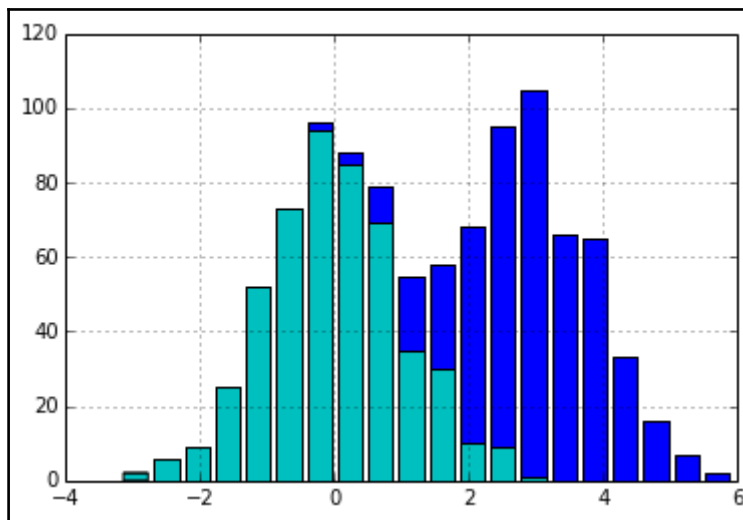
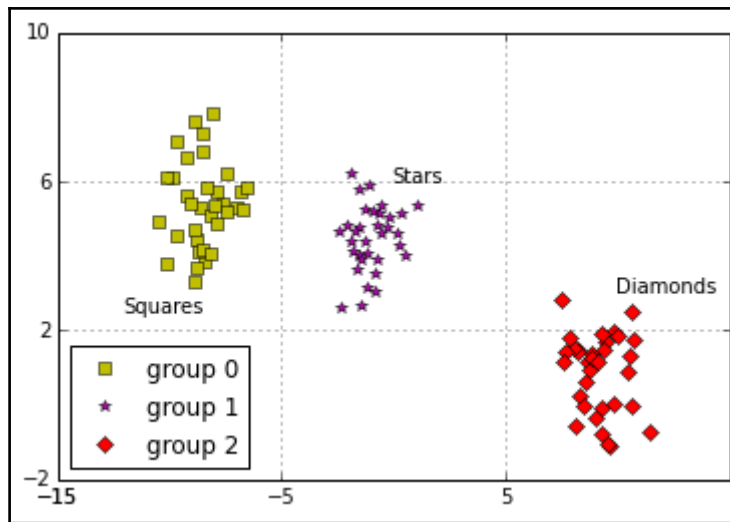






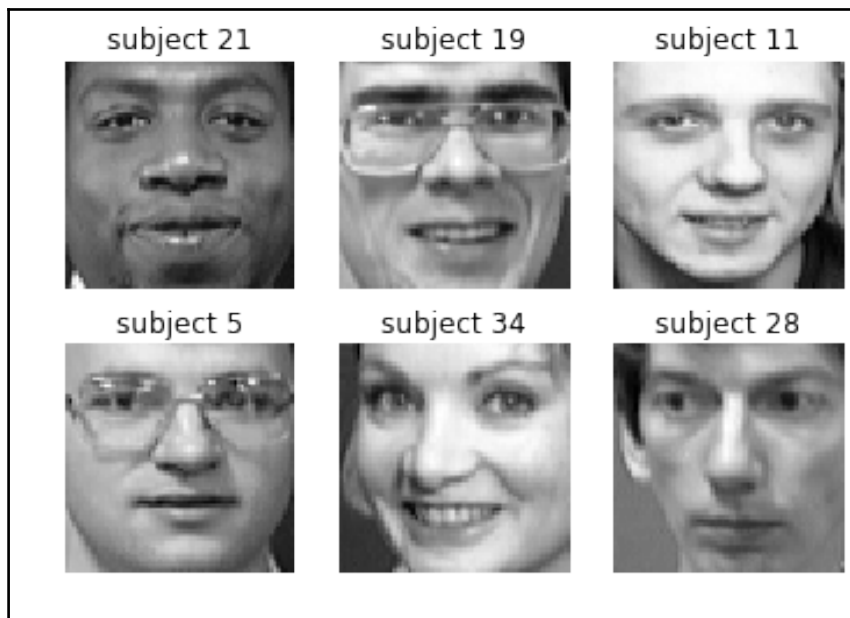
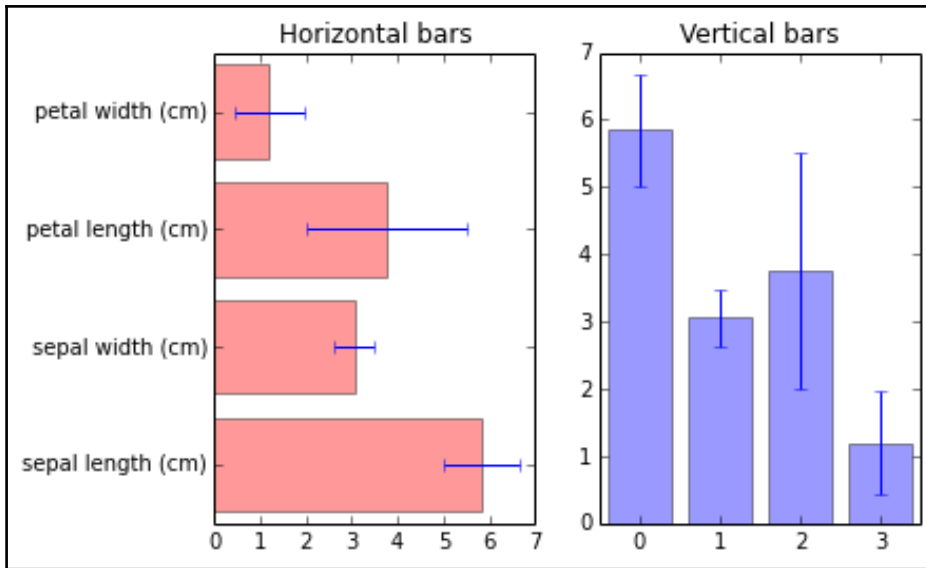
Chapter 6: Visualization, Insights, and Results

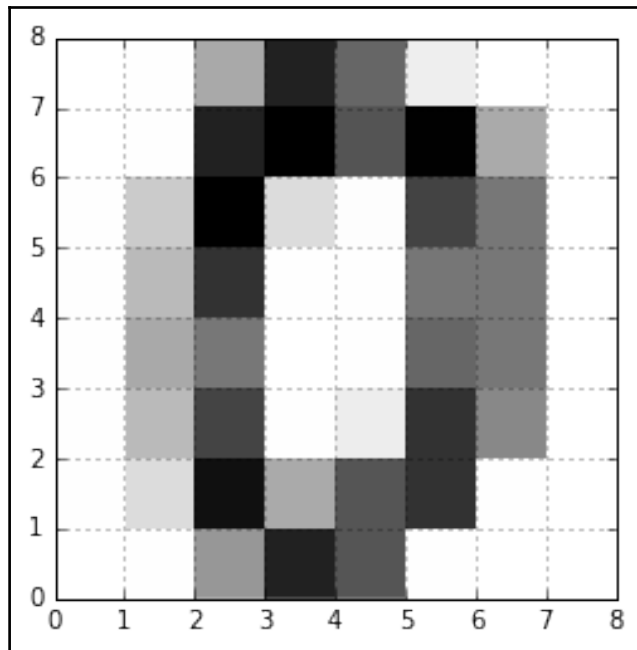
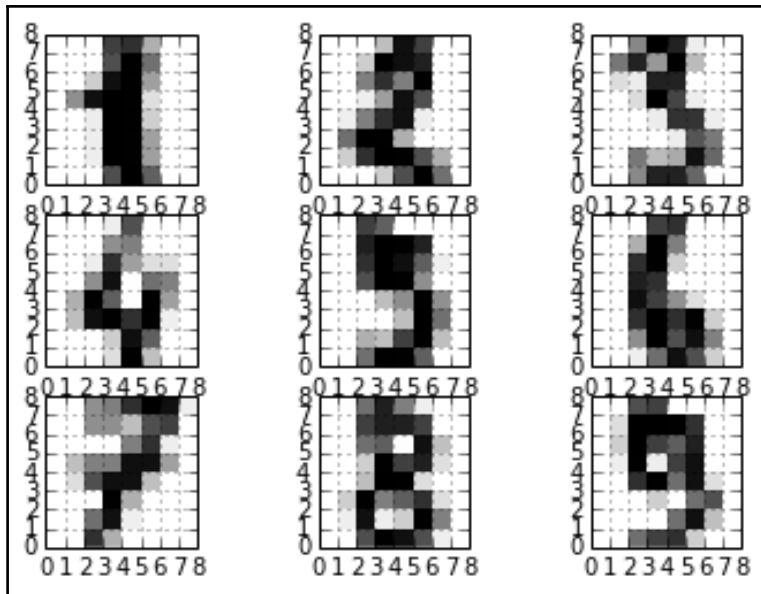


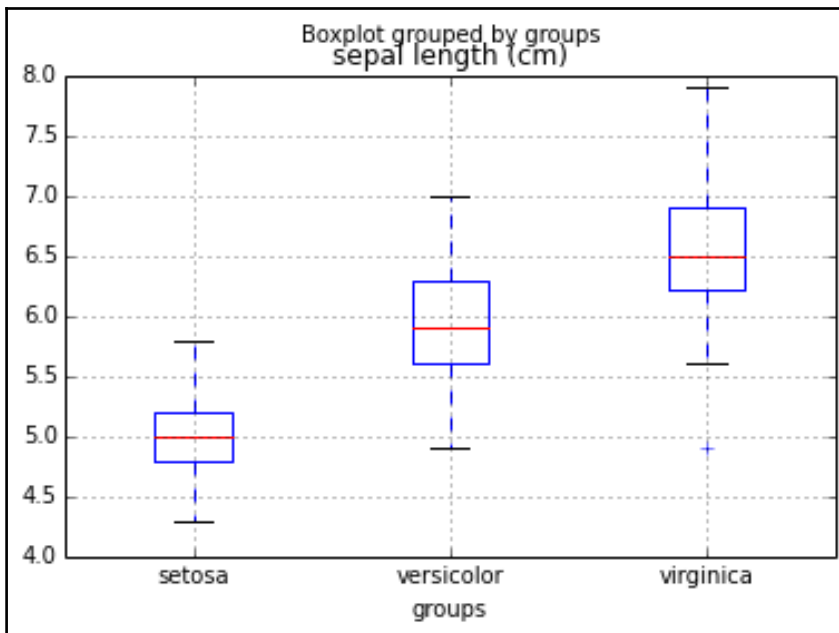
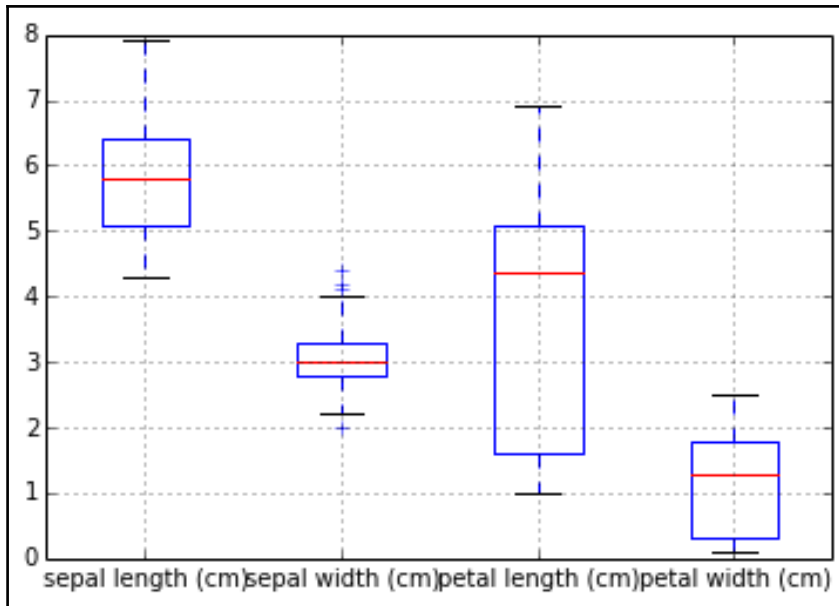


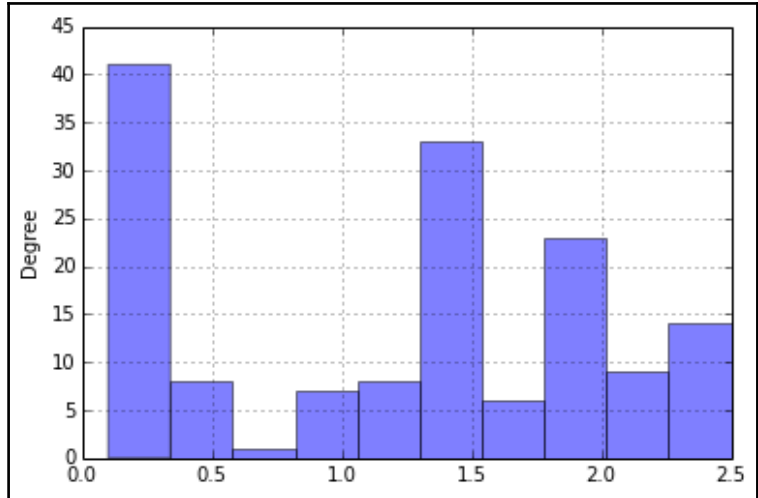
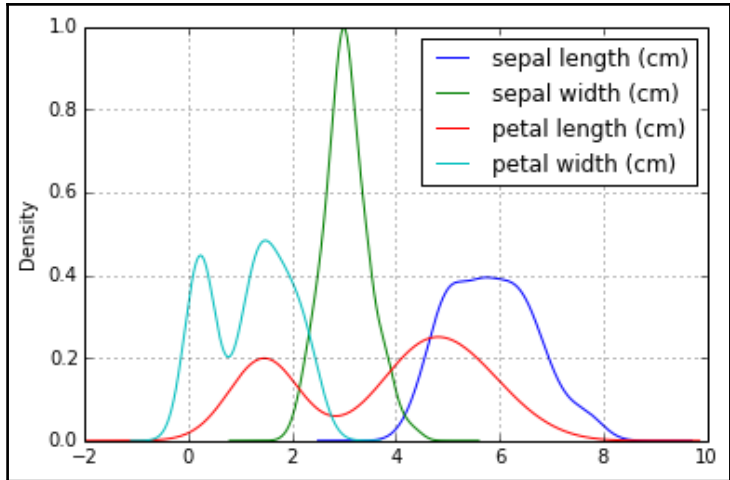
$$h = 2 * IQR * n^{-1/3}$$

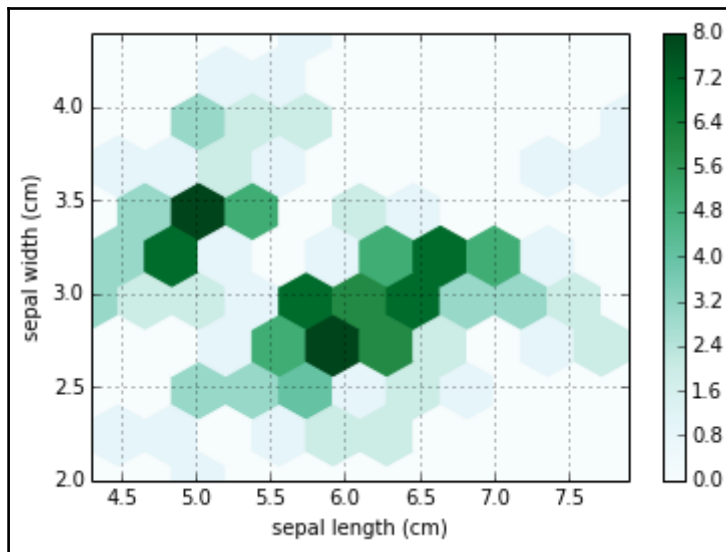
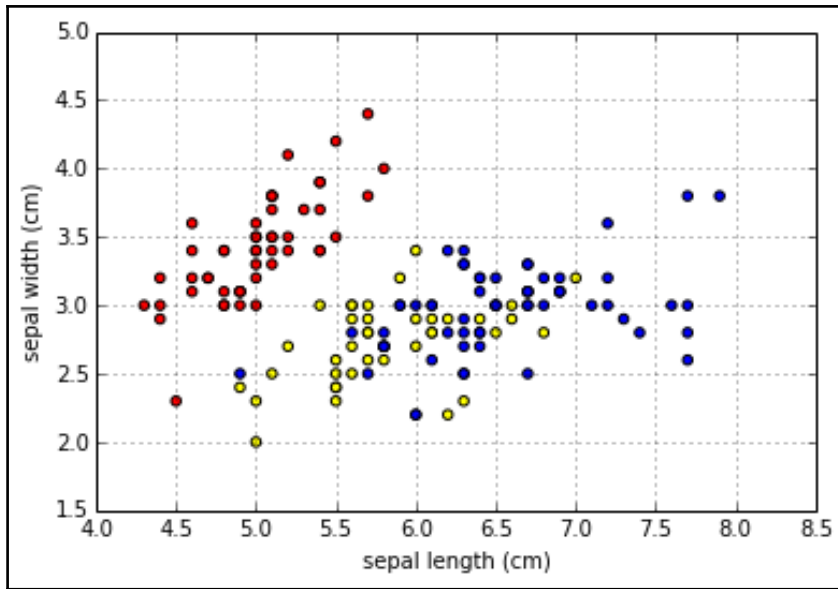
$$bins = (max - min) / h$$

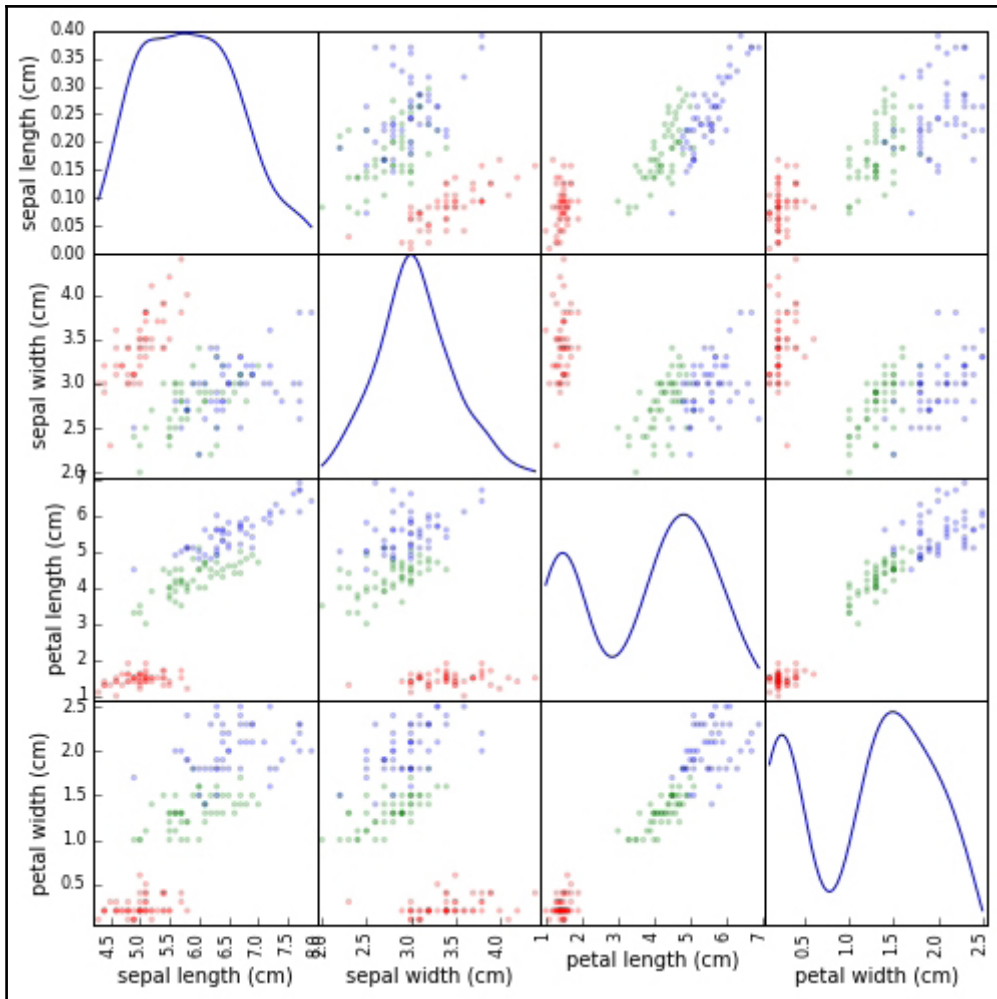


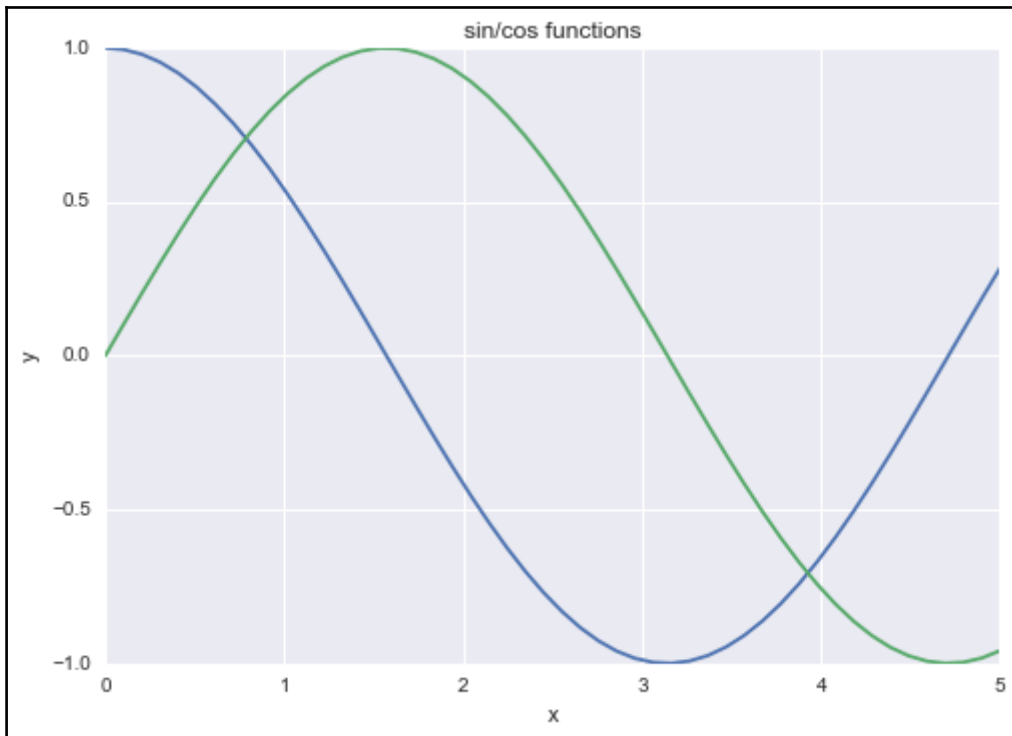
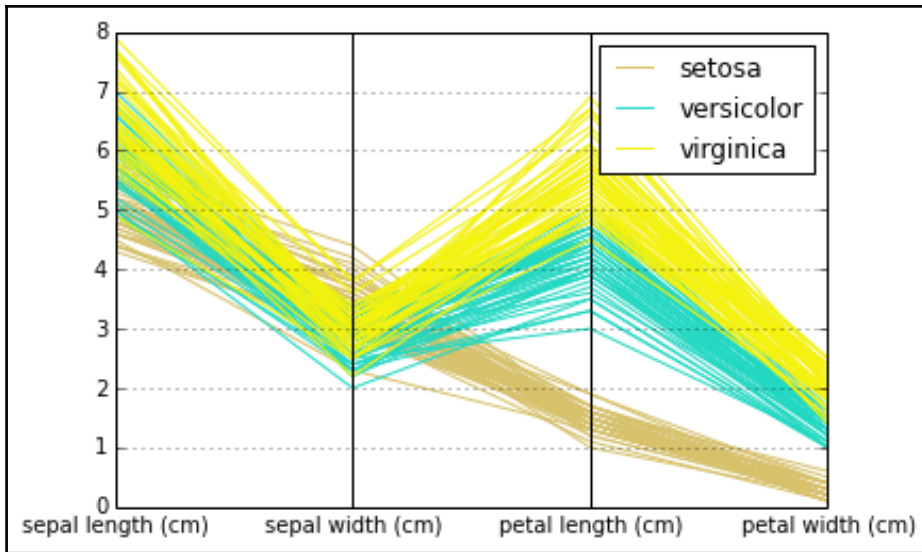


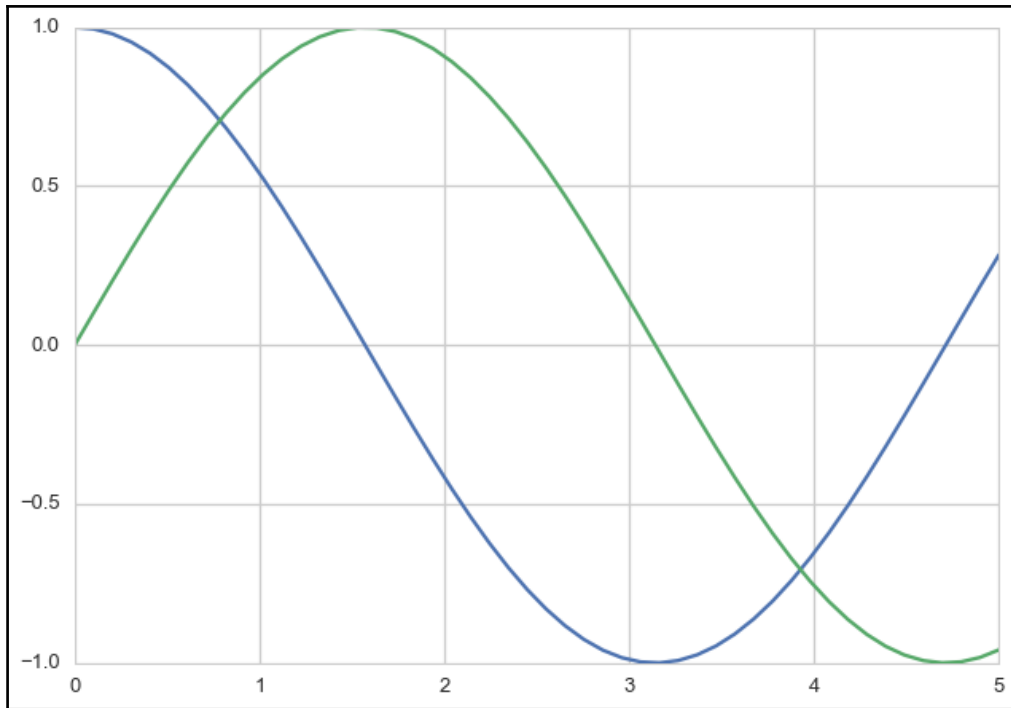






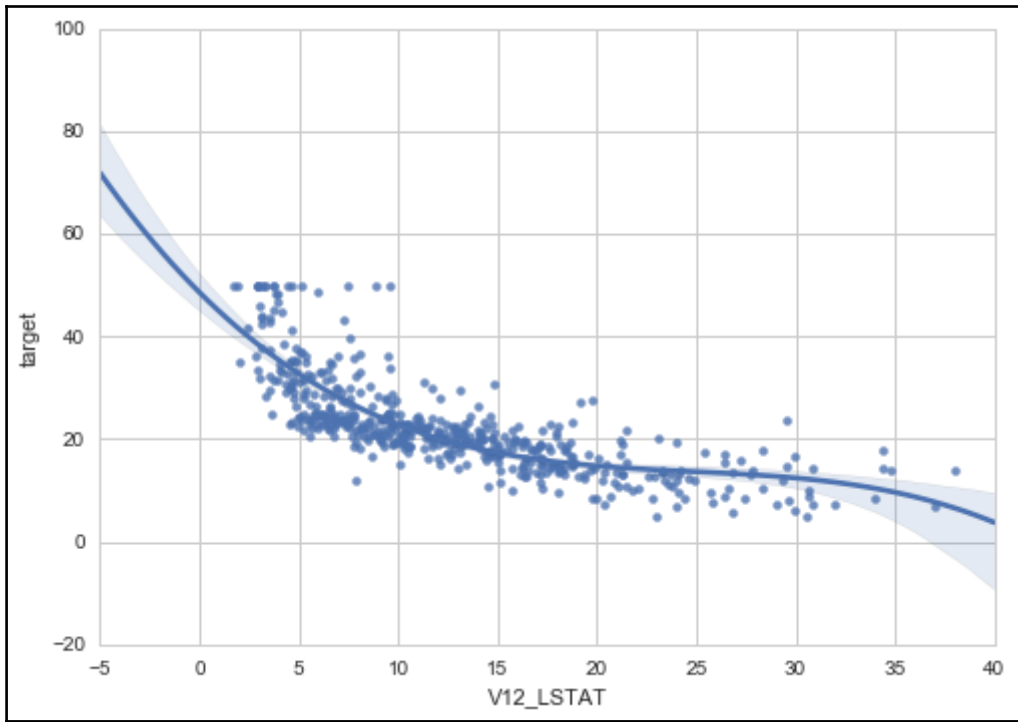
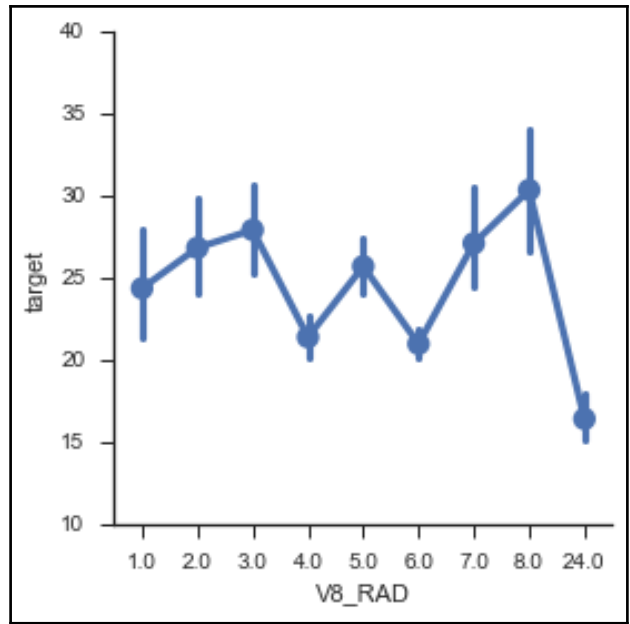


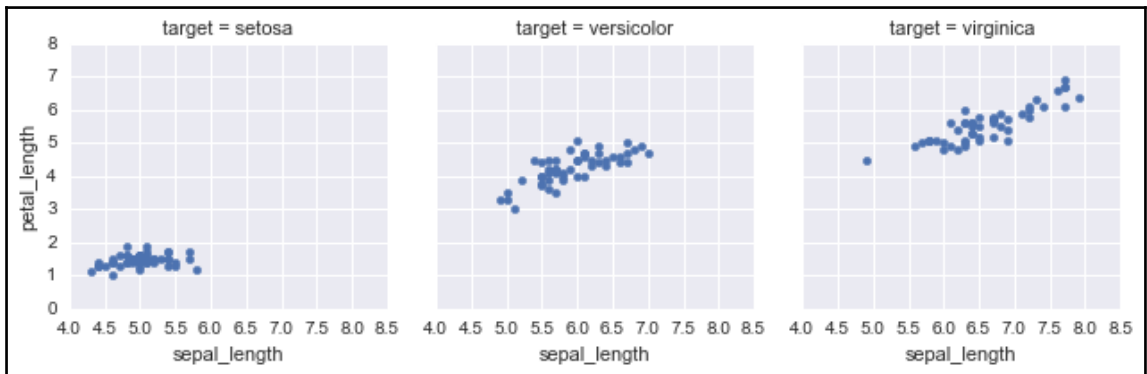
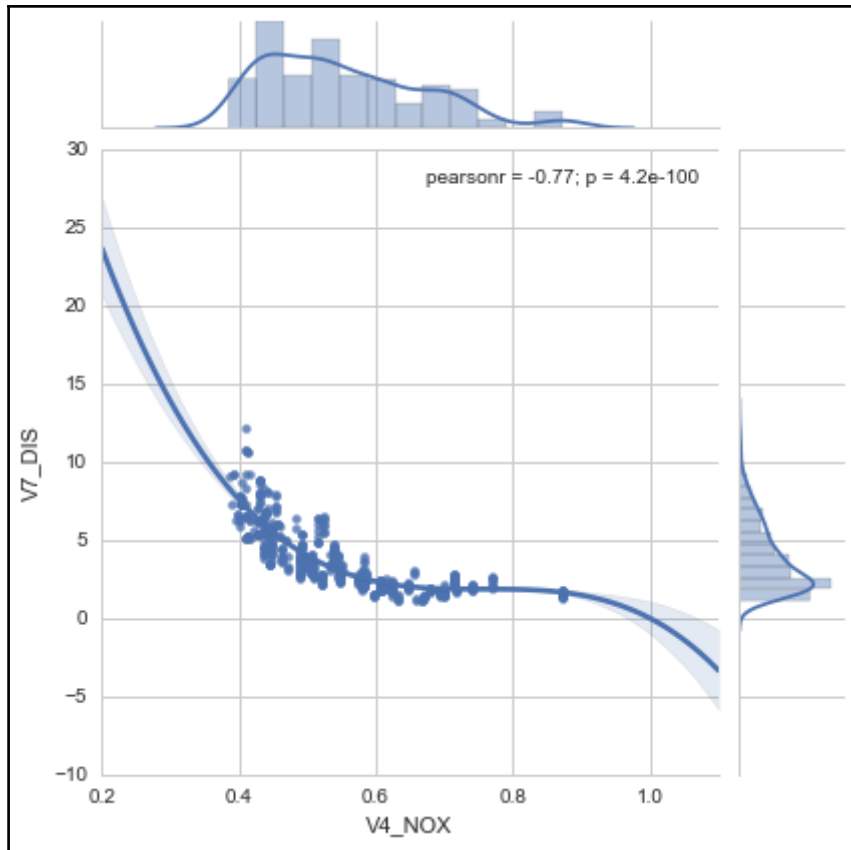


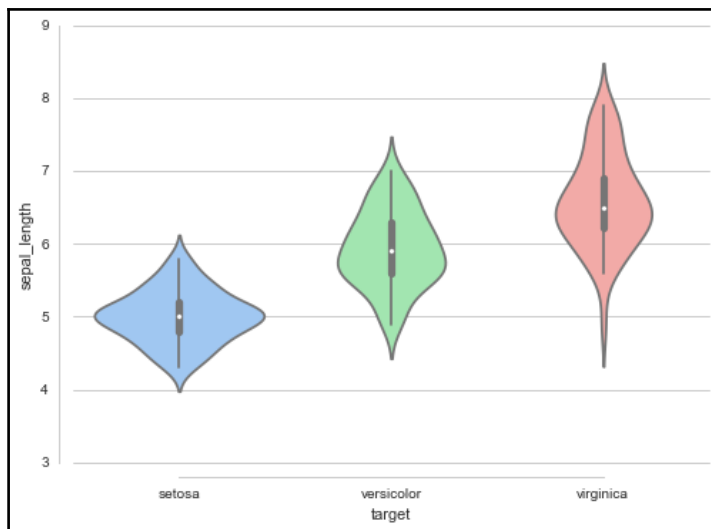
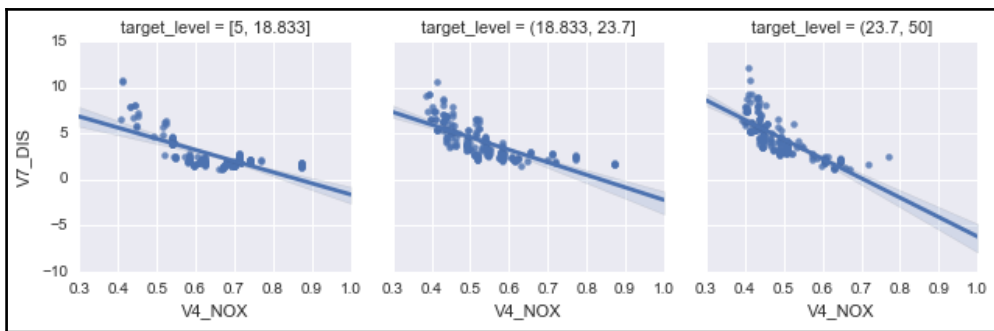
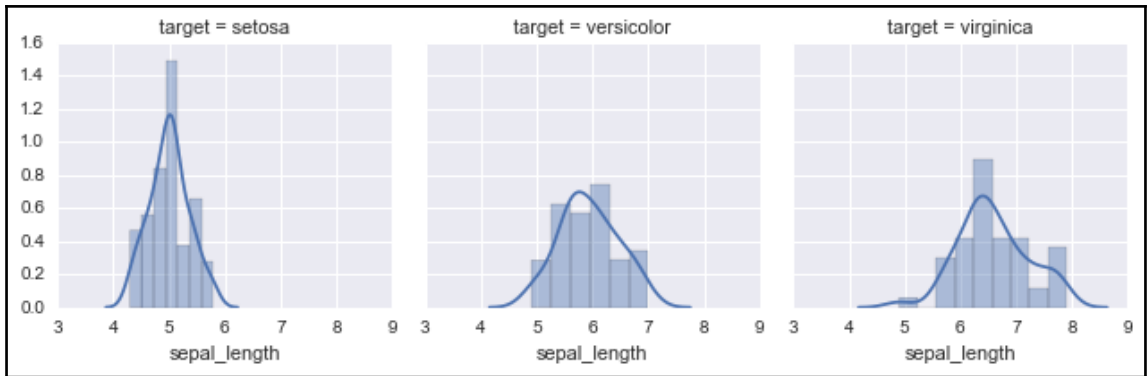


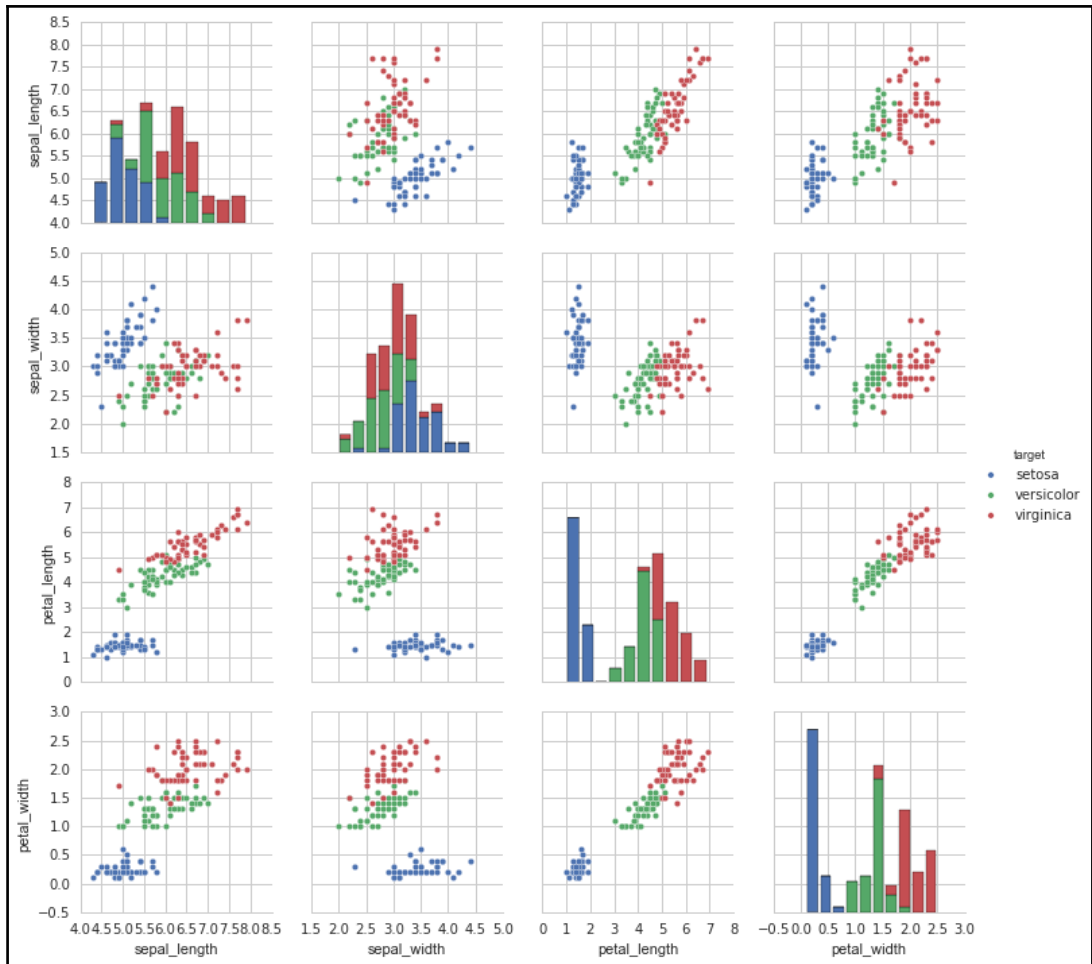
```
[(0.2980392156862745, 0.4470588235294118, 0.6901960784313725), (0.3333333333333333, 0.6588235294117647, 0.40784313725490196), (0.7686274509803922, 0.3058823529411765, 0.3215686274509804), (0.5058823529411764, 0.4470588235294118, 0.6980392156862745), (0.8, 0.7254901960784313, 0.4549019607843137), (0.39215686274509803, 0.7098039215686275, 0.803921568627451)]
```

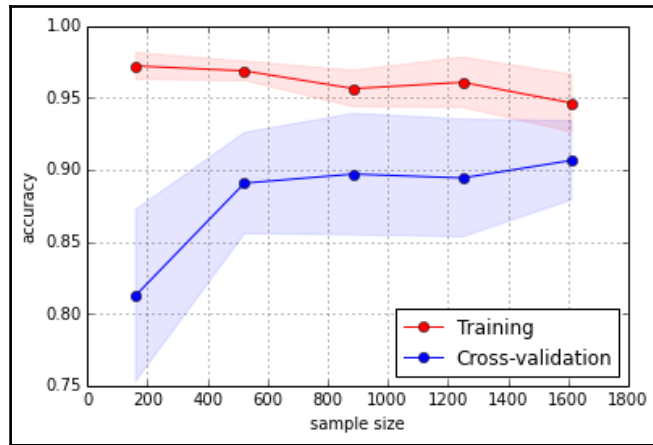
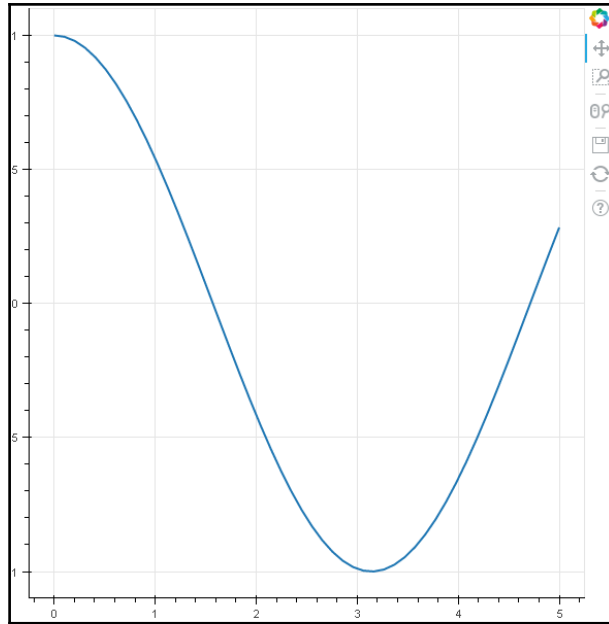


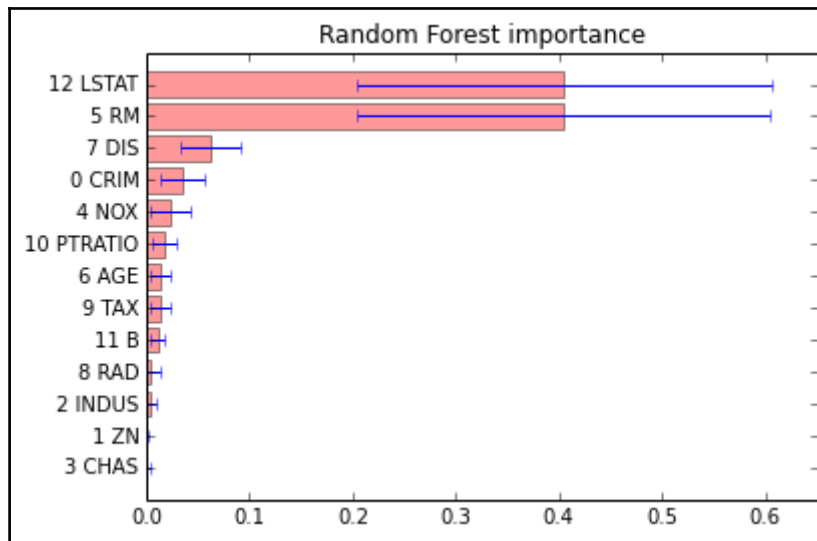
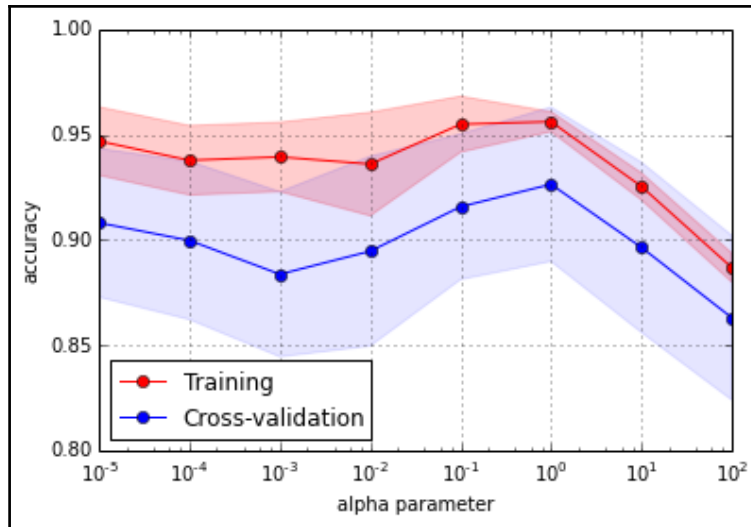


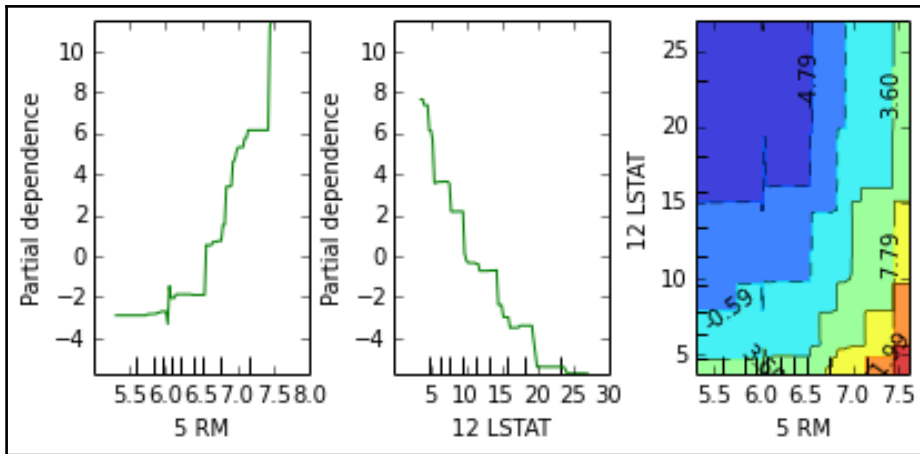












A screenshot of a web browser window showing a prediction interface. The address bar contains 'localhost:9099/predict'. The page title is 'localhost:9099/predict'. The form contains four input fields: 'Sepal length [cm]:', 'Sepal width [cm]:', 'Petal length [cm]:', and 'Petal width [cm]:'. Below the input fields is a 'Predict' button.