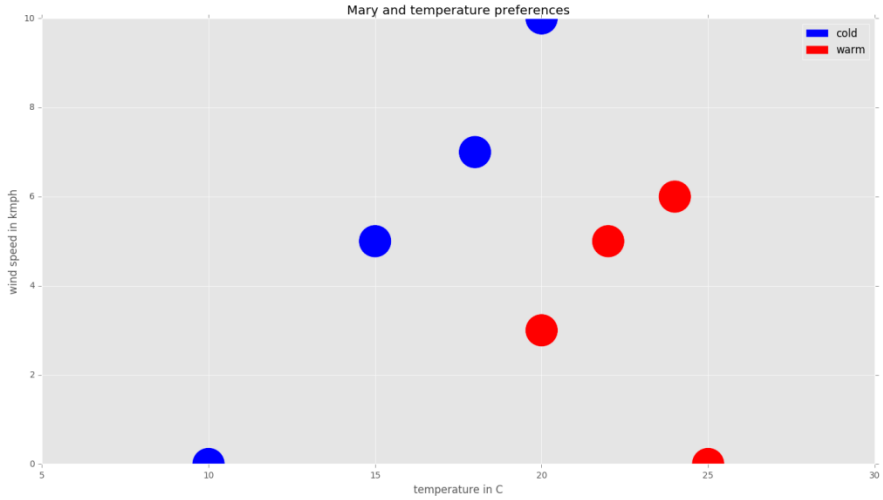
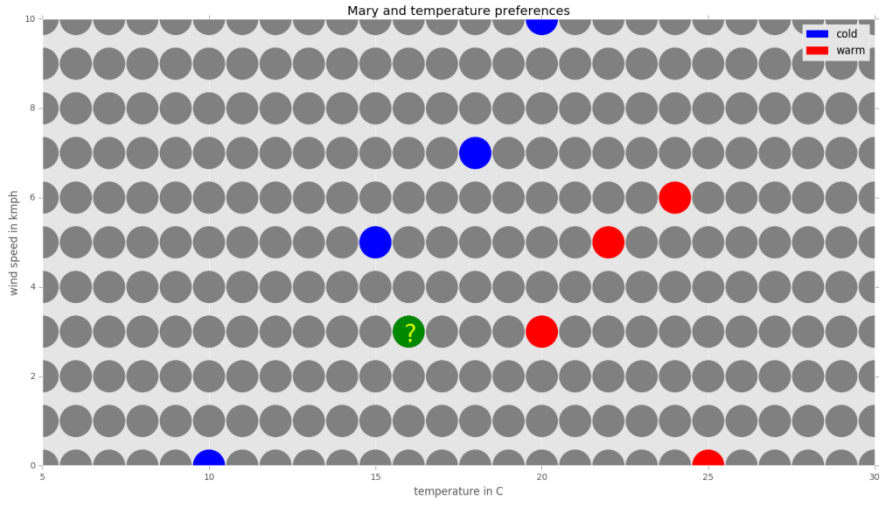
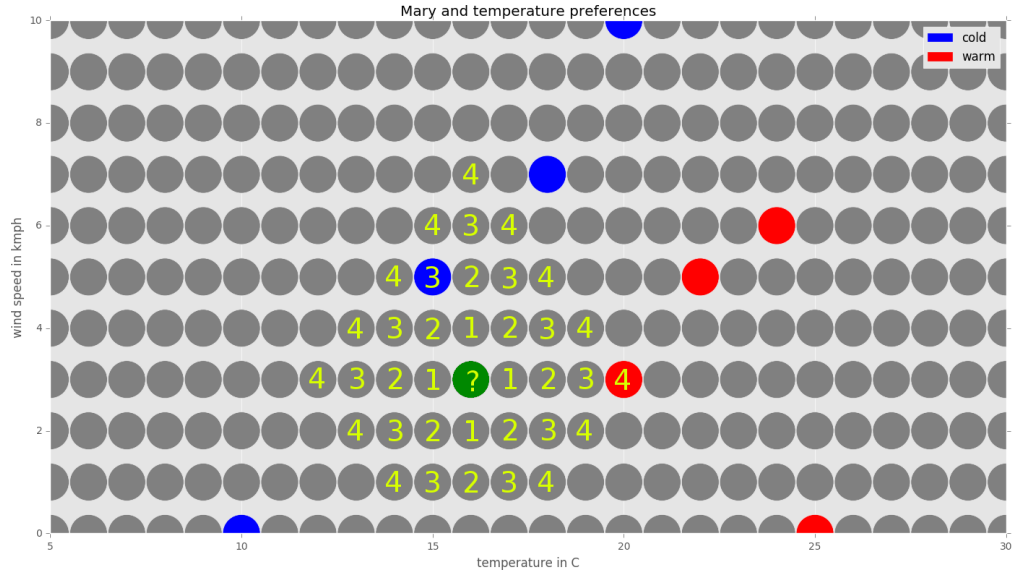
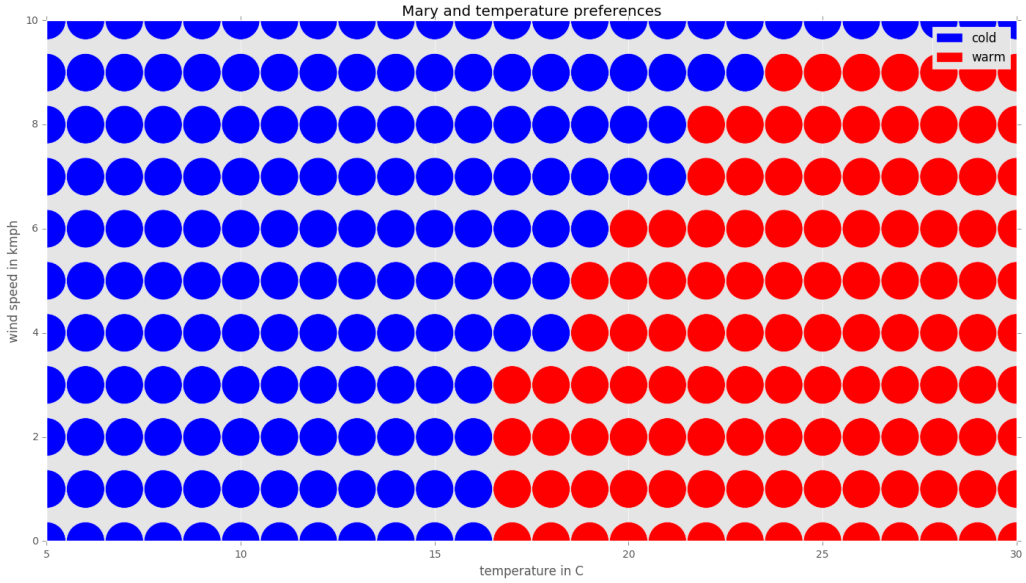
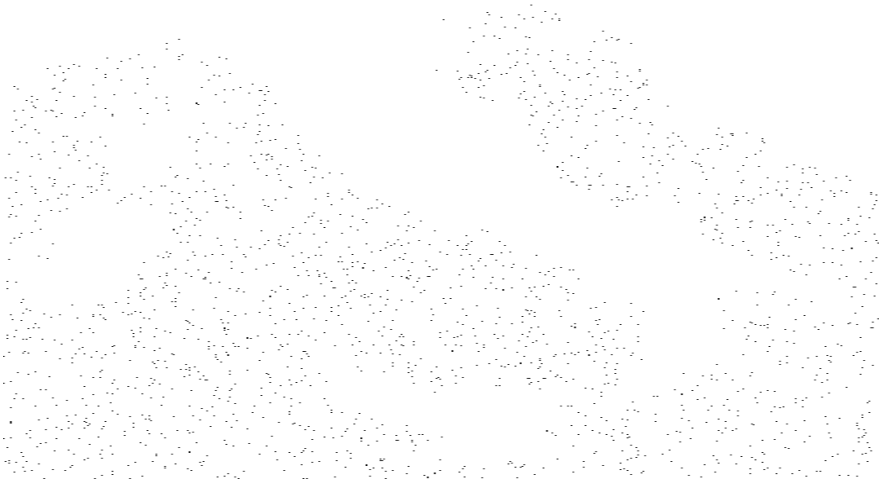
Chapter 1 – Classification Using K Nearest Neighbors

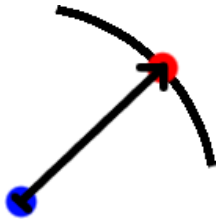


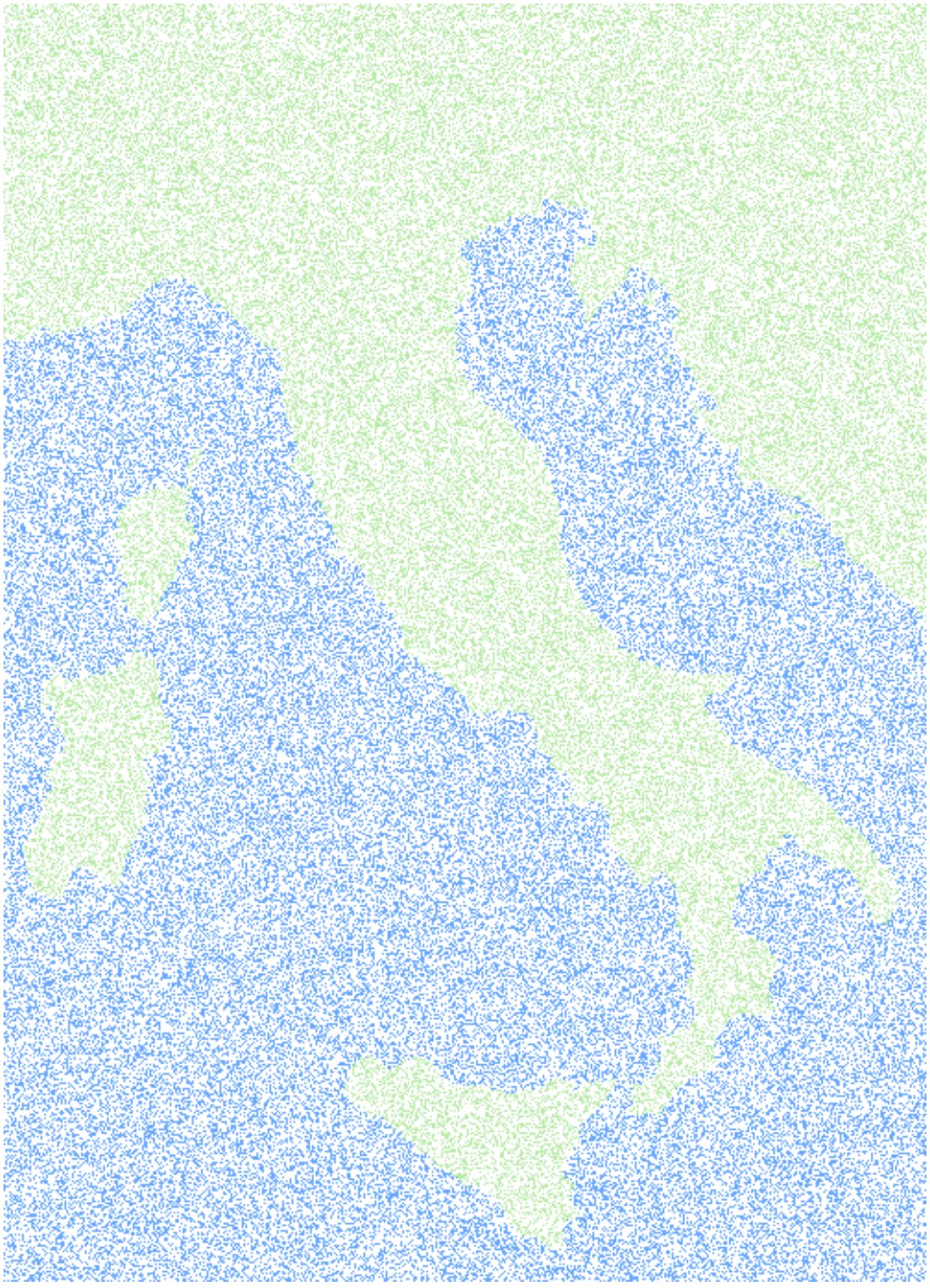


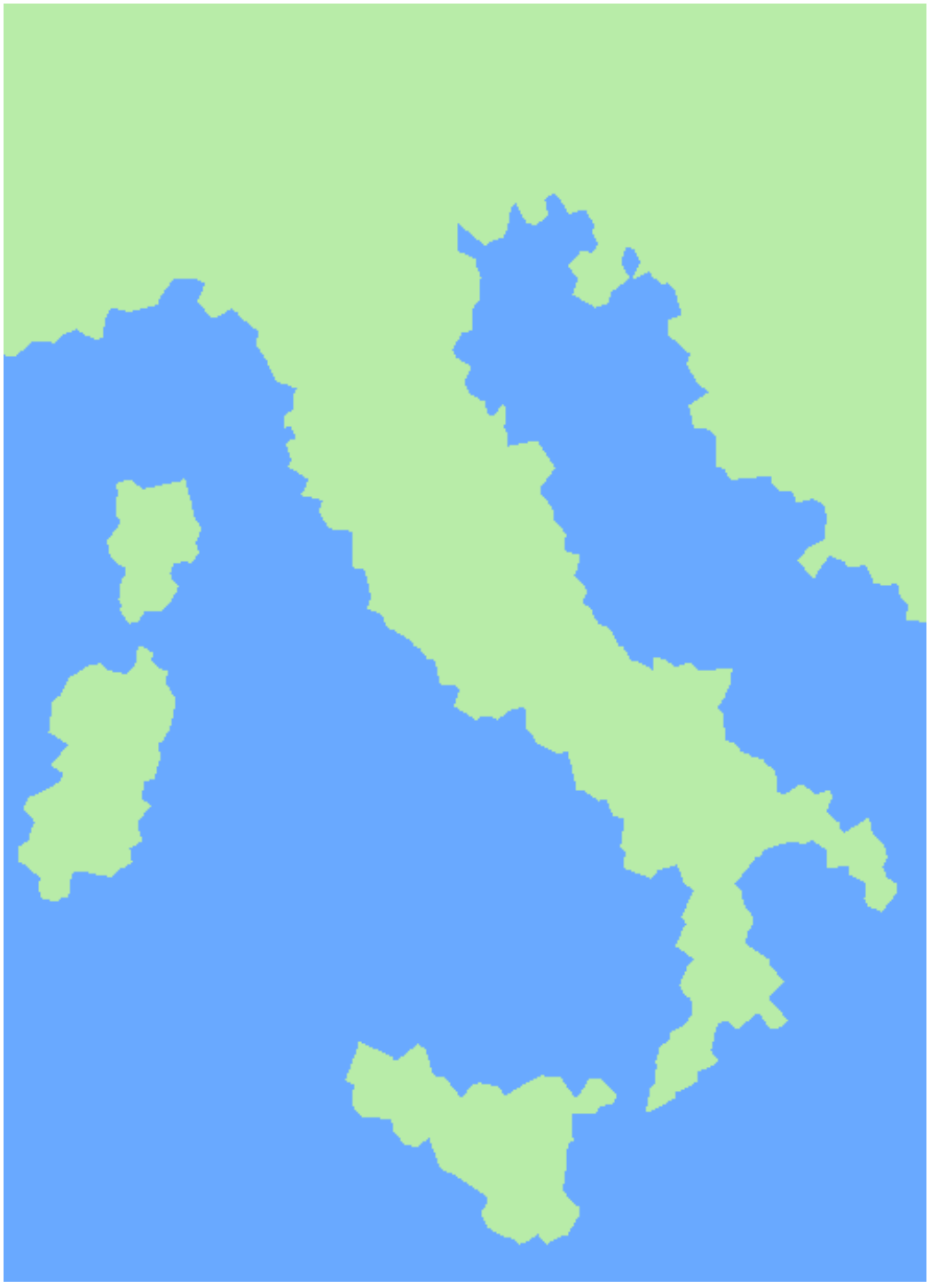

Mary and temperature preferences
cold
warm
wind speed in kmph
temperature in C
0
2
4
6
8
10
5
10
15
20
25
30
4
4 3 4
4 3 2 3 4
4 3 2 1 2 3 4
4 3 2 1 ? 1 2 3 4
4 3 2 1 2 3 4
4 3 2 3 4

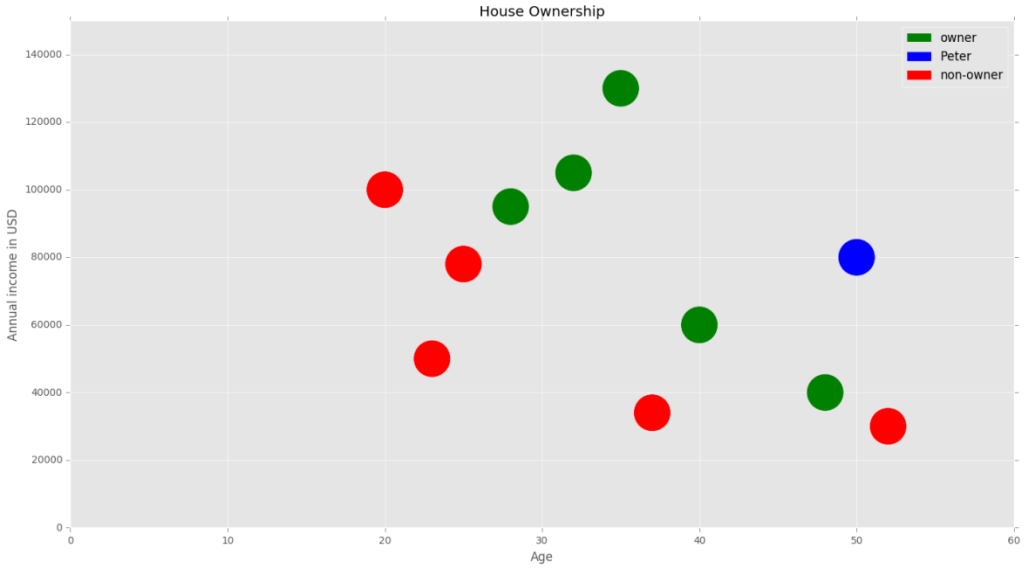

Mary and temperature preferences
cold
warm
wind speed in kmph
temperature in C
0
2
4
6
8
10
5
10
15
20
25
30

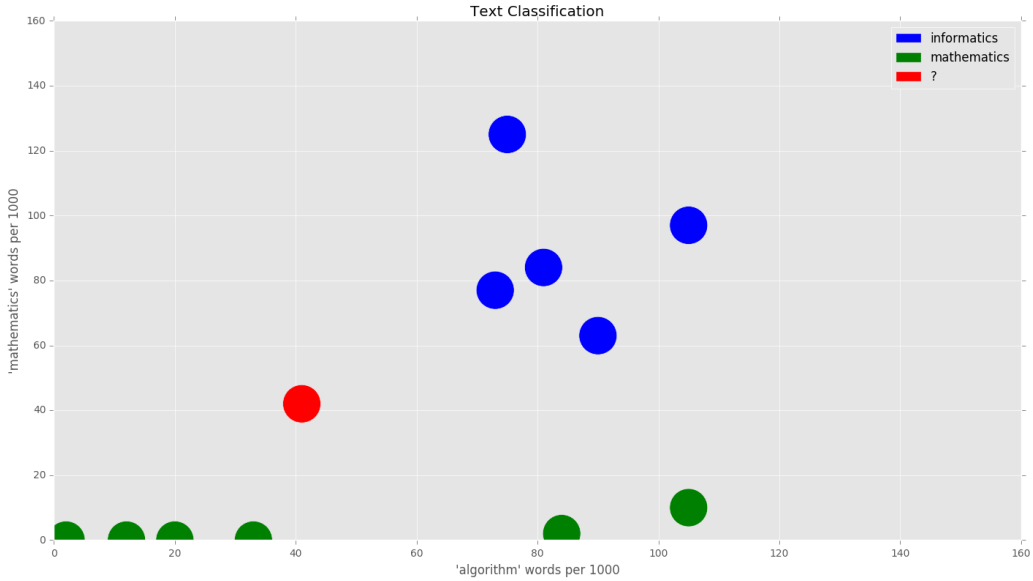


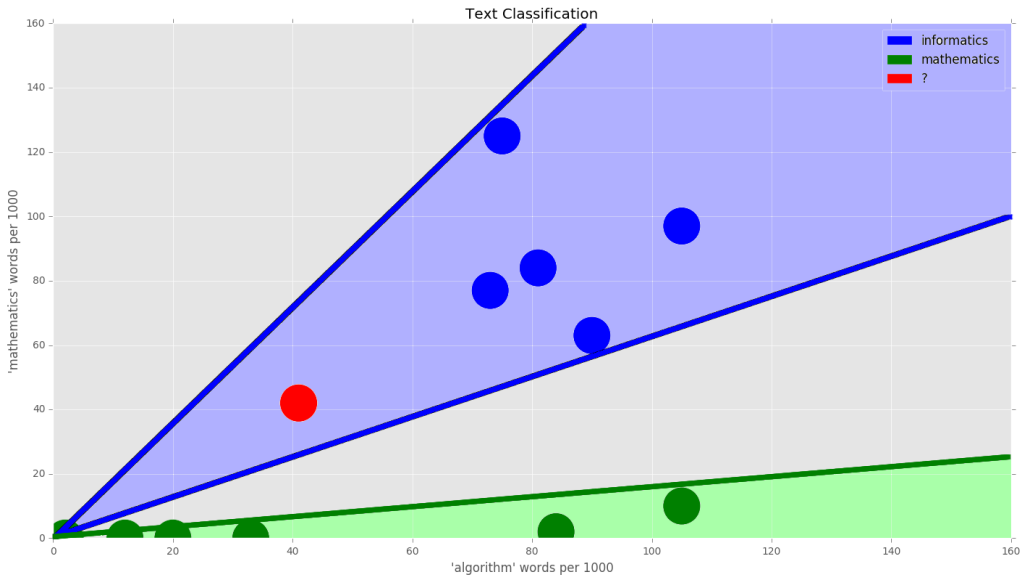




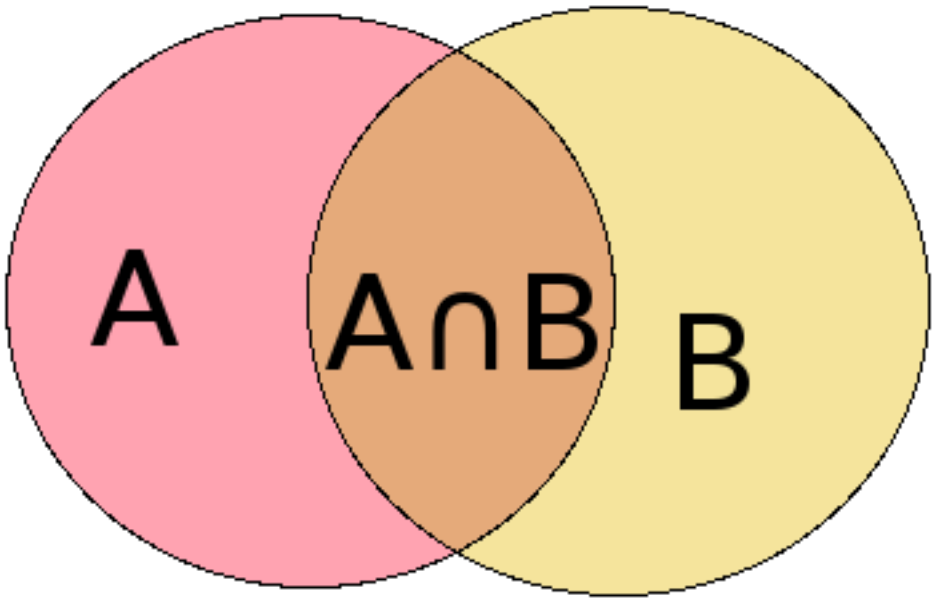



House Ownership
owner
Peter
non-owner
Annual income in USD
Age
0
20000
40000
60000
80000
100000
120000
140000
0
10
20
30
40
50
60

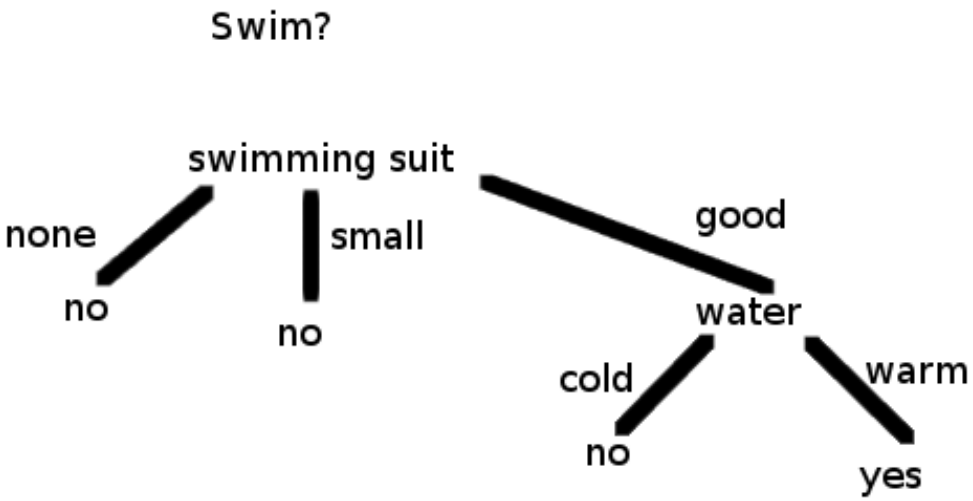

Text Classification
informatics
mathematics
?
'mathematics' words per 1000
'algorithm' words per 1000
0
20
40
60
80
100
120
140
160


Text Classification
informatics
mathematics
?
'mathematics' words per 1000
'algorithm' words per 1000
0
20
40
60
80
100
120
140
160

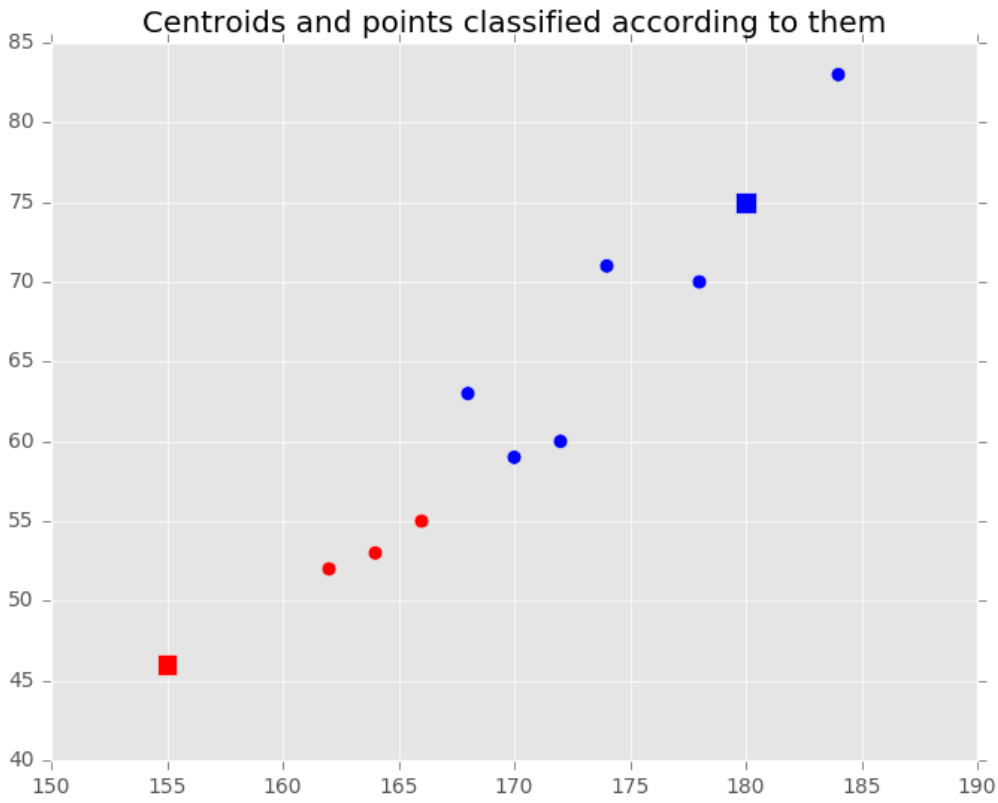
Chapter 2 – Naive Bayes

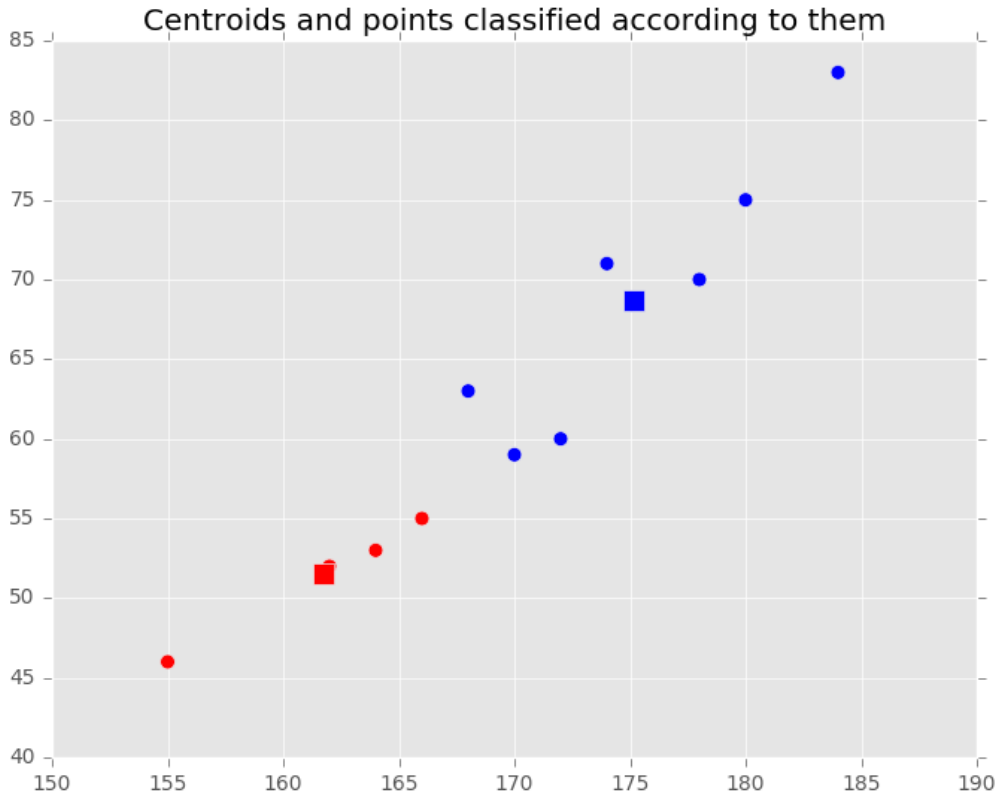


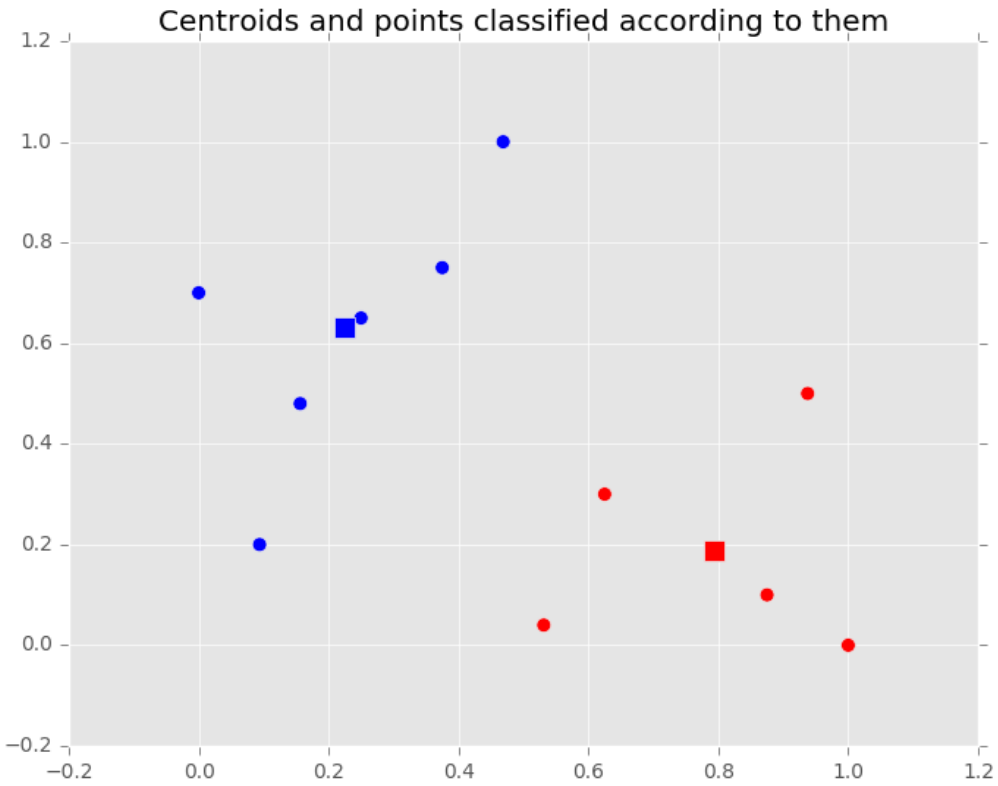
Chapter 3 – Decision Trees

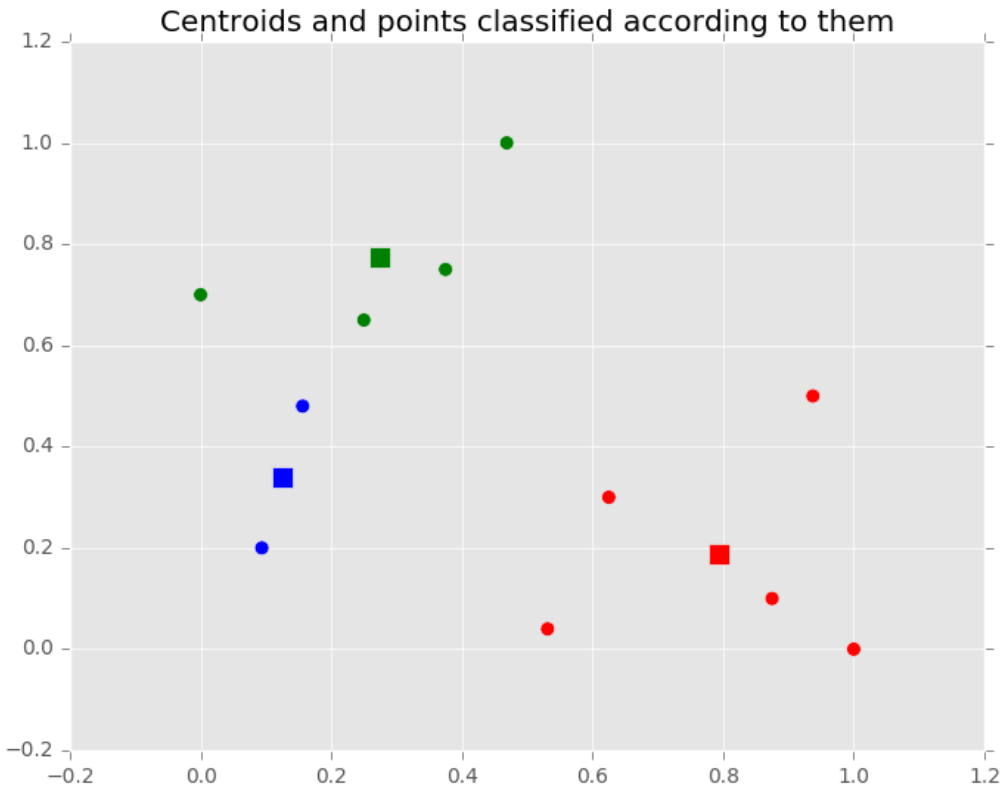


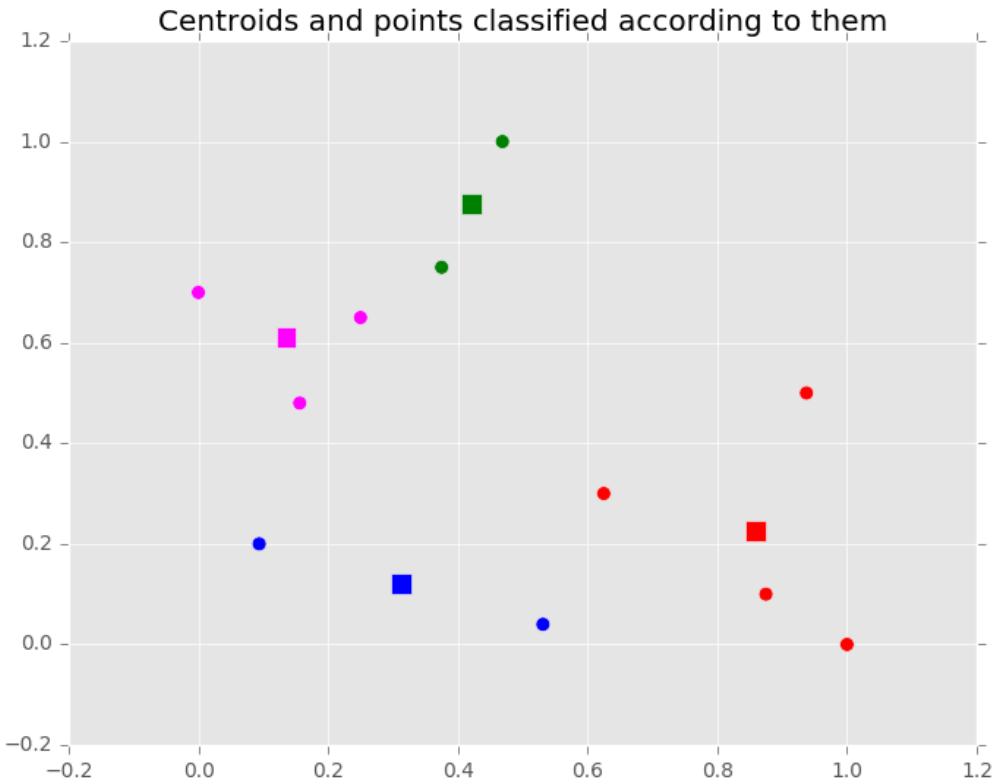
Chapter 5 – Clustering into K Clusters

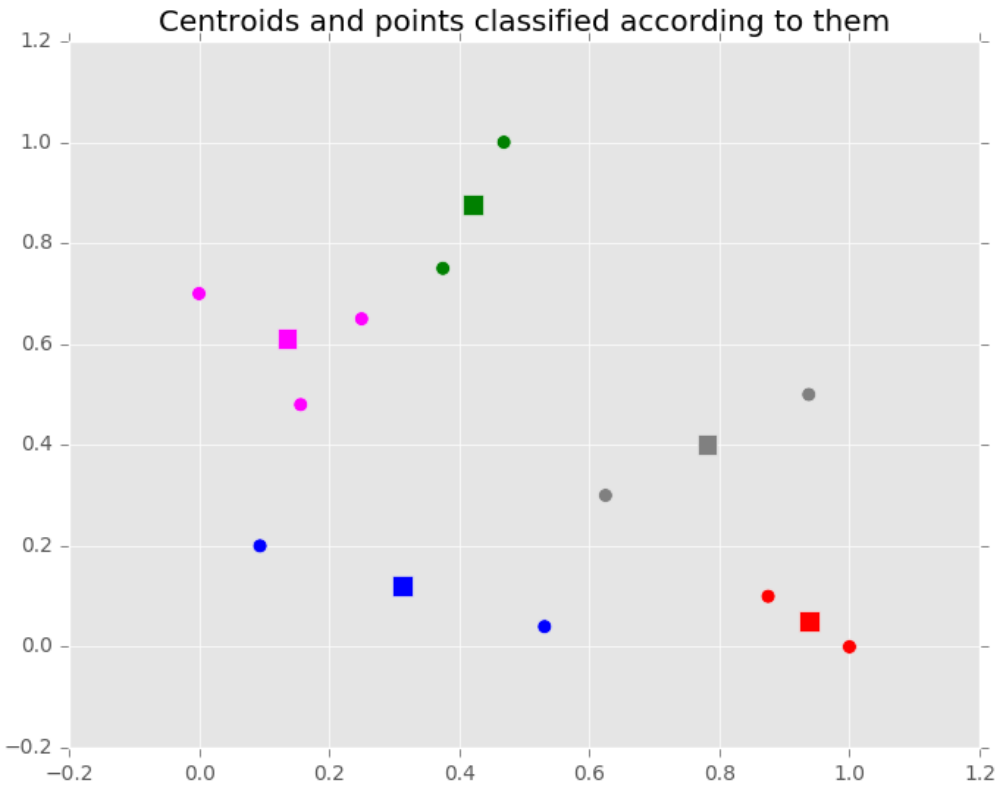


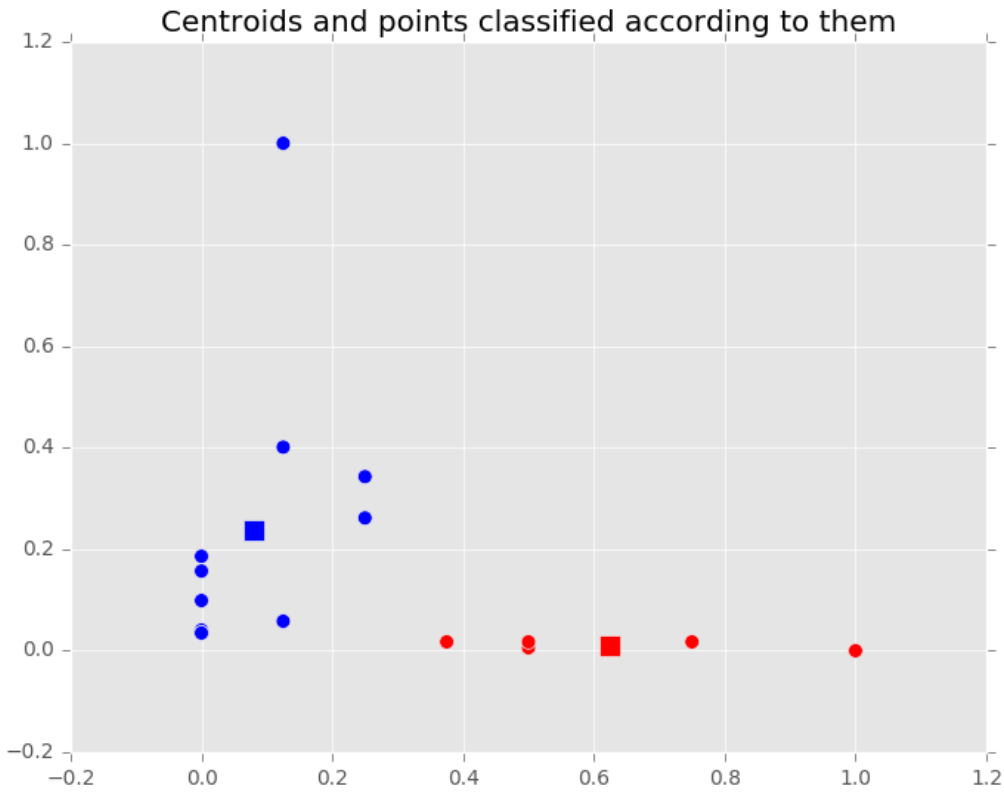

Centroids and points classified according to them
85
80
75
70
65
60
55
50
45
40
150
155
160
165
170
175
180
185
190

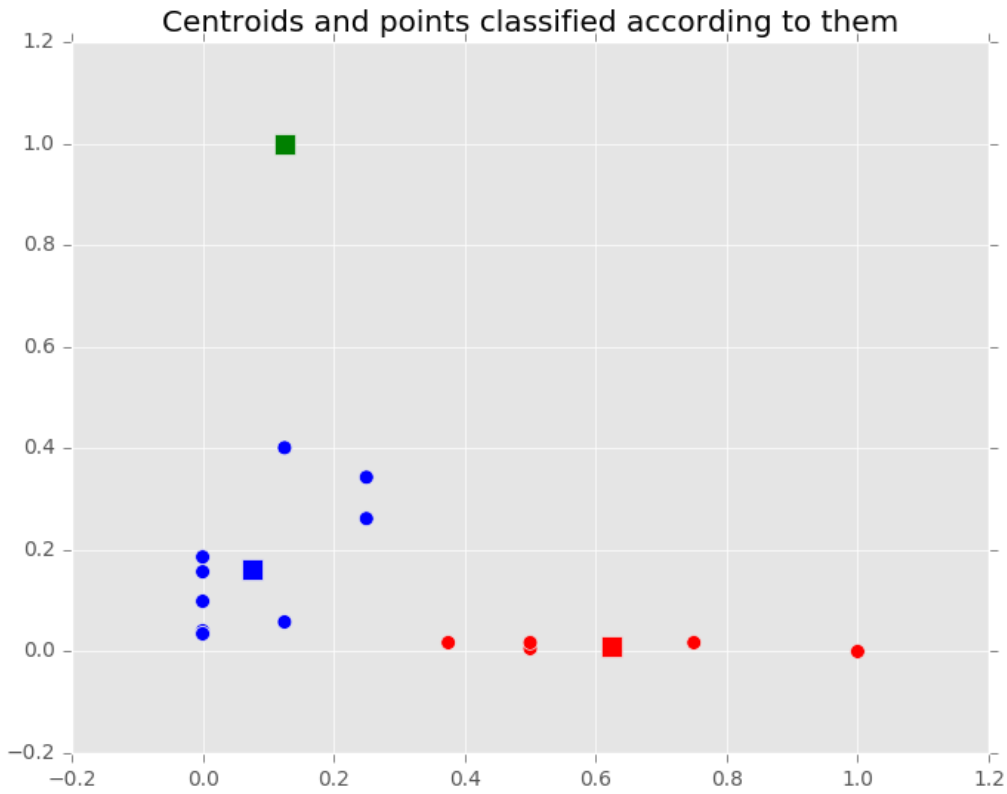

Centroids and points classified according to them
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2

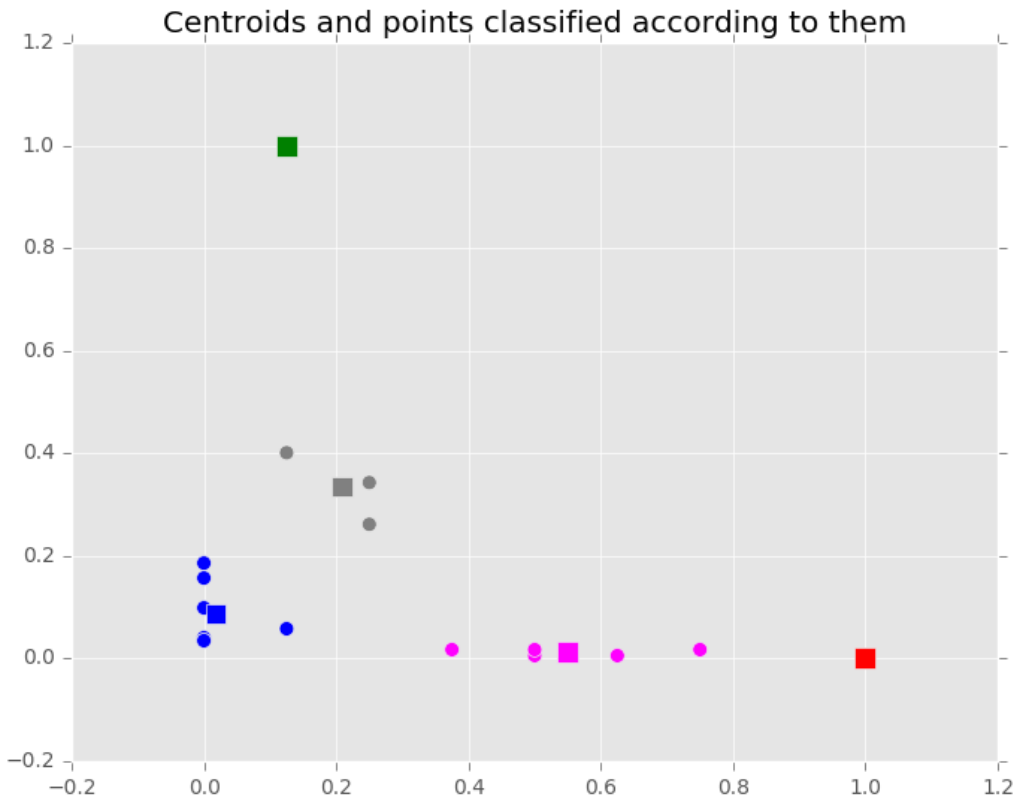

Centroids and points classified according to them
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2


Centroids and points classified according to them
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2

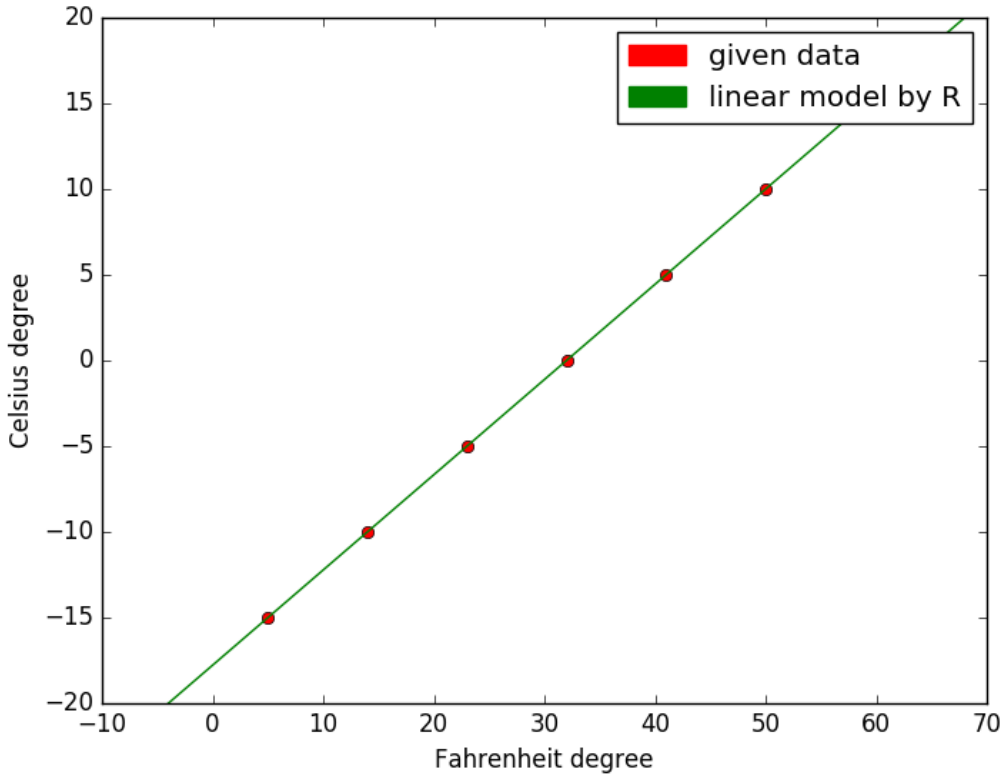

Centroids and points classified according to them
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2

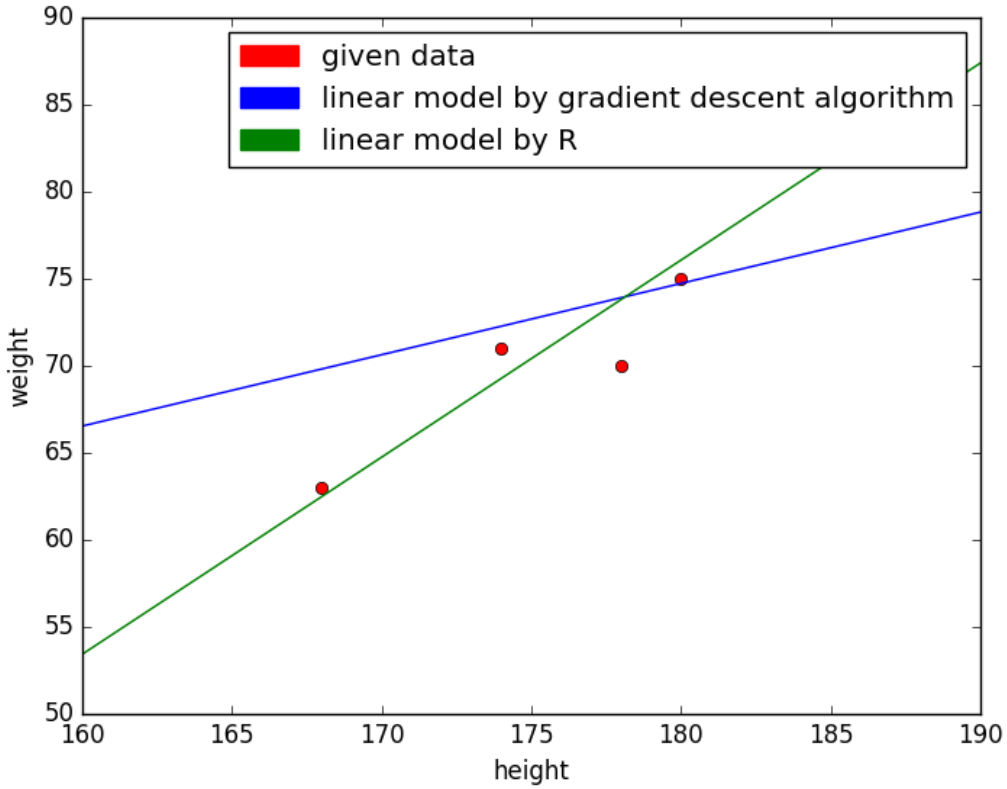

Centroids and points classified according to them
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2

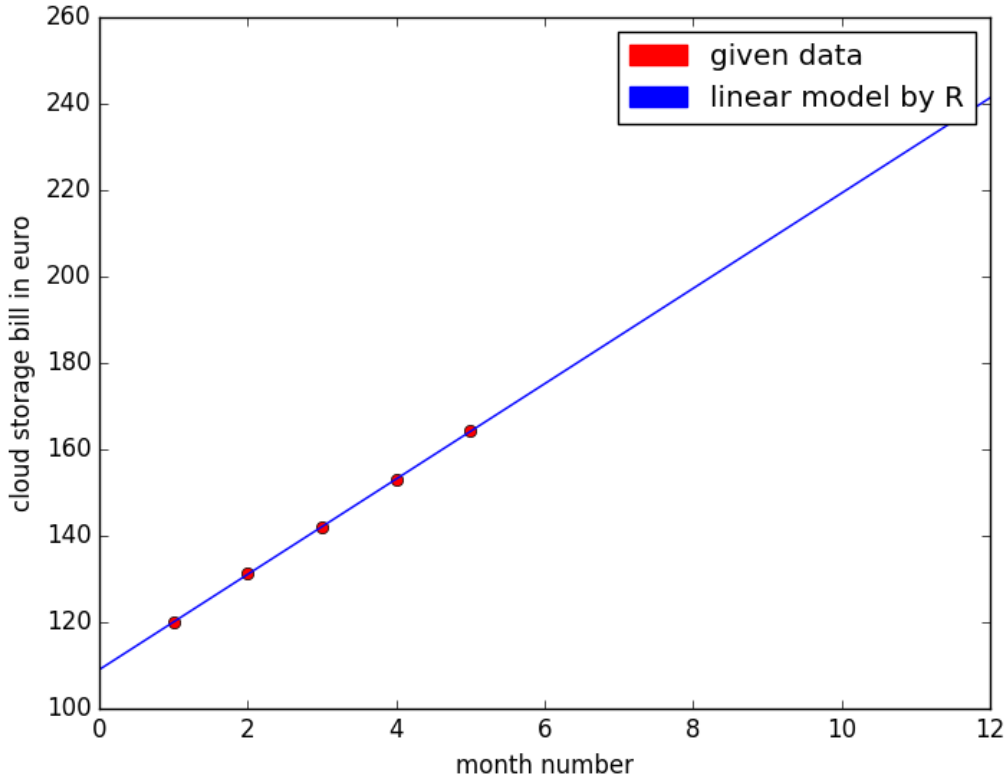

Centroids and points classified according to them
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2


Centroids and points classified according to them
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2

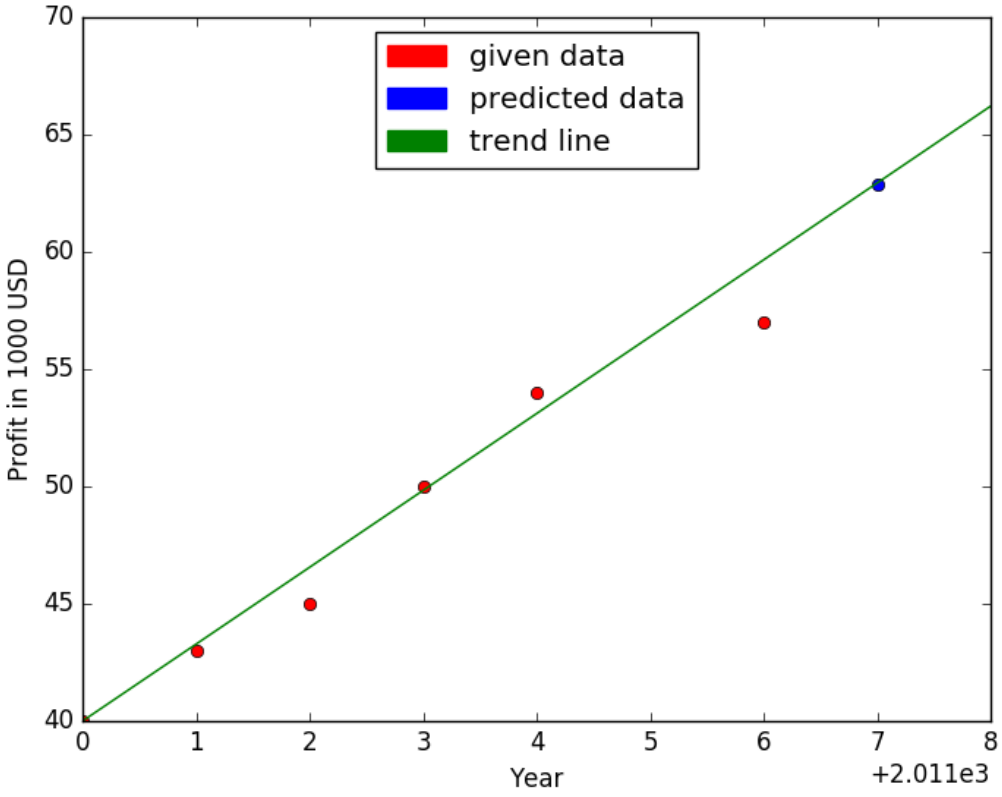
Chapter 6 – Regression

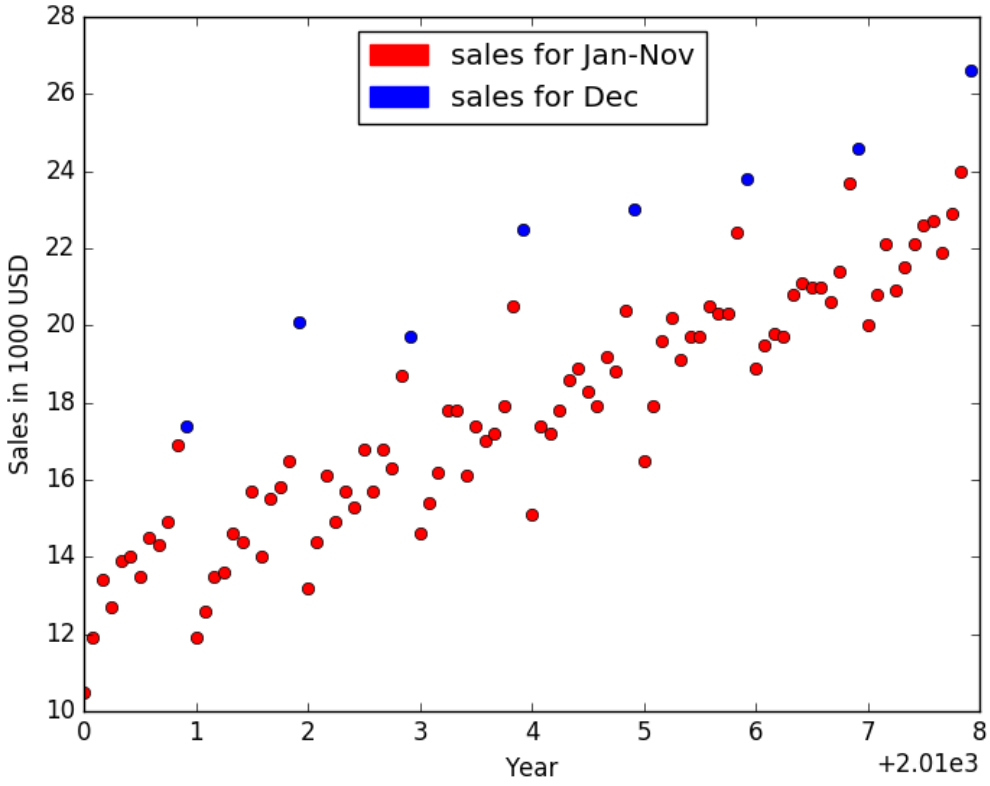


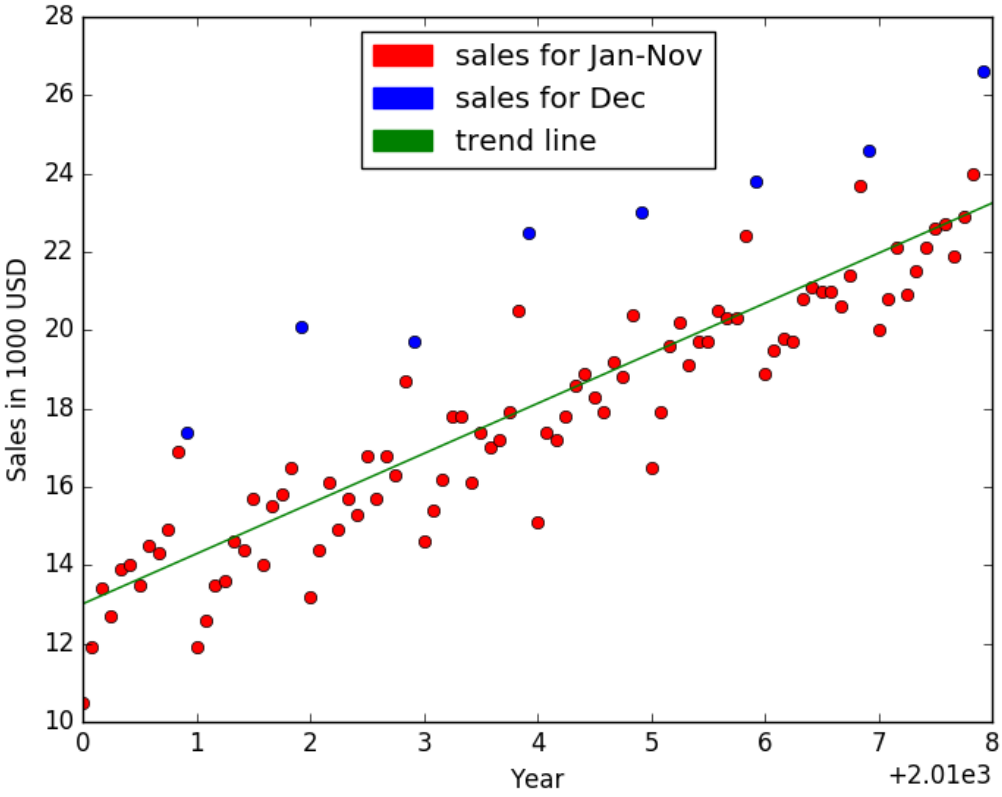

given data
linear model by gradient descent algorithm
linear model by R
weight
height
90
85
80
75
70
65
60
55
50
160
165
170
175
180
185
190

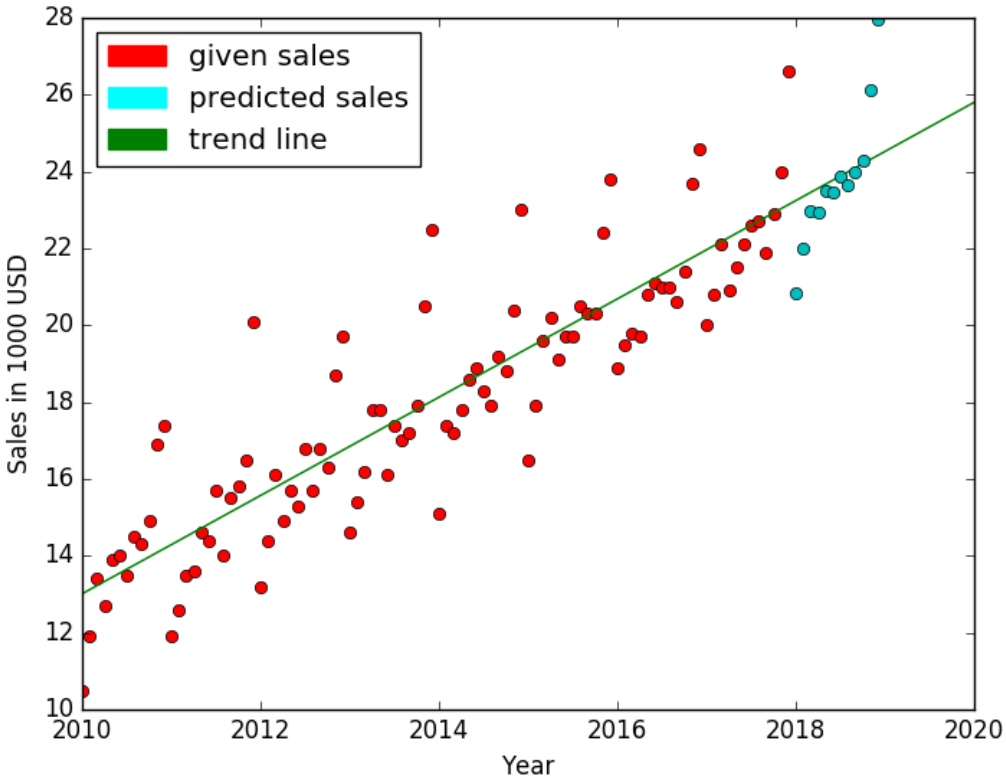

given data
linear model by R
cloud storage bill in euro
month number
260
240
220
200
180
160
140
120
100
0
2
4
6
8
10
12

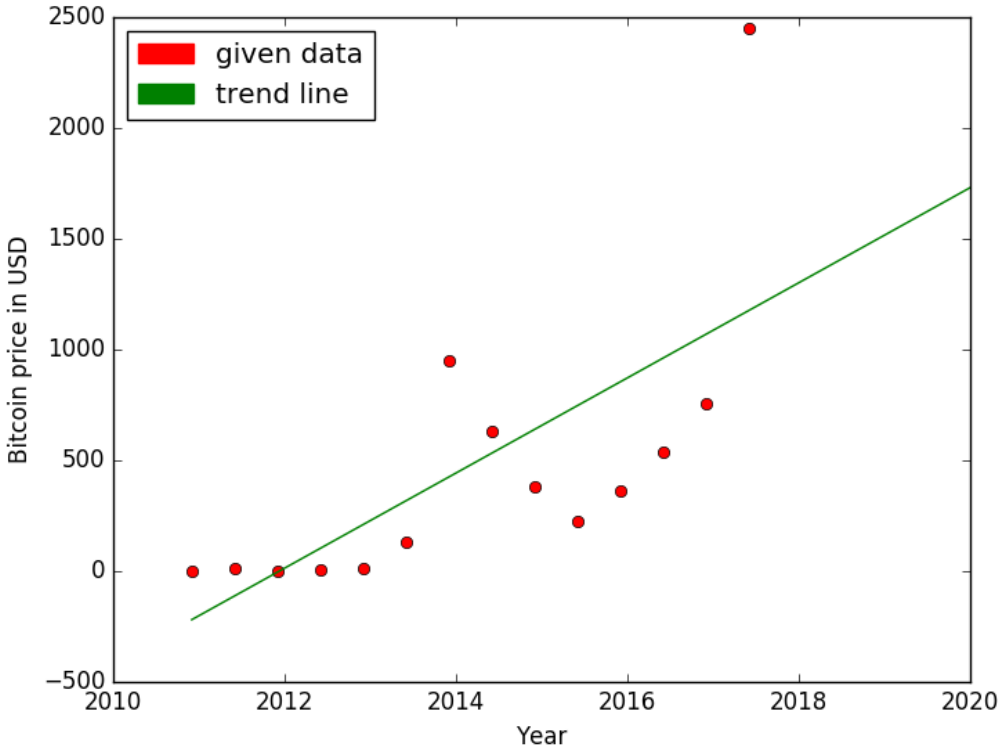
Chapter 7 – Time Series Analysis

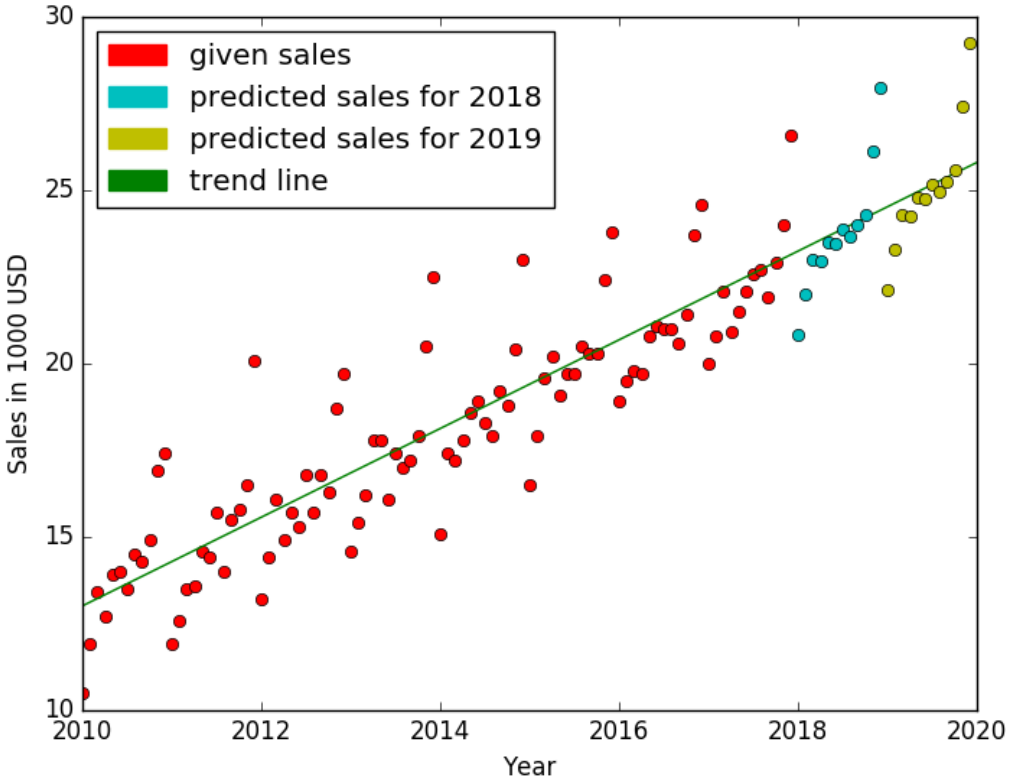
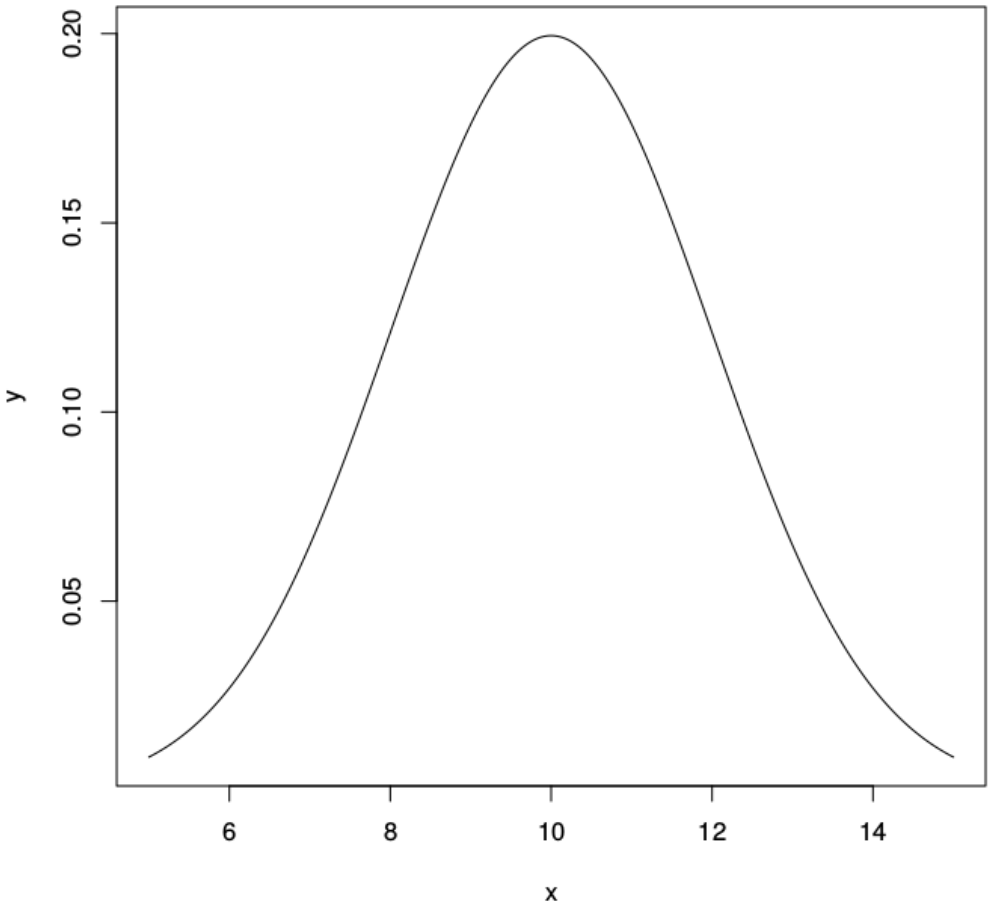



sales for Jan-Nov
sales for Dec
Sales in 1000 USD
Year
+2.01e3
28
26
24
22
20
18
16
14
12
10
0
1
2
3
4
5
6
7
8


sales for Jan-Nov
sales for Dec
trend line
Sales in 1000 USD
Year
+2.01e3
28
26
24
22
20
18
16
14
12
10
0
1
2
3
4
5
6
7
8


given sales
predicted sales
trend line
Sales in 1000 USD
Year
28
26
24
22
20
18
16
14
12
10
2010
2012
2014
2016
2018
2020


given data
trend line
Bitcoin price in USD
Year
2500
2000
1500
1000
500
0
−500
2010
2012
2014
2016
2018
2020


given sales
predicted sales for 2018
predicted sales for 2019
trend line
Sales in 1000 USD
Year
30
25
20
15
10
2010
2012
2014
2016
2018
2020

Appendix A – Statistics