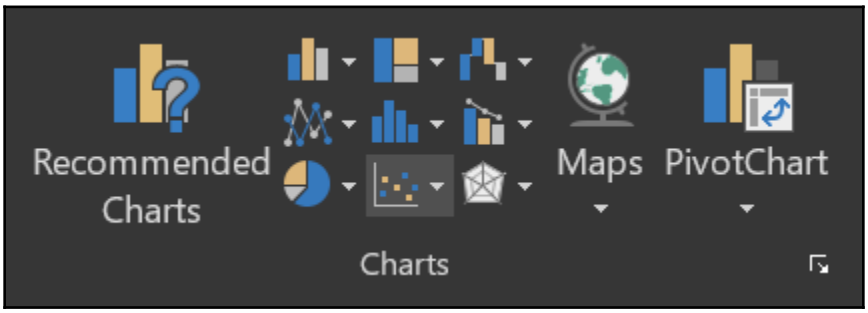
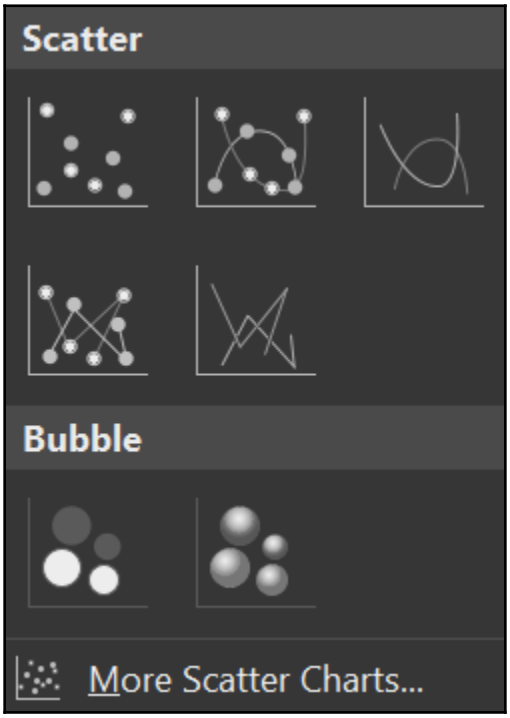
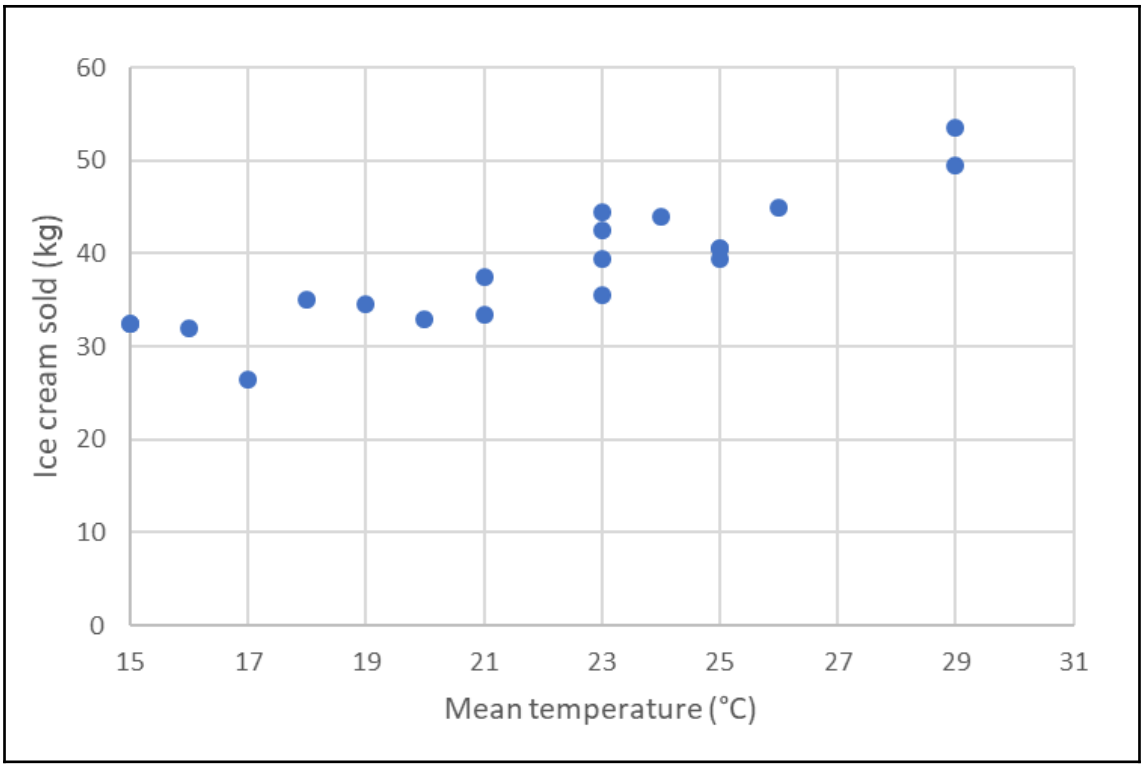
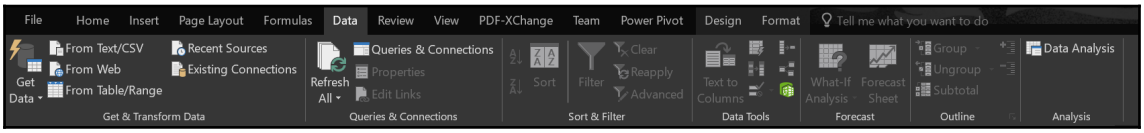
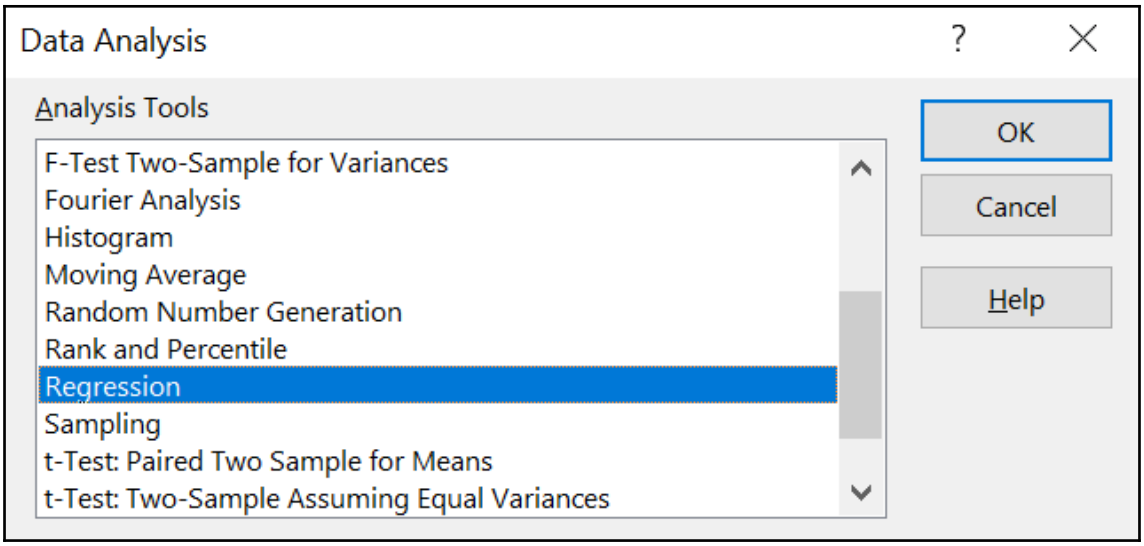
Chapter 1: Implementing Machine Learning Algorithms

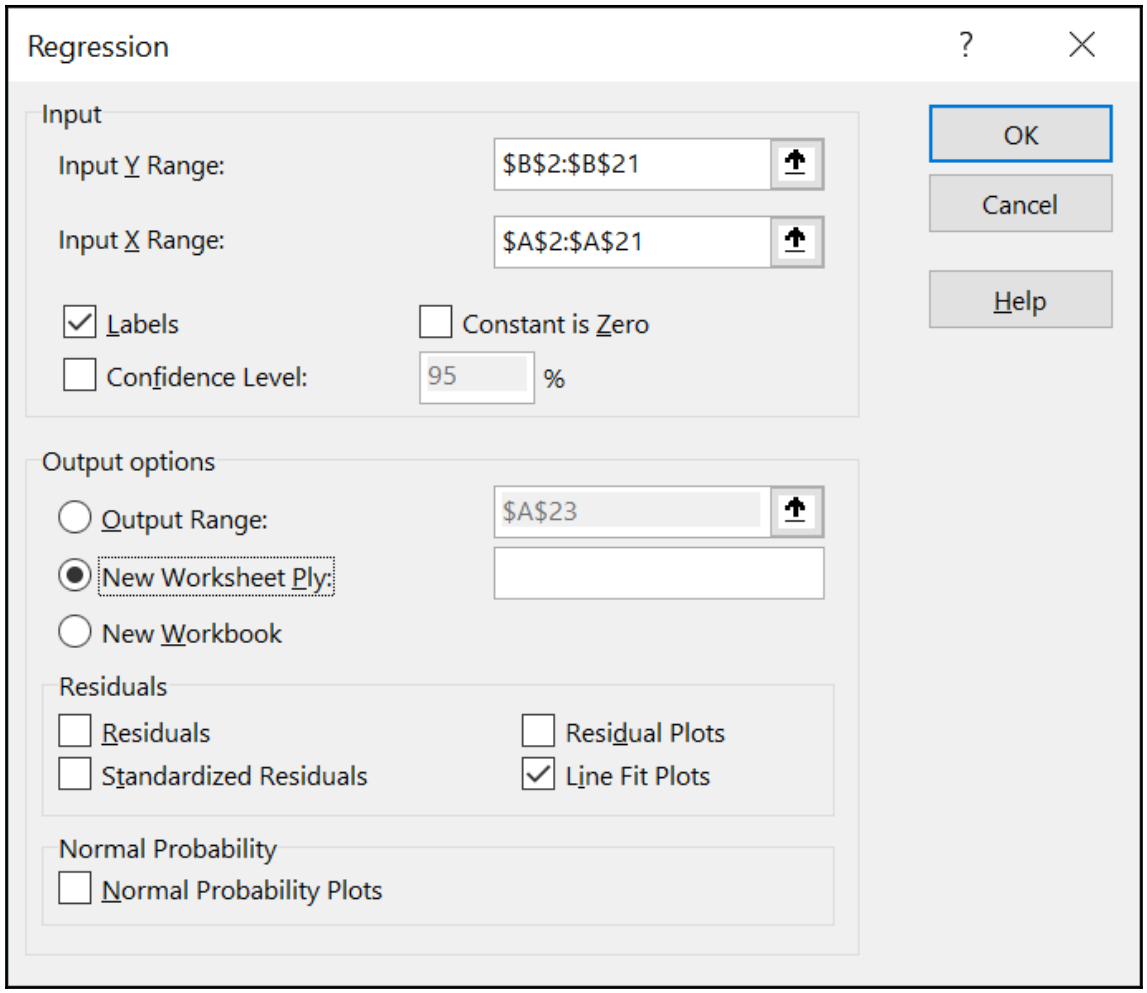


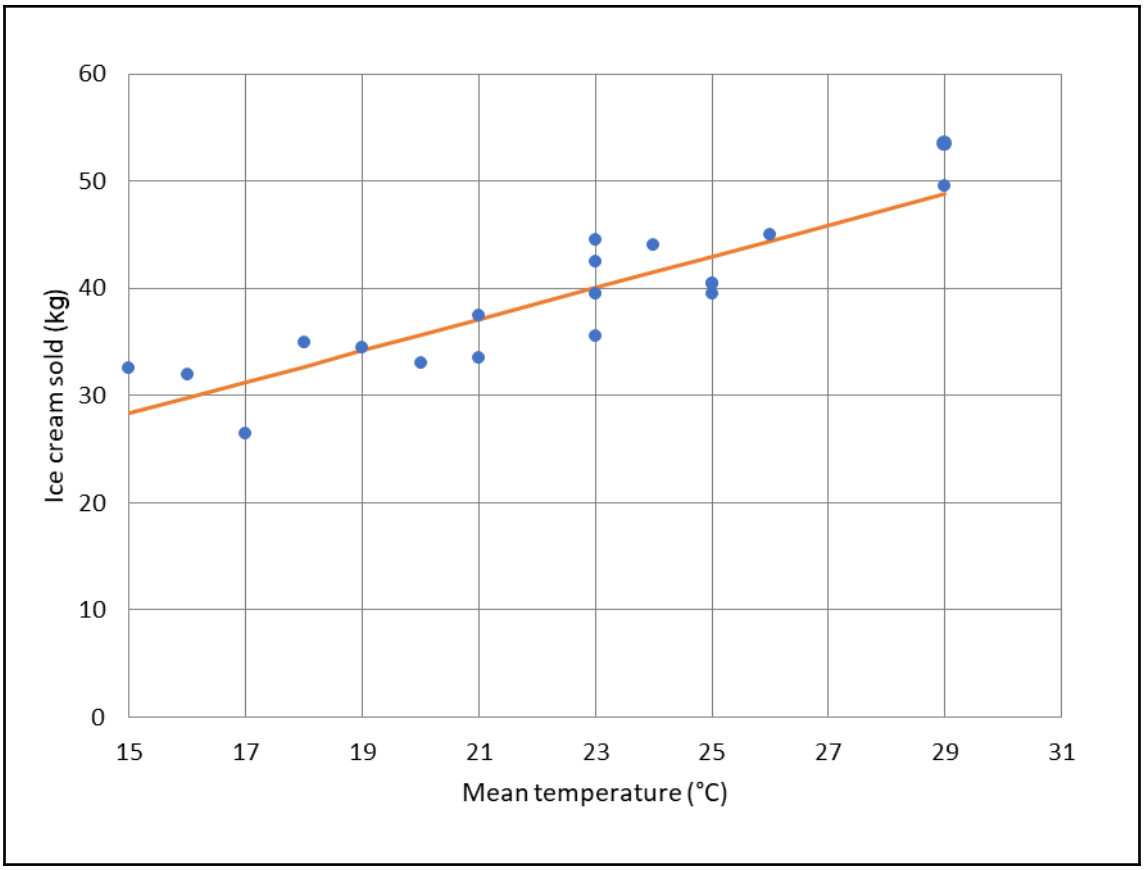


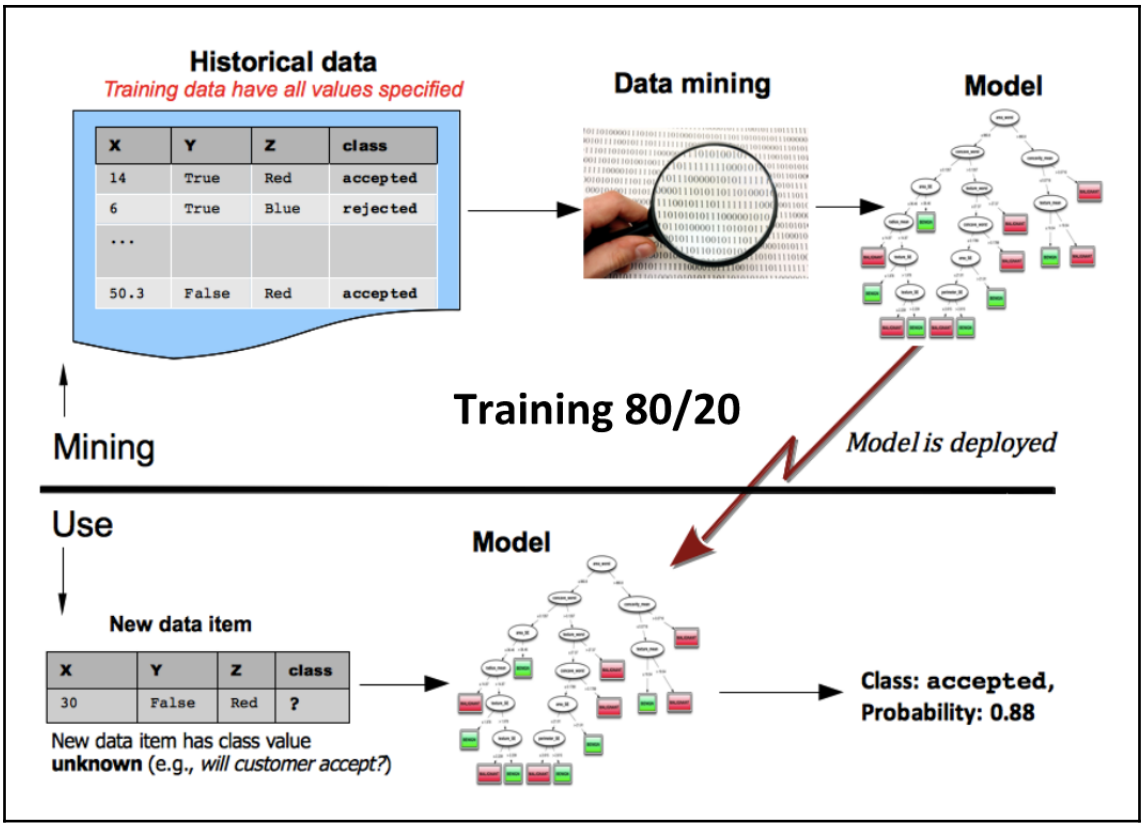

60
50
40
30
20
10
0
Ice cream sold (kg)
15
17
19
21
23
25
27
29
31
Mean temperature (°C)

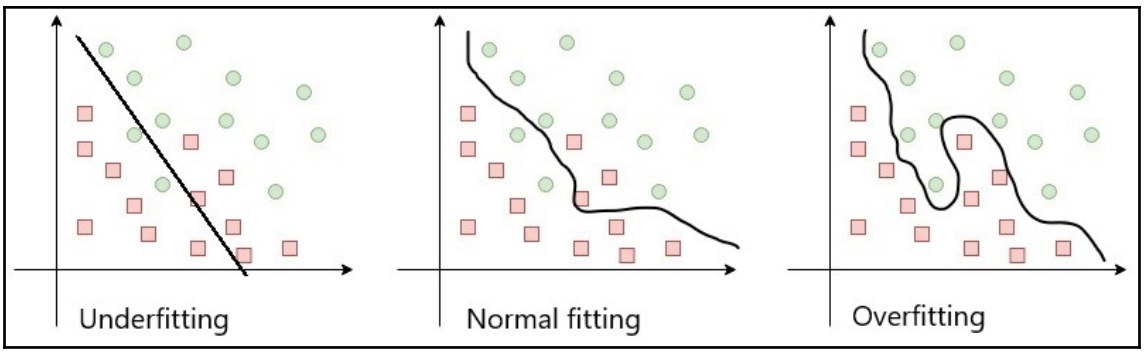

File
Home
Insert
Page Layout
Formulas
Data
Review
View
PDF-XChange
Team
Power Pivot
Design
Format
Tell me what you want to do
Get Data
From Text/CSV
From Web
From Table/Range
Recent Sources
Existing Connections
Get & Transform Data
Refresh All
Queries & Connections
Properties
Edit Links
Queries & Connections
Sort
Filter
Clear
Reapply
Advanced
Sort & Filter
Text to Columns
Data Tools
What-If Analysis
Forecast Sheet
Forecast
Group
Ungroup
Subtotal
Outline
Data Analysis
Analysis

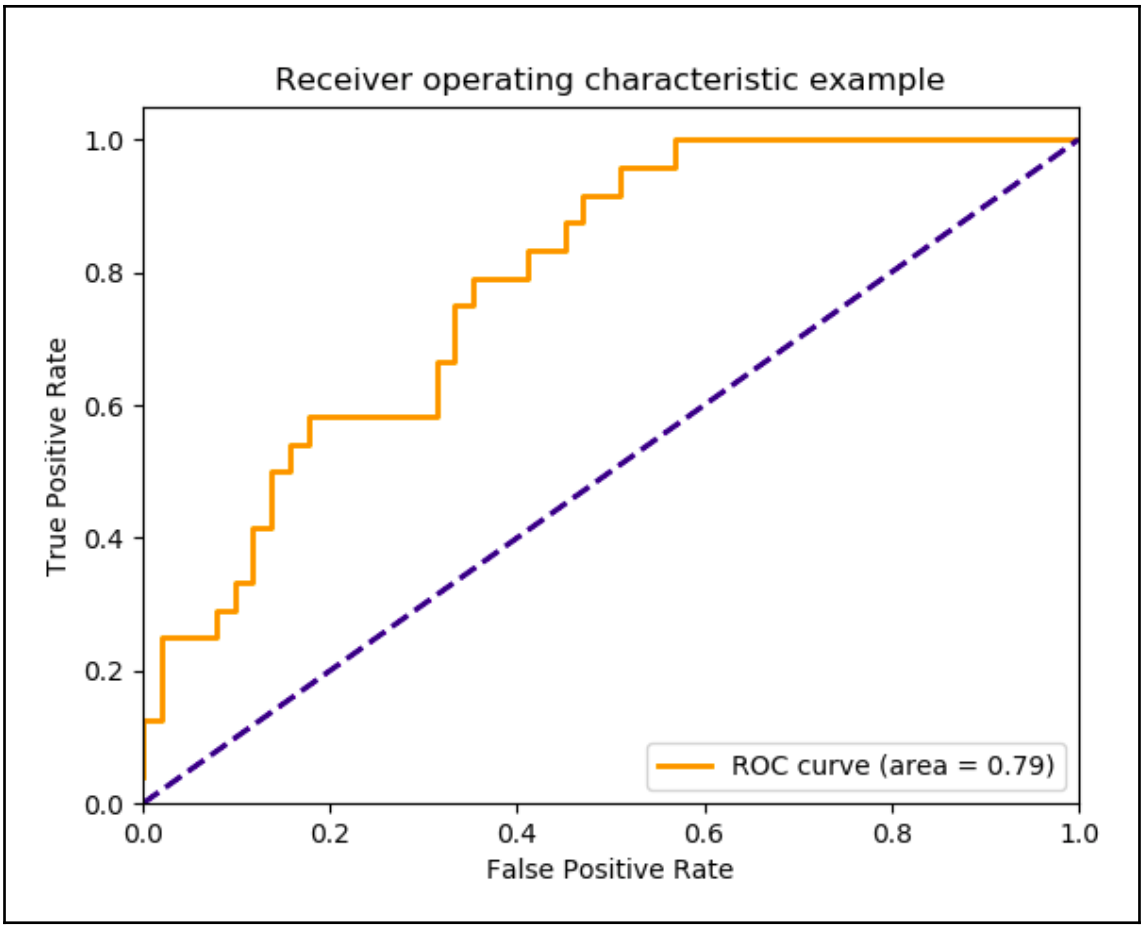

Data Analysis
?
Analysis Tools
F-Test Two-Sample for Variances
Fourier Analysis
Histogram
Moving Average
Random Number Generation
Rank and Percentile
Regression
Sampling
t-Test: Paired Two Sample for Means
t-Test: Two-Sample Assuming Equal Variances
OK
Cancel
Help

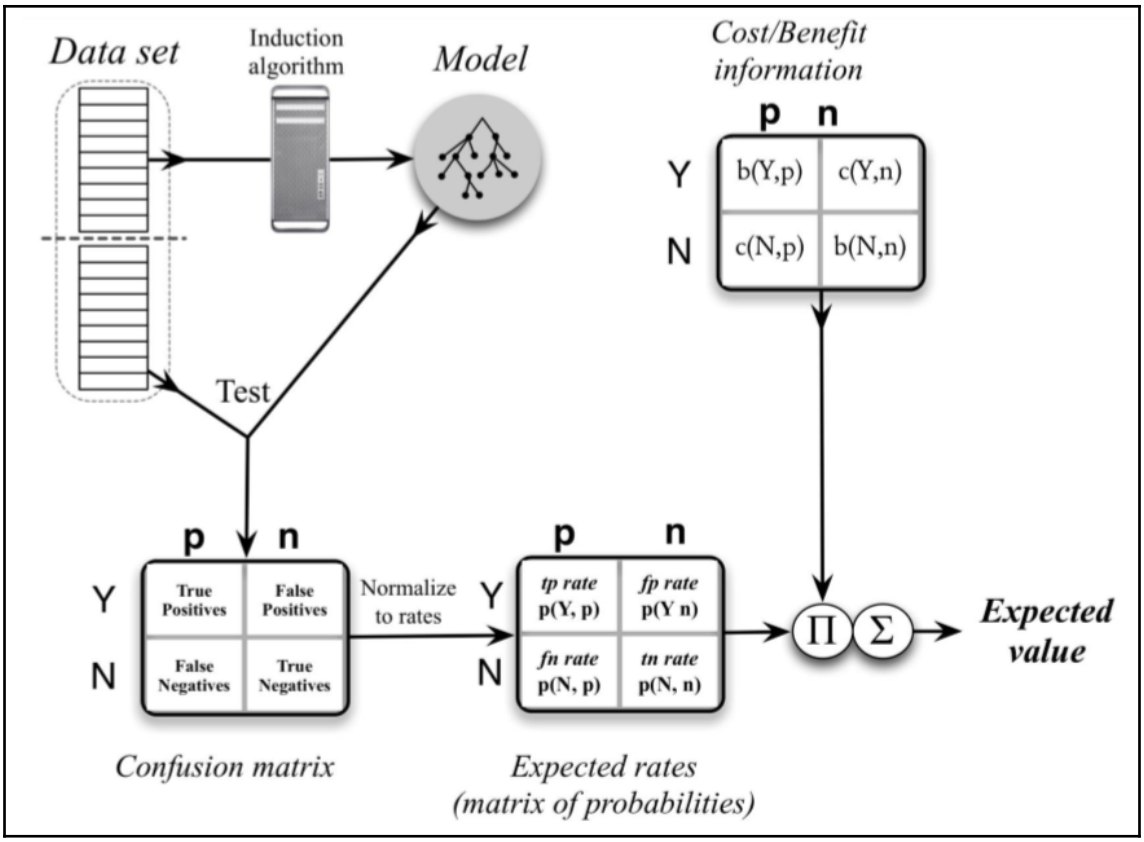

Regression
?
Input
Input Y Range:
$B$2:$B$21
Input X Range:
$A$2:$A$21
Labels
Constant is Zero
Confidence Level:
95
%
OK
Cancel
Help
Output options
Output Range:
$A$23
New Worksheet Ply:
New Workbook
Residuals
Residuals
Residual Plots
Standardized Residuals
Line Fit Plots
Normal Probability
Normal Probability Plots


60
50
40
30
20
10
0
Ice cream sold (kg)
15
17
19
21
23
25
27
29
31
Mean temperature (°C)


Historical data
Training data have all values specified
X | Y | Z | class
14 | True | Red | accepted
6 | True | Blue | rejected
...
50.3 | False | Red | accepted
Data mining
Model
Training 80/20
Mining
Model is deployed
Use
Model
New data item
X | Y | Z | class
30 | False | Red | ?
New data item has class value unknown (e.g., will customer accept?)
Class: accepted,
Probability: 0.88


Underfitting
Normal fitting
Overfitting

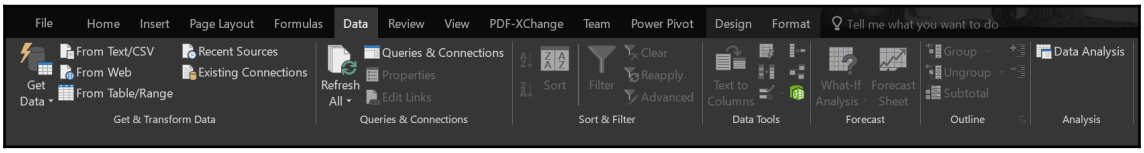

Receiver operating characteristic example
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC curve (area = 0.79)

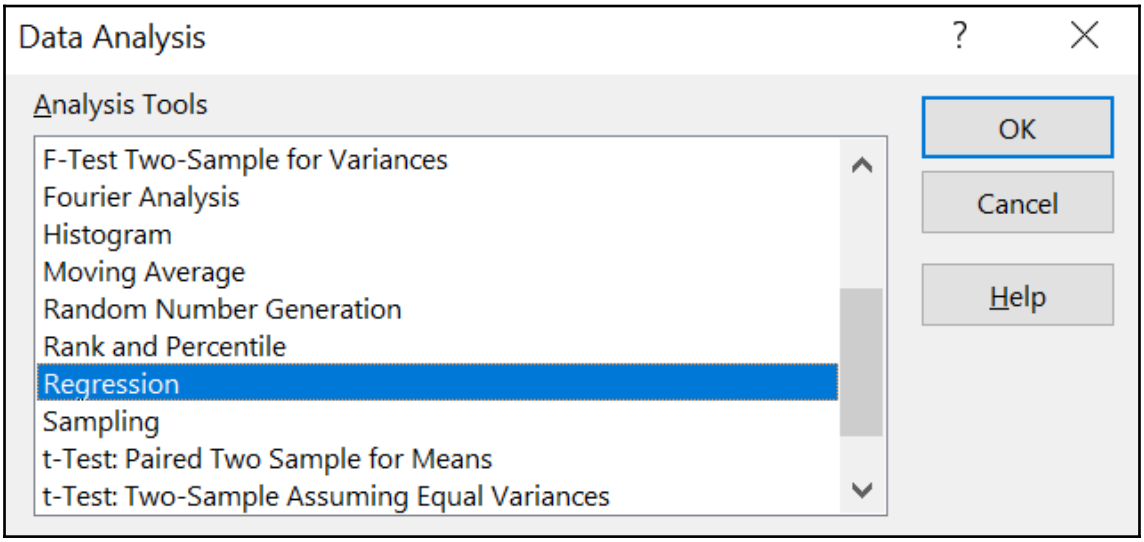

Data set
Induction algorithm
Model
Cost/Benefit information
p
n
Y
N
b(Y,p)
c(Y,n)
c(N,p)
b(N,n)
Test
p
n
Y
N
True Positives
False Positives
False Negatives
True Negatives
Normalize to rates
p
n
Y
N
tp rate p(Y, p)
fp rate p(Y n)
fn rate p(N, p)
tn rate p(N, n)
Π
Σ
Expected value
Confusion matrix
Expected rates (matrix of probabilities)

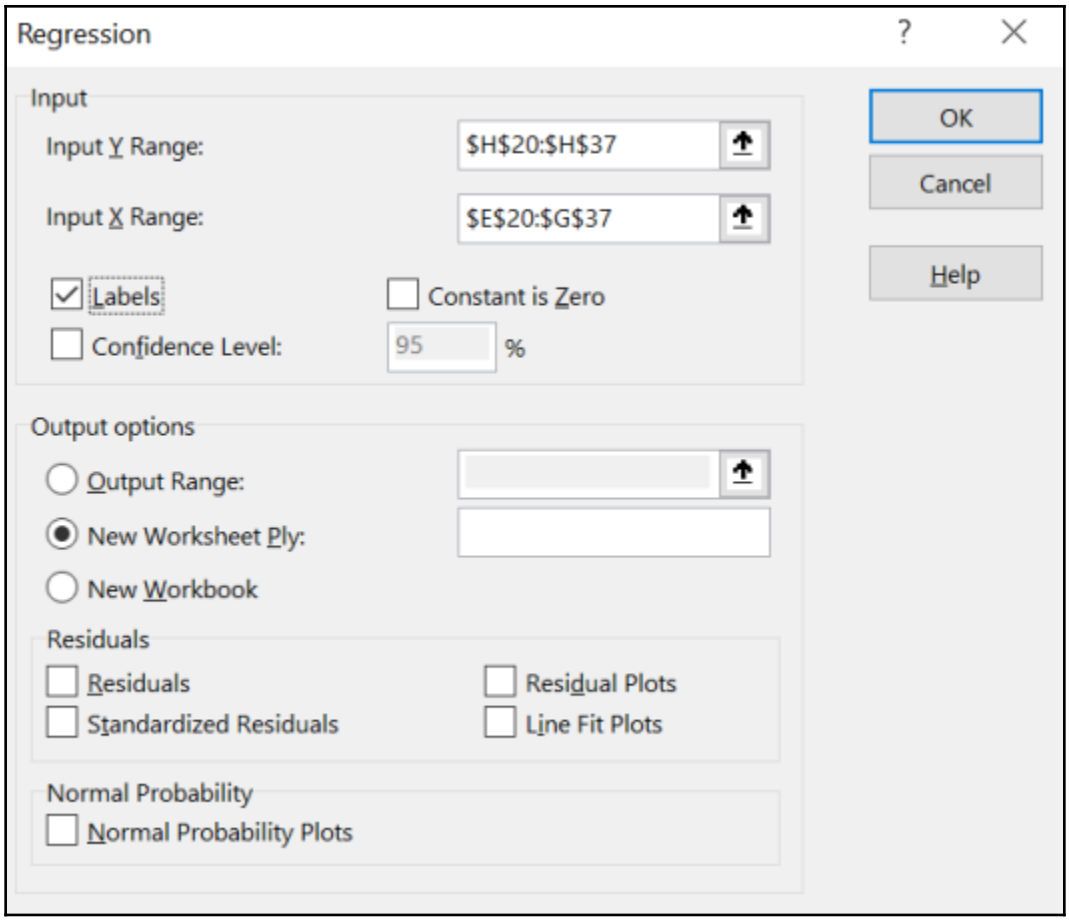
Chapter 2: Hands-On Examples of Machine Learning Models

	A	B	C	D	E	F
1	ID	Gender	Score	Years in company	Division	Salary Increase
2	1	F	11	9	Production	22
3	2	F	89	1	Production	97
4	3	F	21	4	Production	47
5	4	F	81	1	Production	127
6	5	F	31	4	Research	65
7	6	F	71	1	Research	53
8	7	F	11	4	Sales	74
9	8	F	16	7	Production	18
10	9	M	20	6	Research	129
11	10	M	79	3	Sales	475
12	11	M	51	3	Research	342
13	12	M	69	2	Sales	329
14	13	M	30	7	Sales	185
15	14	M	71	7	Sales	332
16	15	M	39	1	Sales	268
17	16	M	89	6	Production	518
18	17	M	50	8	Production	390

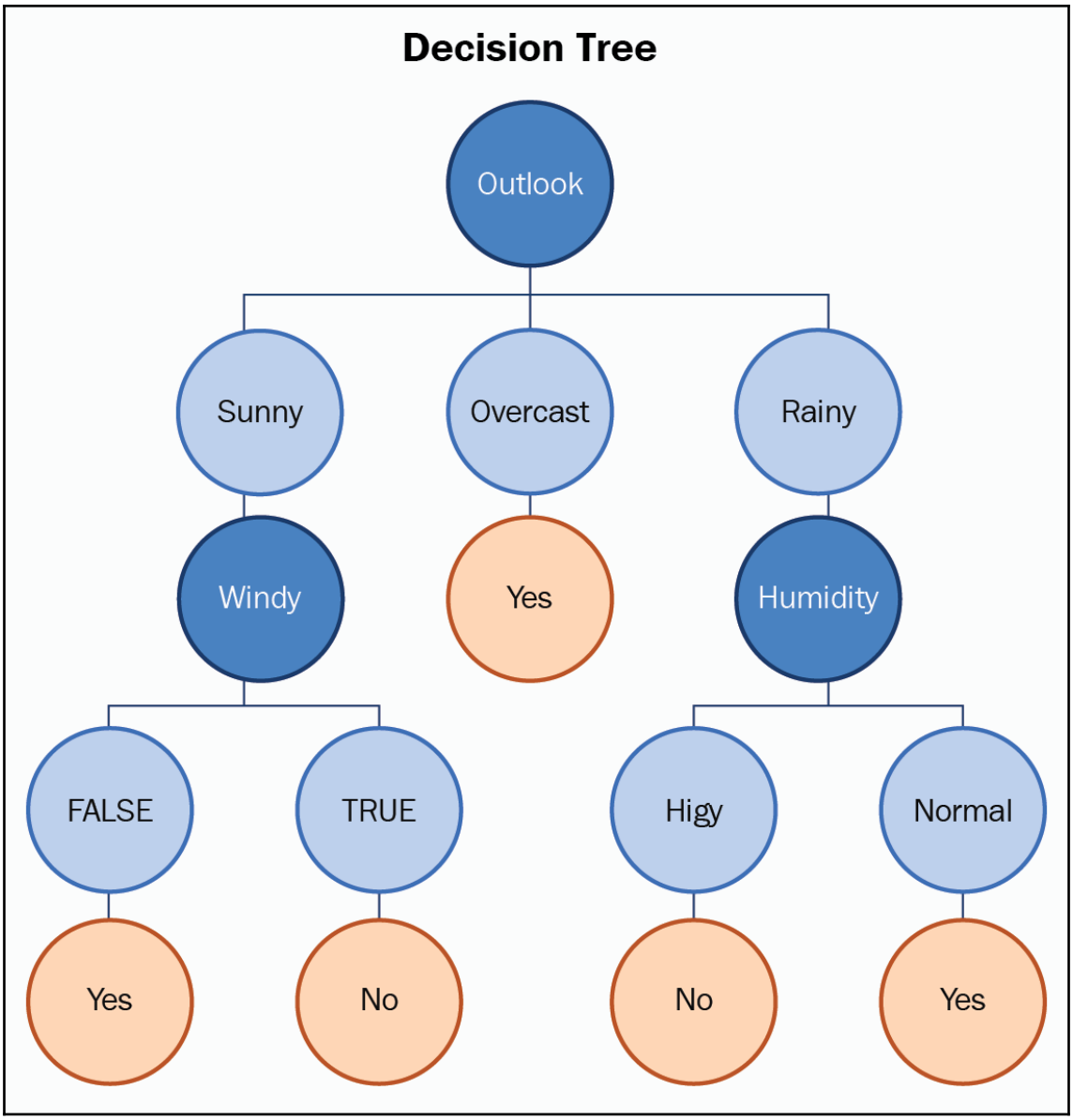
20	ID	Gender	Score	Years in company	IsProduction?	IsResearch?	IsSales?	Salary Increase
21	1	1	11	9	1	0	0	22
22	2	1	89	1	1	0	0	97
23	3	1	21	4	1	0	0	47
24	4	1	81	1	1	0	0	127
25	5	1	31	4	0	1	0	65
26	6	1	71	1	0	1	0	53
27	7	1	11	4	0	0	1	74
28	8	1	16	7	1	0	0	18
29	9	0	20	6	0	1	0	129
30	10	0	79	3	0	0	1	475
31	11	0	51	3	0	1	0	342
32	12	0	69	2	0	0	1	329
33	13	0	30	7	0	0	1	185
34	14	0	71	7	0	0	1	332
35	15	0	39	1	0	0	1	268
36	16	0	89	6	1	0	0	518
37	17	0	50	8	1	0	0	390



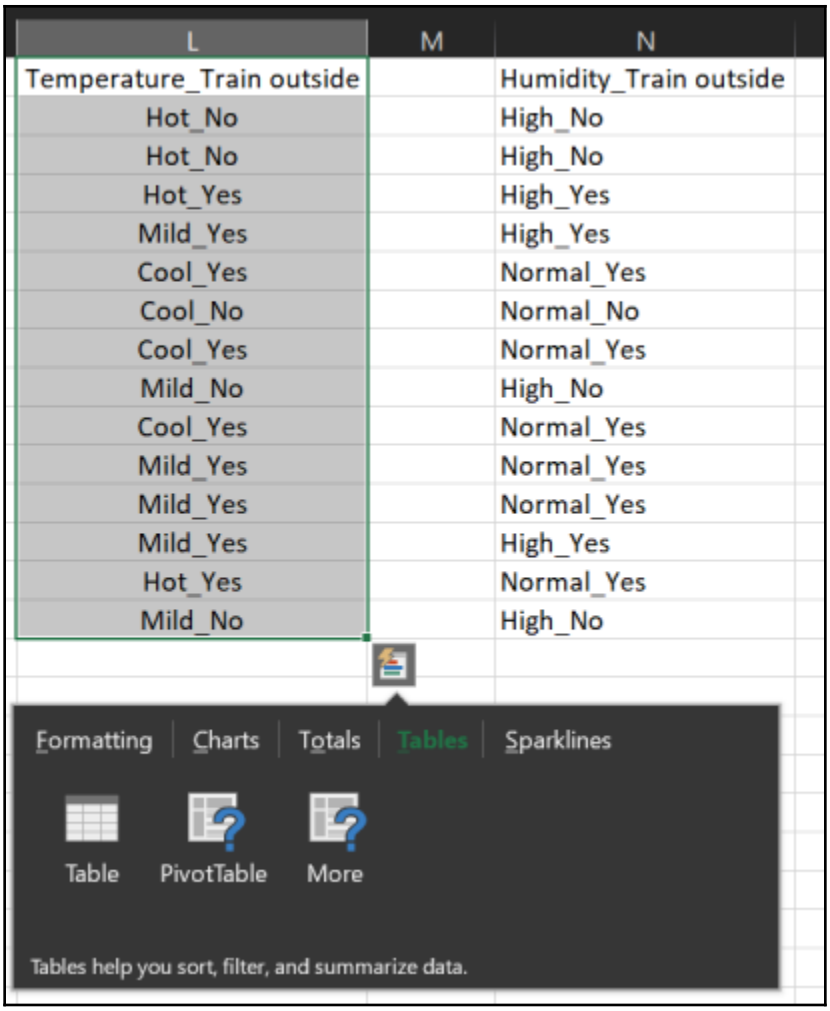



Regression
Input
Input Y Range:
$H$20:$H$37
Input X Range:
$E$20:$G$37
Labels
Constant is Zero
Confidence Level:
95
%
OK
Cancel
Help
Output options
Output Range:
New Worksheet Ply:
New Workbook
Residuals
Residuals
Residual Plots
Standardized Residuals
Line Fit Plots
Normal Probability
Normal Probability Plots

	A	B	C	D	E	F
1	ID	Temperature	Humidity	Windy	Outlook	Train outside
2	1	Hot	High	False	Rainy	No
3	2	Hot	High	True	Rainy	No
4	3	Hot	High	False	Overcast	Yes
5	4	Mild	High	False	Sunny	Yes
6	5	Cool	Normal	False	Sunny	Yes
7	6	Cool	Normal	True	Sunny	No
8	7	Cool	Normal	True	Overcast	Yes
9	8	Mild	High	False	Rainy	No
10	9	Cool	Normal	False	Rainy	Yes
11	10	Mild	Normal	False	Sunny	Yes
12	11	Mild	Normal	True	Rainy	Yes
13	12	Mild	High	True	Overcast	Yes
14	13	Hot	Normal	False	Overcast	Yes
15	14	Mild	High	True	Sunny	No
16						


Decision Tree
Outlook
Sunny
Overcast
Rainy
Windy
Yes
Humidity
FALSE
TRUE
Higy
Normal
Yes
No
No
Yes

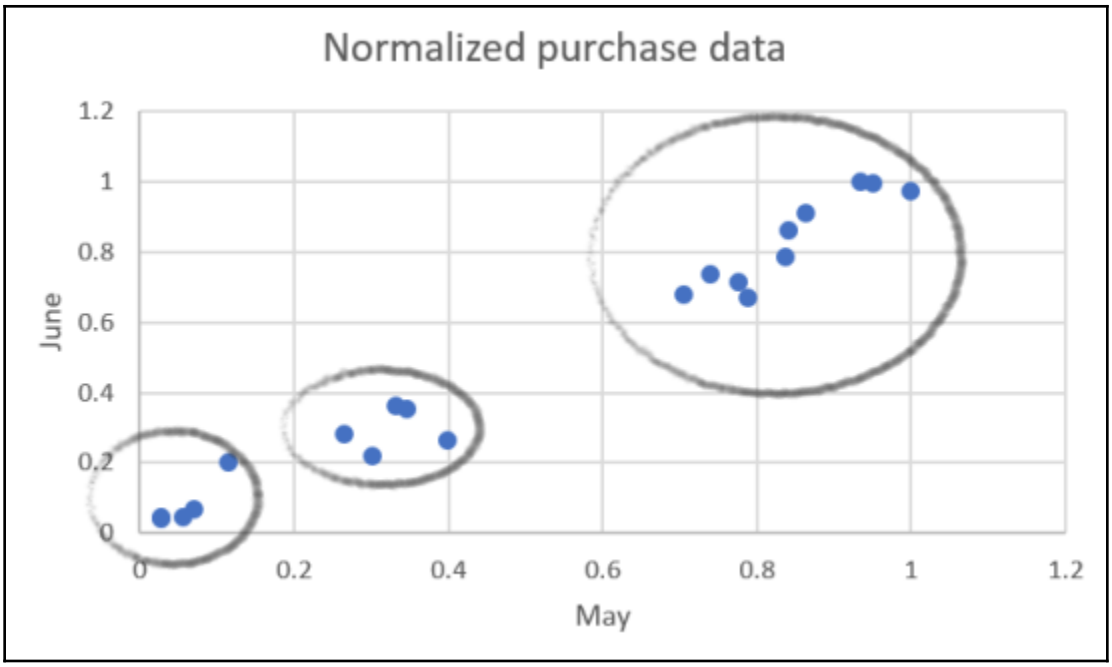
L	M	N	O	P	Q	R
Temperature_Train outside		Humidity_Train outside		Windy_Train outside		Outlook_Train outside
Hot_No		High_No		False_No		Rainy_No
Hot_No		High_No		True_No		Rainy_No
Hot_Yes		High_Yes		False_Yes		Overcast_Yes
Mild_Yes		High_Yes		False_Yes		Sunny_Yes
Cool_Yes		Normal_Yes		False_Yes		Sunny_Yes
Cool_No		Normal_No		True _No		Sunny_No
Cool_Yes		Normal_Yes		True_Yes		Overcast_Yes
Mild_No		High_No		False_No		Rainy_No
Cool_Yes		Normal_Yes		False_Yes		Rainy_Yes
Mild_Yes		Normal_Yes		False_Yes		Sunny_Yes
Mild_Yes		Normal_Yes		True_Yes		Rainy_Yes
Mild_Yes		High_Yes		True_Yes		Overcast_Yes
Hot_Yes		Normal_Yes		False_Yes		Overcast_Yes
Mild_No		High_No		True_No		Sunny_No


L
M
N
Temperature_Train outside
Humidity_Train outside
Hot_No
High_No
Hot_No
High_No
Hot_Yes
High_Yes
Mild_Yes
High_Yes
Cool_Yes
Normal_Yes
Cool_No
Normal_No
Cool_Yes
Normal_Yes
Mild_No
High_No
Cool_Yes
Normal_Yes
Mild_Yes
Normal_Yes
Mild_Yes
Normal_Yes
Mild_Yes
High_Yes
Hot_Yes
Normal_Yes
Mild_No
High_No
Formatting
Charts
Totals
Tables
Sparklines
Table
PivotTable
More
Tables help you sort, filter, and summarize data.

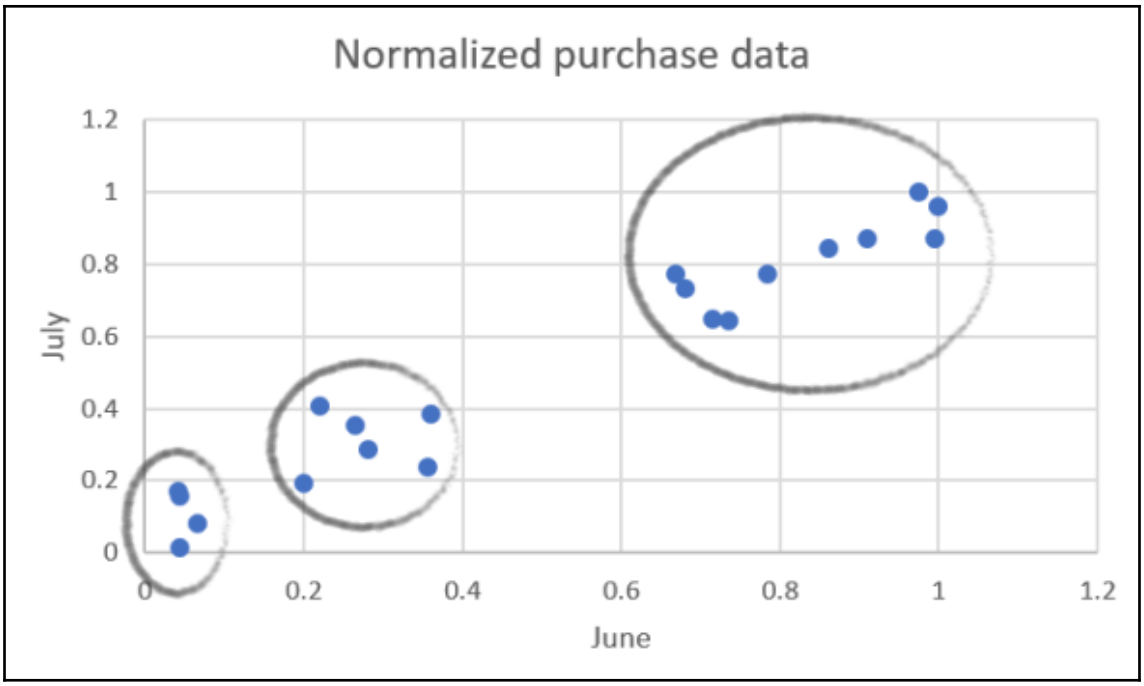
Row Labels	Count of Temperature_Train outside
Cool_No	1
Cool_Yes	3
Hot_No	2
Hot_Yes	2
Mild_No	2
Mild_Yes	4
Grand Total	**14**

	A	B	C	D
1	Customer	May	June	July
2	1	660.69	534.20	1867.02
3	2	841.70	819.17	950.46
4	3	316.89	500.66	2056.10
5	4	662.35	533.69	185.63
6	5	1372.03	2435.76	2289.55
7	6	4113.02	4338.35	2868.91
8	7	3137.67	3436.82	3448.07
9	8	3587.13	2695.09	4878.77
10	9	4730.14	3227.58	4232.90
11	10	3935.28	4404.20	4623.66
12	11	9946.74	9582.46	9239.03
13	12	9231.08	8732.51	7782.24
14	13	8388.63	8314.88	8750.93
15	14	9377.02	8156.16	9262.96
16	15	8786.33	8980.49	7708.27
17	16	11889.66	11902.72	11982.64
18	17	10000.12	10518.74	10096.42
19	18	11296.65	12149.68	10423.90
20	19	10270.80	11107.21	10433.03
21	20	11100.13	12214.41	11475.79

F	G	H	I
Customer	May	June	July
1	0.026653	0.043736	0.15581
2	0.070792	0.067066	0.079319
3	0.026653	0.040989	0.17159
4	0.055708	0.043694	0.015492
5	0.115397	0.199417	0.191073
6	0.345932	0.355183	0.239422
7	0.263899	0.281374	0.287756
8	0.301701	0.220648	0.407153
9	0.397836	0.264244	0.353253
10	0.330984	0.360574	0.385863
11	0.836588	0.784521	0.771035
12	0.776396	0.714935	0.649459
13	0.705539	0.680743	0.730301
14	0.78867	0.667749	0.773031
15	0.738989	0.735237	0.643287
16	1	0.974481	1
17	0.841077	0.861175	0.842588
18	0.950123	0.994701	0.869917
19	0.863843	0.909353	0.870679
20	0.933595	1	0.957702


Normalized purchase data
1.2
1
0.8
0.6
0.4
0.2
0
June
0
0.2
0.4
0.6
0.8
1
1.2
May


Normalized purchase data
1.2
1
0.8
0.6
0.4
0.2
0
July
0
5
10
15
20
25
May


Normalized purchase data
1.2
1
0.8
0.6
0.4
0.2
0
July
0
0.2
0.4
0.6
0.8
1
1.2
June

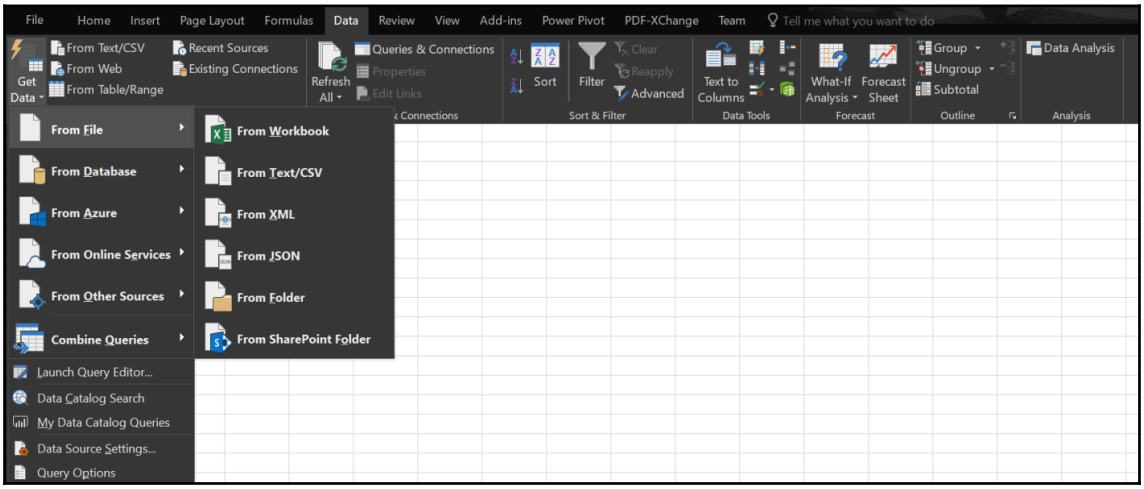
	A	B	C	D	E	F	G	H
1	Customer	May	June	July	D1	D2	D3	Cluster
2	1	0.055568	0.043736	0.15581	0	0.006627	0.001093	1
3	2	0.070792	0.067066	0.079319	0.006627	0	0.011142	2
4	3	0.026653	0.040989	0.17159	0.001093	0.011142	0	3
5	4	0.055708	0.043694	0.015492	0.019689	0.004848	0.025218	2
6	5	0.115397	0.199417	0.191073	0.02906	0.031995	0.033355	1
7	6	0.345932	0.355183	0.239422	0.188301	0.184346	0.205258	2
8	7	0.263899	0.281374	0.287756	0.117283	0.126664	0.127565	1
9	8	0.301701	0.220648	0.407153	0.155053	0.184382	0.163419	1
10	9	0.397836	0.264244	0.353253	0.204755	0.220877	0.220622	1
11	10	0.330984	0.360574	0.385863	0.229165	0.247816	0.240665	1
12	11	0.836588	0.784521	0.771035	1.537256	1.579654	1.568169	1
13	12	0.776396	0.714935	0.649459	1.21379	1.24267	1.244677	1
14	13	0.705539	0.680743	0.730301	1.158281	1.203281	1.182331	1
15	14	0.78867	0.667749	0.773031	1.307793	1.357405	1.335231	1
16	15	0.738989	0.735237	0.643287	1.182872	1.210999	1.211902	1
17	16	1	0.974481	1	2.470895	2.534482	2.505076	1
18	17	0.841077	0.861175	0.842588	1.756893	1.806525	1.786229	1
19	18	0.950123	0.994701	0.869917	2.214512	2.258773	2.250025	1
20	19	0.863843	0.909353	0.870679	1.913637	1.964625	1.943668	1
21	20	0.933595	1	0.957702	2.328403	2.38635	2.360218	1

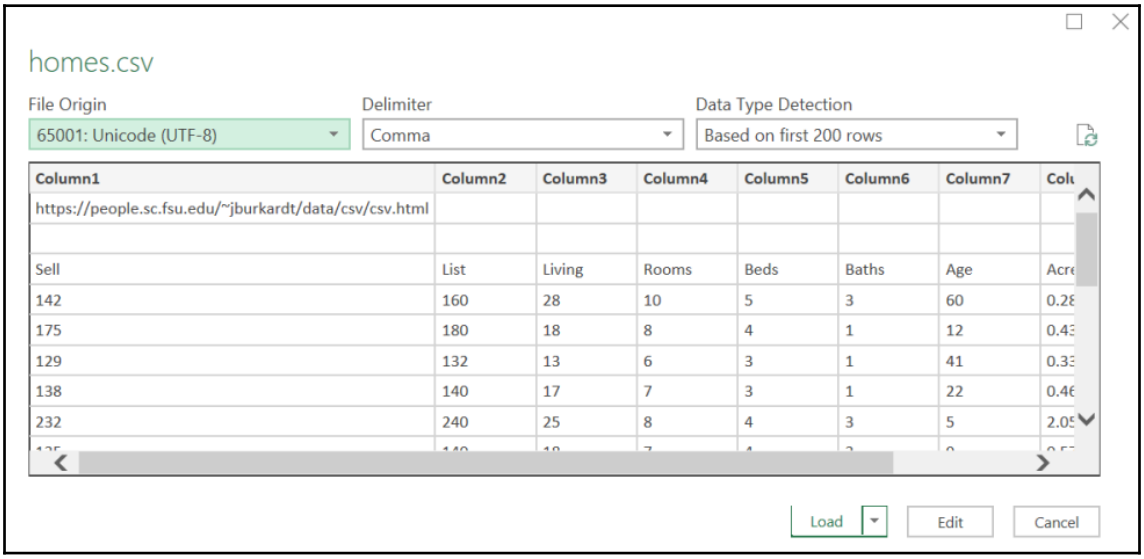
	A	B	C	D	E	F	G	H
1	Customer	May	June	July	D1	D2	D3	Cluster
2	1	0.055568	0.043736	0.15581	0	0.006627	0.001093	1
3	5	0.115397	0.199417	0.191073	0.02906	0.031995	0.033355	1
4	7	0.263899	0.281374	0.287756	0.117283	0.126664	0.127565	1
5	8	0.301701	0.220648	0.407153	0.155053	0.184382	0.163419	1
6	9	0.397836	0.264244	0.353253	0.204755	0.220877	0.220622	1
7	10	0.330984	0.360574	0.385863	0.229165	0.247816	0.240665	1
8	11	0.836588	0.784521	0.771035	1.537256	1.579654	1.568169	1
9	12	0.776396	0.714935	0.649459	1.21379	1.24267	1.244677	1
10	13	0.705539	0.680743	0.730301	1.158281	1.203281	1.182331	1
11	14	0.78867	0.667749	0.773031	1.307793	1.357405	1.335231	1
12	15	0.738989	0.735237	0.643287	1.182872	1.210999	1.211902	1
13	16	1	0.974481	1	2.470895	2.534482	2.505076	1
14	17	0.841077	0.861175	0.842588	1.756893	1.806525	1.786229	1
15	18	0.950123	0.994701	0.869917	2.214512	2.258773	2.250025	1
16	19	0.863843	0.909353	0.870679	1.913637	1.964625	1.943668	1
17	20	0.933595	1	0.957702	2.328403	2.38635	2.360218	1
18	2	0.070792	0.067066	0.079319	0.006627	0	0.011142	2
19	4	0.055708	0.043694	0.015492	0.019689	0.004848	0.025218	2
20	6	0.345932	0.355183	0.239422	0.188301	0.184346	0.205258	2
21	3	0.026653	0.040989	0.17159	0.001093	0.011142	0	3

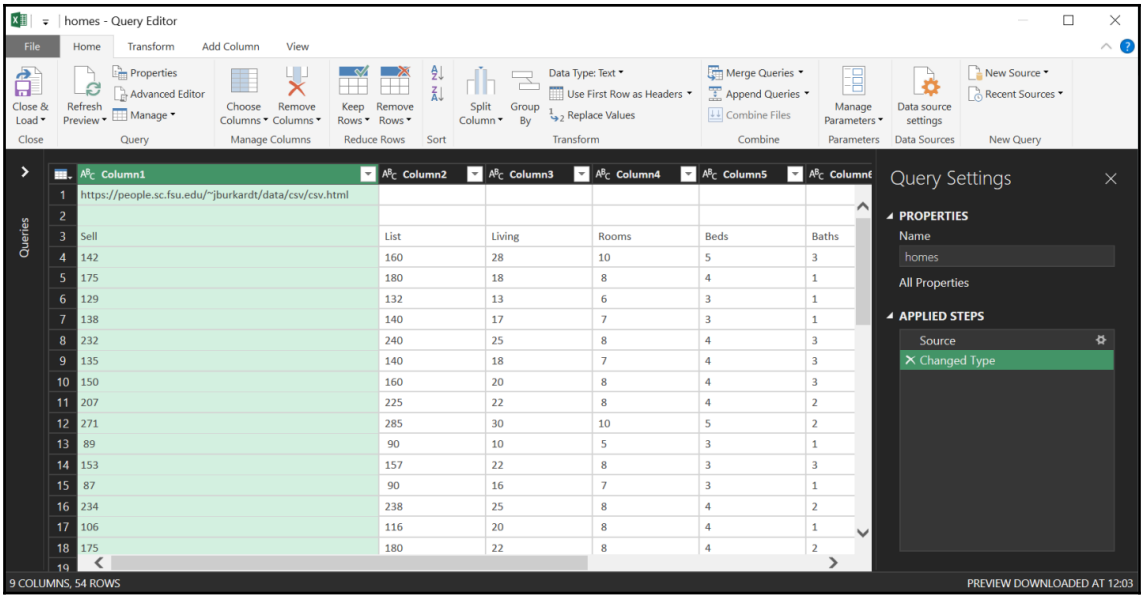
Customer	May	June	July	D1	D2	D3	Cluster
1	0.055568	0.043736	0.15581	0.846783	0.024807	0.001093	3
2	0.070792	0.067066	0.079319	0.88075	0.016332	0.011142	3
3	0.026653	0.040989	0.17159	0.868945	0.033807	0	3
4	0.055708	0.043694	0.015492	0.996086	0.032017	0.025218	3
5	0.115397	0.199417	0.191073	0.600844	0.010062	0.033355	2
6	0.345932	0.355183	0.239422	0.280613	0.091849	0.205258	2
7	0.263899	0.281374	0.287756	0.340283	0.058314	0.127565	2
8	0.301701	0.220648	0.407153	0.293354	0.112533	0.163419	2
9	0.397836	0.264244	0.353253	0.235594	0.128126	0.220622	2
10	0.330984	0.360574	0.385863	0.196869	0.14756	0.240665	2
11	0.836588	0.784521	0.771035	0.102789	1.292196	1.568169	1
12	0.776396	0.714935	0.649459	0.037743	0.985731	1.244677	1
13	0.705539	0.680743	0.730301	0.025745	0.959473	1.182331	1
14	0.78867	0.667749	0.773031	0.056723	1.098735	1.335231	1
15	0.738989	0.735237	0.643287	0.031843	0.957358	1.211902	1
16	1	0.974481	1	0.427144	2.17047	2.505076	1
17	0.841077	0.861175	0.842588	0.165051	1.500166	1.786229	1
18	0.950123	0.994701	0.869917	0.324473	1.908189	2.250025	1
19	0.863843	0.909353	0.870679	0.216023	1.644013	1.943668	1
20	0.933595	1	0.957702	0.369867	2.032061	2.360218	1

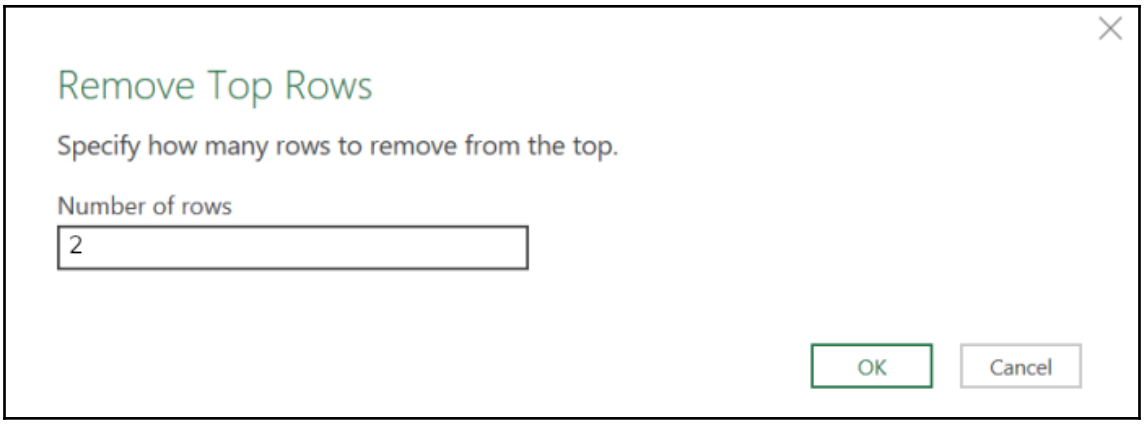
Customer	May	June	July	D1	D2	D3	Cluster
1	0.055568	0.043736	0.15581	1.671637	0.136138	0.002564	3
2	0.070792	0.067066	0.079319	1.71768	0.148215	0.001366	3
3	0.026653	0.040989	0.17159	1.701956	0.147349	0.005075	3
4	0.055708	0.043694	0.015492	1.874987	0.199264	0.00815	3
5	0.115397	0.199417	0.191073	1.314697	0.052266	0.033974	3
6	0.345932	0.355183	0.239422	0.801659	0.013546	0.198038	2
7	0.263899	0.281374	0.287756	0.913	0.001355	0.13208	2
8	0.301701	0.220648	0.407153	0.830561	0.012926	0.182731	2
9	0.397836	0.264244	0.353253	0.730625	0.013132	0.227219	2
10	0.330984	0.360574	0.385863	0.66574	0.013567	0.253464	2
11	0.836588	0.784521	0.771035	0.003911	0.762054	1.599342	1
12	0.776396	0.714935	0.649459	0.044303	0.537715	1.263963	1
13	0.705539	0.680743	0.730301	0.048474	0.506921	1.216451	1
14	0.78867	0.667749	0.773031	0.031504	0.609925	1.370955	1
15	0.738989	0.735237	0.643287	0.048399	0.516842	1.231963	1
16	1	0.974481	1	0.080513	1.457411	2.555153	1
17	0.841077	0.861175	0.842588	0.001851	0.921132	1.825415	1
18	0.950123	0.994701	0.869917	0.041245	1.255422	2.285148	1
19	0.863843	0.909353	0.870679	0.009939	1.035589	1.984642	1
20	0.933595	1	0.957702	0.057827	1.347439	2.407696	1

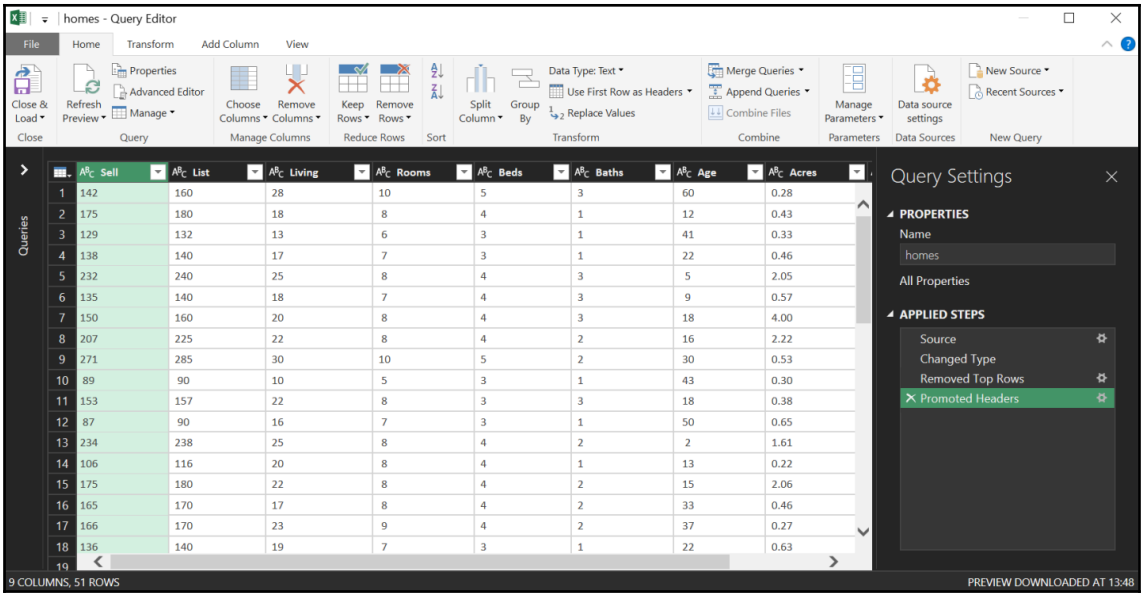
Chapter 3: Importing Data into Excel from Different Data Sources

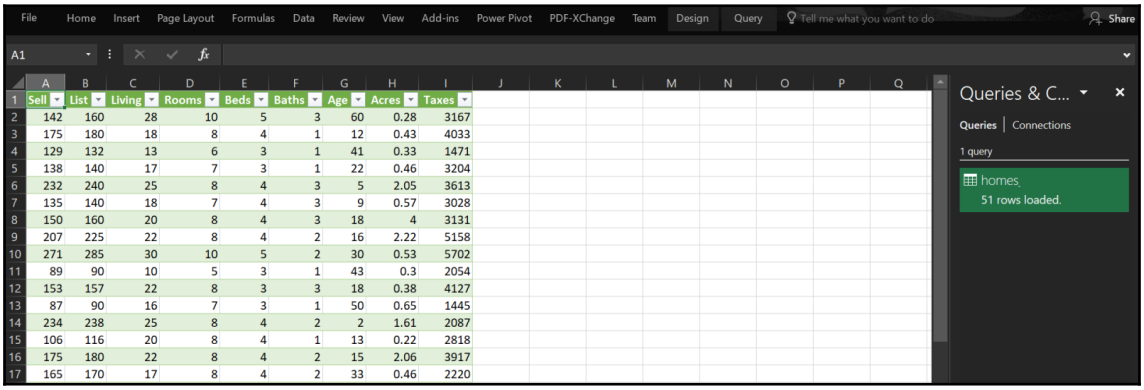


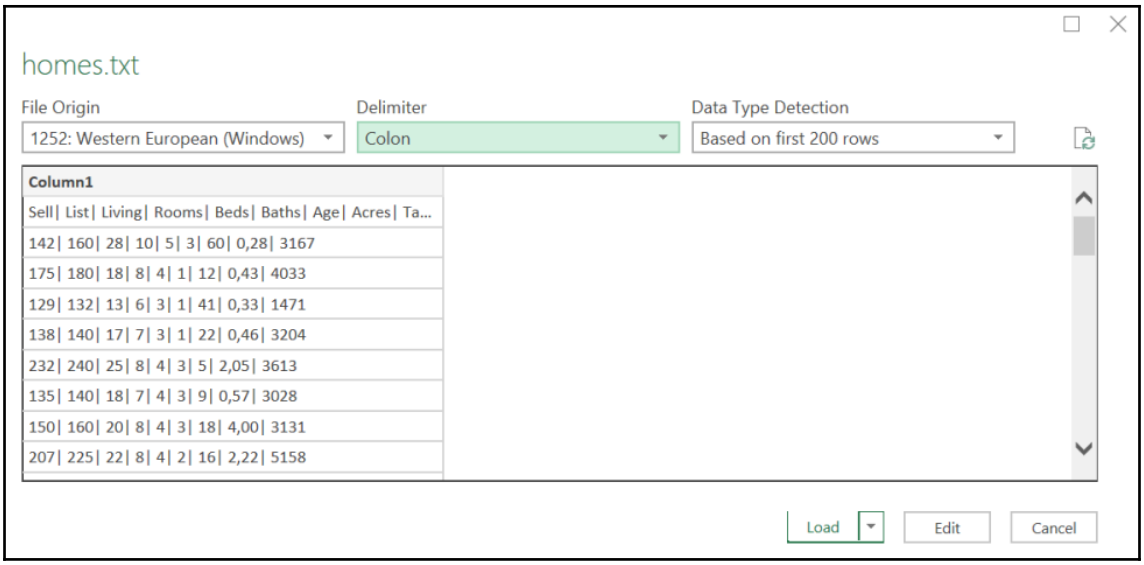


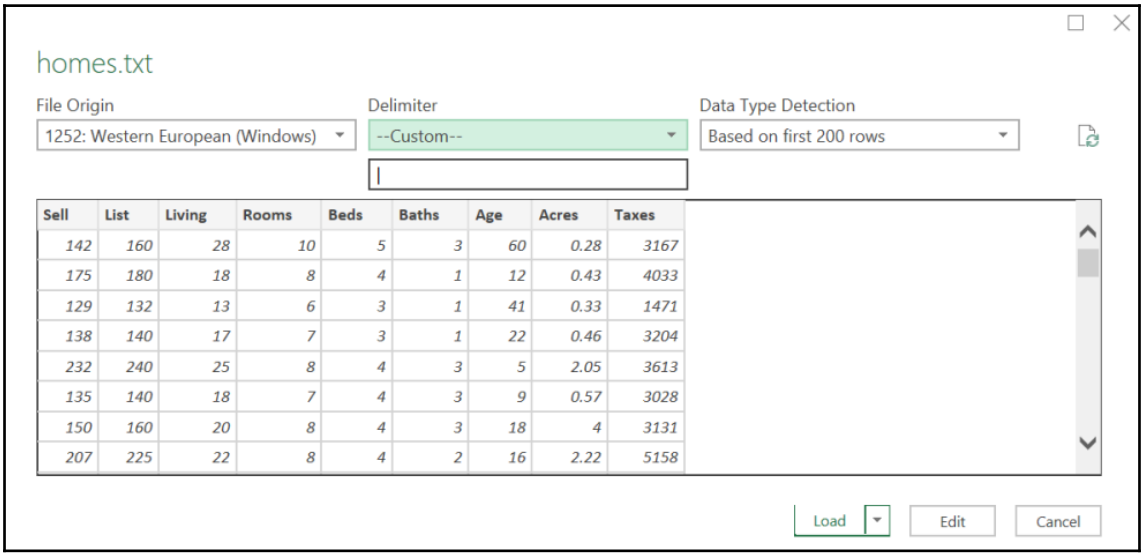

homes - Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Close
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Parameters
Data Sources
New Query
Queries
Column1
Column2
Column3
Column4
Column5
Column6
https://people.sc.fsu.edu/~jburkardt/data/csv/csv.html
Sell
List
Living
Rooms
Beds
Baths
142 160 28 10 5 3
175 180 18 8 4 1
129 132 13 6 3 1
138 140 17 7 3 1
232 240 25 8 4 3
135 140 18 7 4 3
150 160 20 8 4 3
207 225 22 8 4 2
271 285 30 10 5 2
89 90 10 5 3 1
153 157 22 8 3 3
87 90 16 7 3 1
234 238 25 8 4 2
106 116 20 8 4 1
175 180 22 8 4 2
Query Settings
PROPERTIES
Name
homes
All Properties
APPLIED STEPS
Source
Changed Type
9 COLUMNS, 54 ROWS
PREVIEW DOWNLOADED AT 12:03

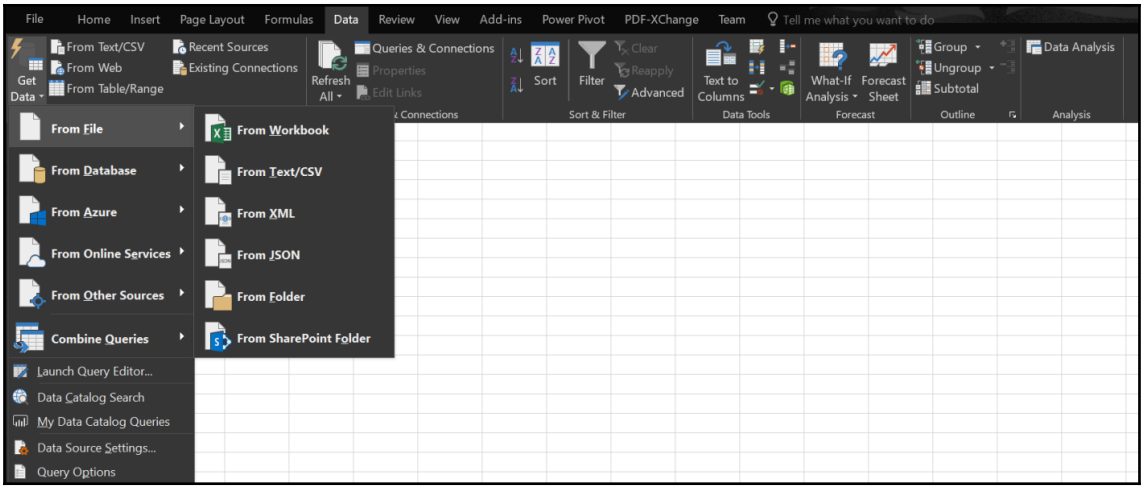

Remove Top Rows
Specify how many rows to remove from the top.
Number of rows
2
OK
Cancel

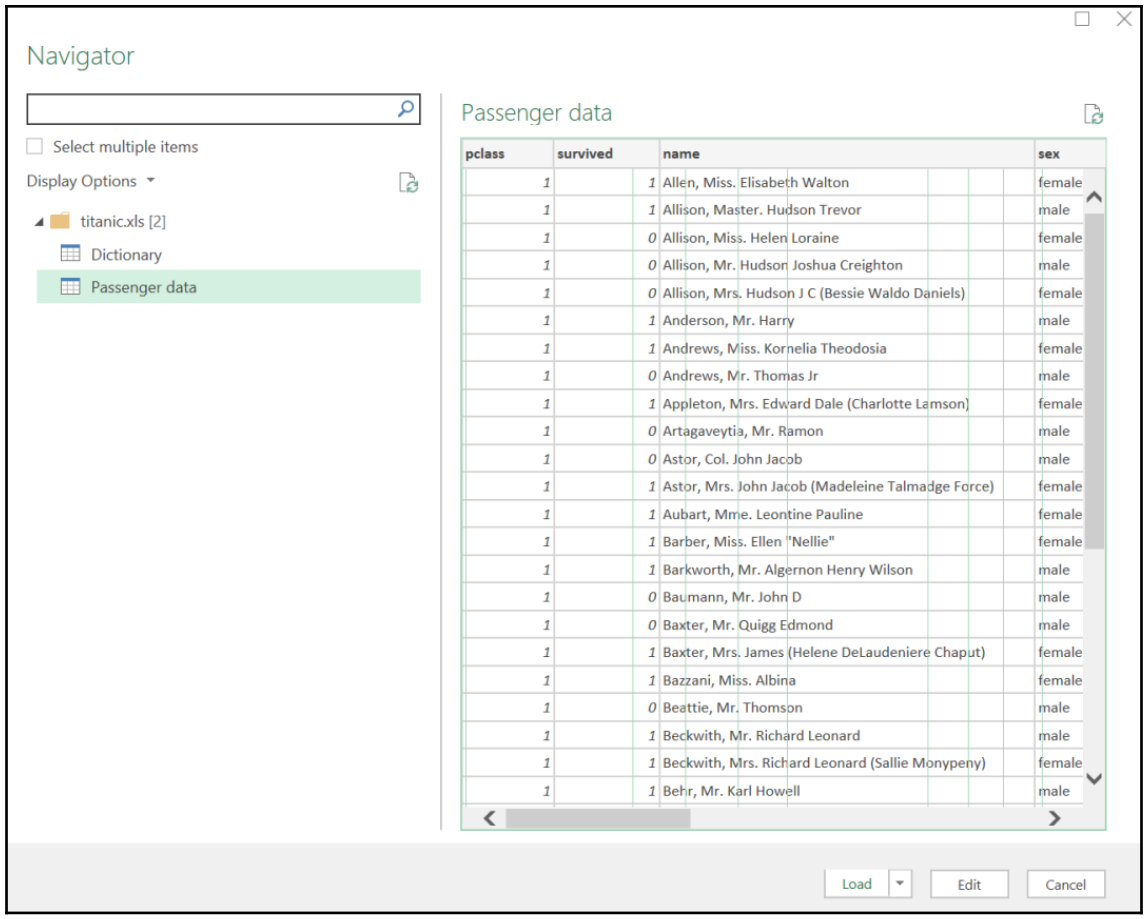

homes - Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Close
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Parameters
Data Sources
New Query
Queries
Sell List Living Rooms Beds Baths Age Acres
1 142 160 28 10 5 3 60 0.28
2 175 180 18 8 4 1 12 0.43
3 129 132 13 6 3 1 41 0.33
4 138 140 17 7 3 1 22 0.46
5 232 240 25 8 4 3 5 2.05
6 135 140 18 7 4 3 9 0.57
7 150 160 20 8 4 3 18 4.00
8 207 225 22 8 4 2 16 2.22
9 271 285 30 10 5 2 30 0.53
10 89 90 10 5 3 1 43 0.30
11 153 157 22 8 3 3 18 0.38
12 87 90 16 7 3 1 50 0.65
13 234 238 25 8 4 2 2 1.61
14 106 116 20 8 4 1 13 0.22
15 175 180 22 8 4 2 15 2.06
16 165 170 17 8 4 2 33 0.46
17 166 170 23 9 4 2 37 0.27
18 136 140 19 7 3 1 22 0.63
19
Query Settings
PROPERTIES
Name
homes
All Properties
APPLIED STEPS
Source
Changed Type
Removed Top Rows
Promoted Headers
9 COLUMNS, 51 ROWS
PREVIEW DOWNLOADED AT 13:48

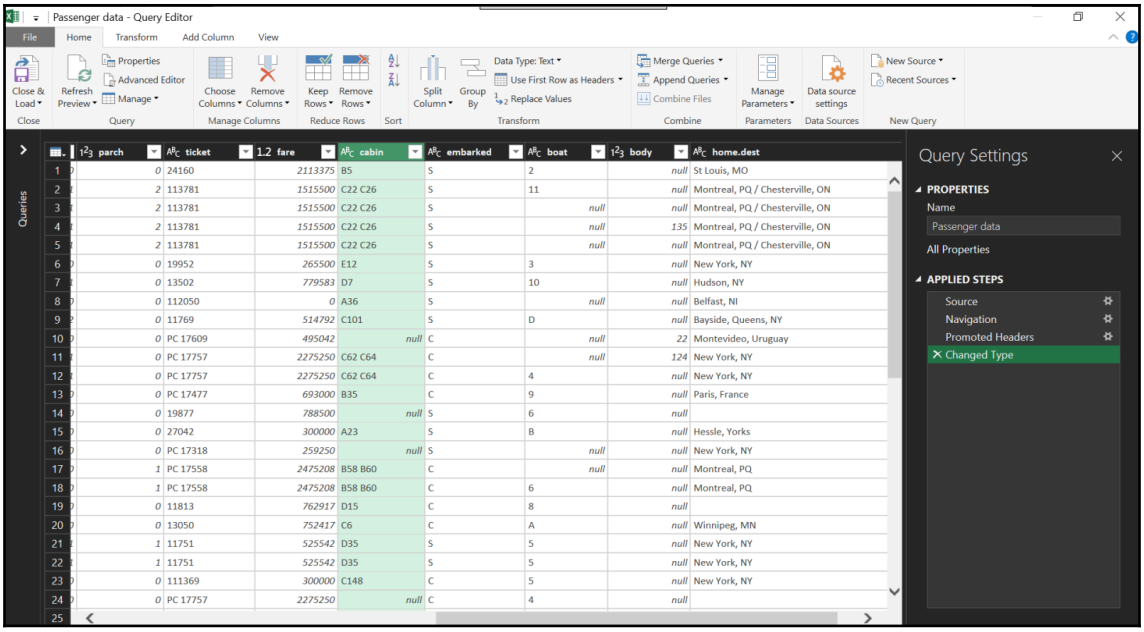

File Home Insert Page Layout Formulas Data Review View Add-ins Power Pivot PDF-XChange Team Design Query
Tell me what you want to do
Share
A1
Sell List Living Rooms Beds Baths Age Acres Taxes
2 142 160 28 10 5 3 60 0.28 3167
3 175 180 18 8 4 1 12 0.43 4033
4 129 132 13 6 3 1 41 0.33 1471
5 138 140 17 7 3 1 22 0.46 3204
6 232 240 25 8 4 3 5 2.05 3613
7 135 140 18 7 4 3 9 0.57 3028
8 150 160 20 8 4 3 18 4 3131
9 207 225 22 8 4 2 16 2.22 5158
10 271 285 30 10 5 2 30 0.53 5702
11 89 90 10 5 3 1 43 0.3 2054
12 153 157 22 8 3 3 18 0.38 4127
13 87 90 16 7 3 1 50 0.65 1445
14 234 238 25 8 4 2 2 1.61 2087
15 106 116 20 8 4 1 13 0.22 2818
16 175 180 22 8 4 2 15 2.06 3917
17 165 170 17 8 4 2 33 0.46 2220
Queries & C...
Queries | Connections
1 query
homes
51 rows loaded.

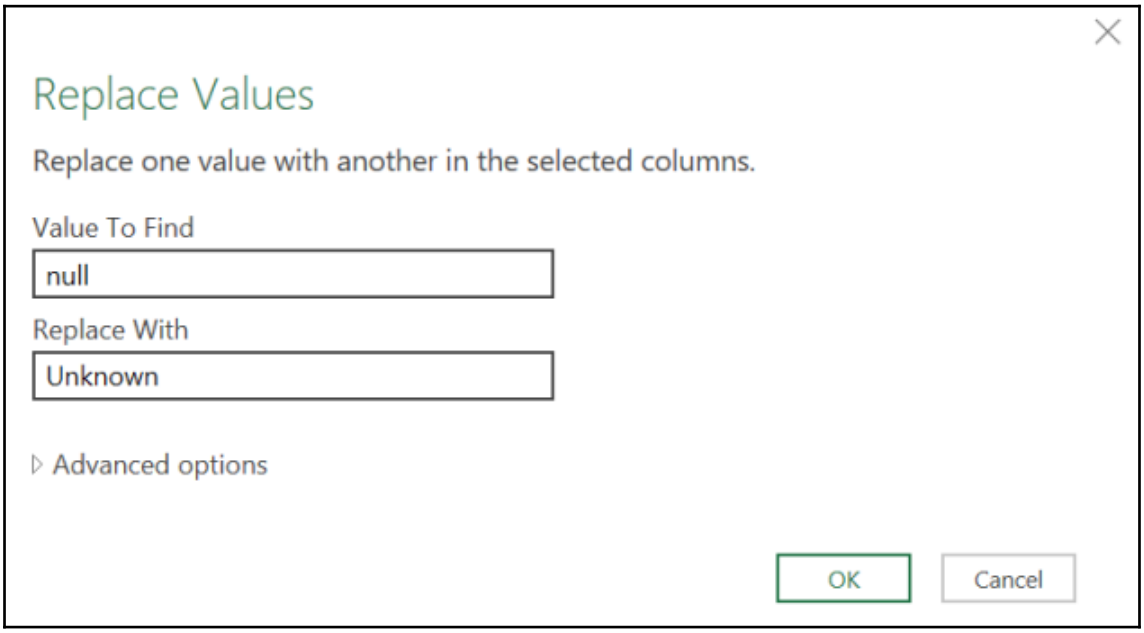

homes.txt
File Origin
1252: Western European (Windows)
Delimiter
Colon
Data Type Detection
Based on first 200 rows
Column1
Sell| List| Living| Rooms| Beds| Baths| Age| Acres| Ta...
142| 160| 28| 10| 5| 3| 60| 0,28| 3167
175| 180| 18| 8| 4| 1| 12| 0,43| 4033
129| 132| 13| 6| 3| 1| 41| 0,33| 1471
138| 140| 17| 7| 3| 1| 22| 0,46| 3204
232| 240| 25| 8| 4| 3| 5| 2,05| 3613
135| 140| 18| 7| 4| 3| 9| 0,57| 3028
150| 160| 20| 8| 4| 3| 18| 4,00| 3131
207| 225| 22| 8| 4| 2| 16| 2,22| 5158
Load
Edit
Cancel

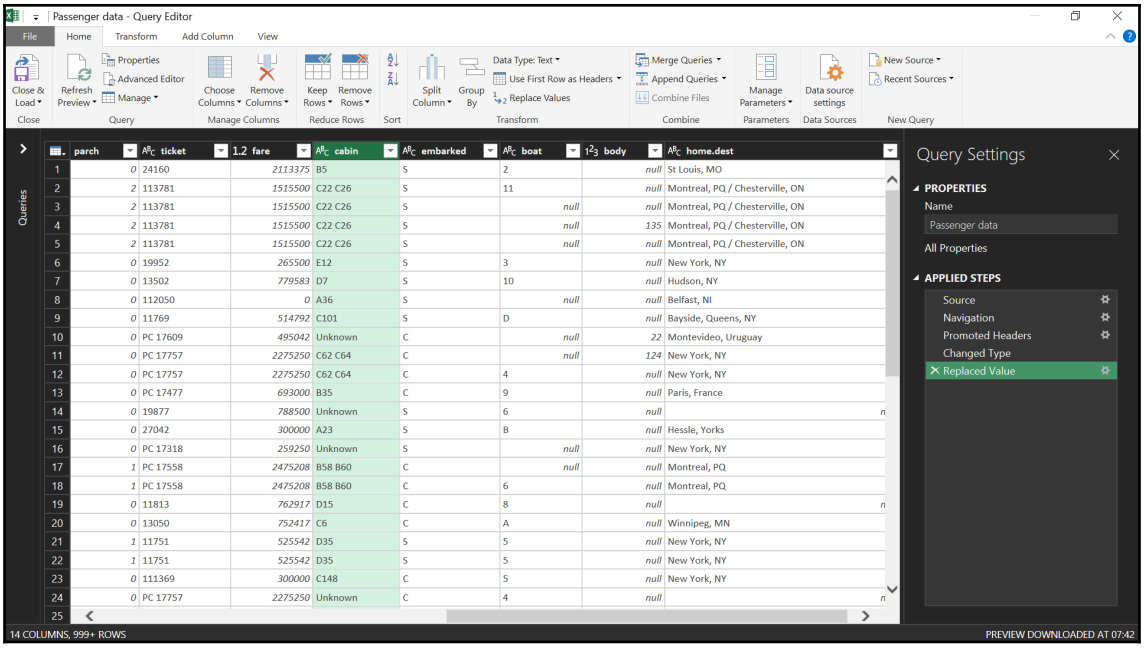

homes.txt
File Origin
1252: Western European (Windows)
Delimiter
--Custom--
Data Type Detection
Based on first 200 rows
Sell List Living Rooms Beds Baths Age Acres Taxes
142 160 28 10 5 3 60 0.28 3167
175 180 18 8 4 1 12 0.43 4033
129 132 13 6 3 1 41 0.33 1471
138 140 17 7 3 1 22 0.46 3204
232 240 25 8 4 3 5 2.05 3613
135 140 18 7 4 3 9 0.57 3028
150 160 20 8 4 3 18 4 3131
207 225 22 8 4 2 16 2.22 5158
Load
Edit
Cancel

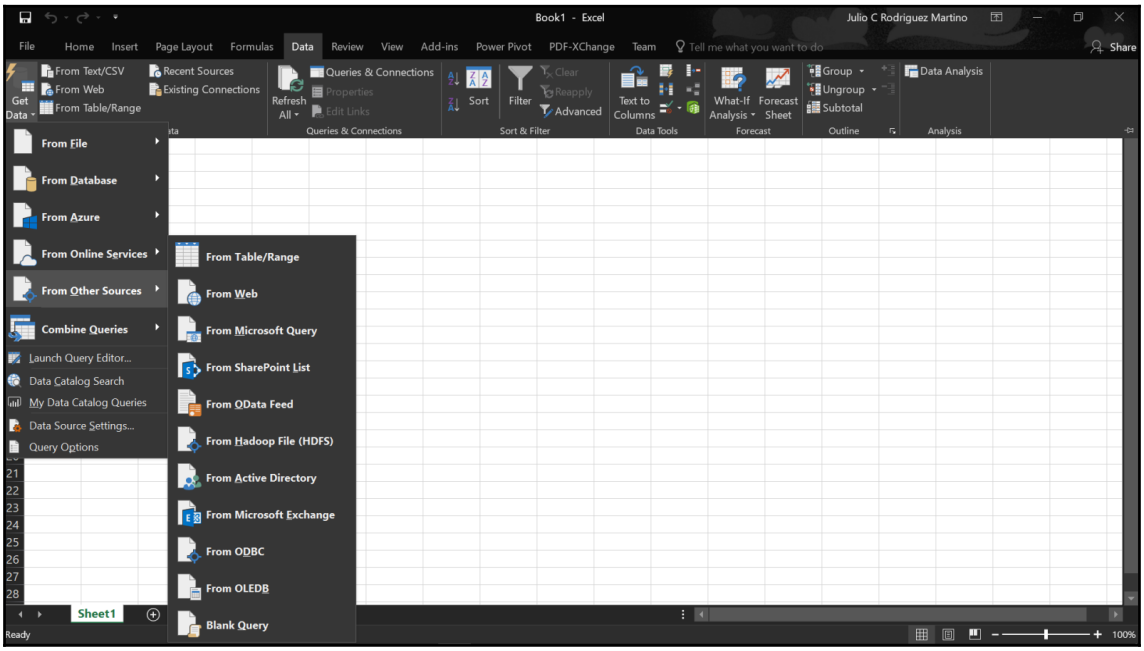

File
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-ins
Power Pivot
PDF-XChange
Team
Tell me what you want to do
Get Data
From Text/CSV
From Web
From Table/Range
Recent Sources
Existing Connections
Refresh All
Queries & Connections
Properties
Edit Links
Sort
Filter
Clear
Reapply
Advanced
Text to Columns
What-If Analysis
Forecast Sheet
Group
Ungroup
Subtotal
Data Analysis
Sort & Filter
Data Tools
Forecast
Outline
Analysis
From File
From Database
From Azure
From Online Services
From Other Sources
Combine Queries
Launch Query Editor...
Data Catalog Search
My Data Catalog Queries
Data Source Settings...
Query Options
From Workbook
From Text/CSV
From XML
From JSON
From Folder
From SharePoint Folder

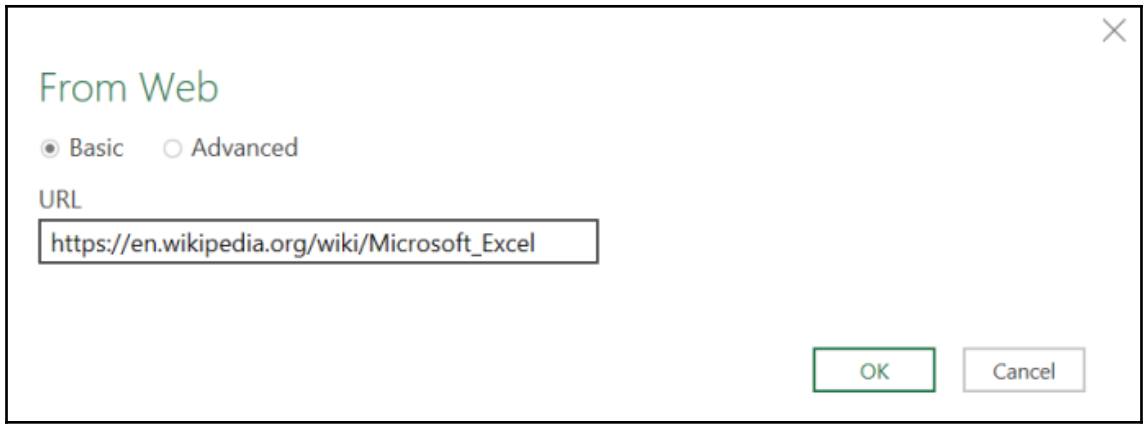

Navigator
Select multiple items
Display Options
titanic.xls [2]
Dictionary
Passenger data
Passenger data
pclass survived name sex
1 1 Allen, Miss. Elisabeth Walton female
1 1 Allison, Master. Hudson Trevor male
1 0 Allison, Miss. Helen Loraine female
1 0 Allison, Mr. Hudson Joshua Creighton male
1 0 Allison, Mrs. Hudson J C (Bessie Waldo Daniels) female
1 1 Anderson, Mr. Harry male
1 1 Andrews, Miss. Kornelia Theodosia female
1 0 Andrews, Mr. Thomas Jr male
1 1 Appleton, Mrs. Edward Dale (Charlotte Lamson) female
1 0 Artagaveytia, Mr. Ramon male
1 0 Astor, Col. John Jacob male
1 1 Astor, Mrs. John Jacob (Madeleine Talmadge Force) female
1 1 Aubart, Mme. Leontine Pauline female
1 1 Barber, Miss. Ellen "Nellie" female
1 1 Barkworth, Mr. Algernon Henry Wilson male
1 0 Baumann, Mr. John D male
1 0 Baxter, Mr. Quigg Edmond male
1 1 Baxter, Mrs. James (Helene DeLaudeniere Chaput) female
1 1 Bazzani, Miss. Albina female
1 0 Beattie, Mr. Thomson male
1 1 Beckwith, Mr. Richard Leonard male
1 1 Beckwith, Mrs. Richard Leonard (Sallie Monypeny) female
1 1 Behr, Mr. Karl Howell male
Load
Edit
Cancel

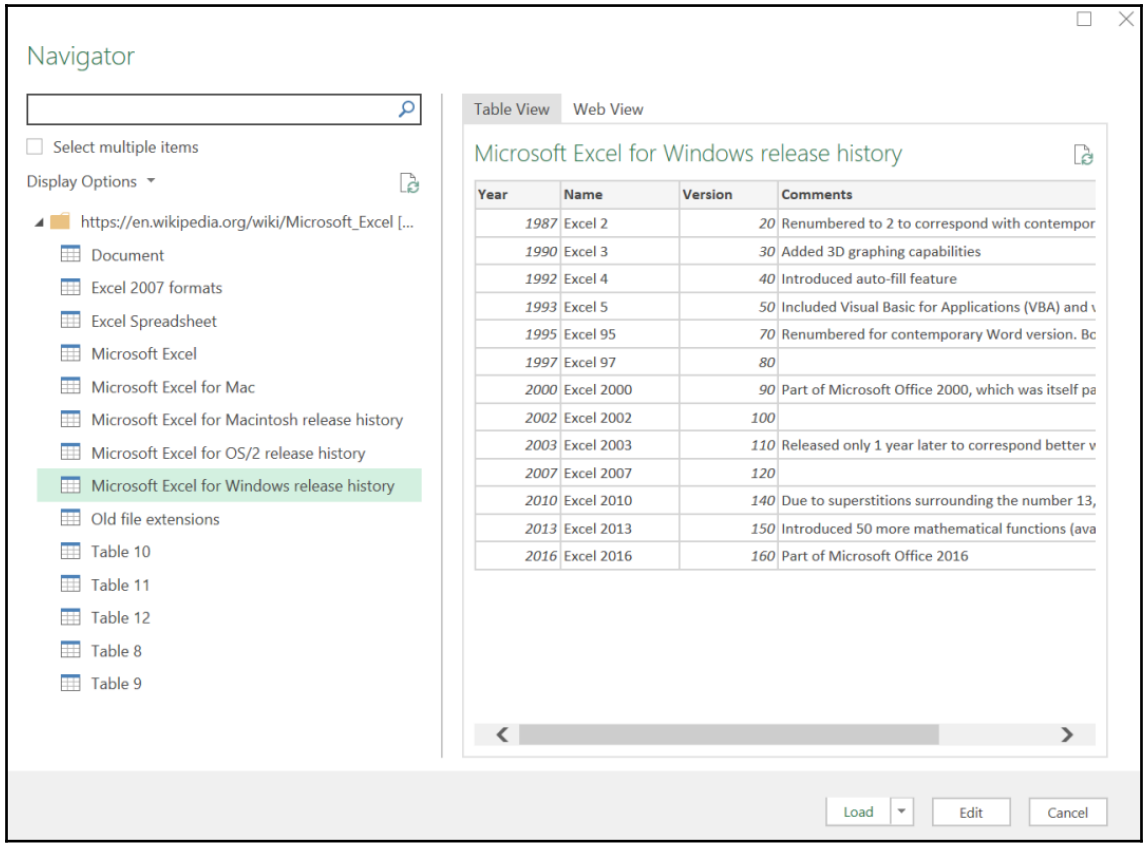

Passenger data - Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Close
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Parameters
Data Sources
New Query
Queries
parch
ticket
fare
cabin
embarked
boat
body
home.dest
1 0 24160 2113375 B5 S 2 null St Louis, MO
2 2 113781 1515500 C22 C26 S 11 null Montreal, PQ / Chesterville, ON
3 2 113781 1515500 C22 C26 S null null Montreal, PQ / Chesterville, ON
4 2 113781 1515500 C22 C26 S null 135 Montreal, PQ / Chesterville, ON
5 2 113781 1515500 C22 C26 S null null Montreal, PQ / Chesterville, ON
6 0 19952 265500 E12 S 3 null New York, NY
7 0 13502 779583 D7 S 10 null Hudson, NY
8 0 112050 0 A36 S null null Belfast, NI
9 0 11769 514792 C101 S D null Bayside, Queens, NY
10 0 PC 17609 495042 null C null 22 Montevideo, Uruguay
11 0 PC 17757 2275250 C62 C64 C null 124 New York, NY
12 0 PC 17757 2275250 C62 C64 C 4 null New York, NY
13 0 PC 17477 693000 B35 C 9 null Paris, France
14 0 19877 788500 null S 6 null
15 0 27042 300000 A23 S B null Hessle, Yorks
16 0 PC 17318 259250 null S null null New York, NY
17 1 PC 17558 2475208 B58 B60 C null null Montreal, PQ
18 1 PC 17558 2475208 B58 B60 C 6 null Montreal, PQ
19 0 11813 762917 D15 C 8 null
20 0 13050 752417 C6 C A null Winnipeg, MN
21 1 11751 525542 D35 S 5 null New York, NY
22 1 11751 525542 D35 S 5 null New York, NY
23 0 111369 300000 C148 C 5 null New York, NY
24 0 PC 17757 2275250 null C 4 null
25
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type

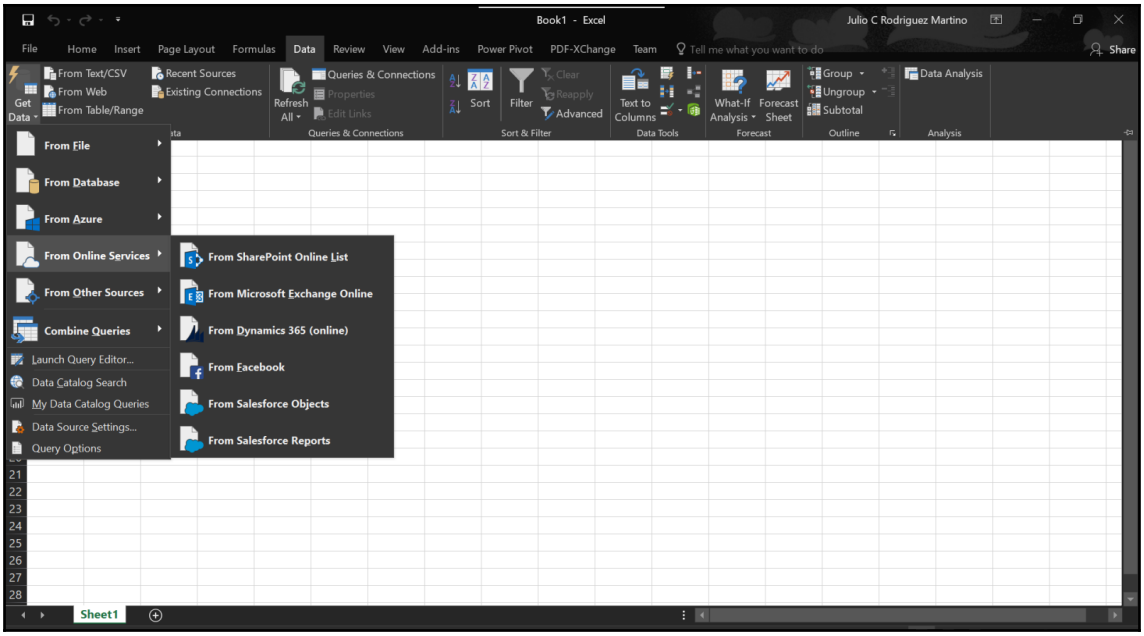

Replace Values
Replace one value with another in the selected columns.
Value To Find
null
Replace With
Unknown
Advanced options
OK
Cancel

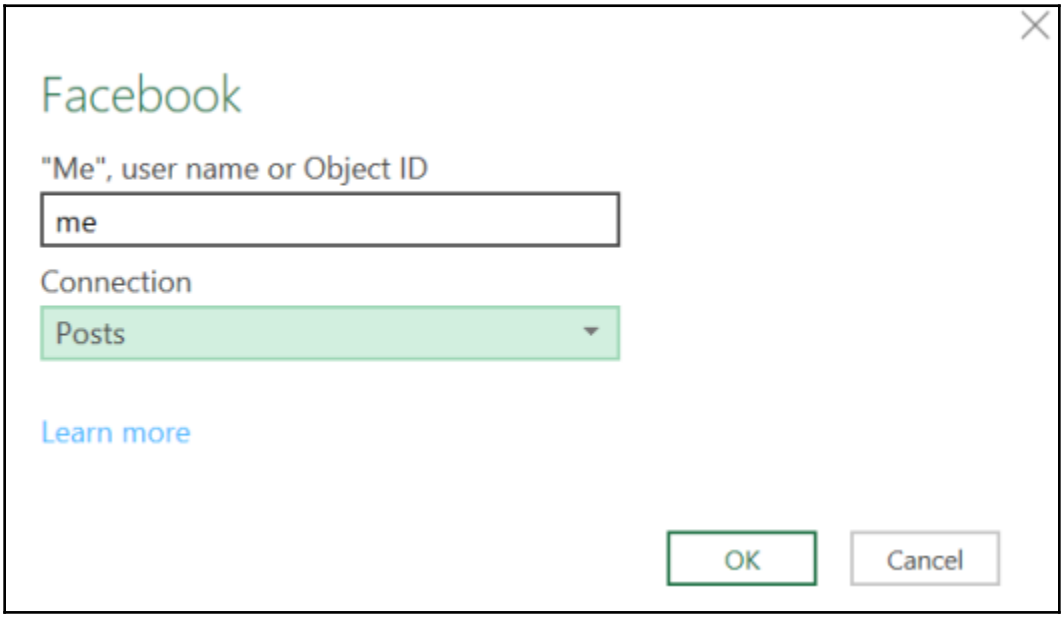

Passenger data - Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Close
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Parameters
Data Sources
New Query
Queries
parch
ticket
fare
cabin
embarked
boat
body
home.dest
1 0 24160 2113375 B5 S 2 null St Louis, MO
2 2 113781 1515500 C22 C26 S 11 null Montreal, PQ / Chesterville, ON
3 2 113781 1515500 C22 C26 S null null Montreal, PQ / Chesterville, ON
4 2 113781 1515500 C22 C26 S null 135 Montreal, PQ / Chesterville, ON
5 2 113781 1515500 C22 C26 S null null Montreal, PQ / Chesterville, ON
6 0 19952 265500 E12 S 3 null New York, NY
7 0 13502 779583 D7 S 10 null Hudson, NY
8 0 112050 0 A36 S null null Belfast, NI
9 0 11769 514792 C101 S D null Bayside, Queens, NY
10 0 PC 17609 495042 Unknown C null 22 Montevideo, Uruguay
11 0 PC 17757 2275250 C62 C64 C null 124 New York, NY
12 0 PC 17757 2275250 C62 C64 C 4 null New York, NY
13 0 PC 17477 693000 B35 C 9 null Paris, France
14 0 19877 788500 Unknown S 6 null n
15 0 27042 300000 A23 S B null Hessle, Yorks
16 0 PC 17318 259250 Unknown S null null New York, NY
17 1 PC 17558 2475208 B58 B60 C null null Montreal, PQ
18 1 PC 17558 2475208 B58 B60 C 6 null Montreal, PQ
19 0 11813 762917 D15 C 8 null n
20 0 13050 752417 C6 C A null Winnipeg, MN
21 1 11751 525542 D35 S 5 null New York, NY
22 1 11751 525542 D35 S 5 null New York, NY
23 0 111369 300000 C148 C 5 null New York, NY
24 0 PC 17757 2275250 Unknown C 4 null n
25
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type
Replaced Value
14 COLUMNS, 999+ ROWS
PREVIEW DOWNLOADED AT 07:42

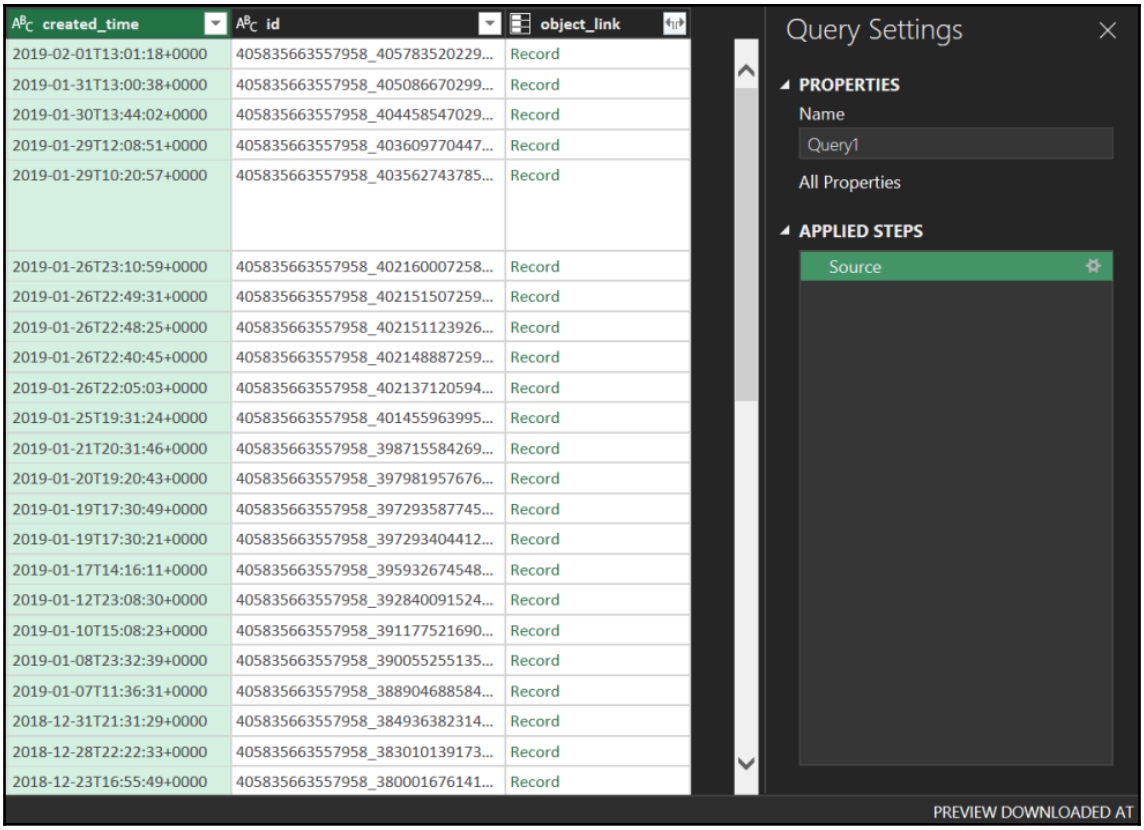

Book1 - Excel
Julio C Rodriguez Martino
File
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-ins
Power Pivot
PDF-XChange
Team
Tell me what you want to do
Share
Get Data
From Text/CSV
From Web
From Table/Range
Recent Sources
Existing Connections
Refresh All
Queries & Connections
Properties
Edit Links
Sort
Filter
Clear
Reapply
Advanced
Text to Columns
What-If Analysis
Forecast Sheet
Group
Ungroup
Subtotal
Data Analysis
Queries & Connections
Sort & Filter
Data Tools
Forecast
Outline
Analysis
From File
From Database
From Azure
From Online Services
From Other Sources
Combine Queries
Launch Query Editor...
Data Catalog Search
My Data Catalog Queries
Data Source Settings...
Query Options
From Table/Range
From Web
From Microsoft Query
From SharePoint List
From OData Feed
From Hadoop File (HDFS)
From Active Directory
From Microsoft Exchange
From ODBC
From OLEDB
Blank Query
Sheet1
Ready
100%

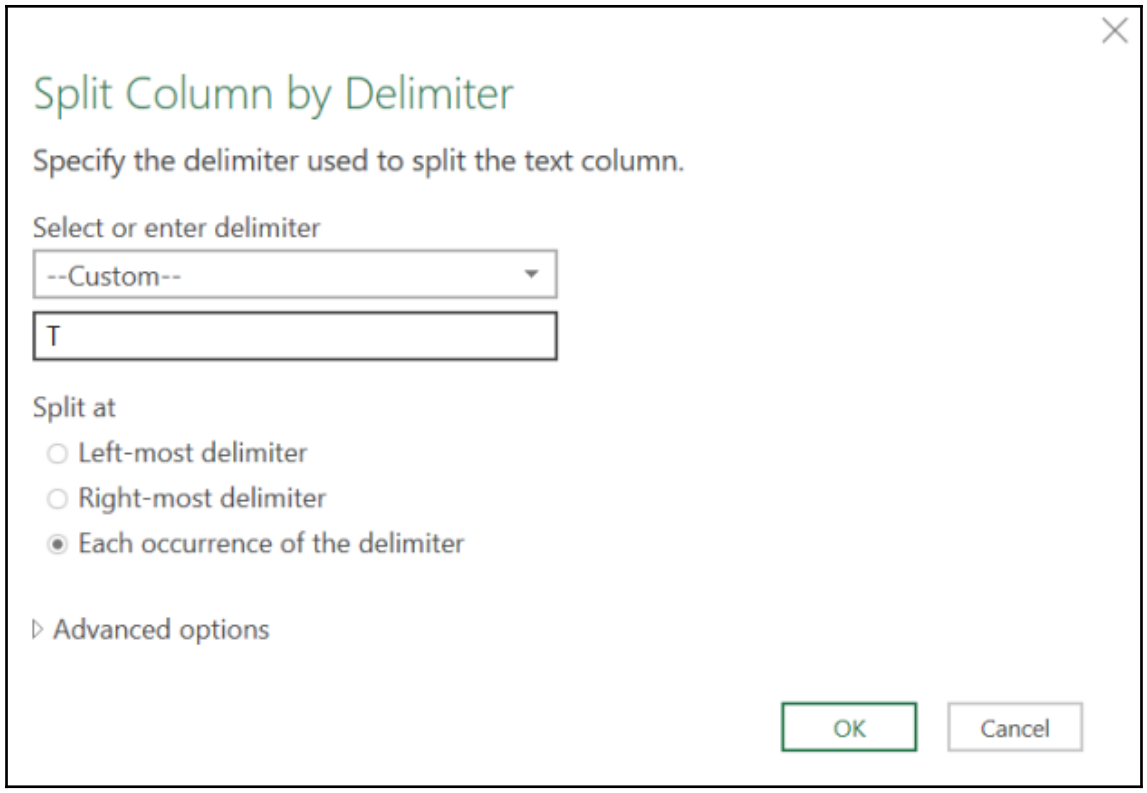

From Web
Basic
Advanced
URL
https://en.wikipedia.org/wiki/Microsoft_Excel
OK
Cancel

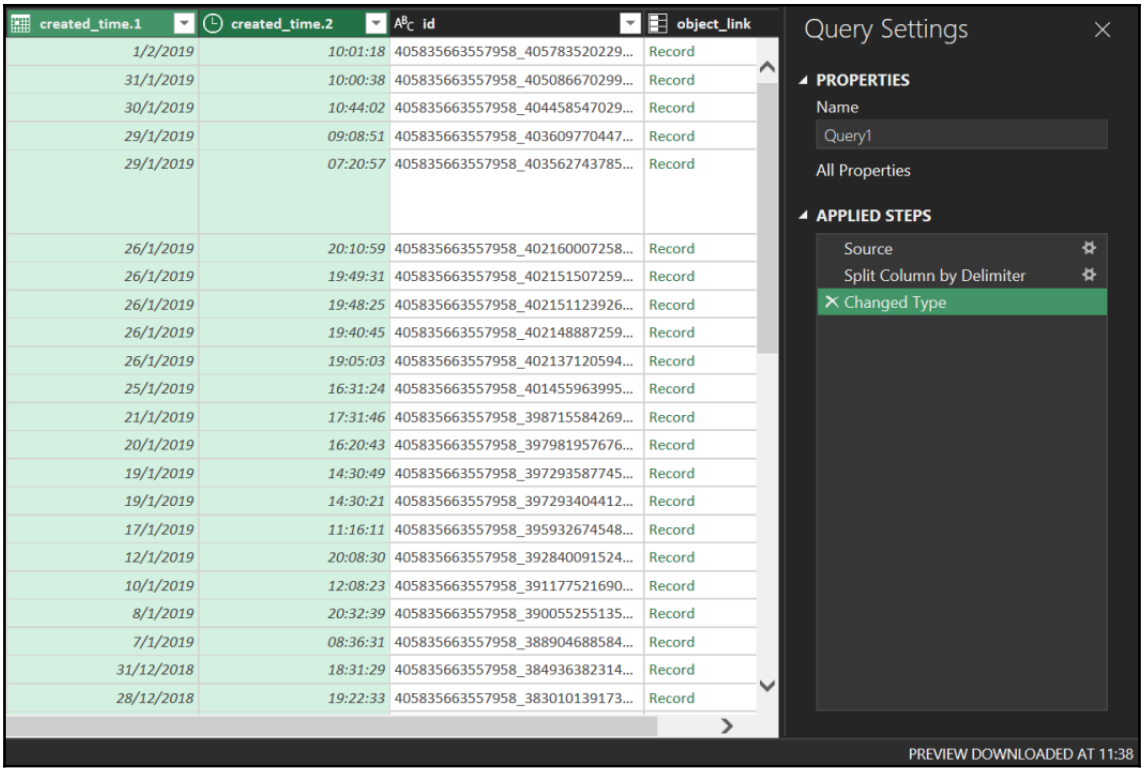

Navigator
Select multiple items
Display Options
https://en.wikipedia.org/wiki/Microsoft_Excel [...
Document
Excel 2007 formats
Excel Spreadsheet
Microsoft Excel
Microsoft Excel for Mac
Microsoft Excel for Macintosh release history
Microsoft Excel for OS/2 release history
Microsoft Excel for Windows release history
Old file extensions
Table 10
Table 11
Table 12
Table 8
Table 9
Table View
Web View
Microsoft Excel for Windows release history
Year Name Version Comments
1987 Excel 2 20 Renumbered to 2 to correspond with contempor
1990 Excel 3 30 Added 3D graphing capabilities
1992 Excel 4 40 Introduced auto-fill feature
1993 Excel 5 50 Included Visual Basic for Applications (VBA) and
1995 Excel 95 70 Renumbered for contemporary Word version. Bc
1997 Excel 97 80
2000 Excel 2000 90 Part of Microsoft Office 2000, which was itself pa
2002 Excel 2002 100
2003 Excel 2003 110 Released only 1 year later to correspond better v
2007 Excel 2007 120
2010 Excel 2010 140 Due to superstitions surrounding the number 13,
2013 Excel 2013 150 Introduced 50 more mathematical functions (ava
2016 Excel 2016 160 Part of Microsoft Office 2016
Load
Edit
Cancel

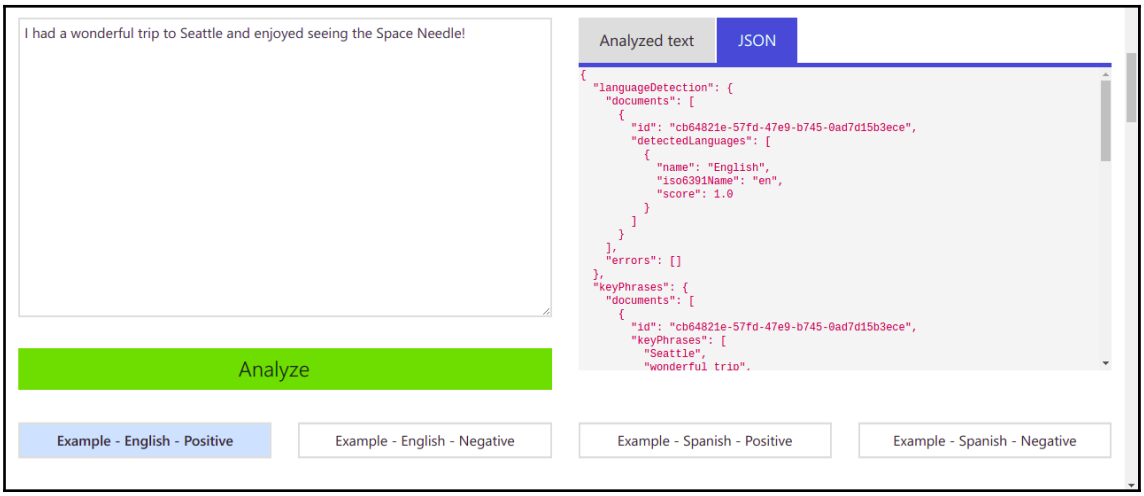

Book1 - Excel
Julio C Rodriguez Martino
File
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-ins
Power Pivot
PDF-XChange
Team
Tell me what you want to do
Share
Get Data
From Text/CSV
From Web
From Table/Range
Recent Sources
Existing Connections
Refresh All
Queries & Connections
Properties
Edit Links
Queries & Connections
Sort
Filter
Clear
Reapply
Advanced
Sort & Filter
Text to Columns
Data Tools
What-If Analysis
Forecast Sheet
Forecast
Group
Ungroup
Subtotal
Outline
Data Analysis
Analysis
From File
From Database
From Azure
From Online Services
From Other Sources
Combine Queries
Launch Query Editor...
Data Catalog Search
My Data Catalog Queries
Data Source Settings...
Query Options
From SharePoint Online List
From Microsoft Exchange Online
From Dynamics 365 (online)
From Facebook
From Salesforce Objects
From Salesforce Reports
Sheet1

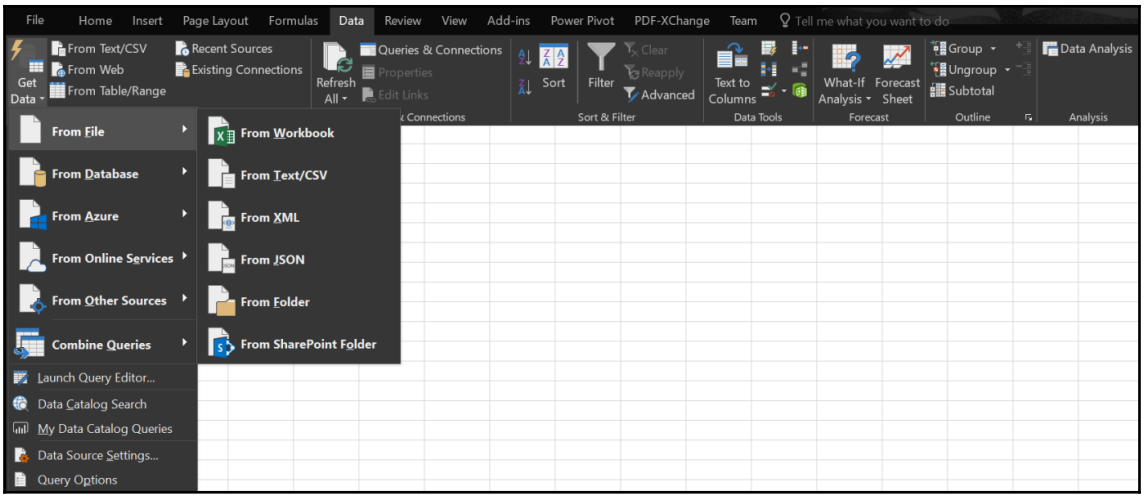

Facebook
"Me", user name or Object ID
me
Connection
Posts
Learn more
OK
Cancel

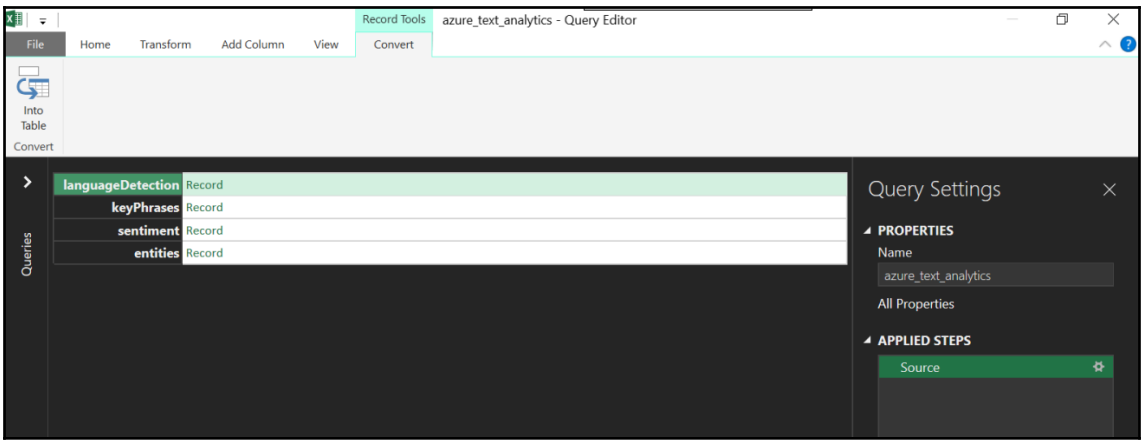

created_time
id
object_link
2019-02-01T13:01:18+0000 405835663557958_405783520229... Record
2019-01-31T13:00:38+0000 405835663557958_405086670299... Record
2019-01-30T13:44:02+0000 405835663557958_404458547029... Record
2019-01-29T12:08:51+0000 405835663557958_403609770447... Record
2019-01-29T10:20:57+0000 405835663557958_403562743785... Record
2019-01-26T23:10:59+0000 405835663557958_402160007258... Record
2019-01-26T22:49:31+0000 405835663557958_402151507259... Record
2019-01-26T22:48:25+0000 405835663557958_402151123926... Record
2019-01-26T22:40:45+0000 405835663557958_402148887259... Record
2019-01-26T22:05:03+0000 405835663557958_402137120594... Record
2019-01-25T19:31:24+0000 405835663557958_401455963995... Record
2019-01-21T20:31:46+0000 405835663557958_398715584269... Record
2019-01-20T19:20:43+0000 405835663557958_397981957676... Record
2019-01-19T17:30:49+0000 405835663557958_397293587745... Record
2019-01-19T17:30:21+0000 405835663557958_397293404412... Record
2019-01-17T14:16:11+0000 405835663557958_395932674548... Record
2019-01-12T23:08:30+0000 405835663557958_392840091524... Record
2019-01-10T15:08:23+0000 405835663557958_391177521690... Record
2019-01-08T23:32:39+0000 405835663557958_390055255135... Record
2019-01-07T11:36:31+0000 405835663557958_388904688584... Record
2018-12-31T21:31:29+0000 405835663557958_384936382314... Record
2018-12-28T22:22:33+0000 405835663557958_383010139173... Record
2018-12-23T16:55:49+0000 405835663557958_380001676141... Record
Query Settings
PROPERTIES
Name
Query1
All Properties
APPLIED STEPS
Source
PREVIEW DOWNLOADED AT


Split Column by Delimiter
Specify the delimiter used to split the text column.
Select or enter delimiter
--Custom--
T
Split at
Left-most delimiter
Right-most delimiter
Each occurrence of the delimiter
Advanced options
OK
Cancel

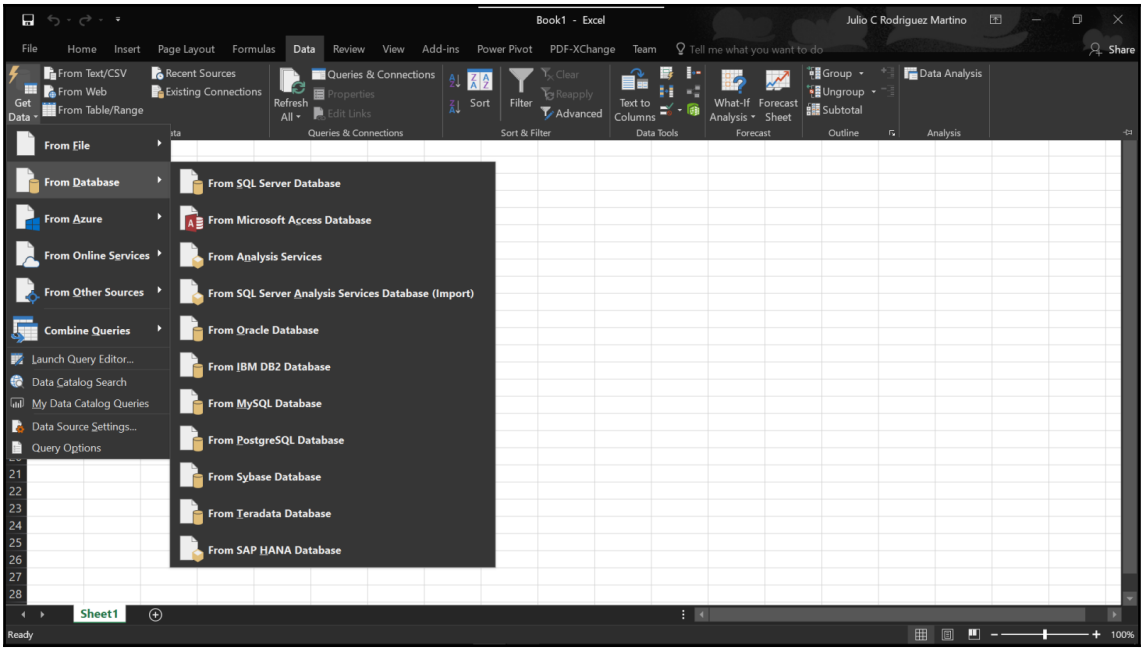

created_time.1
created_time.2
id
object_link
1/2/2019 10:01:18 405835663557958_405783520229... Record
31/1/2019 10:00:38 405835663557958_405086670299... Record
30/1/2019 10:44:02 405835663557958_404458547029... Record
29/1/2019 09:08:51 405835663557958_403609770447... Record
29/1/2019 07:20:57 405835663557958_403562743785... Record
26/1/2019 20:10:59 405835663557958_402160007258... Record
26/1/2019 19:49:31 405835663557958_402151507259... Record
26/1/2019 19:48:25 405835663557958_402151123926... Record
26/1/2019 19:40:45 405835663557958_402148887259... Record
26/1/2019 19:05:03 405835663557958_402137120594... Record
25/1/2019 16:31:24 405835663557958_401455963995... Record
21/1/2019 17:31:46 405835663557958_398715584269... Record
20/1/2019 16:20:43 405835663557958_397981957676... Record
19/1/2019 14:30:49 405835663557958_397293587745... Record
19/1/2019 14:30:21 405835663557958_397293404412... Record
17/1/2019 11:16:11 405835663557958_395932674548... Record
12/1/2019 20:08:30 405835663557958_392840091524... Record
10/1/2019 12:08:23 405835663557958_391177521690... Record
8/1/2019 20:32:39 405835663557958_390055255135... Record
7/1/2019 08:36:31 405835663557958_388904688584... Record
31/12/2018 18:31:29 405835663557958_384936382314... Record
28/12/2018 19:22:33 405835663557958_383010139173... Record
Query Settings
PROPERTIES
Name
Query1
All Properties
APPLIED STEPS
Source
Split Column by Delimiter
Changed Type
PREVIEW DOWNLOADED AT 11:38

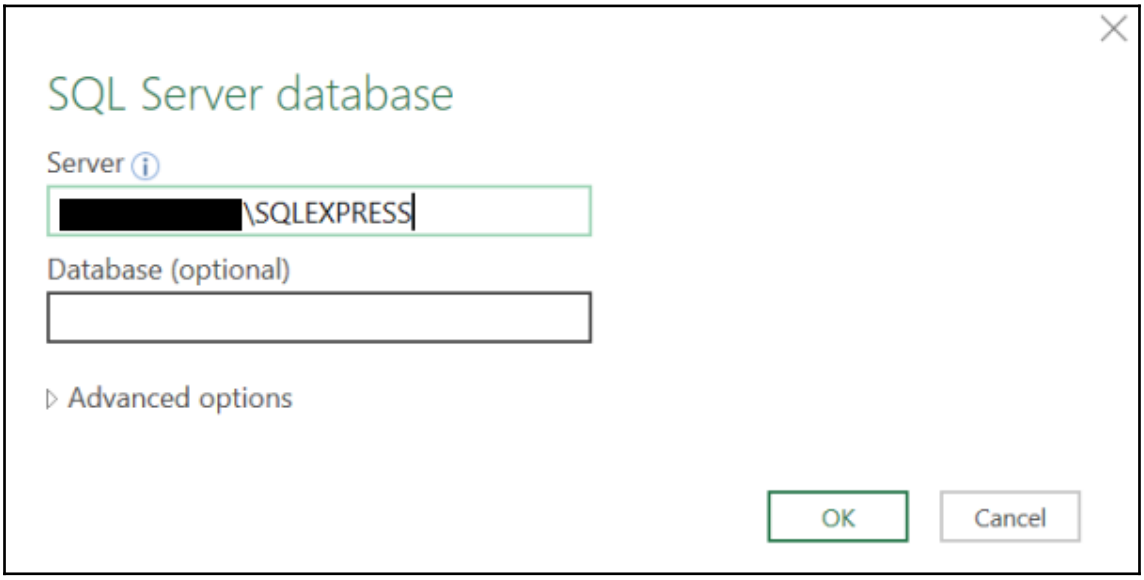

I had a wonderful trip to Seattle and enjoyed seeing the Space Needle!
Analyze
Example - English - Positive
Example - English - Negative
Example - Spanish - Positive
Example - Spanish - Negative
Analyzed text
JSON
{
"languageDetection": {
"documents": [
{
"id": "cb64821e-57fd-47e9-b745-0ad7d15b3ece",
"detectedLanguages": [
{
"name": "English",
"iso6391Name": "en",
"score": 1.0
}
]
}
],
"errors": []
},
"keyPhrases": {
"documents": [
{
"id": "cb64821e-57fd-47e9-b745-0ad7d15b3ece",
"keyPhrases": [
"Seattle",
"wonderful trip".

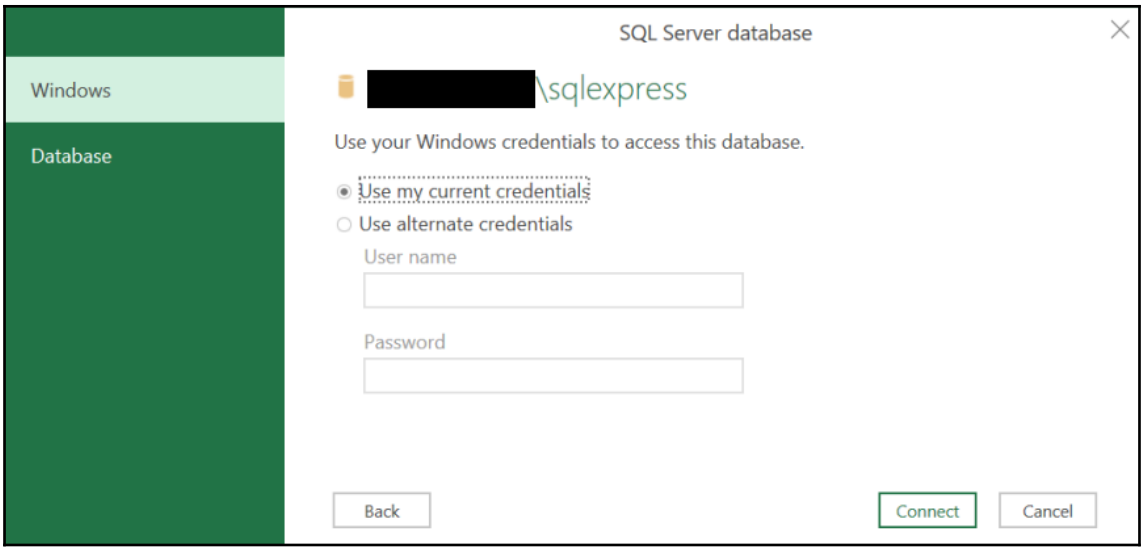


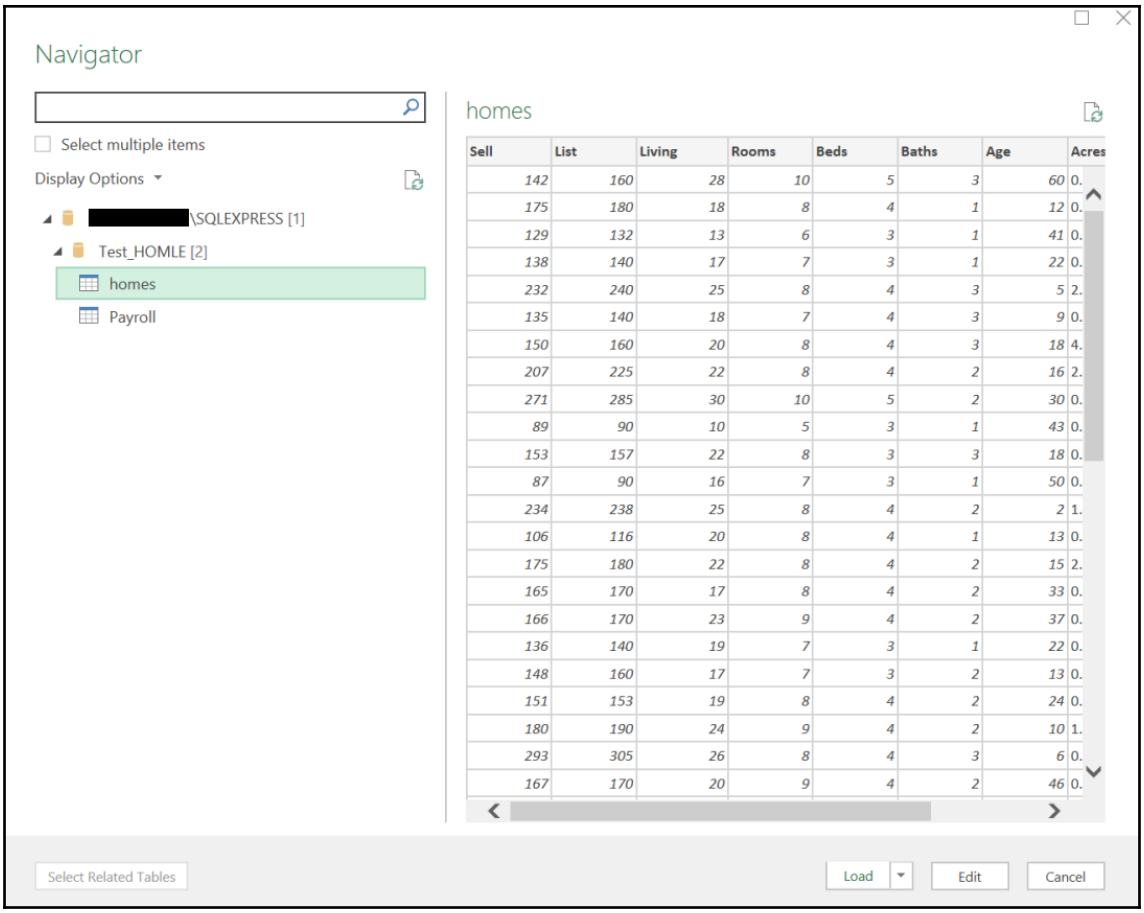


	Name	Value
1	languageDetection	Record
2	keyPhrases	Record
3	sentiment	Record
4	entities	Record

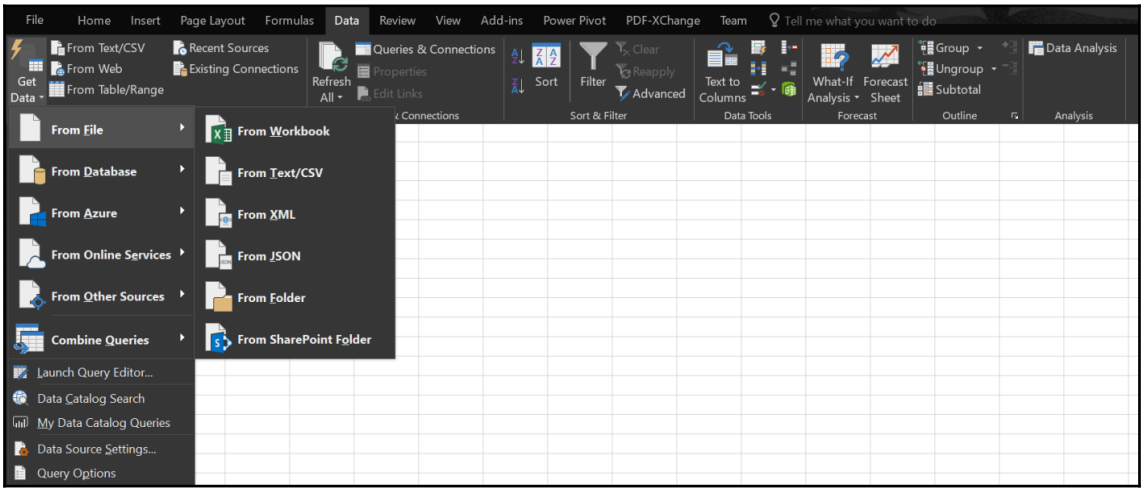

Book1 - Excel
Julio C Rodriguez Martino
File
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-ins
Power Pivot
PDF-XChange
Team
Tell me what you want to do
Share
Get Data
From Text/CSV
From Web
From Table/Range
Recent Sources
Existing Connections
Refresh All
Queries & Connections
Properties
Edit Links
Sort
Filter
Clear
Reapply
Advanced
Text to Columns
What-If Analysis
Forecast Sheet
Group
Ungroup
Subtotal
Data Analysis
Queries & Connections
Sort & Filter
Data Tools
Forecast
Outline
Analysis
From File
From Database
From Azure
From Online Services
From Other Sources
Combine Queries
Launch Query Editor...
Data Catalog Search
My Data Catalog Queries
Data Source Settings...
Query Options
From SQL Server Database
From Microsoft Access Database
From Analysis Services
From SQL Server Analysis Services Database (Import)
From Oracle Database
From IBM DB2 Database
From MySQL Database
From PostgreSQL Database
From Sybase Database
From Teradata Database
From SAP HANA Database
Sheet1
Ready
100%

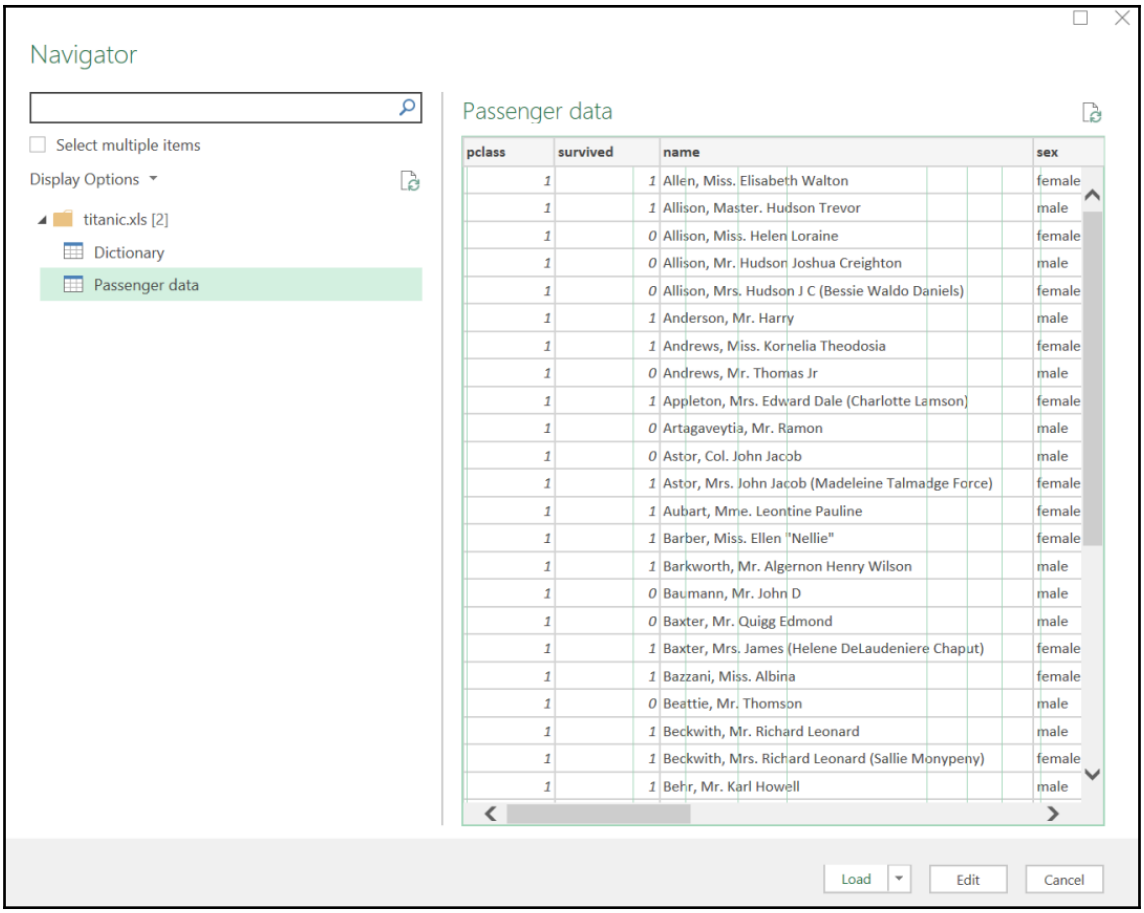

SQL Server database
Server
[redacted]\SQLEXPRESS
Database (optional)
Advanced options
OK
Cancel

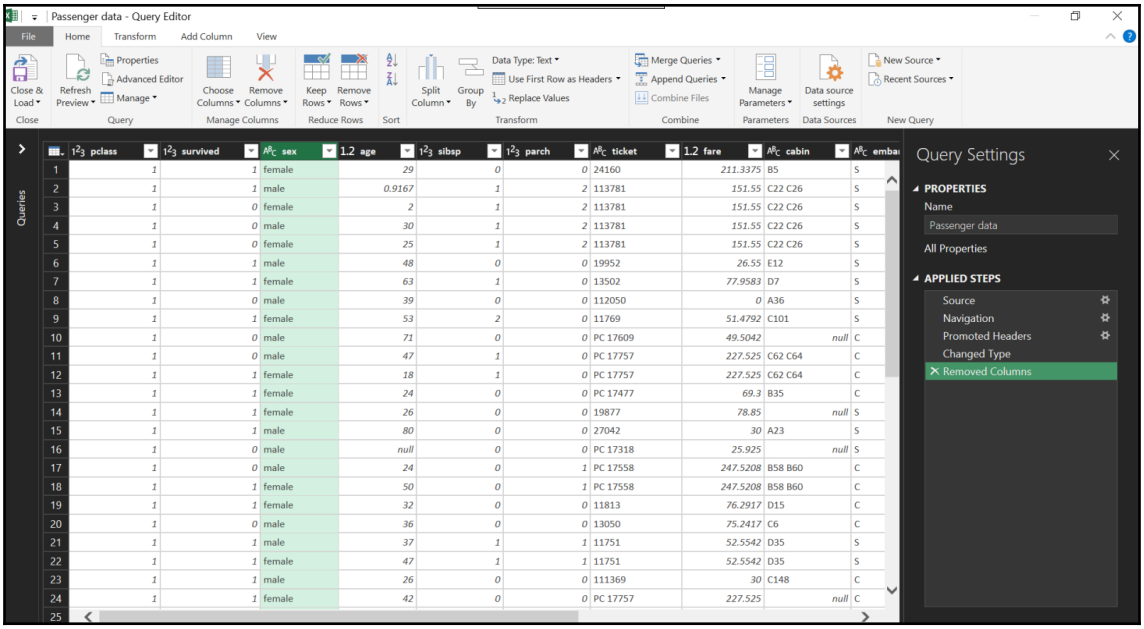

SQL Server database
Windows
Database
[redacted]\sqlexpress
Use your Windows credentials to access this database.
Use my current credentials
Use alternate credentials
User name
Password
Back
Connect
Cancel

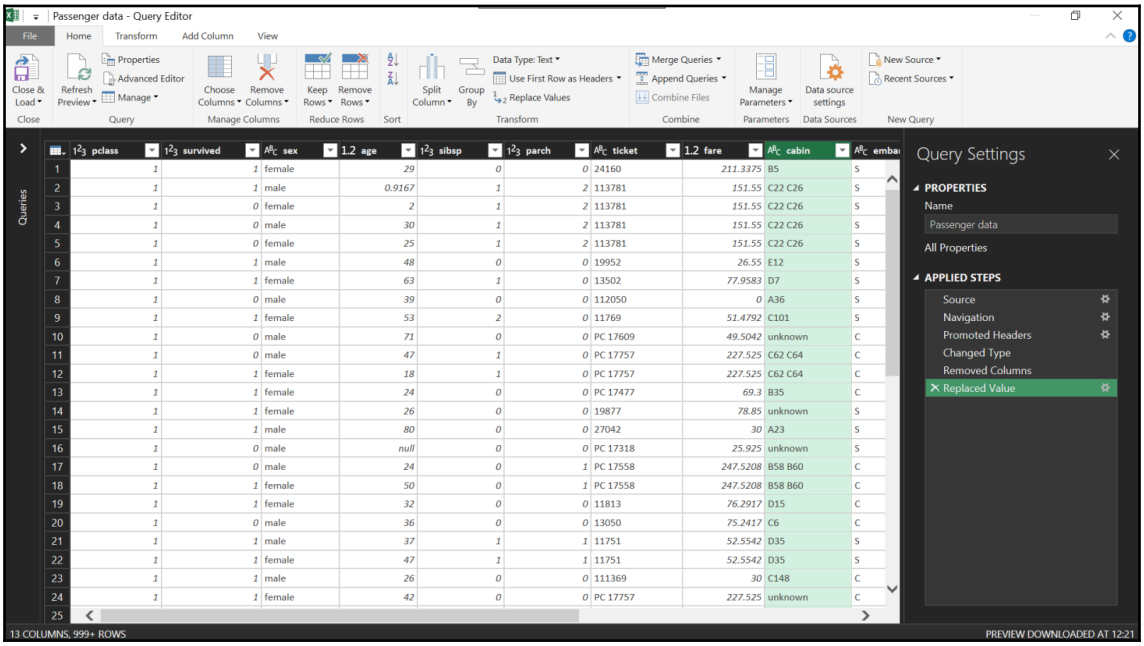

Navigator
Select multiple items
Display Options
[redacted]\SQLEXPRESS [1]
Test_HOMLE [2]
homes
Payroll
homes
Sell List Living Rooms Beds Baths Age Acres
142 160 28 10 5 3 60 0.
175 180 18 8 4 1 12 0.
129 132 13 6 3 1 41 0.
138 140 17 7 3 1 22 0.
232 240 25 8 4 3 5 2.
135 140 18 7 4 3 9 0.
150 160 20 8 4 3 18 4.
207 225 22 8 4 2 16 2.
271 285 30 10 5 2 30 0.
89 90 10 5 3 1 43 0.
153 157 22 8 3 3 18 0.
87 90 16 7 3 1 50 0.
234 238 25 8 4 2 2 1.
106 116 20 8 4 1 13 0.
175 180 22 8 4 2 15 2.
165 170 17 8 4 2 33 0.
166 170 23 9 4 2 37 0.
136 140 19 7 3 1 22 0.
148 160 17 7 3 2 13 0.
151 153 19 8 4 2 24 0.
180 190 24 9 4 2 10 1.
293 305 26 8 4 3 6 0.
167 170 20 9 4 2 46 0.
Select Related Tables
Load
Edit
Cancel

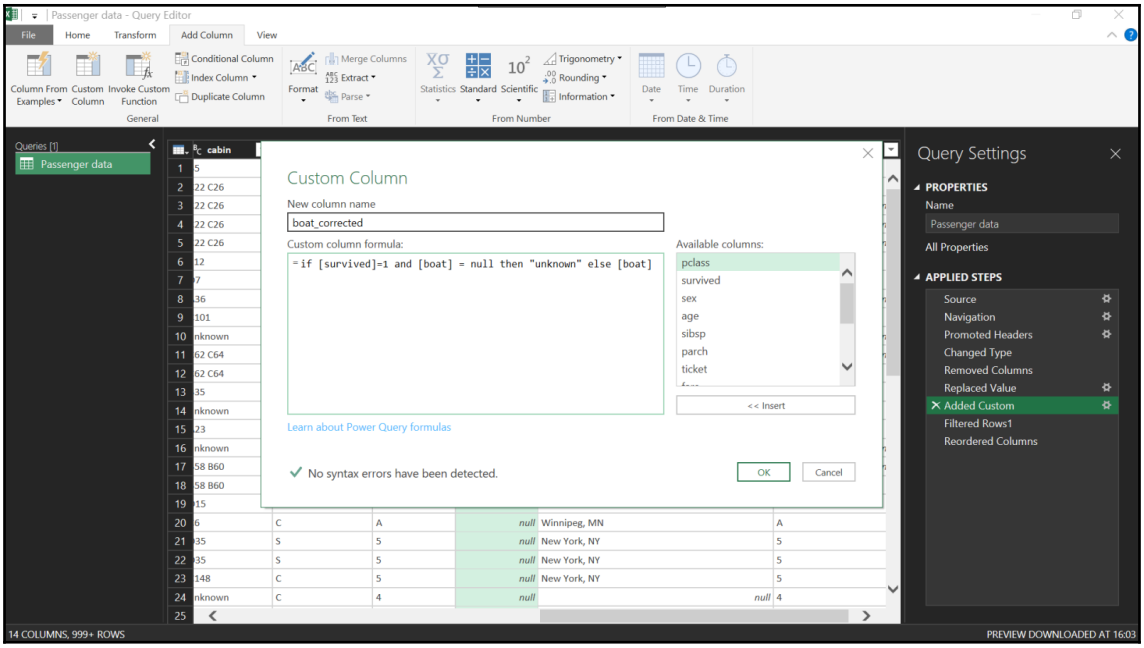
Chapter 4: Data Cleansing and Preliminary Data Analysis

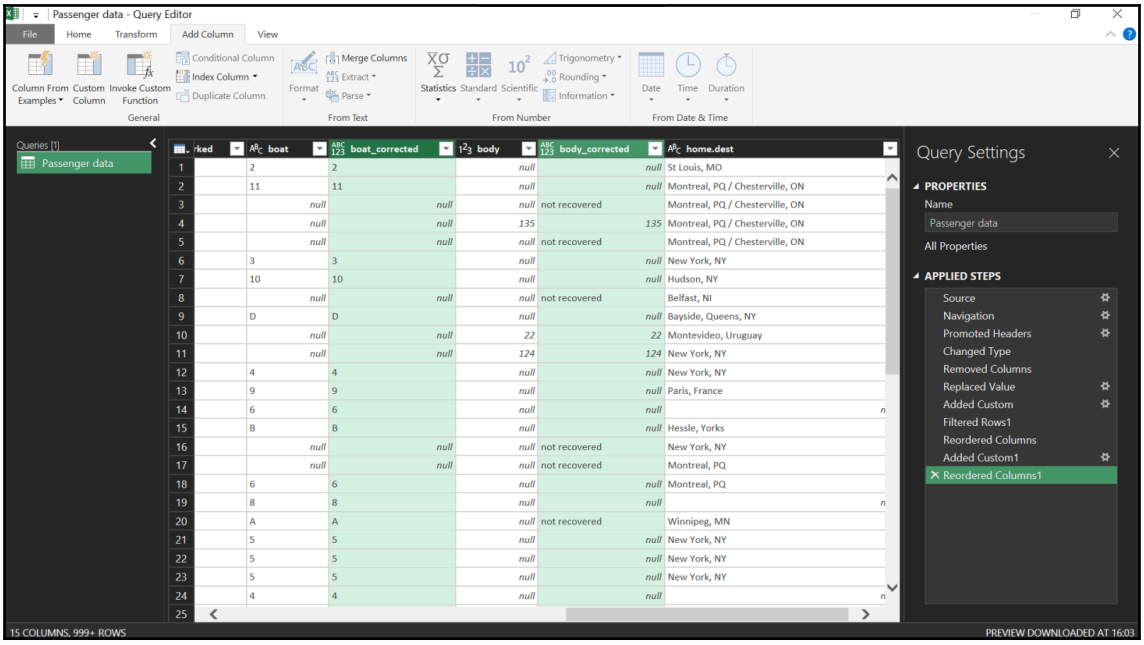


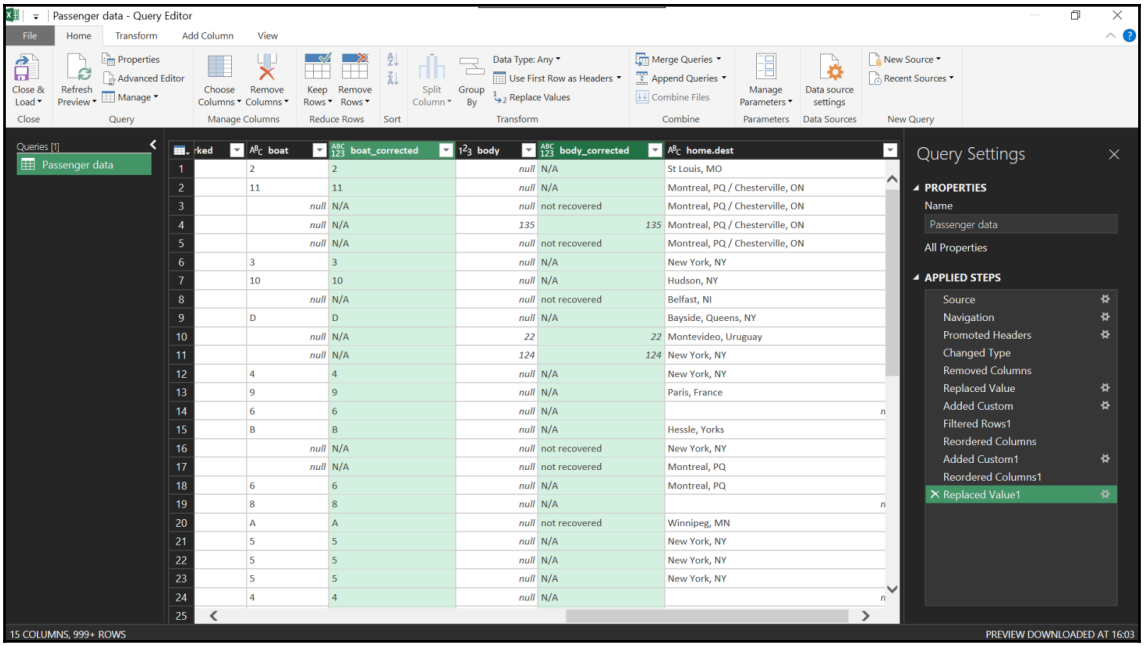

Navigator
Select multiple items
Display Options
titanic.xls [2]
Dictionary
Passenger data
Passenger data
pclass survived name sex
1 1 Allen, Miss. Elisabeth Walton female
1 1 Allison, Master. Hudson Trevor male
1 0 Allison, Miss. Helen Loraine female
1 0 Allison, Mr. Hudson Joshua Creighton male
1 0 Allison, Mrs. Hudson J C (Bessie Waldo Daniels) female
1 1 Anderson, Mr. Harry male
1 1 Andrews, Miss. Kornelia Theodosia female
1 0 Andrews, Mr. Thomas Jr male
1 1 Appleton, Mrs. Edward Dale (Charlotte Lamson) female
1 0 Artagaveytia, Mr. Ramon male
1 0 Astor, Col. John Jacob male
1 1 Astor, Mrs. John Jacob (Madeleine Talmadge Force) female
1 1 Aubart, Mme. Leontine Pauline female
1 1 Barber, Miss. Ellen "Nellie" female
1 1 Barkworth, Mr. Algernon Henry Wilson male
1 0 Baumann, Mr. John D male
1 0 Baxter, Mr. Quigg Edmond male
1 1 Baxter, Mrs. James (Helene DeLaudeniere Chaput) female
1 1 Bazzani, Miss. Albina female
1 0 Beattie, Mr. Thomson male
1 1 Beckwith, Mr. Richard Leonard male
1 1 Beckwith, Mrs. Richard Leonard (Sallie Monypeny) female
1 1 Behr, Mr. Karl Howell male
Load
Edit
Cancel

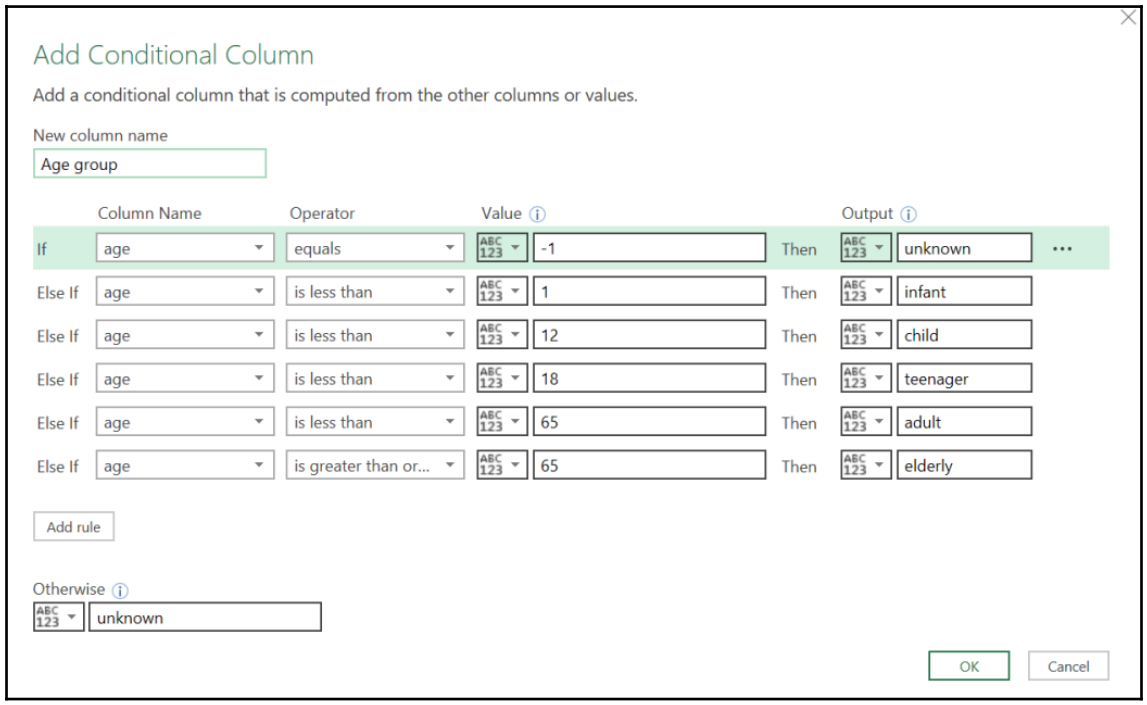

Passenger data - Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Close
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Parameters
Data Sources
New Query
Queries
pclass
survived
sex
age
sibsp
parch
ticket
fare
cabin
embar
1 1 female 29 0 0 24160 211.3375 B5 S
2 1 male 0.9167 1 2 113781 151.55 C22 C26 S
3 1 0 female 2 1 2 113781 151.55 C22 C26 S
4 1 0 male 30 1 2 113781 151.55 C22 C26 S
5 1 0 female 25 1 2 113781 151.55 C22 C26 S
6 1 1 male 48 0 0 19952 26.55 E12 S
7 1 1 female 63 1 0 13502 77.9583 D7 S
8 1 0 male 39 0 0 112050 0 A36 S
9 1 1 female 53 2 0 11769 51.4792 C101 S
10 1 0 male 71 0 0 PC 17609 49.5042 null C
11 1 0 male 47 1 0 PC 17757 227.525 C62 C64 C
12 1 1 female 18 1 0 PC 17757 227.525 C62 C64 C
13 1 1 female 24 0 0 PC 17477 69.3 B35 C
14 1 1 female 26 0 0 19877 78.85 null S
15 1 1 male 80 0 0 27042 30 A23 S
16 1 0 male null 0 0 PC 17318 25.925 null S
17 1 0 male 24 0 1 PC 17558 247.5208 B58 B60 C
18 1 1 female 50 0 1 PC 17558 247.5208 B58 B60 C
19 1 1 female 32 0 0 11813 76.2917 D15 C
20 1 0 male 36 0 0 13050 75.2417 C6 C
21 1 1 male 37 1 1 11751 52.5542 D35 S
22 1 1 female 47 1 1 11751 52.5542 D35 S
23 1 1 male 26 0 0 111369 30 C148 C
24 1 1 female 42 0 0 PC 17757 227.525 null C
25
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type
Removed Columns

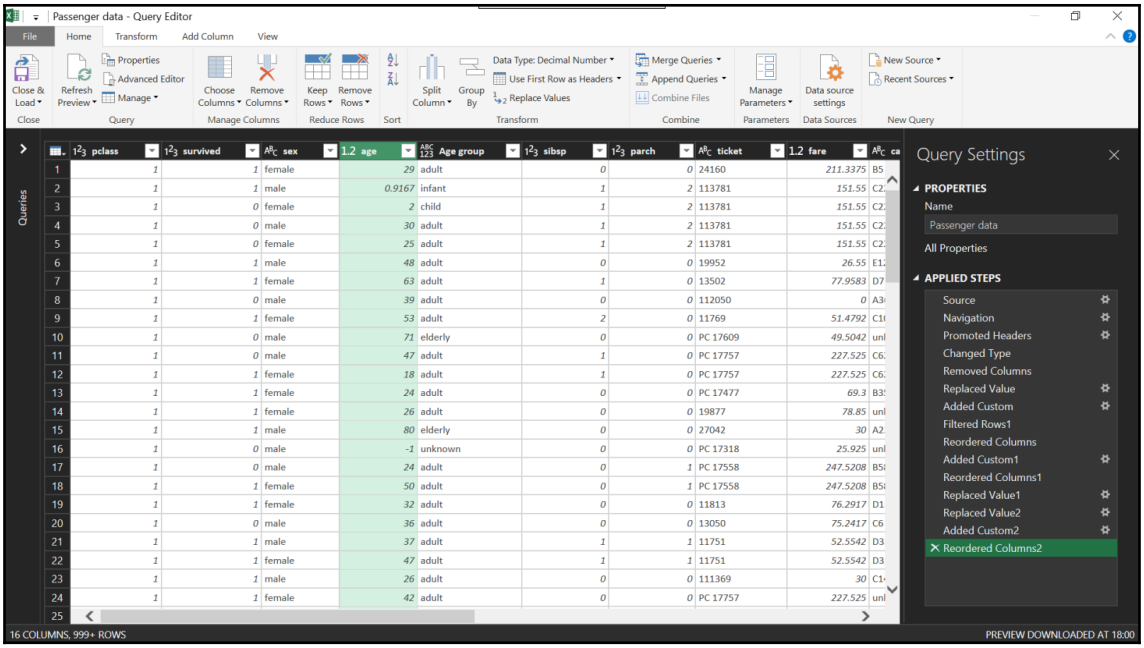

Passenger data - Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Close
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Parameters
Data Sources
New Query
Queries
pclass
survived
sex
age
sibsp
parch
ticket
fare
cabin
embar
1 1 female 29 0 0 24160 211.3375 B5 S
2 1 male 0.9167 1 2 113781 151.55 C22 C26 S
3 1 0 female 2 1 2 113781 151.55 C22 C26 S
4 1 0 male 30 1 2 113781 151.55 C22 C26 S
5 1 0 female 25 1 2 113781 151.55 C22 C26 S
6 1 1 male 48 0 0 19952 26.55 E12 S
7 1 1 female 63 1 0 13502 77.9583 D7 S
8 1 0 male 39 0 0 112050 0 A36 S
9 1 1 female 53 2 0 11769 51.4792 C101 S
10 1 0 male 71 0 0 PC 17609 49.5042 unknown C
11 1 0 male 47 1 0 PC 17757 227.525 C62 C64 C
12 1 1 female 18 1 0 PC 17757 227.525 C62 C64 C
13 1 1 female 24 0 0 PC 17477 69.3 B35 C
14 1 1 female 26 0 0 19877 78.85 unknown S
15 1 1 male 80 0 0 27042 30 A23 S
16 1 0 male null 0 0 PC 17318 25.925 unknown S
17 1 0 male 24 0 1 PC 17558 247.5208 B58 B60 C
18 1 1 female 50 0 1 PC 17558 247.5208 B58 B60 C
19 1 1 female 32 0 0 11813 76.2917 D15 C
20 1 0 male 36 0 0 13050 75.2417 C6 C
21 1 1 male 37 1 1 11751 52.5542 D35 S
22 1 1 female 47 1 1 11751 52.5542 D35 S
23 1 1 male 26 0 0 111369 30 C148 C
24 1 1 female 42 0 0 PC 17757 227.525 unknown C
25
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type
Removed Columns
Replaced Value
13 COLUMNS, 999+ ROWS
PREVIEW DOWNLOADED AT 12:21

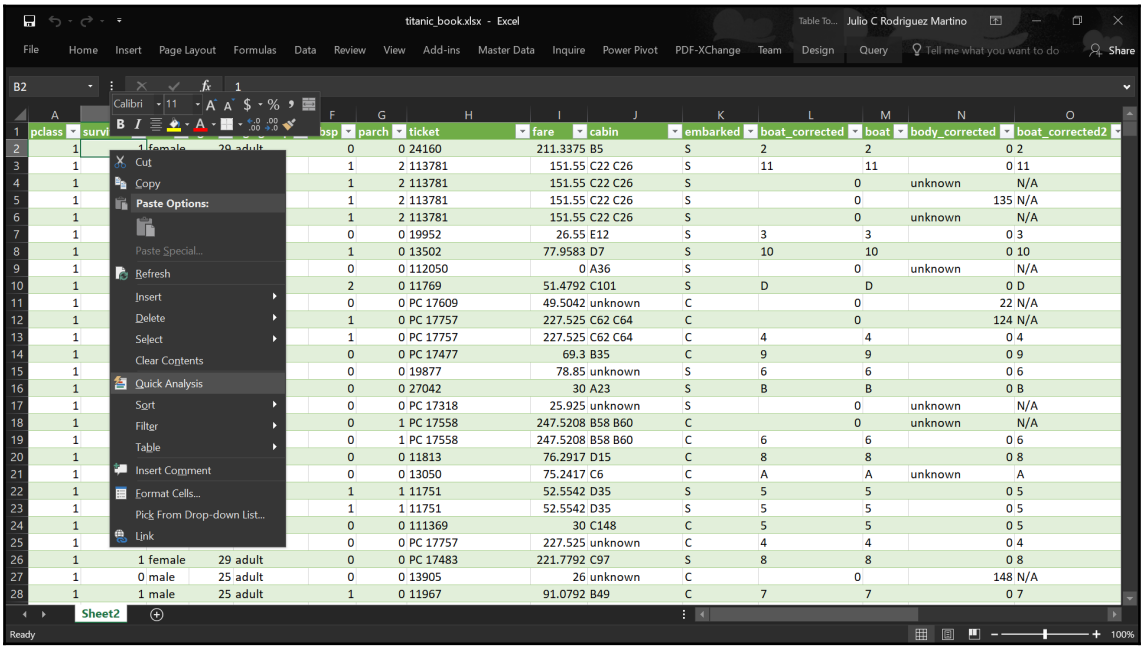

Passenger data - Query Editor
File
Home
Transform
Add Column
View
Column From Examples
Custom Column
Invoke Custom Function
Conditional Column
Index Column
Duplicate Column
General
Format
Merge Columns
Extract
Parse
From Text
Statistics
Standard
Scientific
Trigonometry
Rounding
Information
From Number
Date
Time
Duration
From Date & Time
Queries [1]
Passenger data
Custom Column
New column name
boat_corrected
Custom column formula:
=if [survived]=1 and [boat] = null then "unknown" else [boat]
Available columns:
pclass
survived
sex
age
sibsp
parch
ticket
<< Insert
Learn about Power Query formulas
No syntax errors have been detected.
OK
Cancel
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type
Removed Columns
Replaced Value
Added Custom
Filtered Rows1
Reordered Columns
14 COLUMNS, 999+ ROWS
PREVIEW DOWNLOADED AT 16:03

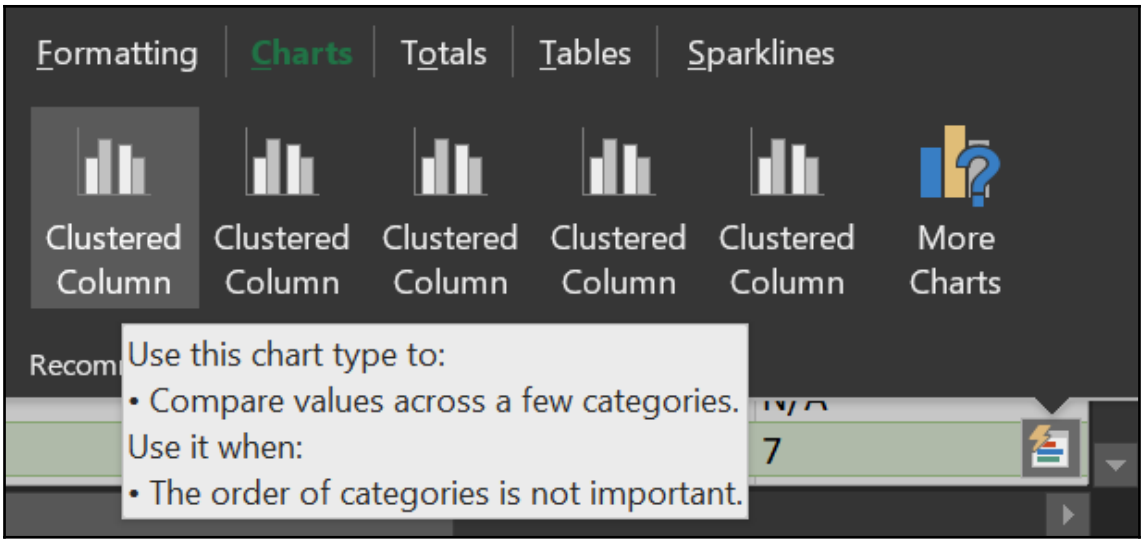

Passenger data - Query Editor
File
Home
Transform
Add Column
View
Column From Examples
Custom Column
Invoke Custom Function
Conditional Column
Index Column
Duplicate Column
General
Format
Merge Columns
Extract
Parse
From Text
Statistics
Standard
Scientific
Trigonometry
Rounding
Information
From Number
Date
Time
Duration
From Date & Time
Queries [1]
Passenger data
boat
boat_corrected
body
body_corrected
home.dest
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type
Removed Columns
Replaced Value
Added Custom
Filtered Rows1
Reordered Columns
Added Custom1
Reordered Columns1
15 COLUMNS, 999+ ROWS
PREVIEW DOWNLOADED AT 16:03

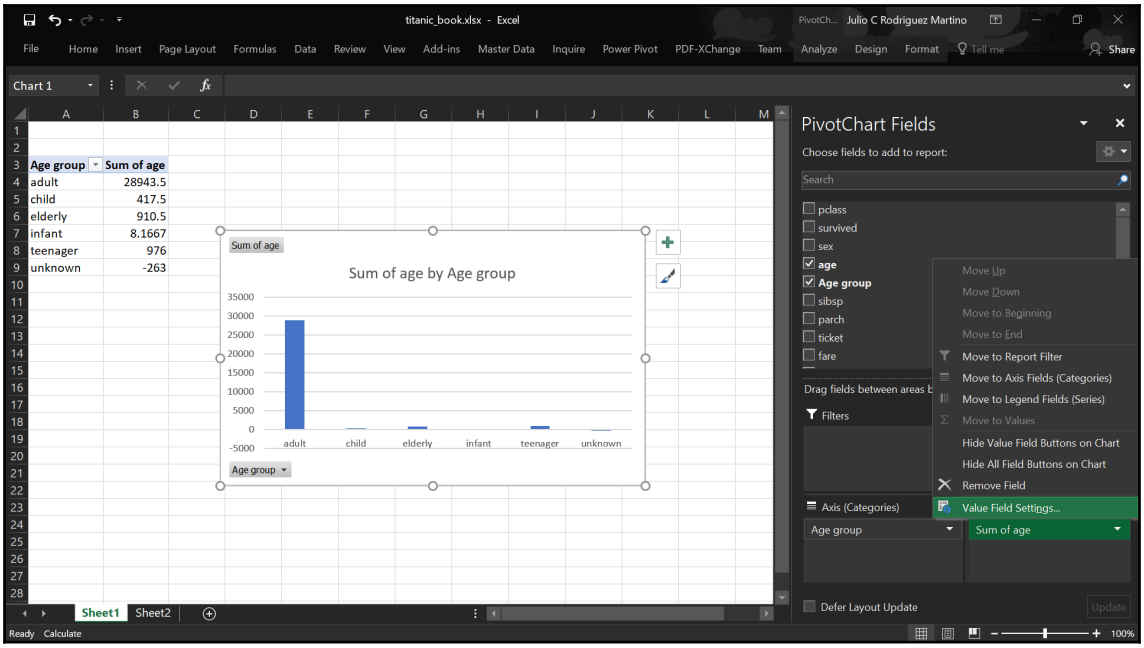

Passenger data - Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Any
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Close
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Parameters
Data Sources
New Query
Queries [1]
Passenger data
rked
boat
boat_corrected
body
body_corrected
home.dest
1 2 2 null N/A St Louis, MO
2 11 11 null N/A Montreal, PQ / Chesterville, ON
3 null N/A null not recovered Montreal, PQ / Chesterville, ON
4 null N/A 135 135 Montreal, PQ / Chesterville, ON
5 null N/A null not recovered Montreal, PQ / Chesterville, ON
6 3 3 null N/A New York, NY
7 10 10 null N/A Hudson, NY
8 null N/A null not recovered Belfast, NI
9 D D null N/A Bayside, Queens, NY
10 null N/A 22 22 Montevideo, Uruguay
11 null N/A 124 124 New York, NY
12 4 4 null N/A New York, NY
13 9 9 null N/A Paris, France
14 6 6 null N/A
15 B B null N/A Hessle, Yorks
16 null N/A null not recovered New York, NY
17 null N/A null not recovered Montreal, PQ
18 6 6 null N/A Montreal, PQ
19 8 8 null N/A
20 A A null not recovered Winnipeg, MN
21 5 5 null N/A New York, NY
22 5 5 null N/A New York, NY
23 5 5 null N/A New York, NY
24 4 4 null N/A
25
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type
Removed Columns
Replaced Value
Added Custom
Filtered Rows1
Reordered Columns
Added Custom1
Reordered Columns1
Replaced Value1
15 COLUMNS, 999+ ROWS
PREVIEW DOWNLOADED AT 16:03

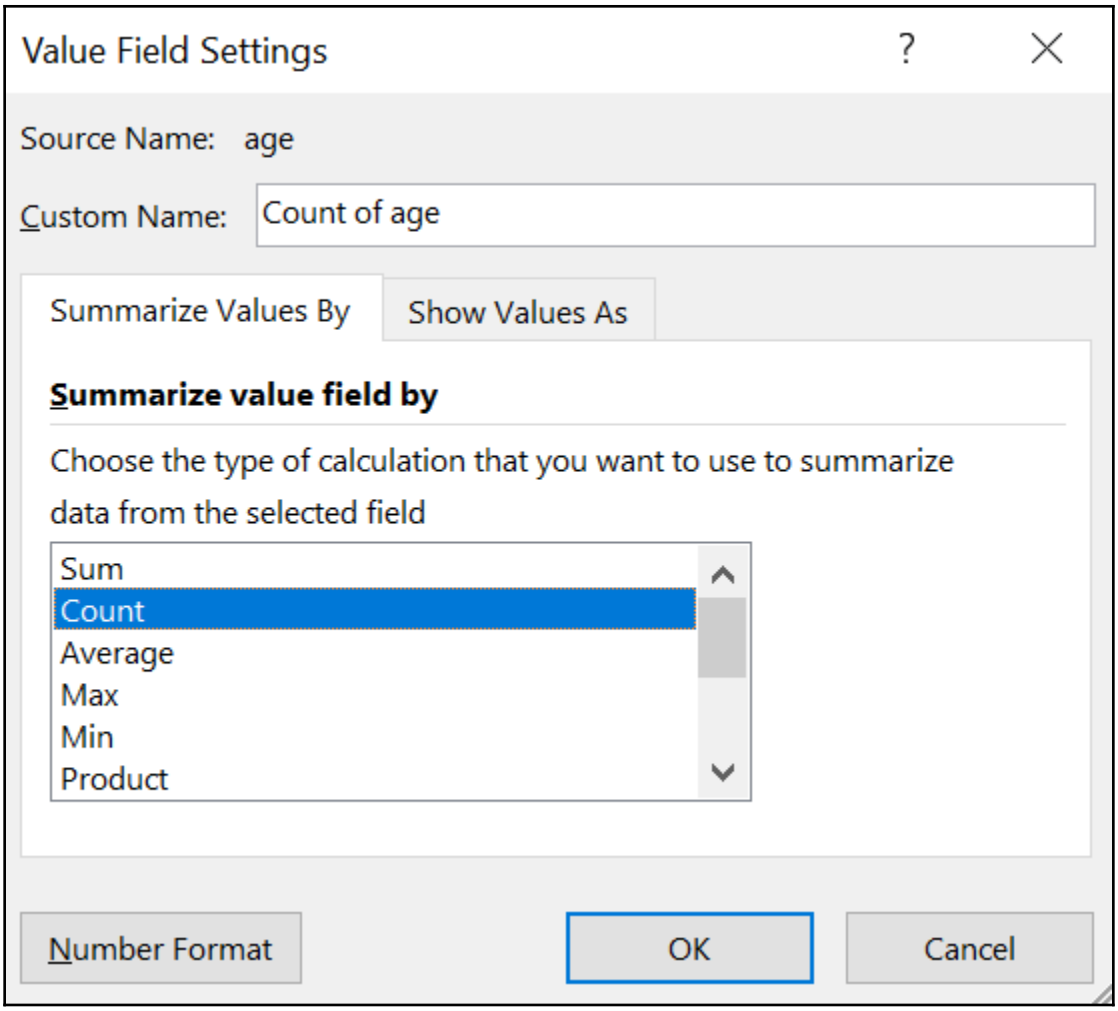

Add Conditional Column
Add a conditional column that is computed from the other columns or values.
New column name
Age group
Column Name
Operator
Value
Output
If age equals -1 Then unknown
Else If age is less than 1 Then infant
Else If age is less than 12 Then child
Else If age is less than 18 Then teenager
Else If age is less than 65 Then adult
Else If age is greater than or... 65 Then elderly
Add rule
Otherwise
unknown
OK
Cancel

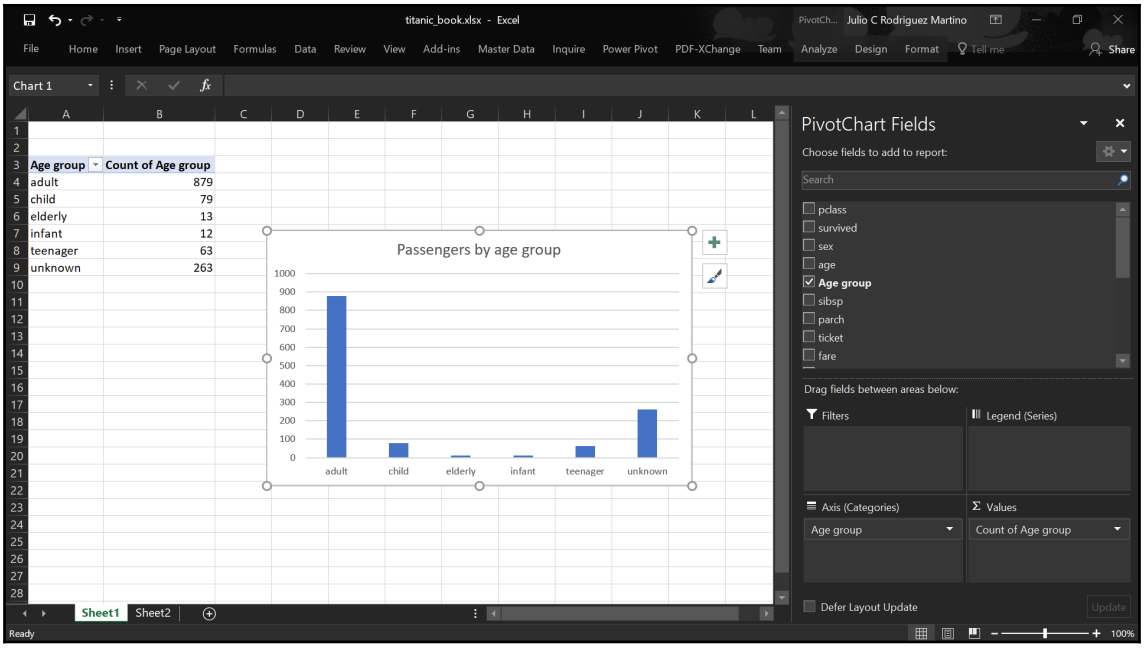

Passenger data - Query Editor
Query Settings
PROPERTIES
Name
Passenger data
All Properties
APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type
Removed Columns
Replaced Value
Added Custom
Filtered Rows1
Reordered Columns
Added Custom1
Reordered Columns1
Replaced Value1
Replaced Value2
Added Custom2
Reordered Columns2
16 COLUMNS, 999+ ROWS
PREVIEW DOWNLOADED AT 18:00

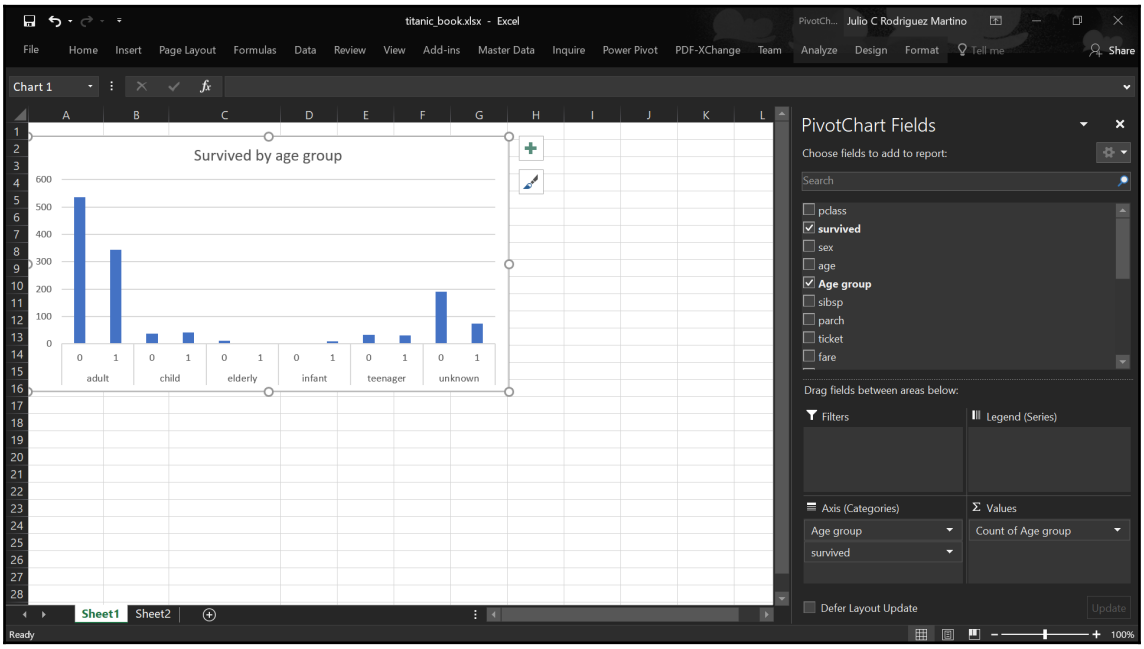

titanic_book.xlsx - Excel
Sheet2
Ready

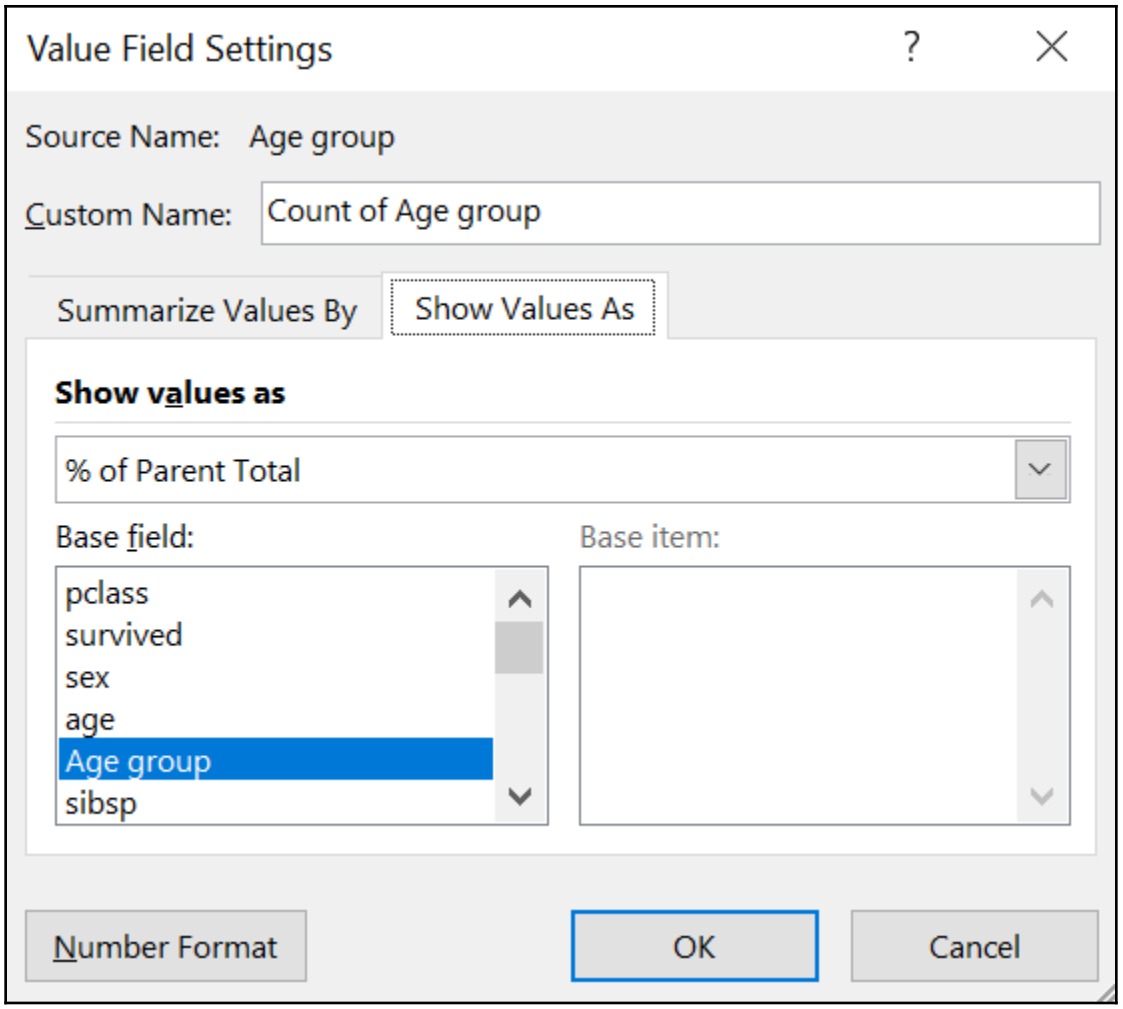

Formatting
Charts
Totals
Tables
Sparklines
Clustered Column
Clustered Column
Clustered Column
Clustered Column
Clustered Column
More Charts
Recom
Use this chart type to:
• Compare values across a few categories.
Use it when:
• The order of categories is not important.

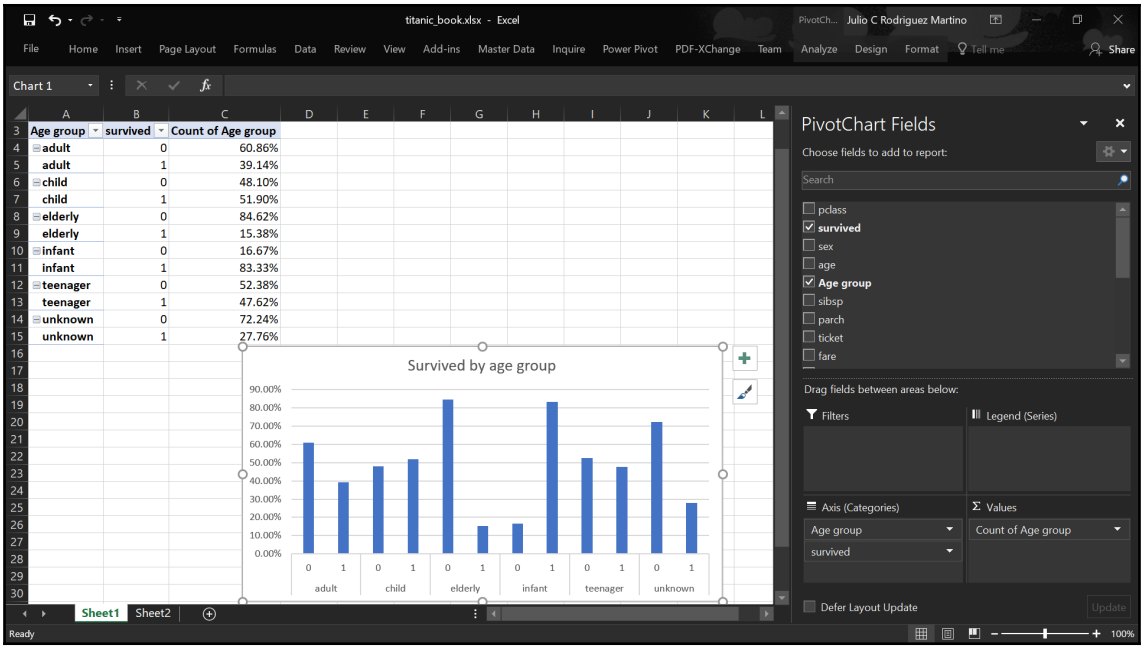

titanic_book.xlsx - Excel
Chart 1
Age group
Sum of age
adult 28943.5
child 417.5
elderly 910.5
infant 8.1667
teenager 976
unknown -263
Sum of age by Age group
PivotChart Fields
Choose fields to add to report:
Sheet1
Sheet2
Ready
Calculate


Value Field Settings
?
Source Name: age
Custom Name: Count of age
Summarize Values By
Show Values As
Summarize value field by
Choose the type of calculation that you want to use to summarize data from the selected field
Sum
Count
Average
Max
Min
Product
Number Format
OK
Cancel

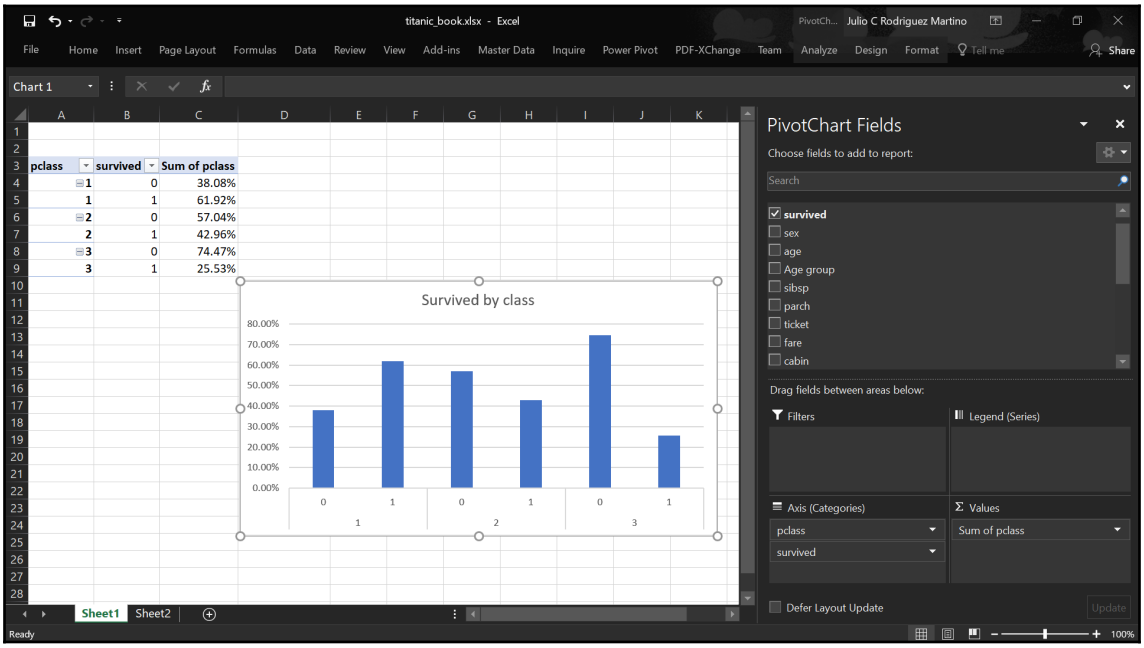

titanic_book.xlsx - Excel
Chart 1
Age group
Count of Age group
adult 879
child 79
elderly 13
infant 12
teenager 63
unknown 263
Passengers by age group
1000
900
800
700
600
500
400
300
200
100
0
adult
child
elderly
infant
teenager
unknown
PivotChart Fields
Choose fields to add to report:
Search
pclass
survived
sex
age
Age group
sibsp
parch
ticket
fare
Drag fields between areas below:
Filters
Legend (Series)
Axis (Categories)
Values
Age group
Count of Age group
Defer Layout Update
Update
Sheet1
Sheet2
Ready
100%


titanic_book.xlsx - Excel
Chart 1
Survived by age group
600
500
400
300
200
100
0
0 1 0 1 0 1 0 1 0 1 0 1
adult
child
elderly
infant
teenager
unknown
PivotChart Fields
Choose fields to add to report:
Search
pclass
survived
sex
age
Age group
sibsp
parch
ticket
fare
Drag fields between areas below:
Filters
Legend (Series)
Axis (Categories)
Values
Age group
survived
Count of Age group
Defer Layout Update
Update
Sheet1
Sheet2
Ready
100%

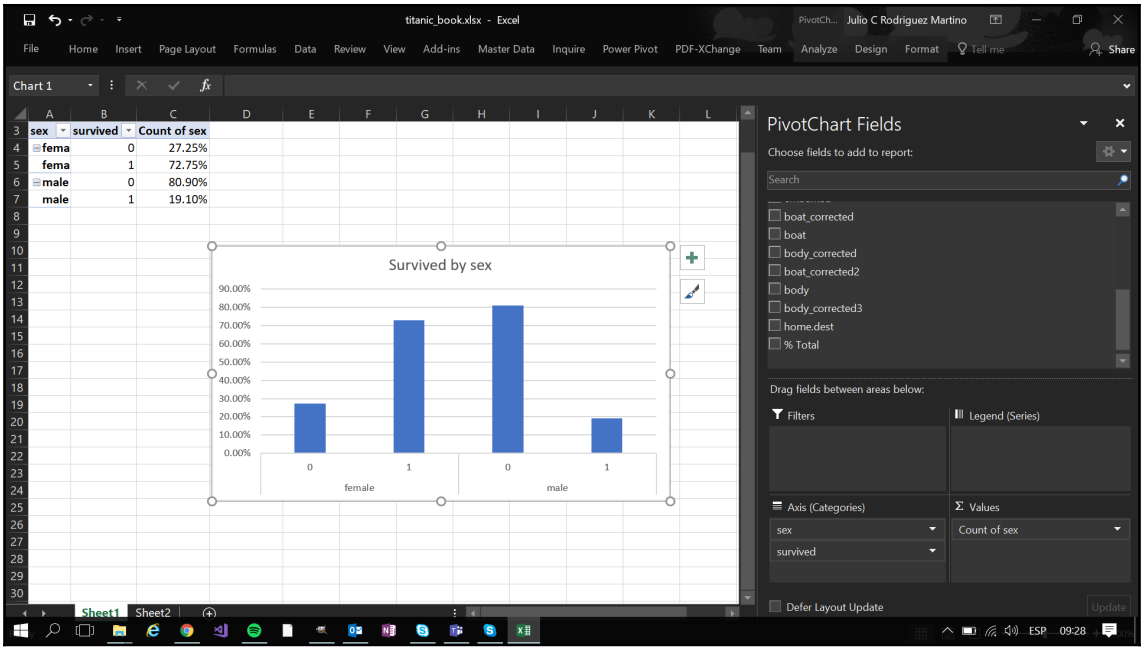

Value Field Settings
Source Name: Age group
Custom Name: Count of Age group
Summarize Values By
Show Values As
Show values as
% of Parent Total
Base field:
Base item:
pclass
survived
sex
age
Age group
sibsp
Number Format
OK
Cancel

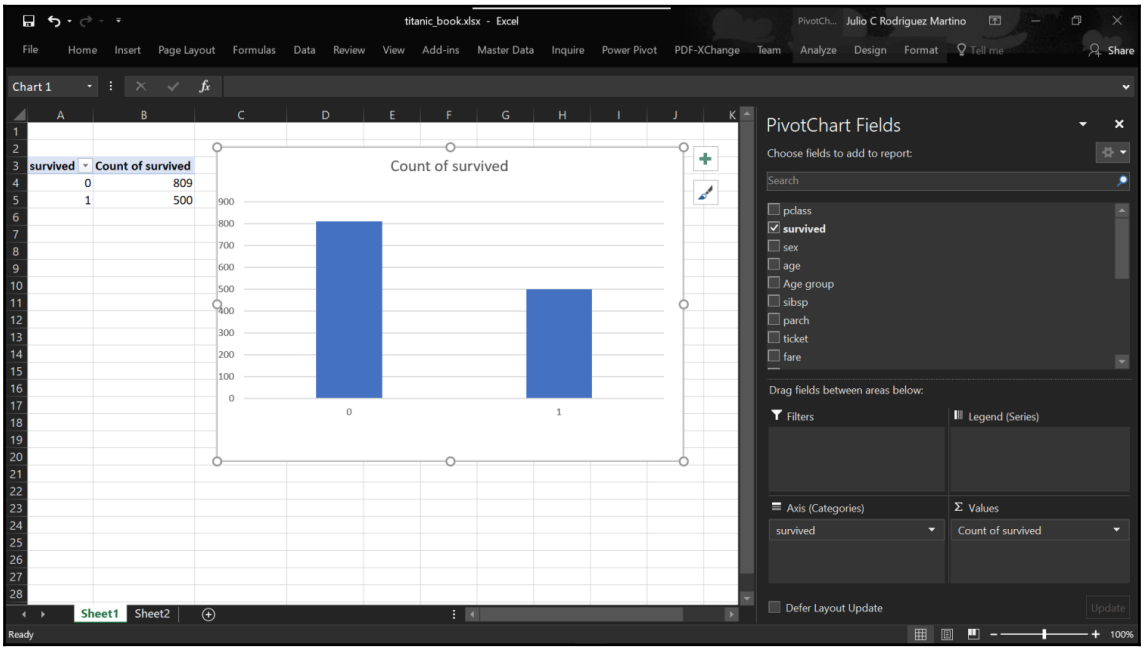


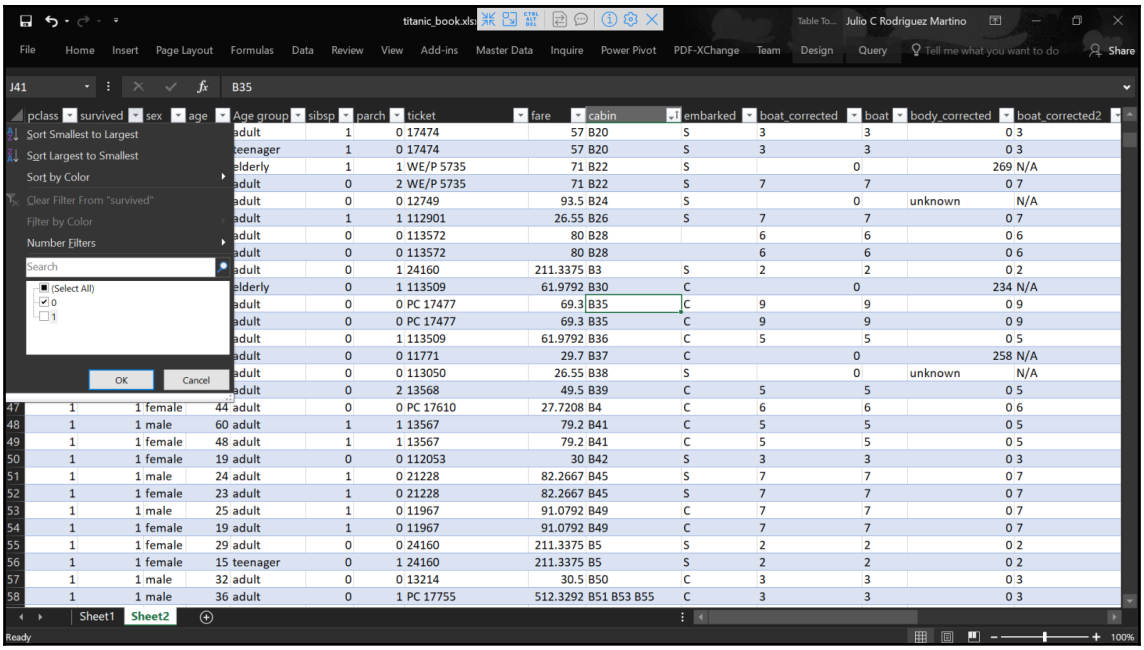
Age group	survived	Count of Age group
adult	0	60.86%
adult	1	39.14%
child	0	48.10%
child	1	51.90%
elderly	0	84.62%
elderly	1	15.38%
infant	0	16.67%
infant	1	83.33%
teenager	0	52.38%
teenager	1	47.62%
unknown	0	72.24%
unknown	1	27.76%

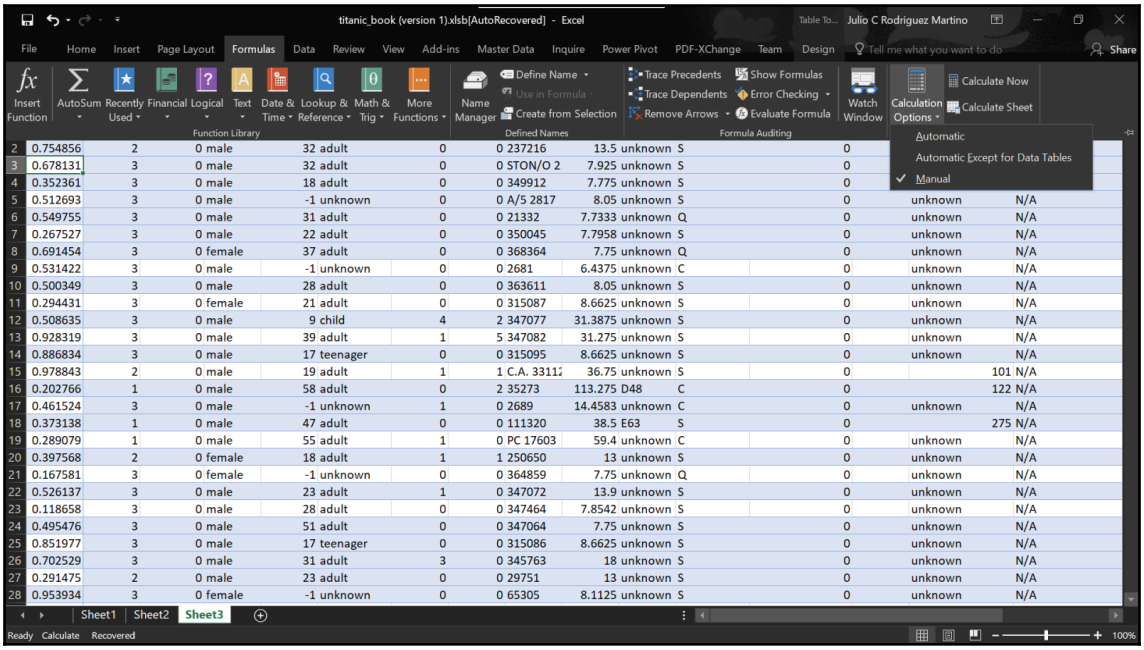


pclass	survived	Sum of pclass
1	0	38.08%
1	1	61.92%
2	0	57.04%
2	1	42.96%
3	0	74.47%
3	1	25.53%

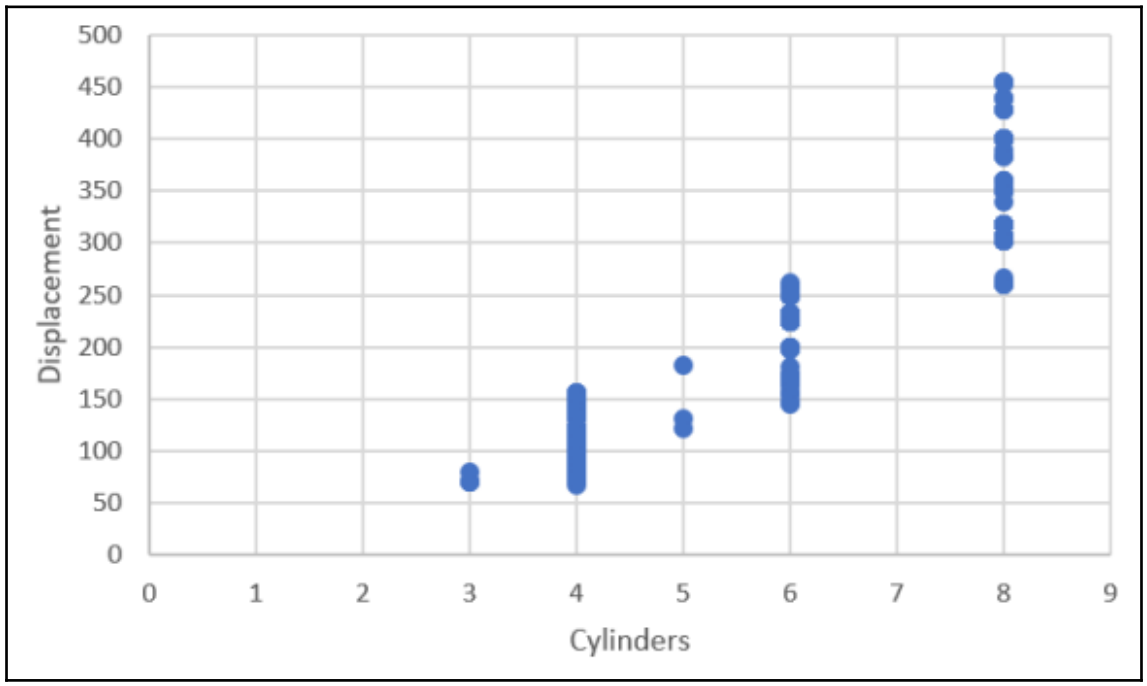

titanic_book.xlsx - Excel
Julio C Rodriguez Martino
Chart 1
sex survived Count of sex
fema 0 27.25%
fema 1 72.75%
male 0 80.90%
male 1 19.10%
Survived by sex
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
0
1
0
1
female
male
PivotChart Fields
Choose fields to add to report:
Search
boat_corrected
boat
body_corrected
boat_corrected2
body
body_corrected3
home.dest
% Total
Drag fields between areas below:
Filters
Legend (Series)
Axis (Categories)
Values
sex
survived
Count of sex
Defer Layout Update
Update
Sheet1
Sheet2
ESP
09:28

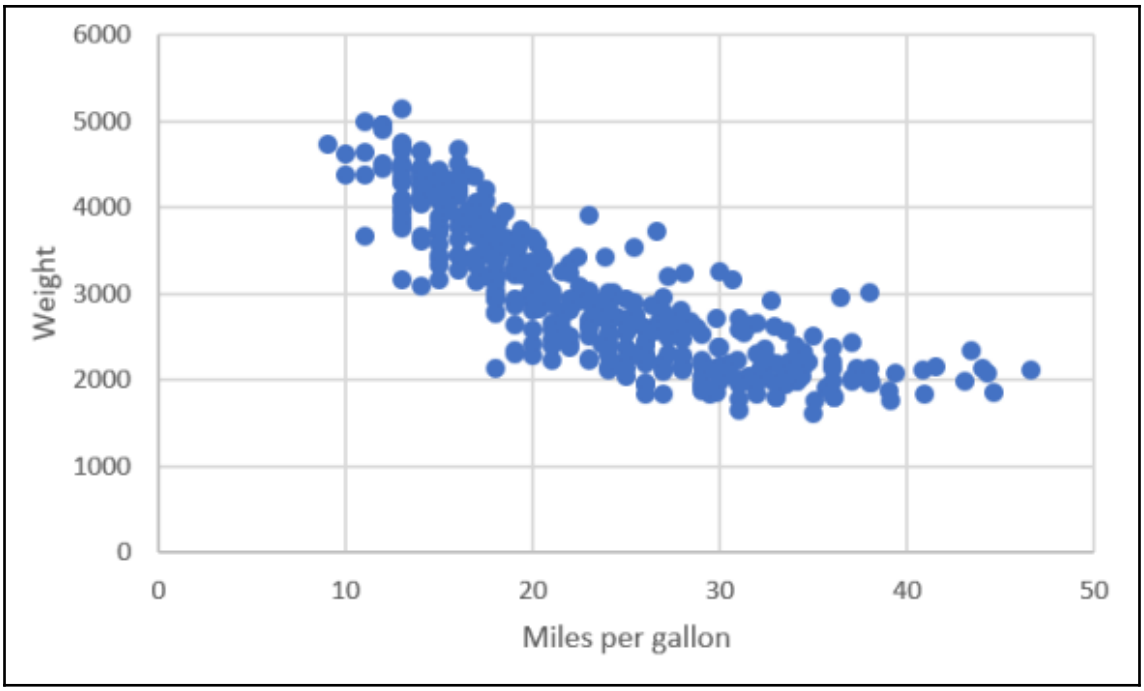

titanic_book.xlsx - Excel
Julio C Rodriguez Martino
Chart 1
survived Count of survived
0 809
1 500
Count of survived
900
800
700
600
500
400
300
200
100
0
0
1
PivotChart Fields
Choose fields to add to report:
Search
pclass
survived
sex
age
Age group
sibsp
parch
ticket
fare
Drag fields between areas below:
Filters
Legend (Series)
Axis (Categories)
Values
survived
Count of survived
Defer Layout Update
Update
Sheet1
Sheet2
Ready
100%

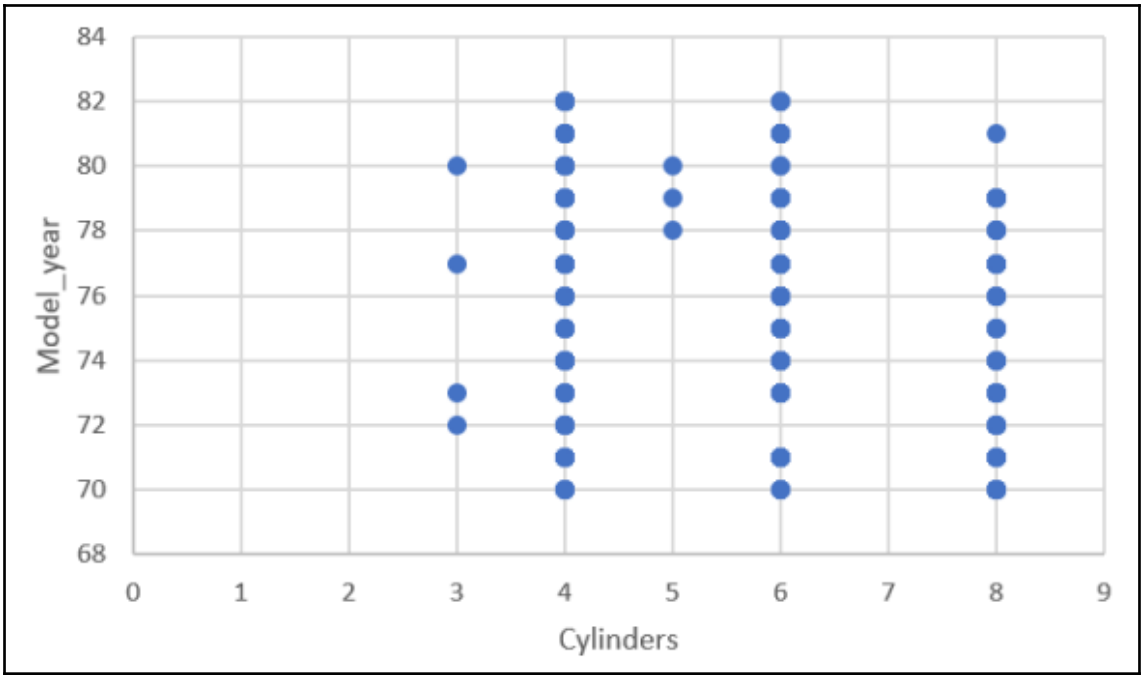

titanic_book.xlsb
Julio C Rodriguez Martino
File Home Insert Page Layout Formulas Data Review View Add-ins Master Data Inquire Power Pivot PDF-XChange Team Design Query
J41
B35
pclass survived sex age Age group sibsp parch ticket fare cabin embarked boat_corrected boat body_corrected boat_corrected2
Sort Smallest to Largest
Sort Largest to Smallest
Sort by Color
Clear Filter From "survived"
Filter by Color
Number Filters
Search
(Select All)
0
1
OK
Cancel
Sheet1 Sheet2
Ready

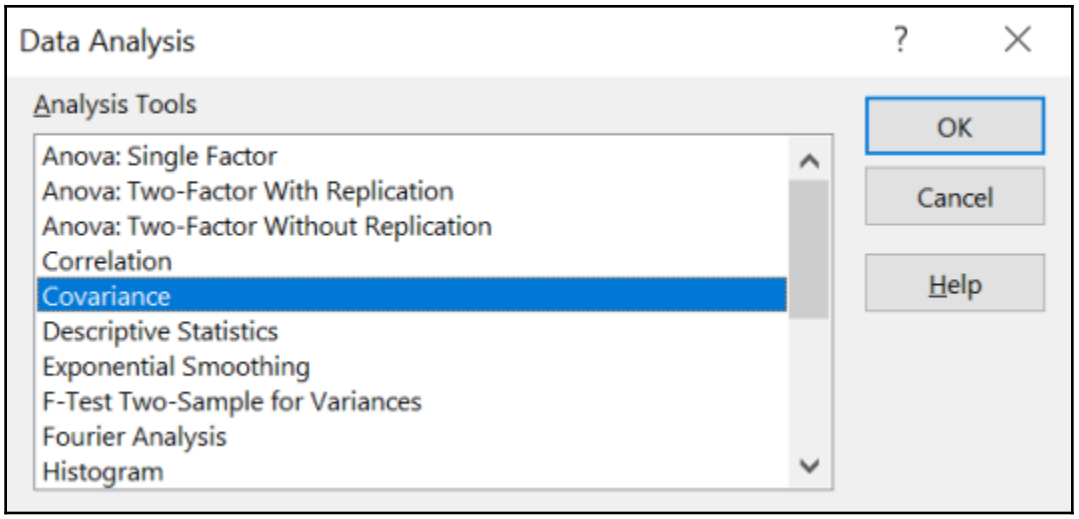

titanic_book (version 1).xlsb[AutoRecovered] - Excel
Julio C Rodriguez Martino
File Home Insert Page Layout Formulas Data Review View Add-ins Master Data Inquire Power Pivot PDF-XChange Team Design
Insert Function AutoSum Recently Used Financial Logical Text Date & Time Lookup & Reference Math & Trig More Functions
Function Library
Name Manager Define Name Use in Formula Create from Selection
Defined Names
Trace Precedents Trace Dependents Remove Arrows Show Formulas Error Checking Evaluate Formula
Formula Auditing
Watch Window
Calculation Options Calculate Now Calculate Sheet
Automatic
Automatic Except for Data Tables
Manual
Sheet1 Sheet2 Sheet3
Ready Calculate Recovered

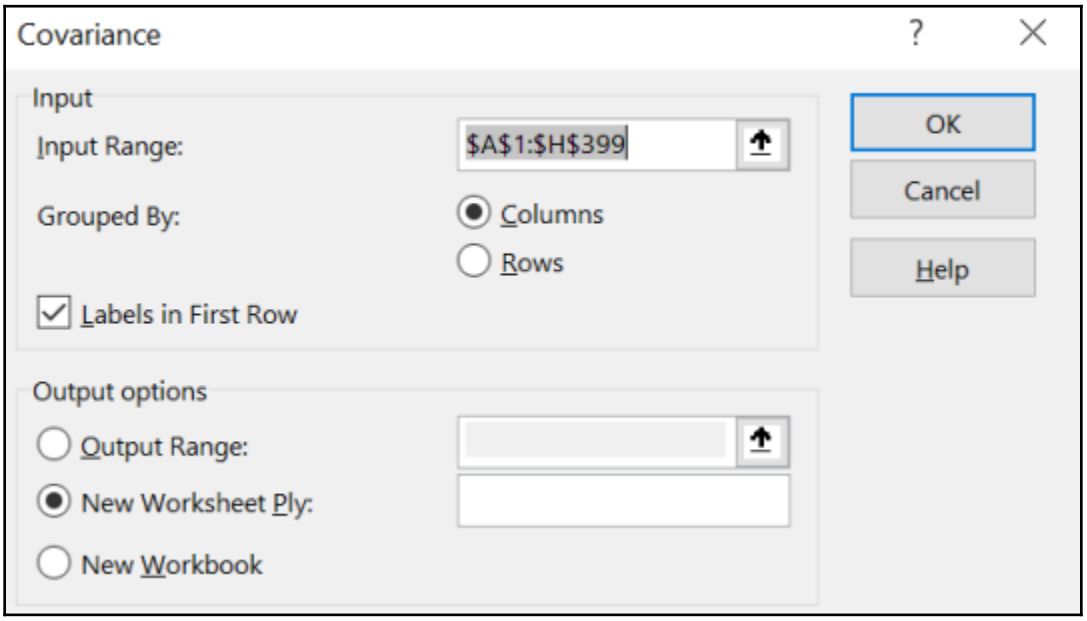
Chapter 5: Correlations and the Importance of Variables




6000
5000
4000
3000
2000
1000
0
Weight
0
10
20
30
40
50
Miles per gallon


84
82
80
78
76
74
72
70
68
Model_year
0
1
2
3
4
5
6
7
8
9
Cylinders


Data Analysis
Analysis Tools
Anova: Single Factor
Anova: Two-Factor With Replication
Anova: Two-Factor Without Replication
Correlation
Covariance
Descriptive Statistics
Exponential Smoothing
F-Test Two-Sample for Variances
Fourier Analysis
Histogram
OK
Cancel
Help


Covariance
Input
Input Range:
$A$1:$H$399
Grouped By:
Columns
Rows
Labels in First Row
Output options
Output Range:
New Worksheet Ply:
New Workbook
OK
Cancel
Help

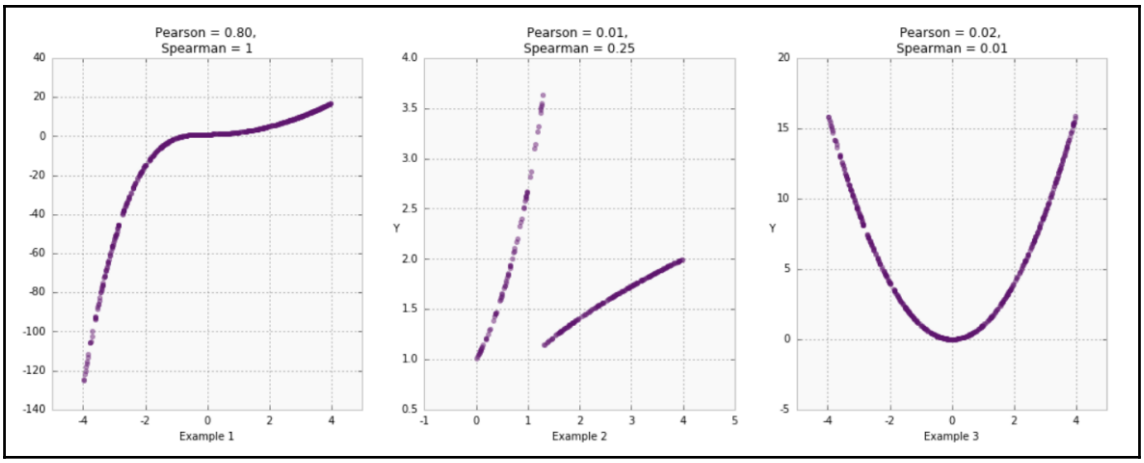
	A	B	C	D	E	F	G	H	I
1		*mpg*	*cylinders*	*displacement*	*horsepower*	*weight*	*acceleration*	*model_year*	*origin*
2	mpg	60.93611929							
3	cylinders	-10.28300927	2.886146						
4	displacem	-653.7555781	168.1995	10844.88207					
5	horsepow	-233.2613494	55.20705	3604.81427	1477.789879				
6	weight	-5491.379555	1287.453	82161.4674	28193.51406	715339.1287			
7	acceleratic	9.036168531	-2.36489	-155.9401792	-73.00026551	-972.4495158	7.585740575		
8	model_ye	16.69909977	-2.18799	-142.3585516	-58.88582882	-957.5344183	2.930722709	13.63808995	
9	origin	3.523310017	-0.76555	-50.83693594	-14.07673886	-393.647774	0.45420949	0.534443575	0.641676
10									

	A	B	C	D	E	F	G	H	I
1		*mpg*	*cylinders*	*displacement*	*horsepower*	*weight*	*acceleration*	*model_year*	*origin*
2	mpg	1							
3	cylinders	-0.7754	1						
4	displacement	-0.8042	0.950721	1					
5	horsepower	-0.77843	0.842983	0.897257002	1				
6	weight	-0.83174	0.896017	0.932824147	0.864537738	1			
7	acceleration	0.420289	-0.50542	-0.543684084	-0.68919551	-0.41745732	1		
8	model_year	0.579267	-0.34875	-0.370164161	-0.416361477	-0.306564334	0.288136954	1	
9	origin	0.56345	-0.56254	-0.609409399	-0.455171453	-0.581023914	0.205873007	0.180662195	1
10									

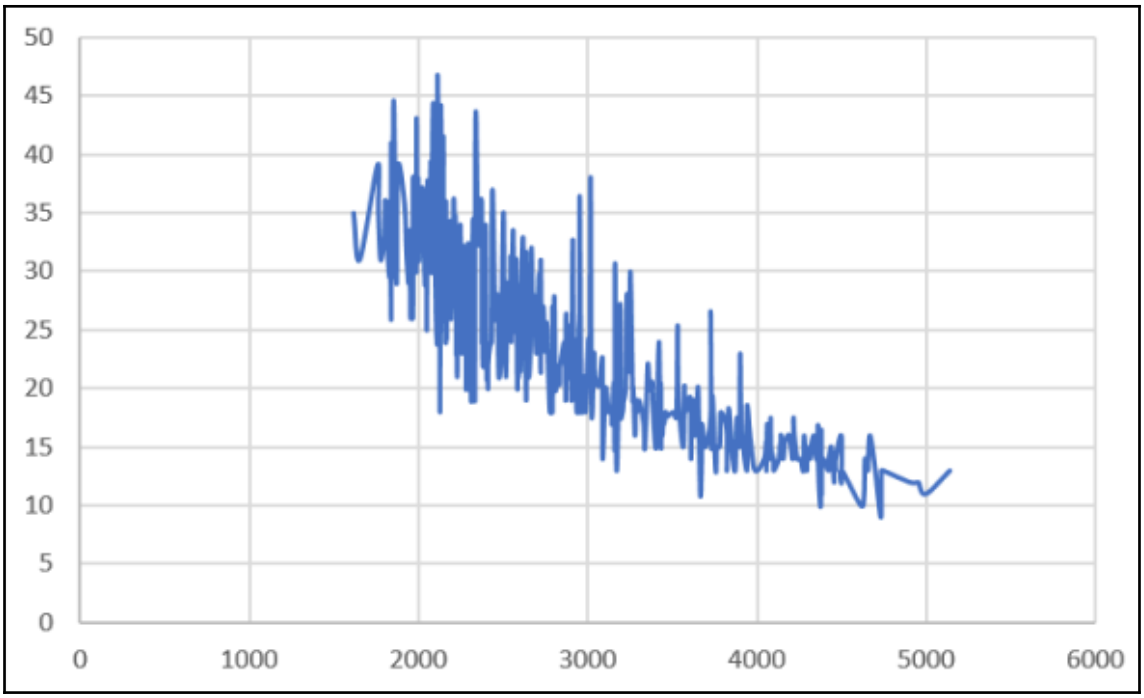
	A	B	C	D	E	F	G	H
1	Rank_mpg	Rank_cylinders	Rank_displacement	Rank_horsepower	Rank_weight	Rank_acceleration	Rank_model_year	Rank_origin
2	283	52	75	94	109	362.5	384	274
3	337.5	52	46.5	35.5	90	372	384	274
4	283	52	65	56.5	115	384	384	274
5	318	52	84	56.5	116	362.5	384	274
6	303	52	93	81	112	388	384	274
7	337.5	52	8	12.5	34	390.5	384	274
8	356	52	4	5	33	395	384	274
9	356	52	5.5	7	37	396.5	384	274
10	356	52	2	3	25	390.5	384	274
11	337.5	52	23	16	76	396.5	384	274
12	337.5	52	24.5	30	104	390.5	384	274
13	356	52	56	38.5	100	398	384	274
14	337.5	52	16	56.5	84	393.5	384	274
15	356	52	2	3	159	390.5	384	274
16	179	292.5	274	188.5	272	224.5	384	40
17	209.5	145.5	170	188.5	196	195	384	274
18	283	145.5	167.5	174	204	195	384	274
19	223.5	145.5	162.5	256	237	162.5	384	274
20	129	292.5	331	235	325.5	258	384	40
21	145.5	292.5	331	391.5	385.5	19	384	114.5
22	164	292.5	281	245.5	218	89.5	384	114.5
23	179	292.5	289	214.5	259	258	384	114.5
24	164	292.5	299	188.5	271	89.5	384	114.5
25	145.5	292.5	246	113	295	350.5	384	114.5
26	223.5	145.5	167.5	214.5	224	224.5	384	274
27	396.5	52	27.5	7	16	288.5	384	274
28	396.5	52	75	11	30	224.5	384	274

... | Data | Covariance | Pearson | **Ranks** | Spearman | Sheet8

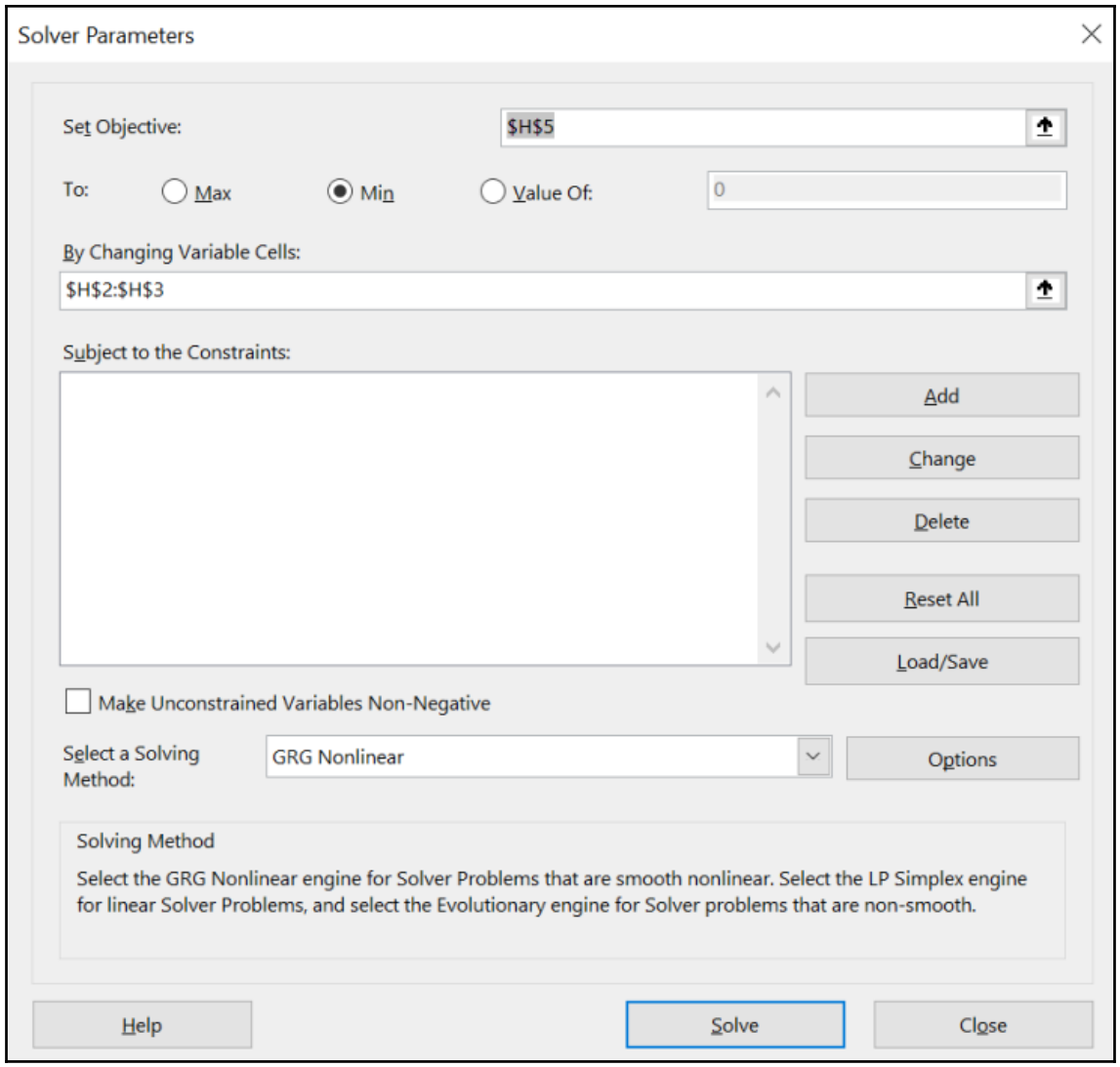
	A	B	C	D	E	F	G	H	I
1		*Rank_mpg*	*Rank_cylinders*	*Rank_displacement*	*Rank_horsepower*	*Rank_weight*	*Rank_acceleration*	*Rank_model_year*	*Rank_origin*
2	Rank_mpg	1							
3	Rank_cylinders	-0.821864491	1						
4	Rank_displacement	-0.855692012	0.911875915	1					
5	Rank_horsepower	-0.853320216	0.815689638	0.875770352	1				
6	Rank_weight	-0.874947398	0.873313559	0.945985564	0.878284909	1			
7	Rank_acceleration	0.43867748	-0.474189066	-0.496511921	-0.657631236	-0.404550372	1		
8	Rank_model_year	0.573468703	-0.335012387	-0.30525727	-0.389975332	-0.277014582	0.274632098	1	
9	Rank_origin	0.580693694	-0.604550452	-0.707196539	-0.509090776	-0.628434003	0.220573847	0.166551172	1
10									

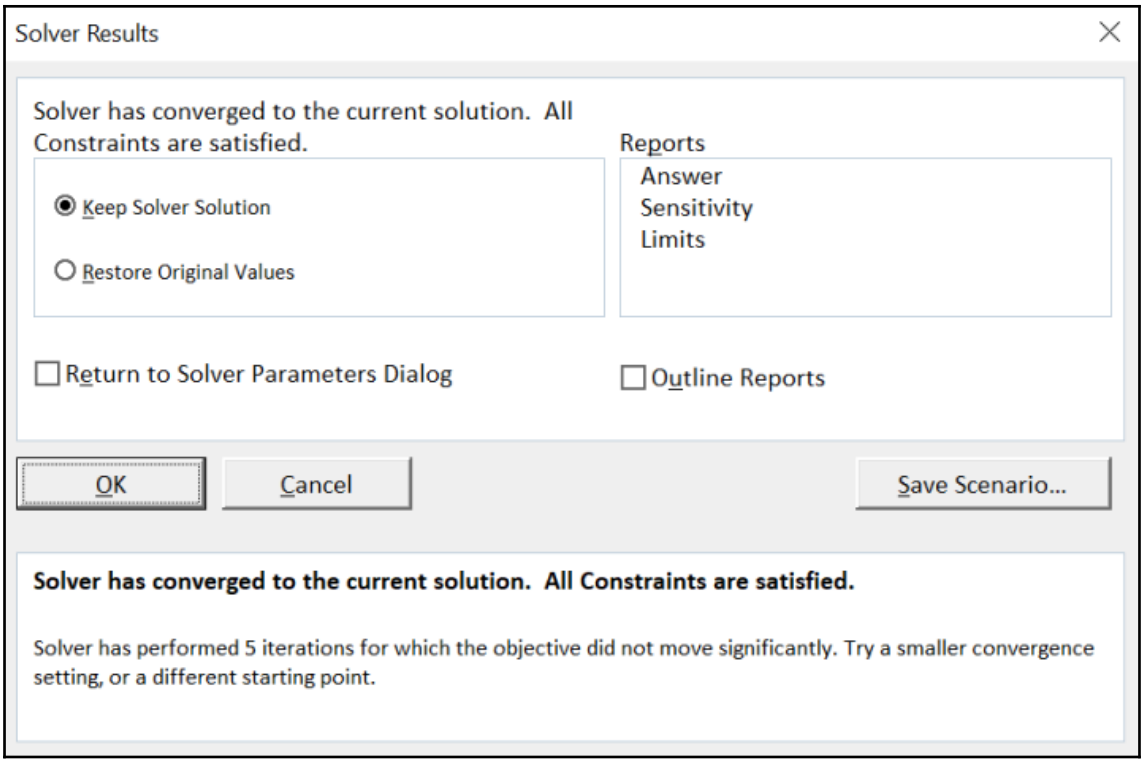

Pearson = 0.80,
Spearman = 1
Example 1
Pearson = 0.01,
Spearman = 0.25
Y
Example 2
Pearson = 0.02,
Spearman = 0.01
Y
Example 3

	A	B
1	weight	mpg
2	1613	35
3	1649	31
4	1755	39.1
5	1760	35.1
6	1773	31
7	1795	33
8	1795	33
9	1800	36.1
10	1800	36.1
11	1825	29.5
12	1825	36
13	1834	27
14	1835	26
15	1835	40.9
16	1836	32
17	1845	29.8
18	1850	44.6
19	1867	29
20	1875	39
21	1915	35.7
22	1925	31.9
23	1937	29

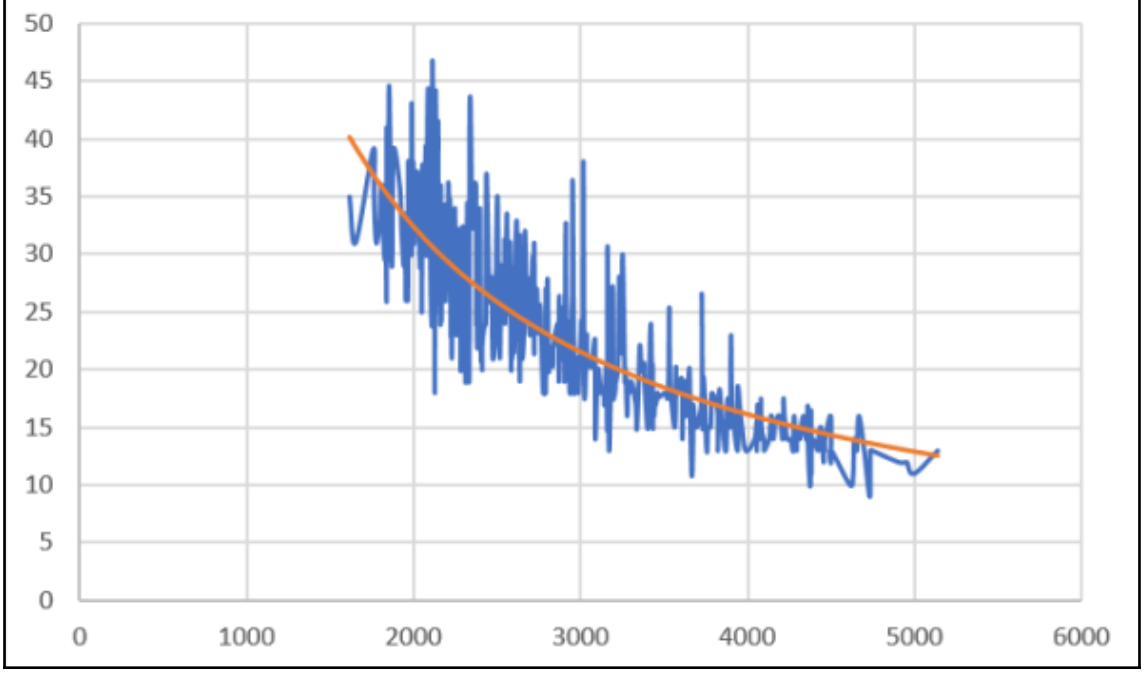

50
45
40
35
30
25
20
15
10
5
0
0
1000
2000
3000
4000
5000
6000

	A	B	C
1	weight	mpg	predict
2	1613	35	1.493943
3	1649	31	1.477546
4	1755	39.1	1.43223
5	1760	35.1	1.430194
6	1773	31	1.424941
7	1795	33	1.416182
8	1795	33	1.416182
9	1800	36.1	1.414214
10	1800	36.1	1.414214
11	1825	29.5	1.404494
12	1825	36	1.404494
13	1834	27	1.401043
14	1835	26	1.400662
15	1835	40.9	1.400662
16	1836	32	1.40028
17	1845	29.8	1.396861
18	1850	44.6	1.394972
19	1867	29	1.388606
20	1875	39	1.385641


Solver Parameters
Set Objective:
$H$5
To:
Max
Min
Value Of:
0
By Changing Variable Cells:
$H$2:$H$3
Subject to the Constraints:
Add
Change
Delete
Reset All
Load/Save
Make Unconstrained Variables Non-Negative
Select a Solving Method:
GRG Nonlinear
Options
Solving Method
Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.
Help
Solve
Close

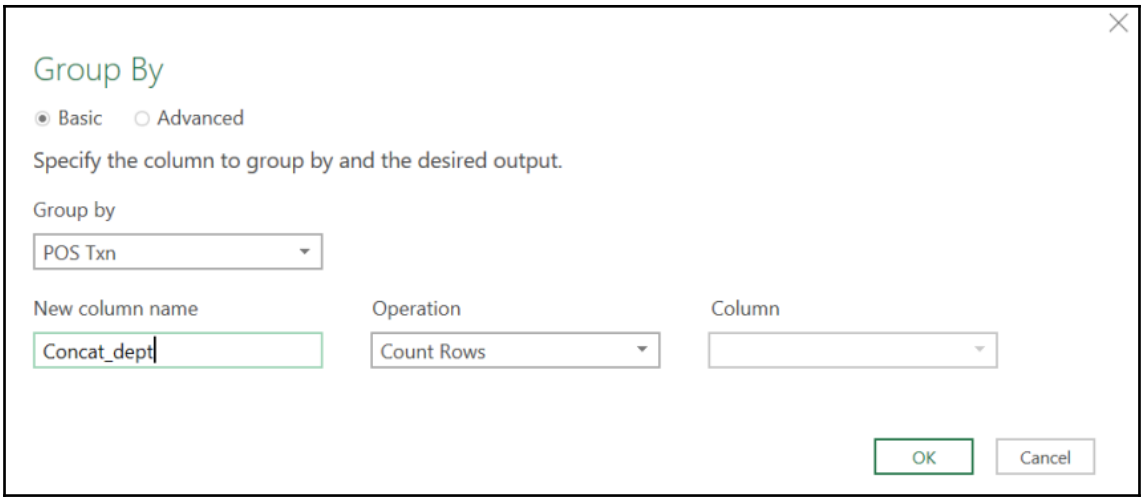

Solver Results
Solver has converged to the current solution. All Constraints are satisfied.
Keep Solver Solution
Restore Original Values
Reports
Answer
Sensitivity
Limits
Return to Solver Parameters Dialog
Outline Reports
OK
Cancel
Save Scenario...
Solver has converged to the current solution. All Constraints are satisfied.
Solver has performed 5 iterations for which the objective did not move significantly. Try a smaller convergence setting, or a different starting point.

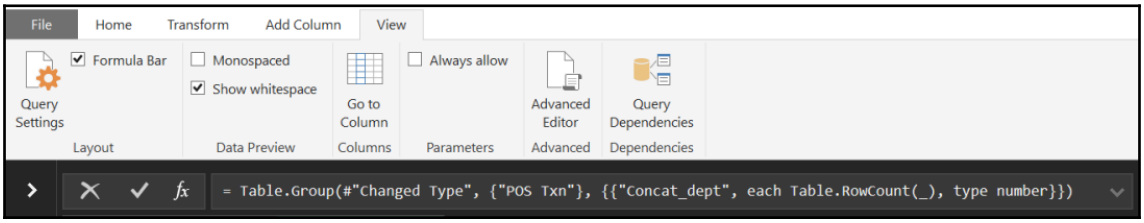
	a	68563.9126
	b	-1.0074493
	SSE	7117.53396



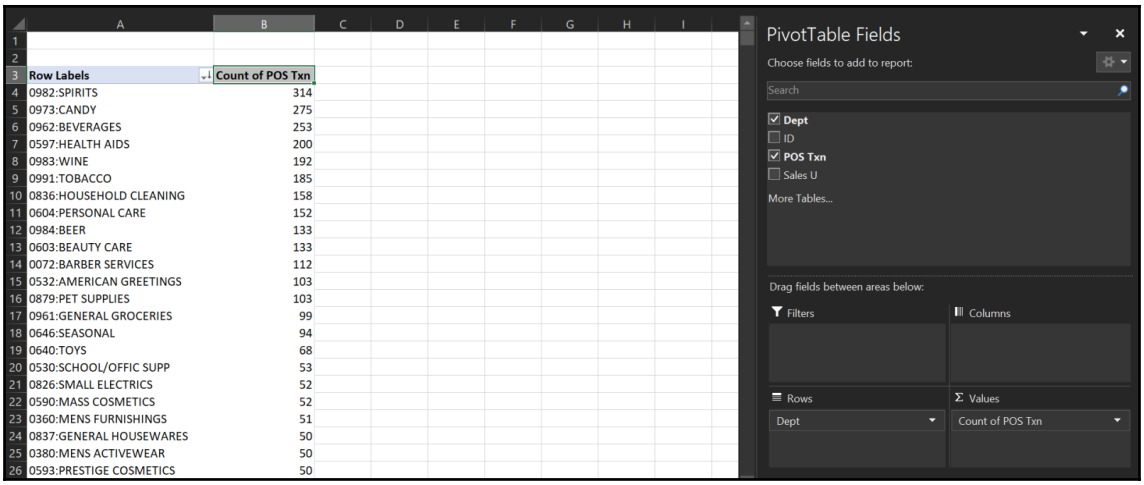
Chapter 6: Data Mining Models in Excel Hands-On Examples

	A	B	C	D
1	POS Txn	Dept	ID	Sales U
2	16120100160021008773	0261:HOSIERY	250	2
3	16120100160021008773	0634:VITAMINS & HLTH AIDS	102	1
4	16120100160021008773	0879:PET SUPPLIES	158	2
5	16120100160021008773	0973:CANDY	175	2
6	16120100160021008773	0982:SPIRITS	176	1
7	16120100160021008773	0983:WINE	177	4
8	16120100160021008773	0991:TOBACCO	179	2
9	16120100160021008774	0597:HEALTH AIDS	93	1
10	16120100160021008774	0604:PERSONAL CARE	100	5
11	16120100160021008775	0819:PRE-RECORDED A/V	135	1
12	16120100160021008775	0826:SMALL ELECTRICS	138	1
13	16120100160021008775	0982:SPIRITS	176	1
14	16120100160021008776	0961:GENERAL GROCERIES	169	3
15	16120100160021008777	0982:SPIRITS	176	2
16	16120100160021008778	0982:SPIRITS	176	4
17	16120100160021008778	0991:TOBACCO	179	1
18	16120100160021008779	0879:PET SUPPLIES	158	16
19	16120100160021008779	0982:SPIRITS	176	1
20	16120100160021008779	0983:WINE	177	2
21	16120100160021008779	0984:BEER	178	1
22	16120100160021008780	0530:SCHOOL/OFFIC SUPP	70	1
23	16120100160021008780	0597:HEALTH AIDS	93	1
24	16120100160021008780	0601:VALUE ZONE	97	1
25	16120100160021008780	0634:VITAMINS & HLTH AIDS	102	1





	A	B
1	POS Txn	Attributes
2	16120100160021008773	0261:HOSIERY | 0634:VITAMINS & HLTH AIDS | 0879:PET SUPPLIES | 0973:CANDY | 0982:SPIRITS | 0983:WINE | 0991:TOBACCO
3	16120100160021008774	0597:HEALTH AIDS | 0604:PERSONAL CARE
4	16120100160021008775	0819:PRE-RECORDED A/V | 0826:SMALL ELECTRICS | 0982:SPIRITS
5	16120100160021008776	0961:GENERAL GROCERIES
6	16120100160021008777	0982:SPIRITS
7	16120100160021008778	0982:SPIRITS | 0991:TOBACCO
8	16120100160021008779	0879:PET SUPPLIES | 0982:SPIRITS | 0983:WINE | 0984:BEER
9	16120100160021008780	0530:SCHOOL/OFFIC SUPP | 0597:HEALTH AIDS | 0601:VALUE ZONE | 0634:VITAMINS & HLTH AIDS | 0836:HOUSEHOLD CLEANING
10	16120100160021008781	0593:PRESTIGE COSMETICS | 0597:HEALTH AIDS | 0598:BABY CARE | 0836:HOUSEHOLD CLEANING | 0965:PERISHABLES | 0973:CANDY | 0983:WINE
11	16120100160021008782	0837:GENERAL HOUSEWARES | 0982:SPIRITS
12	16120100160021008783	0879:PET SUPPLIES | 0973:CANDY | 0984:BEER
13	16120100160021008784	0983:WINE
14	16120100160021008785	0962:BEVERAGES | 0982:SPIRITS
15	16120100160021008786	0982:SPIRITS | 0983:WINE
16	16120100160021008787	0982:SPIRITS
17	16120100160021008788	0638:GEN SPORTING GOODS | 0961:GENERAL GROCERIES | 0973:CANDY | 0991:TOBACCO
18	16120100160021008789	0646:SEASONAL | 0991:TOBACCO
19	16120100160021008790	0962:BEVERAGES | 0982:SPIRITS | 0983:WINE
20	16120100160021008791	0982:SPIRITS
21	16120100160021008792	0982:SPIRITS | 0983:WINE
22	16120100160021008793	0962:BEVERAGES | 0982:SPIRITS



Row Labels	Count of POS Txn
0982:SPIRITS	314
0973:CANDY	275
0962:BEVERAGES	253
0597:HEALTH AIDS	200
0983:WINE	192
0991:TOBACCO	185
0836:HOUSEHOLD CLEANING	158
0604:PERSONAL CARE	152
0984:BEER	133
0603:BEAUTY CARE	133
0072:BARBER SERVICES	112
0532:AMERICAN GREETINGS	103
0879:PET SUPPLIES	103
0961:GENERAL GROCERIES	99
0646:SEASONAL	94
0640:TOYS	68
0530:SCHOOL/OFFIC SUPP	53
0826:SMALL ELECTRICS	52
0590:MASS COSMETICS	52
0360:MENS FURNISHINGS	51
0837:GENERAL HOUSEWARES	50
0380:MENS ACTIVEWEAR	50
0593:PRESTIGE COSMETICS	50

PivotTable Fields

Choose fields to add to report:

Search

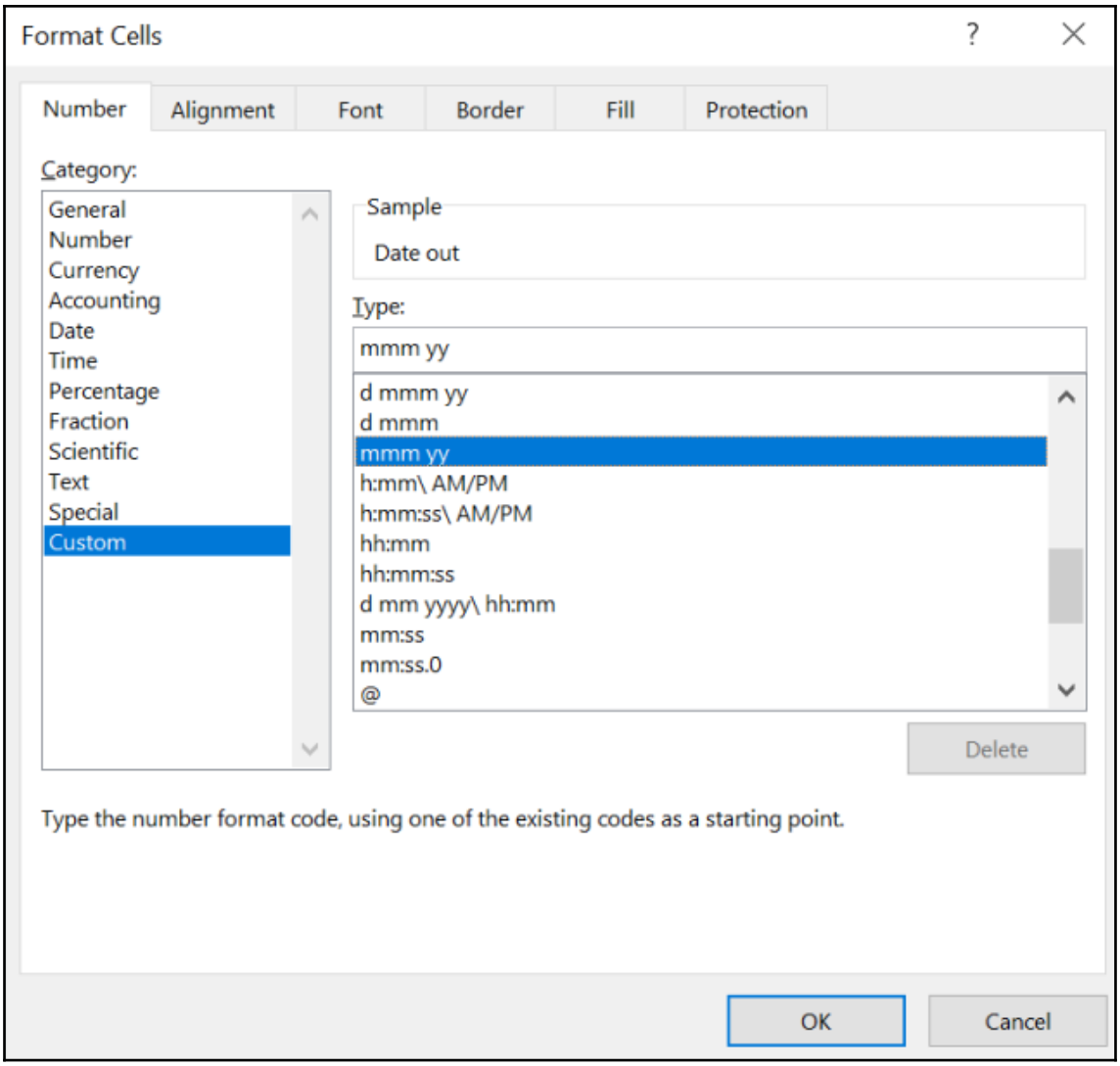
Dept
ID
POS Txn
Sales U
More Tables...

Drag fields between areas below:

Filters | Columns | Rows: Dept | Values: Count of POS Txn

C	D
X	Y
0982:SPIRITS	0973:CANDY
0982:SPIRITS	0962:BEVERAGES
0982:SPIRITS	0597:HEALTH AIDS
0982:SPIRITS	0983:WINE
0982:SPIRITS	0991:TOBACCO
0982:SPIRITS	0836:HOUSEHOLD CLEANING
0982:SPIRITS	0604:PERSONAL CARE
0982:SPIRITS	0603:BEAUTY CARE
0982:SPIRITS	0984:BEER

C	D	E	F	G	H	I
X	Y	X&Y	N	Support	X	Confidence
0982:SPIRITS	0973:CANDY	31	2064	2%	314	10%
0982:SPIRITS	0962:BEVERAGES	49	2064	2%	314	16%
0982:SPIRITS	0597:HEALTH AIDS	24	2064	1%	314	8%
0982:SPIRITS	0983:WINE	77	2064	4%	314	25%
0982:SPIRITS	0991:TOBACCO	52	2064	3%	314	17%
0982:SPIRITS	0836:HOUSEHOLD CLEANING	22	2064	1%	314	7%
0982:SPIRITS	0604:PERSONAL CARE	15	2064	1%	314	5%
0982:SPIRITS	0603:BEAUTY CARE	16	2064	1%	314	5%
0982:SPIRITS	0984:BEER	50	2064	2%	314	16%
0973:CANDY	0962:BEVERAGES	67	2064	3%	275	24%
0973:CANDY	0597:HEALTH AIDS	45	2064	2%	275	16%
0973:CANDY	0983:WINE	37	2064	2%	275	13%
0973:CANDY	0991:TOBACCO	25	2064	1%	275	9%
0973:CANDY	0836:HOUSEHOLD CLEANING	35	2064	2%	275	13%
0973:CANDY	0604:PERSONAL CARE	30	2064	1%	275	11%
0973:CANDY	0603:BEAUTY CARE	31	2064	2%	275	11%
0973:CANDY	0984:BEER	16	2064	1%	275	6%
0962:BEVERAGES	0597:HEALTH AIDS	36	2064	2%	253	14%
0962:BEVERAGES	0983:WINE	24	2064	1%	253	9%
0962:BEVERAGES	0991:TOBACCO	32	2064	2%	253	13%
0962:BEVERAGES	0836:HOUSEHOLD CLEANING	27	2064	1%	253	11%
0962:BEVERAGES	0604:PERSONAL CARE	30	2064	1%	253	12%
0962:BEVERAGES	0603:BEAUTY CARE	25	2064	1%	253	10%
0962:BEVERAGES	0984:BEER	18	2064	1%	253	7%
0597:HEALTH AIDS	0983:WINE	29	2064	1%	200	15%


Format Cells
?
Number
Alignment
Font
Border
Fill
Protection
Category:
General
Number
Currency
Accounting
Date
Time
Percentage
Fraction
Scientific
Text
Special
Custom
Sample
Date out
Type:
mmm yy
d mmm yy
d mmm
mmm yy
h:mm\ AM/PM
h:mm:ss\ AM/PM
hh:mm
hh:mm:ss
d mm yyyy\ hh:mm
mm:ss
mm:ss.0
@
Delete
Type the number format code, using one of the existing codes as a starting point.
OK
Cancel

	A	B	C	D	E	F
1	Id	Date in	Date out	Mean Monthly Spend	Cohort	Active months
2	236503	Feb-15	May-16	$ 203.90	Feb-15	15
3	236508	Feb-15	Aug-16	$ 547.80	Feb-15	23
4	236574	Feb-15	Jun-15	$ 865.80	Feb-15	22
5	236584	Feb-15	Aug-16	$ 408.20	Feb-15	16
6	236593	Feb-15	Nov-16	$ 455.40	Feb-15	21
7	236622	Feb-15	Sep-15	$ 387.60	Feb-15	14
8	236630	Feb-15	Jan-17	$ 156.90	Feb-15	23
9	236661	Feb-15	Oct-15	$ 941.20	Feb-15	7
10	236667	Feb-15	Apr-16	$ 195.80	Feb-15	7
11	236677	Feb-15	Jun-15	$ 869.60	Feb-15	9
12	236692	Feb-15	Jun-16	$ 692.10	Feb-15	6
13	236712	Feb-15	Jul-16	$ 878.30	Feb-15	11
14	236742	Feb-15	Jan-17	$ 918.80	Feb-15	20
15	236749	Feb-15	Apr-16	$ 452.30	Feb-15	8
16	236768	Feb-15	Nov-16	$ 181.10	Feb-15	3
17	236881	Feb-15	Jan-17	$ 121.40	Feb-15	5
18	236951	Feb-15	Jun-16	$ 279.90	Feb-15	23
19	236996	Feb-15	Aug-15	$ 334.80	Feb-15	14
20	237071	Feb-15	Sep-15	$ 774.40	Feb-15	11
21	237073	Feb-15	Jun-16	$ 304.20	Feb-15	16
22	237077	Feb-15	Aug-15	$ 646.30	Feb-15	3

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Cohorts		Feb-15	Mar-15	Apr-15	May-15	Jun-15	Jul-15	Aug-15	Sep-15	Oct-15	Nov-15	Dec-15	Jan-16	Feb-16	Mar-16
2	Month	0														
3	Month	1														
4	Month	2														
5	Month	3														
6	Month	4														
7	Month	5														
8	Month	6														
9	Month	7														
10	Month	8														
11	Month	9														
12	Month	10														
13	Month	11														
14	Month	12														
15	Month	13														
16	Month	14														
17	Month	15														
18	Month	16														
19	Month	17														
20	Month	18														
21	Month	19														
22	Month	20														

C52 =COUNTIF('Customer data'!$E$2:$E$751;"="&C$1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
46	Month	44															
47	Month	45															
48	Month	46															
49	Month	47															
50	Month	48															
51																	
52	Total		31	31	25	28	30	32	30	22	38	25	31	33	30	25	35
53																	

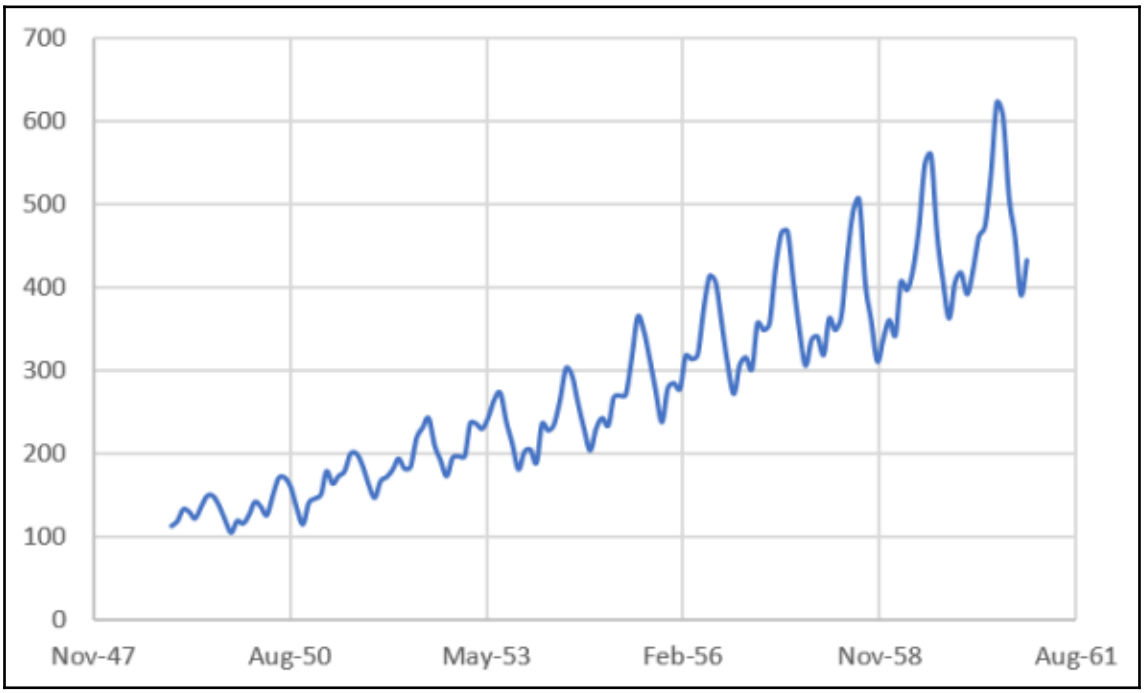
C2 =COUNTIFS('Customer data'!$E$2:$E$751;"="&C$1;'Customer data'!$C$2:$C$751;">"&EOMONTH(C$1;$B1))/C$52

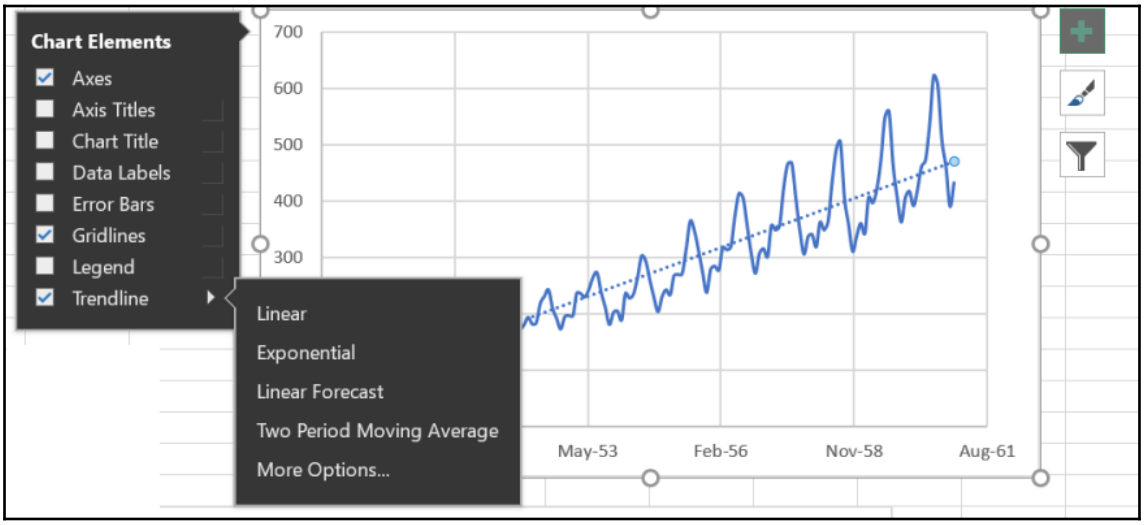
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Cohorts		Feb-15	Mar-15	Apr-15	May-15	Jun-15	Jul-15	Aug-15	Sep-15	Oct-15	Nov-15	Dec-15	Jan-16	Feb-16	Mar-16	Apr-16	May-16	Jun-16
2	Month	0	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	Month	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	Month	2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
5	Month	3	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6	Month	4	100.00%	96.77%	100.00%	92.86%	96.67%	96.88%	100.00%	90.91%	92.11%	100.00%	90.32%	87.88%	93.33%	96.00%	97.14%	96.43%	85.00%
7	Month	5	90.32%	93.55%	96.00%	89.29%	80.00%	90.63%	100.00%	90.91%	81.58%	100.00%	77.42%	84.85%	86.67%	92.00%	88.57%	92.86%	85.00%
8	Month	6	90.32%	90.32%	92.00%	78.57%	76.67%	87.50%	93.33%	90.91%	78.95%	92.00%	74.19%	81.82%	83.33%	92.00%	77.14%	82.14%	80.00%
9	Month	7	83.87%	83.87%	84.00%	78.57%	70.00%	84.38%	86.67%	90.91%	71.05%	88.00%	70.97%	75.76%	80.00%	88.00%	77.14%	78.57%	75.00%
10	Month	8	74.19%	80.65%	80.00%	78.57%	60.00%	81.25%	86.67%	90.91%	65.79%	80.00%	64.52%	72.73%	73.33%	88.00%	68.57%	75.00%	75.00%
11	Month	9	67.74%	74.19%	72.00%	71.43%	56.67%	75.00%	80.00%	77.27%	60.53%	76.00%	64.52%	72.73%	70.00%	80.00%	62.86%	64.29%	65.00%
12	Month	10	67.74%	70.97%	64.00%	67.86%	53.33%	71.88%	70.00%	72.73%	55.26%	68.00%	54.84%	69.70%	63.33%	68.00%	57.14%	64.29%	65.00%
13	Month	11	67.74%	67.74%	60.00%	64.29%	43.33%	65.63%	66.67%	68.18%	50.00%	64.00%	51.61%	69.70%	60.00%	52.00%	54.29%	60.71%	55.00%
14	Month	12	61.29%	61.29%	60.00%	57.14%	43.33%	65.63%	60.00%	68.18%	47.37%	60.00%	38.71%	60.61%	56.67%	52.00%	54.29%	57.14%	55.00%
15	Month	13	61.29%	51.61%	52.00%	50.00%	43.33%	59.38%	56.67%	59.09%	44.74%	52.00%	35.48%	54.55%	50.00%	48.00%	51.43%	50.00%	55.00%
16	Month	14	58.06%	48.39%	48.00%	46.43%	43.33%	56.25%	46.67%	45.45%	39.47%	52.00%	35.48%	54.55%	43.33%	44.00%	48.57%	46.43%	55.00%
17	Month	15	48.39%	45.16%	44.00%	39.29%	36.67%	53.13%	40.00%	40.91%	39.47%	44.00%	32.26%	45.45%	40.00%	44.00%	40.00%	46.43%	50.00%
18	Month	16	45.16%	45.16%	36.00%	35.71%	33.33%	50.00%	36.67%	36.36%	36.84%	40.00%	25.81%	39.39%	33.33%	32.00%	34.29%	46.43%	35.00%
19	Month	17	35.48%	45.16%	36.00%	28.57%	30.00%	43.75%	36.67%	22.73%	34.21%	36.00%	25.81%	33.33%	30.00%	32.00%	31.43%	42.86%	35.00%
20	Month	18	29.03%	45.16%	32.00%	21.43%	26.67%	40.63%	36.67%	22.73%	28.95%	28.00%	22.58%	30.30%	26.67%	28.00%	25.71%	35.71%	35.00%
21	Month	19	22.58%	41.94%	28.00%	14.29%	23.33%	34.38%	23.33%	13.64%	28.95%	24.00%	16.13%	24.24%	20.00%	28.00%	17.14%	32.14%	35.00%
22	Month	20	19.35%	32.26%	16.00%	10.71%	20.00%	34.38%	16.67%	13.64%	26.32%	20.00%	16.13%	21.21%	16.67%	24.00%	17.14%	32.14%	35.00%

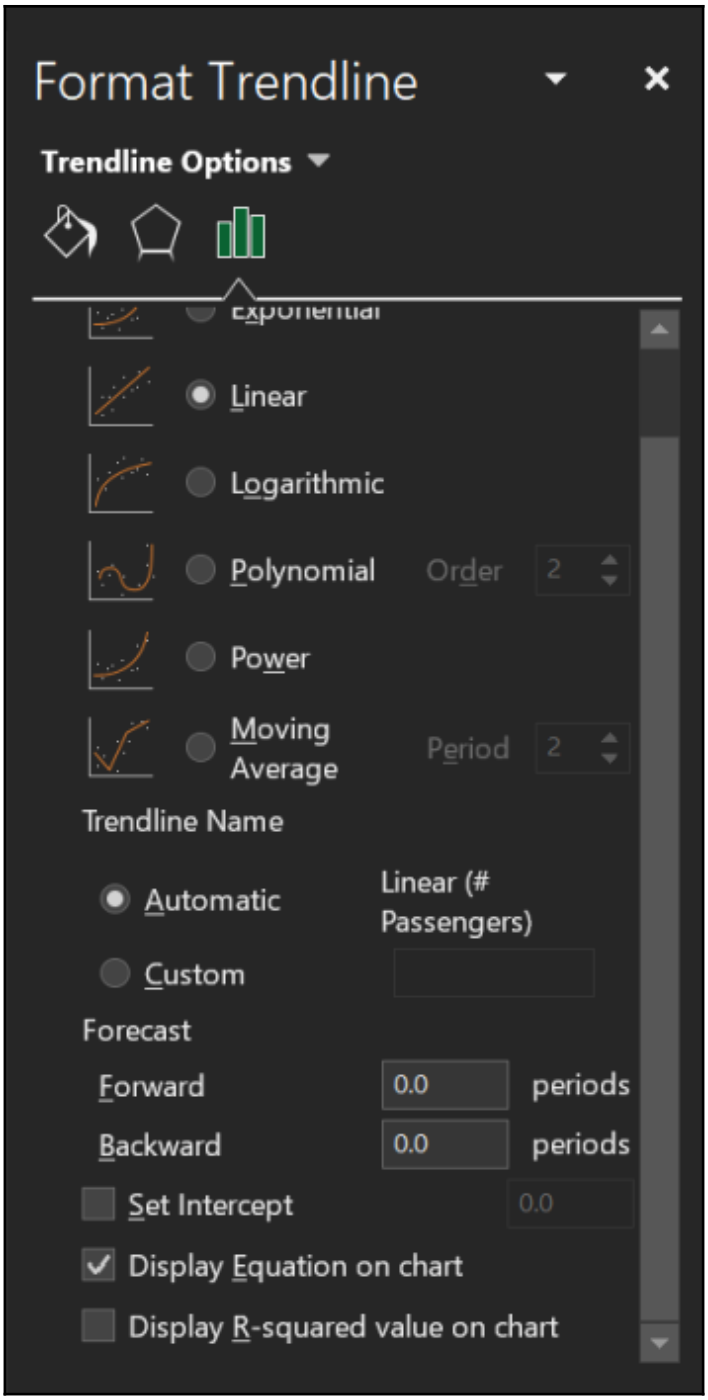
C2 =SUMIFS('Customer data'!$D$2:$D$751;'Customer data'!$E$2:$E$751;"="&C$1;'Customer data'!$C$2:$C$751;">"&EOMONTH(C$1;$B1))

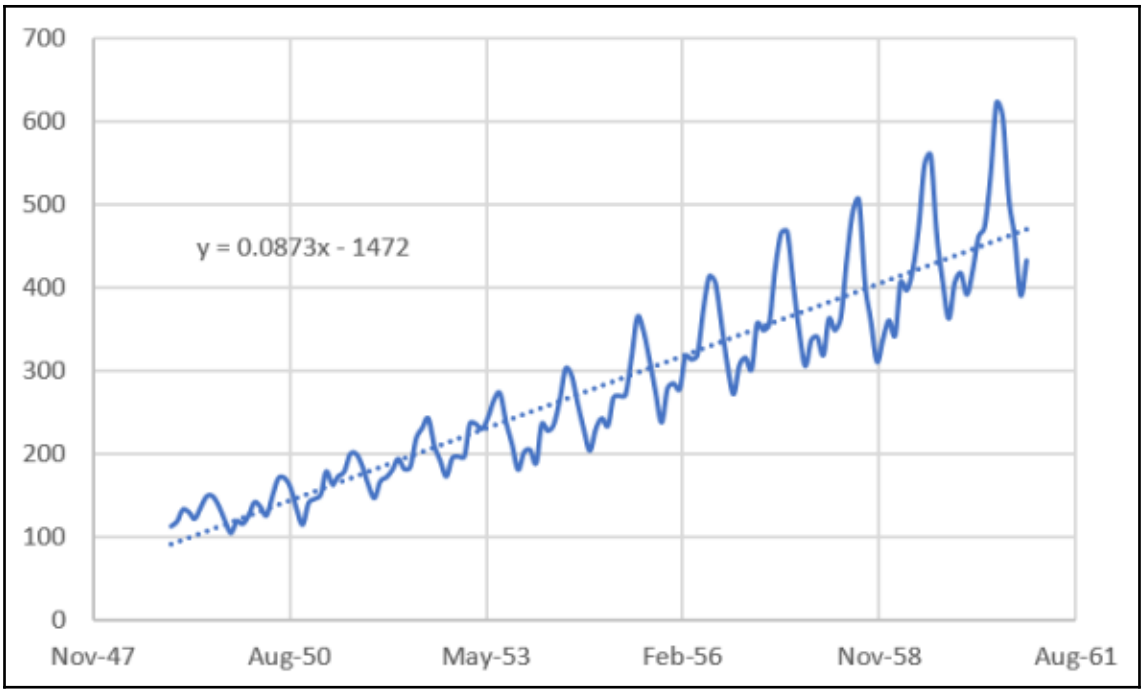
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Cohorts		Feb-15	Mar-15	Apr-15	May-15	Jun-15	Jul-15	Aug-15	Sep-15	Oct-15	Nov-15	Dec-15	Jan-16	Feb-16	Mar-16	Apr-16
2	Month	0	$ 16,532.50	$ 15,449.20	$ 12,329.30	$ 15,661.80	$ 15,636.40	$ 18,377.30	$ 15,741.20	$ 12,170.70	$ 23,526.00	$ 12,506.00	$ 16,640.30	$ 20,427.70	$ 16,234.40	$ 16,372.70	$ 20,611.20
3	Month	1	$ 16,532.50	$ 15,449.20	$ 12,329.30	$ 15,661.80	$ 15,636.40	$ 18,377.30	$ 15,741.20	$ 12,170.70	$ 23,526.00	$ 12,506.00	$ 16,640.30	$ 20,427.70	$ 16,234.40	$ 16,372.70	$ 20,611.20
4	Month	2	$ 16,532.50	$ 15,449.20	$ 12,329.30	$ 15,661.80	$ 15,636.40	$ 18,377.30	$ 15,741.20	$ 12,170.70	$ 23,526.00	$ 12,506.00	$ 16,640.30	$ 20,427.70	$ 16,234.40	$ 16,372.70	$ 20,611.20
5	Month	3	$ 16,532.50	$ 15,449.20	$ 12,329.30	$ 15,661.80	$ 15,636.40	$ 18,377.30	$ 15,741.20	$ 12,170.70	$ 23,526.00	$ 12,506.00	$ 16,640.30	$ 20,427.70	$ 16,234.40	$ 16,372.70	$ 20,611.20
6	Month	4	$ 16,532.50	$ 15,141.50	$ 12,329.30	$ 14,303.40	$ 15,251.30	$ 17,513.50	$ 15,741.20	$ 11,423.00	$ 22,145.60	$ 12,506.00	$ 14,828.30	$ 17,951.90	$ 15,095.30	$ 15,537.60	$ 20,448.80
7	Month	5	$ 13,906.20	$ 14,800.50	$ 11,539.90	$ 14,016.00	$ 12,388.50	$ 16,211.30	$ 15,741.20	$ 11,423.00	$ 19,253.10	$ 12,506.00	$ 12,197.10	$ 17,287.40	$ 14,155.70	$ 14,655.00	$ 18,635.00
8	Month	6	$ 13,906.20	$ 14,547.90	$ 11,324.70	$ 12,258.00	$ 11,636.40	$ 15,358.30	$ 14,443.30	$ 11,423.00	$ 18,620.90	$ 11,258.40	$ 11,849.80	$ 16,880.40	$ 13,506.00	$ 14,655.00	$ 16,609.50
9	Month	7	$ 12,925.10	$ 13,140.20	$ 10,741.40	$ 12,258.00	$ 10,891.80	$ 14,832.60	$ 13,439.40	$ 11,423.00	$ 17,542.30	$ 10,561.20	$ 11,214.20	$ 15,955.80	$ 13,158.60	$ 13,861.70	$ 16,609.50
10	Month	8	$ 11,316.80	$ 12,350.80	$ 9,872.80	$ 12,258.00	$ 9,568.50	$ 14,128.10	$ 13,439.40	$ 11,423.00	$ 15,692.00	$ 9,470.90	$ 10,339.20	$ 15,726.40	$ 11,471.80	$ 13,861.70	$ 15,275.40
11	Month	9	$ 9,637.10	$ 11,109.80	$ 9,287.60	$ 11,365.70	$ 8,672.80	$ 13,215.40	$ 12,881.50	$ 9,942.00	$ 14,364.50	$ 8,877.80	$ 10,339.20	$ 15,726.40	$ 10,474.40	$ 12,548.10	$ 13,973.30
12	Month	10	$ 9,637.10	$ 10,538.00	$ 8,551.70	$ 10,950.00	$ 8,530.50	$ 12,218.90	$ 11,263.80	$ 9,706.20	$ 12,764.50	$ 7,603.30	$ 8,677.70	$ 15,012.60	$ 9,579.80	$ 11,088.10	$ 13,011.20
13	Month	11	$ 9,637.10	$ 10,330.70	$ 7,732.60	$ 10,513.70	$ 6,277.00	$ 10,488.20	$ 10,304.20	$ 8,977.70	$ 11,824.70	$ 7,351.60	$ 8,084.60	$ 15,012.60	$ 9,166.50	$ 8,347.30	$ 12,469.40
14	Month	12	$ 8,353.70	$ 9,052.70	$ 7,732.60	$ 9,462.90	$ 6,277.00	$ 10,488.20	$ 9,375.40	$ 8,977.70	$ 11,224.70	$ 7,031.70	$ 5,090.80	$ 12,839.00	$ 8,492.40	$ 8,347.30	$ 12,469.40
15	Month	13	$ 8,353.70	$ 7,718.00	$ 6,625.40	$ 7,981.00	$ 6,277.00	$ 9,825.60	$ 9,194.60	$ 7,279.20	$ 10,791.30	$ 6,462.80	$ 4,152.20	$ 11,241.80	$ 7,329.00	$ 7,963.10	$ 12,242.00
16	Month	14	$ 7,996.60	$ 7,031.80	$ 6,011.70	$ 7,058.50	$ 6,277.00	$ 9,538.50	$ 7,653.80	$ 5,735.90	$ 9,802.40	$ 6,462.80	$ 4,152.20	$ 11,241.80	$ 6,920.30	$ 7,090.20	$ 11,523.40
17	Month	15	$ 7,035.50	$ 6,609.90	$ 5,049.10	$ 5,952.30	$ 5,878.80	$ 8,738.80	$ 7,340.50	$ 5,020.70	$ 9,802.40	$ 5,134.50	$ 3,954.30	$ 9,477.40	$ 6,217.80	$ 7,090.20	$ 9,350.90
18	Month	16	$ 6,831.60	$ 6,609.90	$ 3,720.70	$ 5,124.20	$ 5,739.00	$ 7,940.40	$ 6,575.50	$ 4,438.70	$ 9,443.40	$ 4,508.20	$ 2,958.60	$ 7,752.50	$ 5,540.10	$ 4,971.70	$ 7,514.70
19	Month	17	$ 5,555.40	$ 6,609.90	$ 3,720.70	$ 4,520.30	$ 4,798.90	$ 6,544.40	$ 6,575.50	$ 3,601.90	$ 8,741.60	$ 4,344.10	$ 2,958.60	$ 6,729.20	$ 5,335.50	$ 4,971.70	$ 7,080.40
20	Month	18	$ 3,824.40	$ 6,609.90	$ 3,443.80	$ 3,452.50	$ 4,300.70	$ 6,312.30	$ 6,575.50	$ 3,601.90	$ 7,564.50	$ 3,617.10	$ 1,959.50	$ 6,431.60	$ 4,885.20	$ 4,052.80	$ 6,281.70
21	Month	19	$ 2,868.40	$ 6,476.90	$ 3,326.60	$ 2,111.00	$ 3,558.10	$ 5,186.10	$ 3,468.20	$ 2,343.20	$ 7,564.50	$ 3,181.00	$ 1,431.50	$ 5,537.40	$ 3,510.50	$ 4,052.80	$ 3,501.90
22	Month	20	$ 2,373.40	$ 5,175.10	$ 1,686.70	$ 1,633.00	$ 3,367.00	$ 5,186.10	$ 1,999.10	$ 2,343.20	$ 6,750.50	$ 2,872.90	$ 1,431.50	$ 4,717.20	$ 2,667.10	$ 3,262.60	$ 3,501.90

Chapter 7: Implementing Time Series

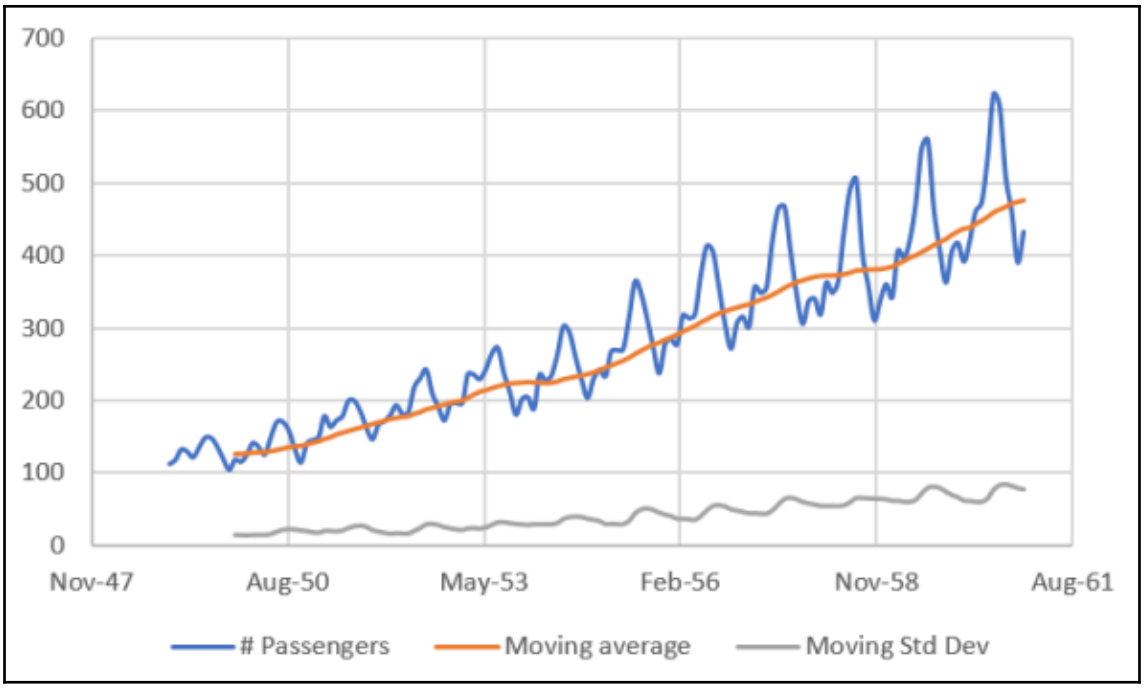


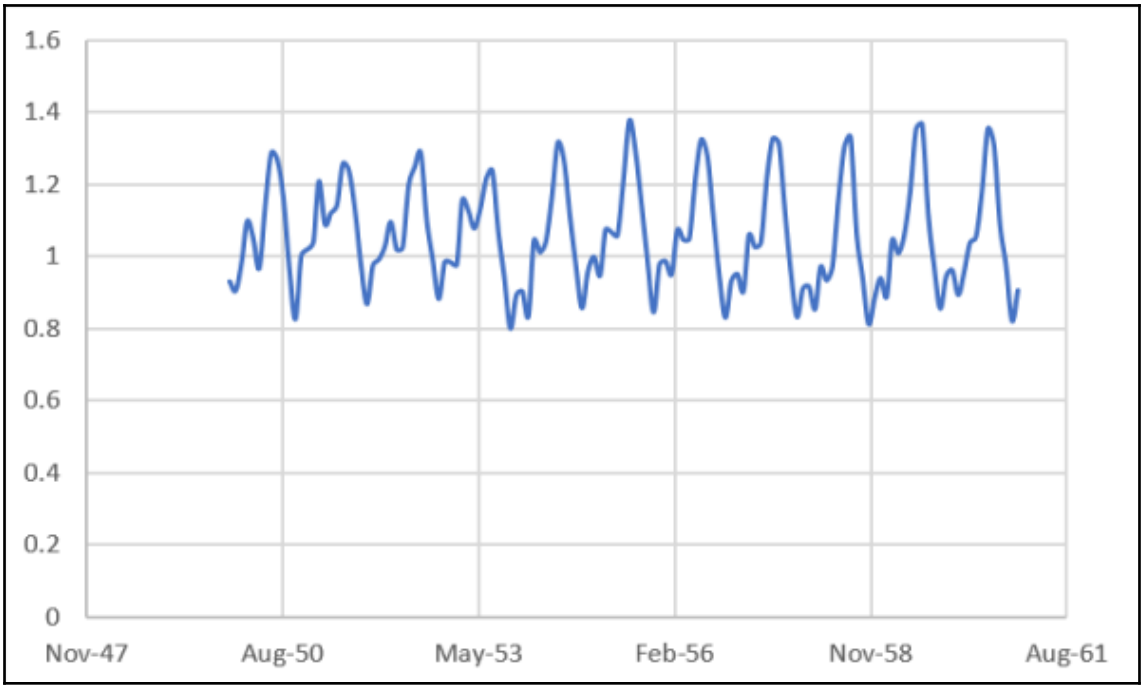


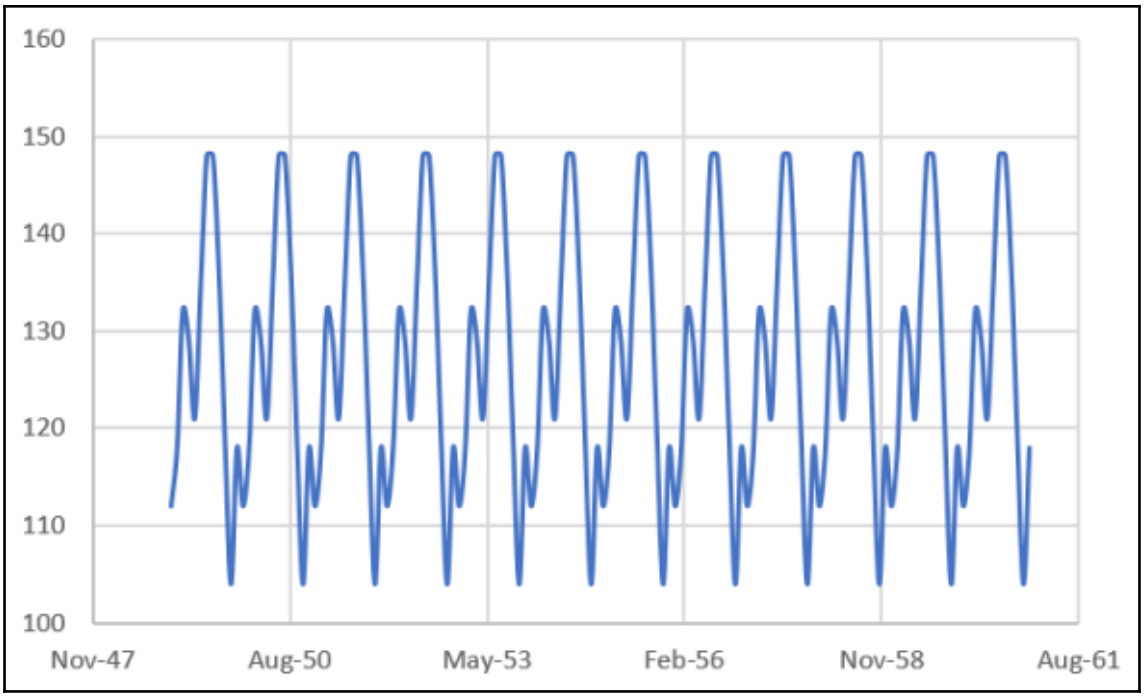

Format Trendline
Trendline Options
Linear
Logarithmic
Polynomial
Order
2
Power
Moving Average
Period
2
Trendline Name
Automatic
Linear (# Passengers)
Custom
Forecast
Forward
0.0
periods
Backward
0.0
periods
Set Intercept
0.0
Display Equation on chart
Display R-squared value on chart

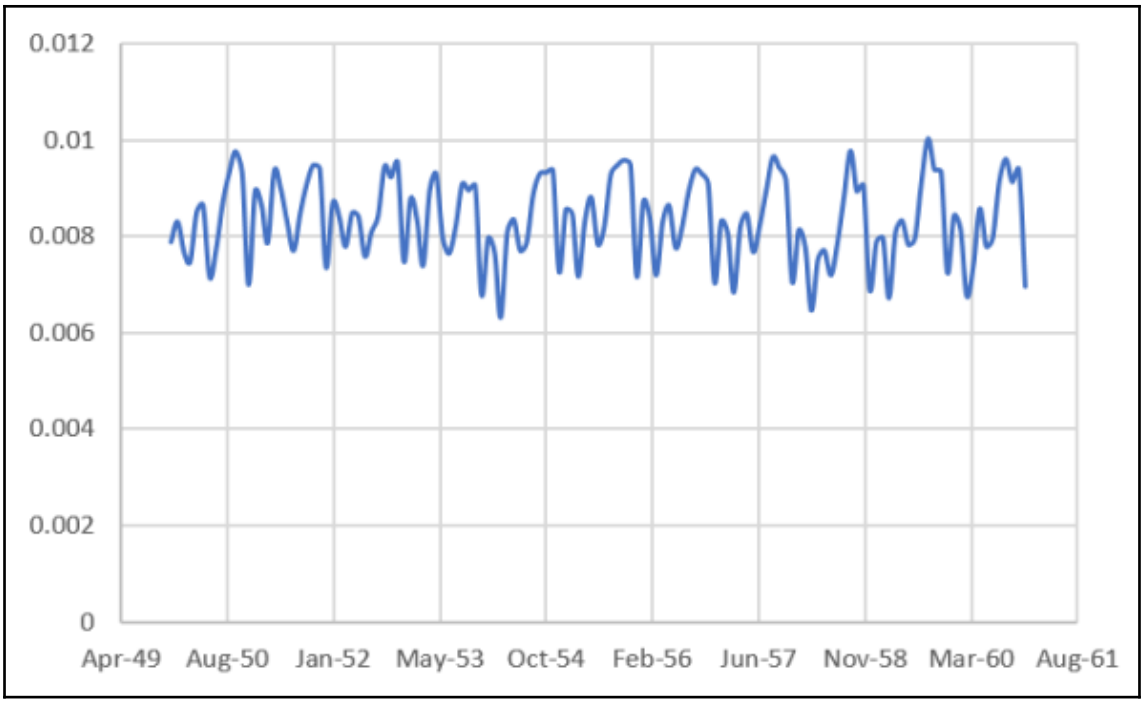

700
600
500
400
300
200
100
0
y = 0.0873x - 1472
Nov-47
Aug-50
May-53
Feb-56
Nov-58
Aug-61

TravelDate	# Passengers	Moving average	Moving Std Dev
Jan-49	112		
Feb-49	118		
Mar-49	132		
Apr-49	129		
May-49	121		
Jun-49	135		
Jul-49	148		
Aug-49	148		
Sep-49	136		
Oct-49	119		
Nov-49	104		
Dec-49	118	126.6666667	13.72014666
Jan-50	115	126.9166667	13.45334249
Feb-50	126	127.5833333	13.16647487
Mar-50	141	128.3333333	13.68697678
Apr-50	135	128.8333333	13.82246744
May-50	125	129.1666667	13.66370995
Jun-50	149	130.3333333	14.76071773

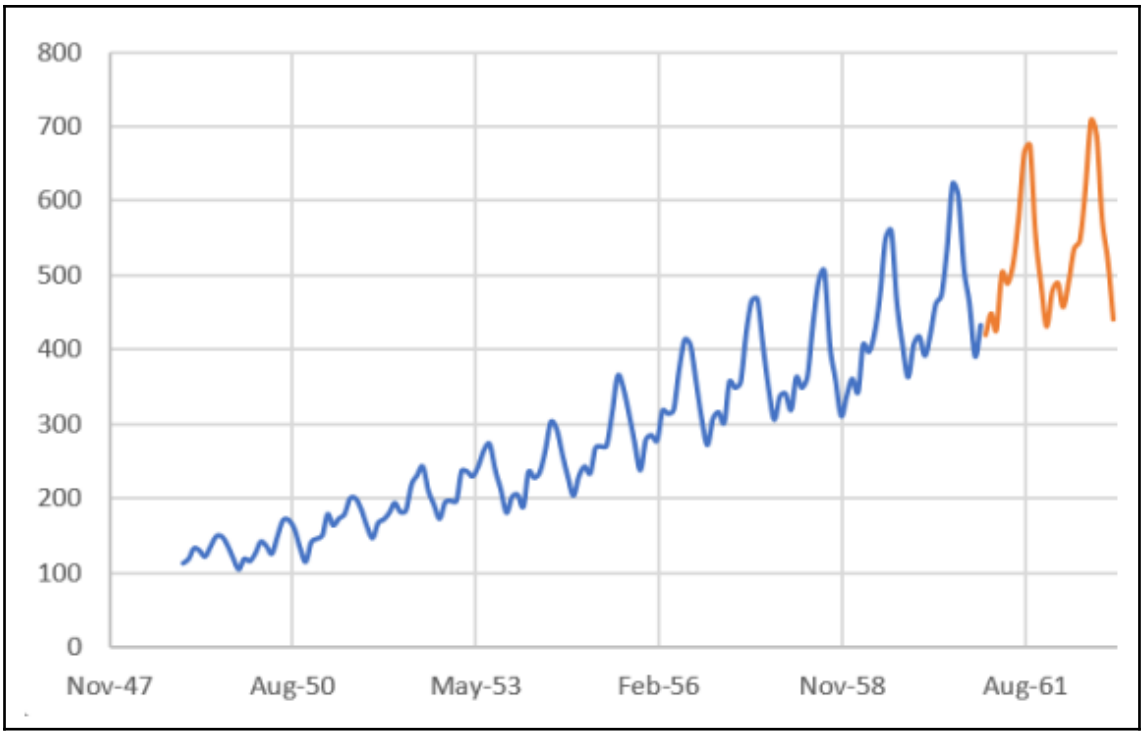

700
600
500
400
300
200
100
0
Nov-47
Aug-50
May-53
Feb-56
Nov-58
Aug-61
Passengers
Moving average
Moving Std Dev


1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Nov-47
Aug-50
May-53
Feb-56
Nov-58
Aug-61

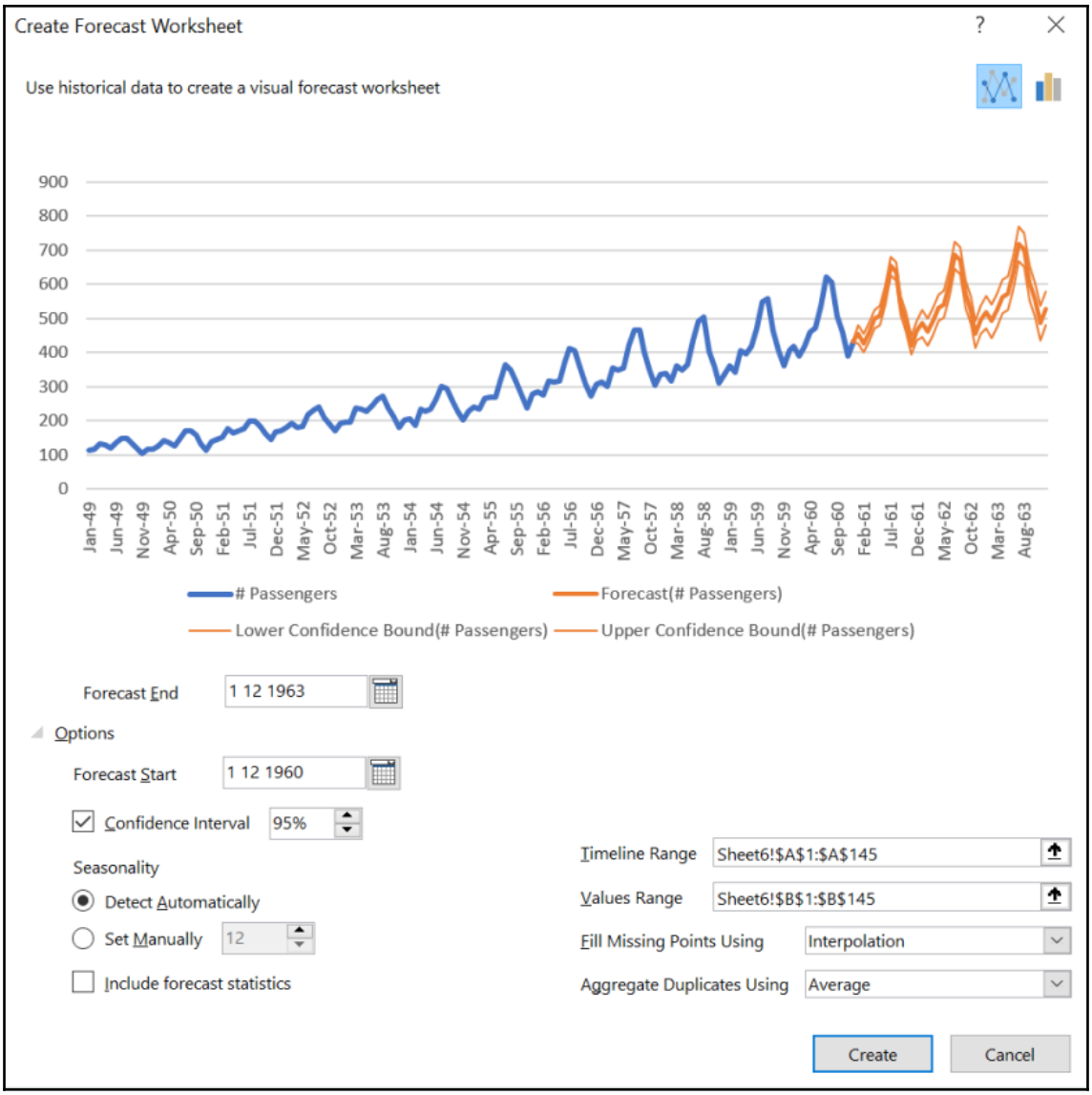

160
150
140
130
120
110
100
Nov-47
Aug-50
May-53
Feb-56
Nov-58
Aug-61


0.012
0.01
0.008
0.006
0.004
0.002
0
Apr-49
Aug-50
Jan-52
May-53
Oct-54
Feb-56
Jun-57
Nov-58
Mar-60
Aug-61

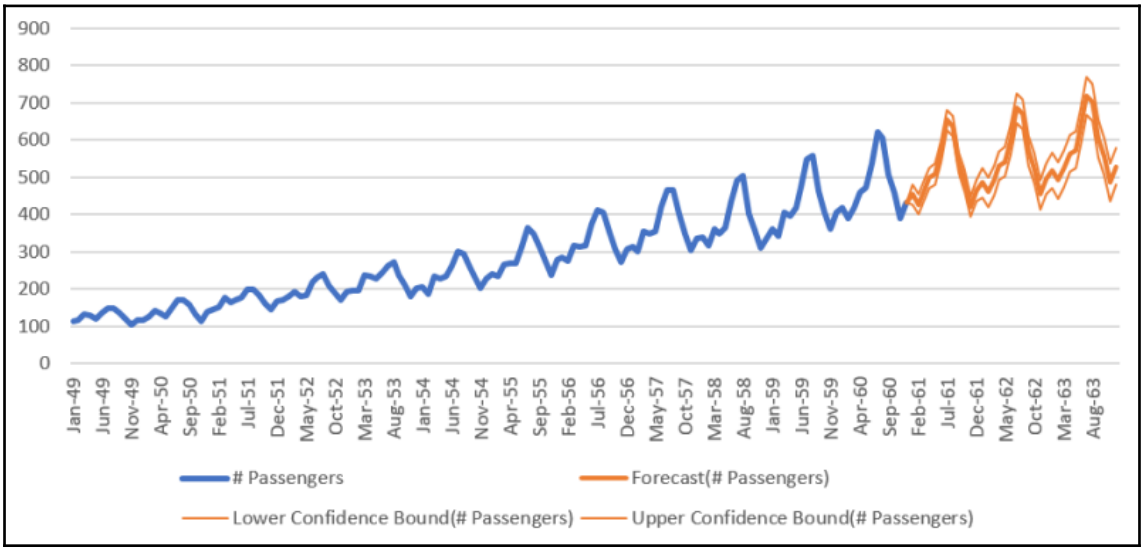
	A	B	C	D	E
144	Nov-60	390			
145	Dec-60	432			
146	Jan-61	418.5687	112	0.007897	473.22
147	Feb-61	447.7342	118	0.007973	475.92
148	Mar-61	425.3093	132	0.006735	478.37
149	Apr-61	502.9614	129	0.008105	481.08
150	May-61	488.2151	121	0.008342	483.69
151	Jun-61	514.4723	135	0.007835	486.40
152	Jul-61	576.8031	148	0.00797	489.02
153	Aug-61	665.485	148	0.009144	491.73
154	Sep-61	675.077	136	0.010039	494.43
155	Oct-61	555.4341	119	0.00939	497.05
156	Nov-61	486.2185	104	0.009355	499.76
157	Dec-61	430.2691	118	0.007258	502.38
158	Jan-62	477.5688	112	0.008442	505.08
159	Feb-62	488.9317	118	0.00816	507.79
160	Mar-62	456.3509	132	0.006776	510.23
161	Apr-62	490.4094	129	0.007411	512.94
162	May-62	535.7012	121	0.008587	515.56
163	Jun-62	546.0296	135	0.007804	518.27
164	Jul-62	614.8332	148	0.007975	520.88
165	Aug-62	708.8846	148	0.009148	523.59
166	Sep-62	688.351	136	0.009617	526.30
167	Oct-62	575.2492	119	0.009139	528.92
168	Nov-62	519.6916	104	0.0094	531.62
169	Dec-62	439.6429	118	0.006974	534.24


800
700
600
500
400
300
200
100
0
Nov-47
Aug-50
May-53
Feb-56
Nov-58
Aug-61


What-If Analysis
Forecast Sheet
Forecast
Group
Ungroup
Subtotal
Outline

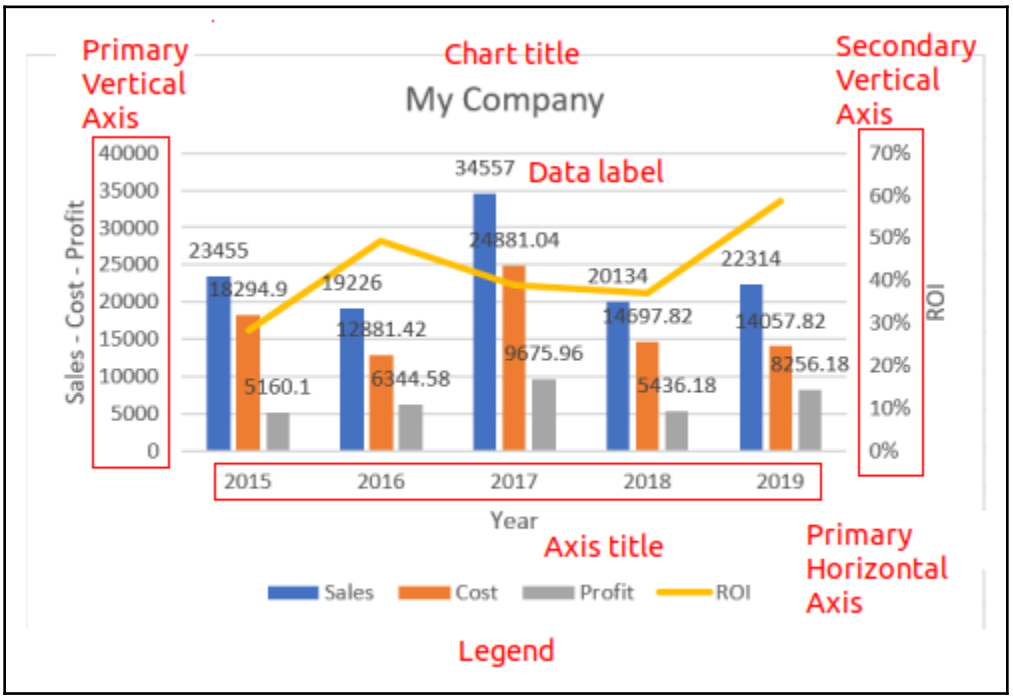

Create Forecast Worksheet
Use historical data to create a visual forecast worksheet
900
800
700
600
500
400
300
200
100
0
Jan-49
Jun-49
Nov-49
Apr-50
Sep-50
Feb-51
Jul-51
Dec-51
May-52
Oct-52
Mar-53
Aug-53
Jan-54
Jun-54
Nov-54
Apr-55
Sep-55
Feb-56
Jul-56
Dec-56
May-57
Oct-57
Mar-58
Aug-58
Jan-59
Jun-59
Nov-59
Apr-60
Sep-60
Feb-61
Jul-61
Dec-61
May-62
Oct-62
Mar-63
Aug-63
# Passengers
Forecast(# Passengers)
Lower Confidence Bound(# Passengers)
Upper Confidence Bound(# Passengers)
Forecast End
1 12 1963
Options
Forecast Start
1 12 1960
Confidence Interval
95%
Seasonality
Detect Automatically
Set Manually
12
Include forecast statistics
Timeline Range
Sheet6!$A$1:$A$145
Values Range
Sheet6!$B$1:$B$145
Fill Missing Points Using
Interpolation
Aggregate Duplicates Using
Average
Create
Cancel

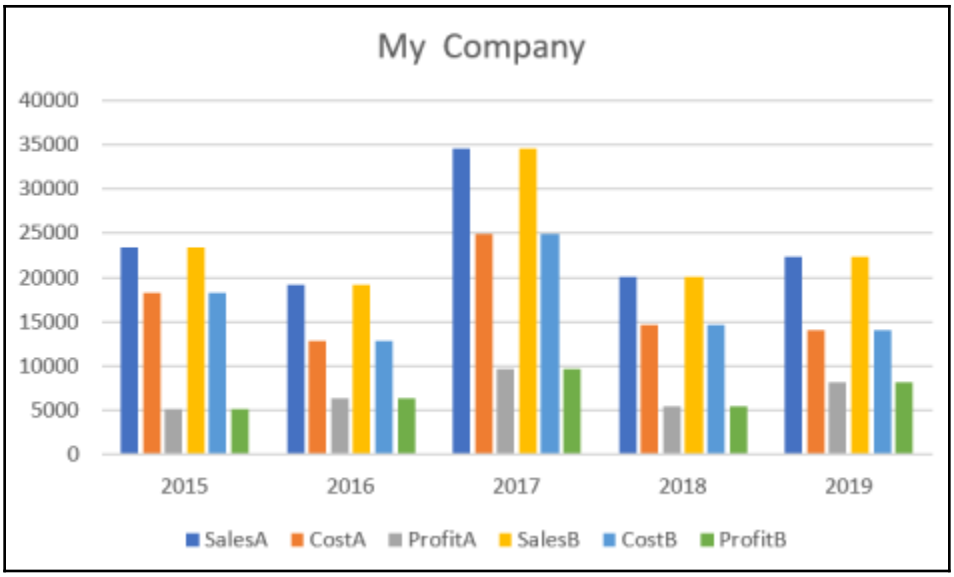
	TravelDate	# Passengers	Forecast(# Passengers)	Lower Confidence Bound(# Passengers)	Upper Confidence Bound(# Passengers)
142	Sep-60	508			
143	Oct-60	461			
144	Nov-60	390			
145	Dec-60	432	432	432.00	432.00
146	Jan-61		453.8122736	427.08	480.54
147	Feb-61		427.8263121	400.96	454.69
148	Mar-61		459.262672	432.26	486.27
149	Apr-61		498.909199	471.76	526.05
150	May-61		508.8944168	481.61	536.18
151	Jun-61		569.8482444	542.42	597.28
152	Jul-61		653.6777865	626.10	681.25
153	Aug-61		637.3598214	609.63	665.08
154	Sep-61		539.05814	511.18	566.93
155	Oct-61		490.8222267	462.80	518.85
156	Nov-61		421.258682	393.08	449.44
157	Dec-61		464.425986	436.09	492.76
158	Jan-62		486.2264429	447.07	525.38
159	Feb-62		460.2404815	420.97	499.51
160	Mar-62		491.6768414	452.28	531.07
161	Apr-62		531.3233684	491.81	570.84
162	May-62		541.3085862	501.68	580.94
163	Jun-62		602.2624138	562.51	642.02

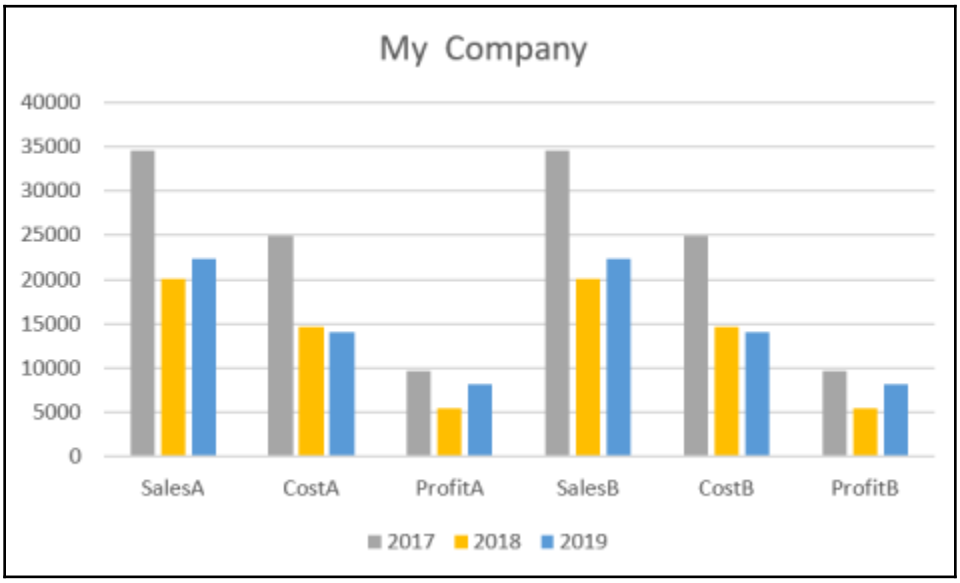


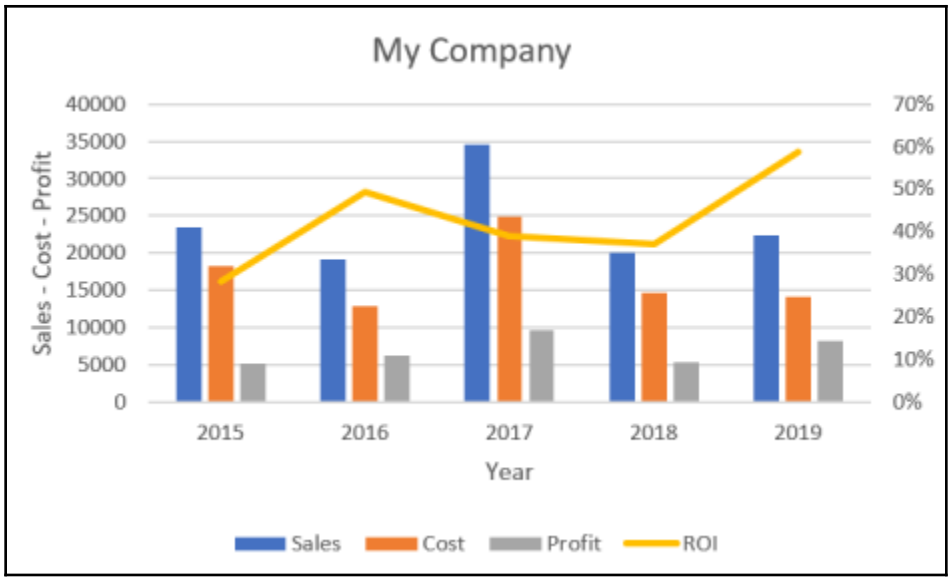
Statistic	Value
Alpha	0.10
Beta	0.00
Gamma	0.90
MASE	0.69
SMAPE	0.03
MAE	14.58
RMSE	17.07

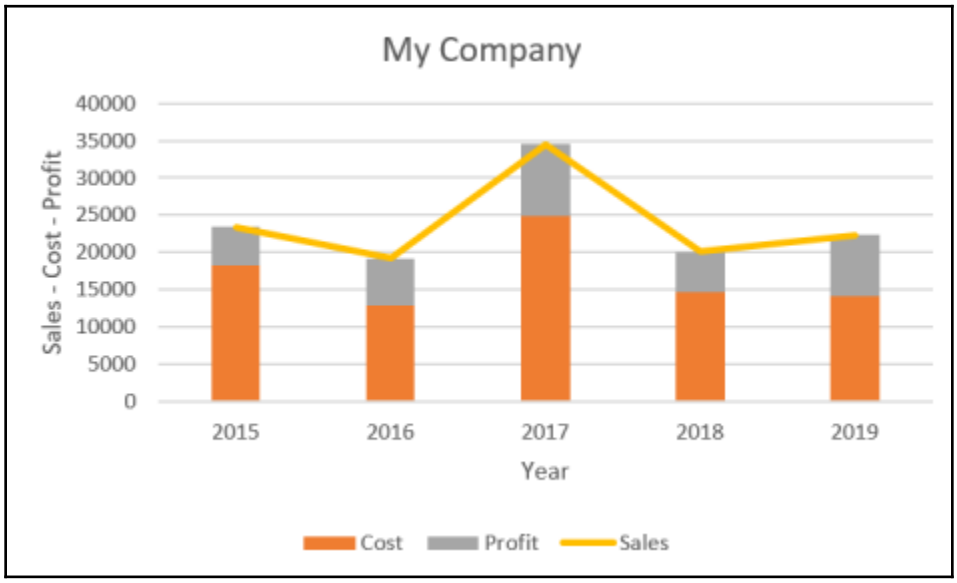
Chapter 8: Visualizing Data in Diagrams, Histograms, and Maps

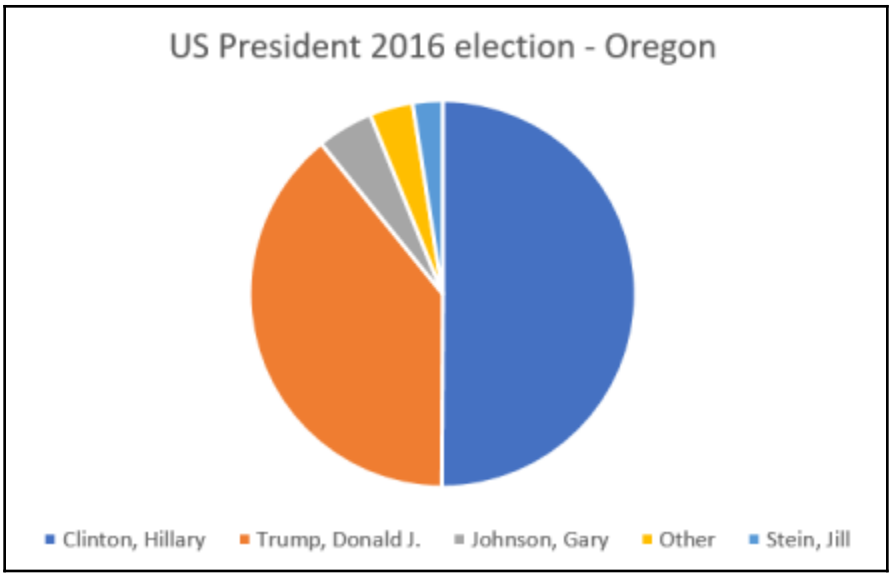


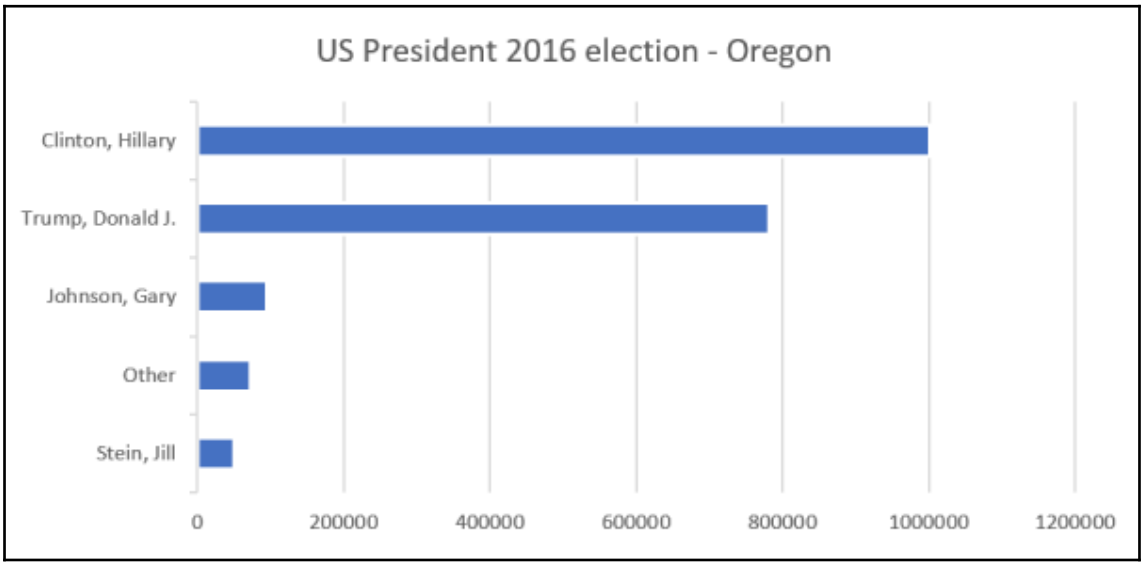

My Company
40000
35000
30000
25000
20000
15000
10000
5000
0
2015
2016
2017
2018
2019
SalesA
CostA
ProfitA
SalesB
CostB
ProfitB

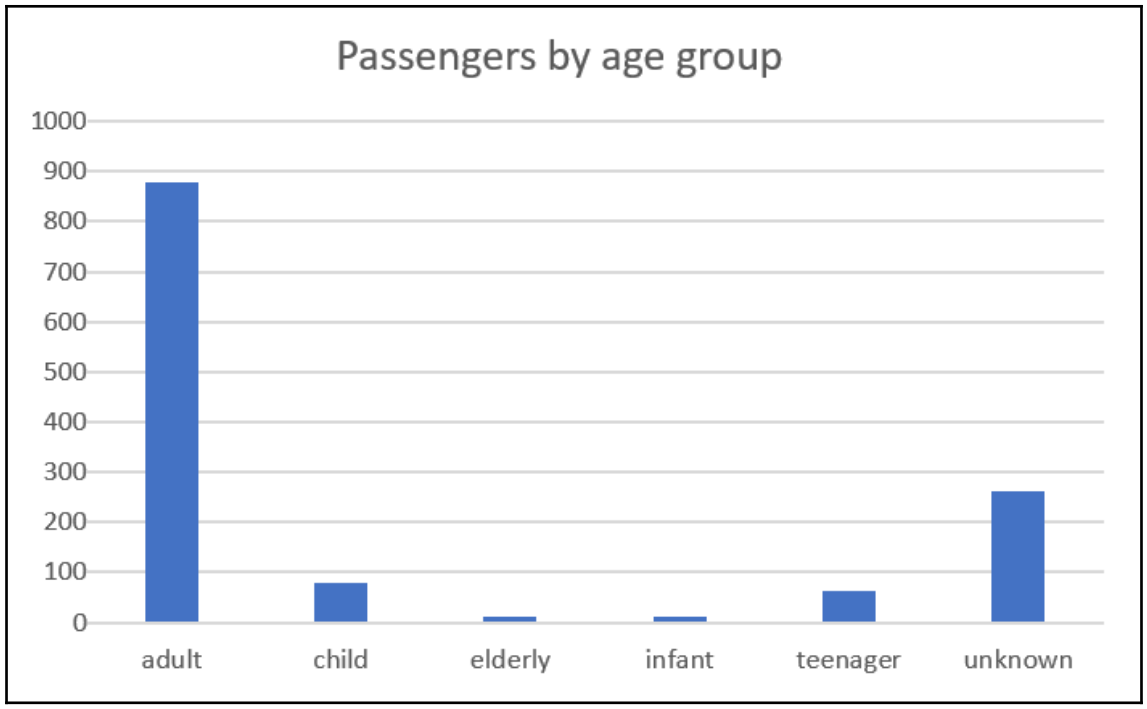

My Company
40000
35000
30000
25000
20000
15000
10000
5000
0
SalesA
CostA
ProfitA
SalesB
CostB
ProfitB
2017
2018
2019

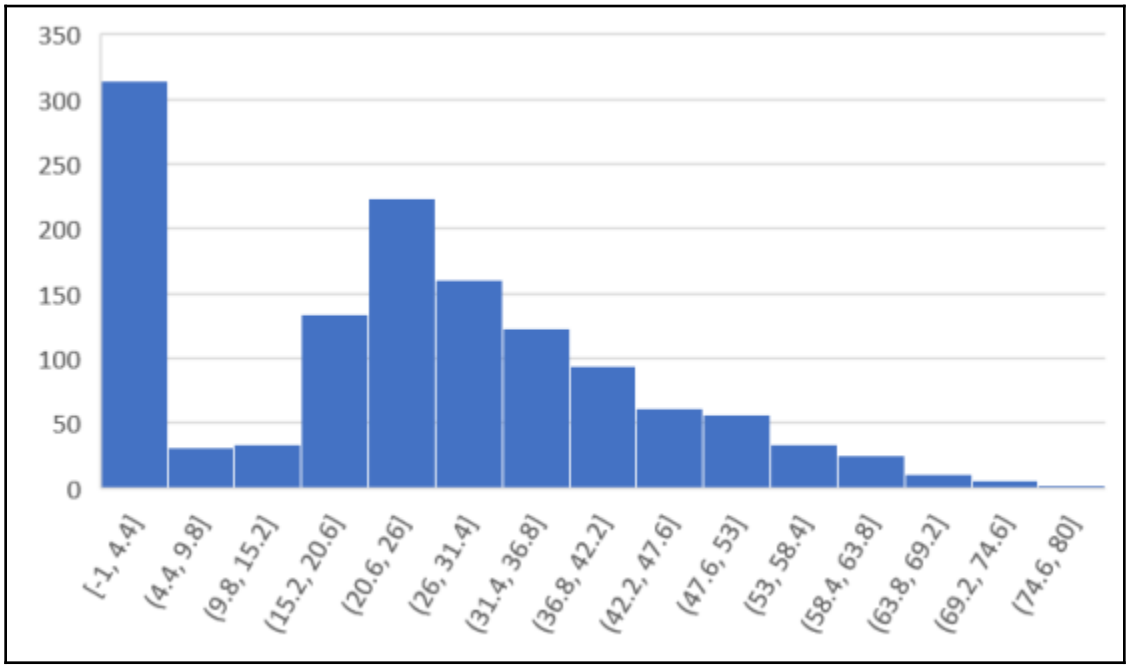

My Company
Sales - Cost - Profit
40000
35000
30000
25000
20000
15000
10000
5000
0
70%
60%
50%
40%
30%
20%
10%
0%
2015
2016
2017
2018
2019
Year
Sales
Cost
Profit
ROI

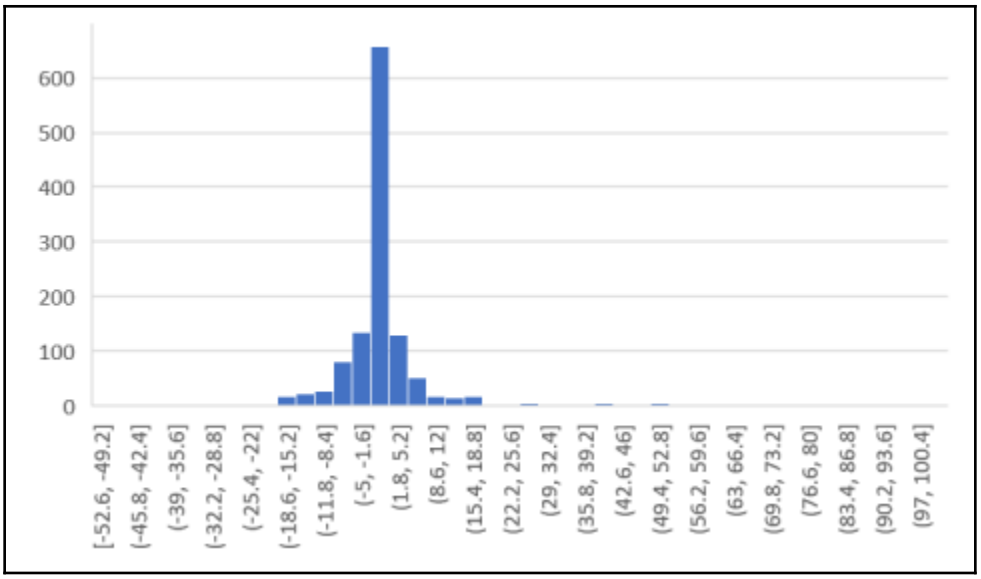

My Company
Sales - Cost - Profit
40000
35000
30000
25000
20000
15000
10000
5000
0
2015
2016
2017
2018
2019
Year
Cost
Profit
Sales


US President 2016 election - Oregon
Clinton, Hillary
Trump, Donald J.
Johnson, Gary
Other
Stein, Jill

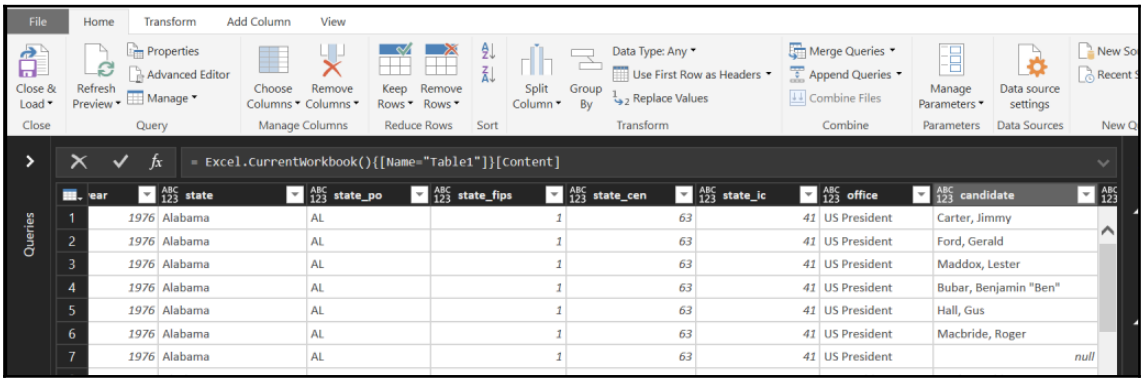

US President 2016 election - Oregon
Clinton, Hillary
Trump, Donald J.
Johnson, Gary
Other
Stein, Jill
0
200000
400000
600000
800000
1000000
1200000


Passengers by age group
1000
900
800
700
600
500
400
300
200
100
0
adult
child
elderly
infant
teenager
unknown

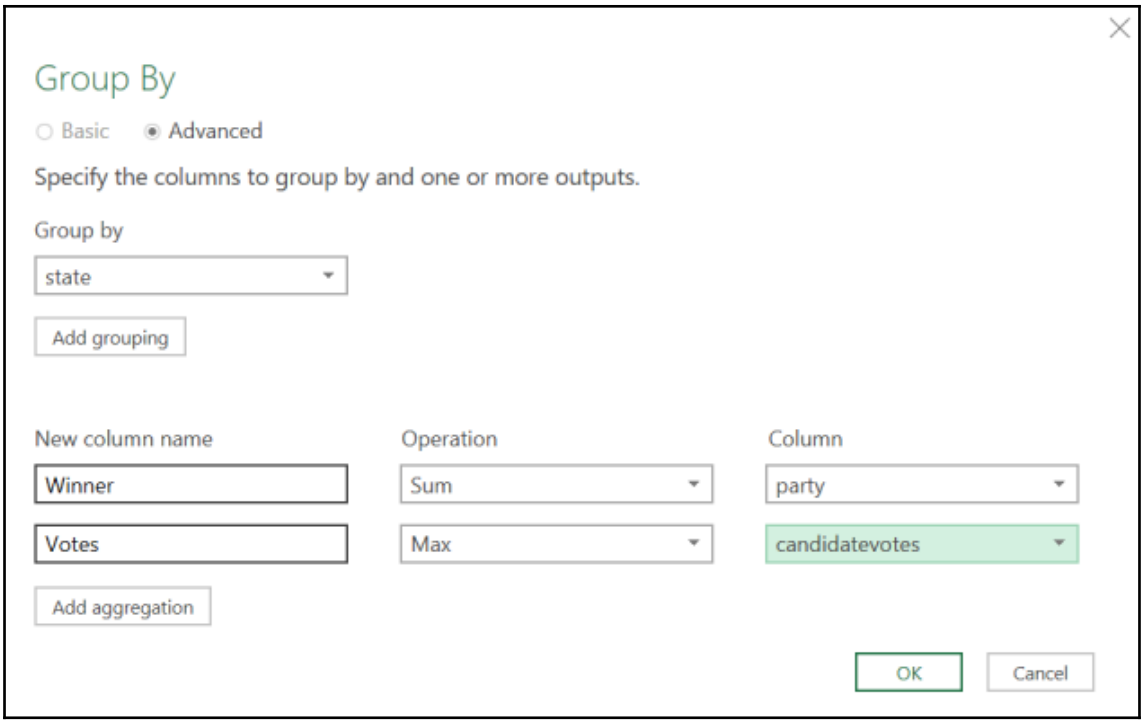

350
300
250
200
150
100
50
0
[-1, 4.4]
(4.4, 9.8]
(9.8, 15.2]
(15.2, 20.6]
(20.6, 26]
(26, 31.4]
(31.4, 36.8]
(36.8, 42.2]
(42.2, 47.6]
(47.6, 53]
(53, 58.4]
(58.4, 63.8]
(63.8, 69.2]
(69.2, 74.6]
(74.6, 80]

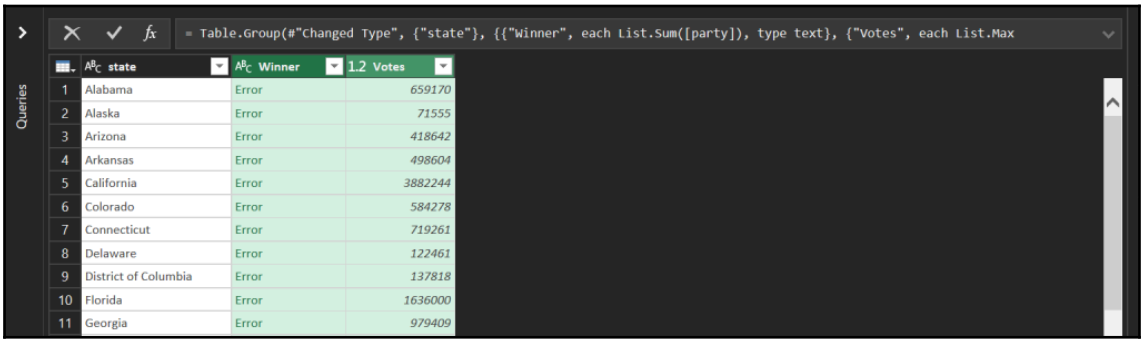


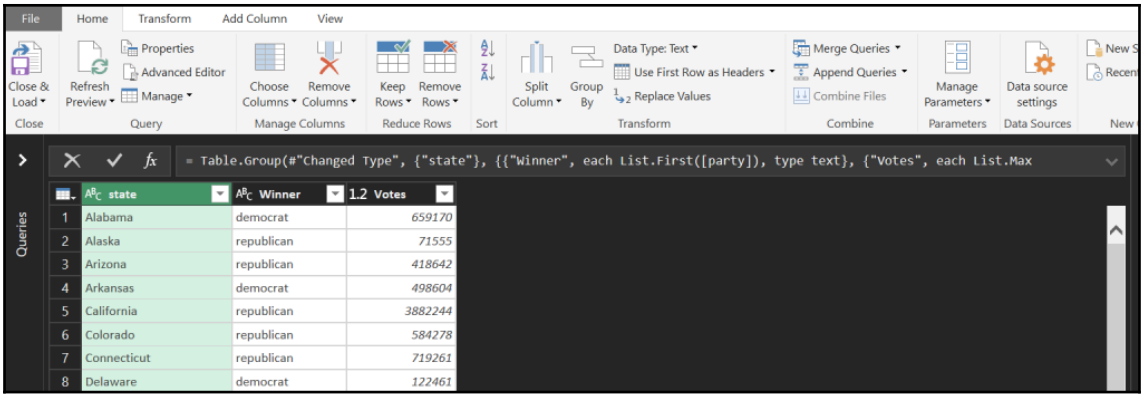
year	state	state_po	state_fips	state_cen	state_ic	office	candidate	party	writein	candidatevotes	totalvotes	version
1976	Alabama	AL	1	63	41	US President	Carter, Jimmy	democrat	FALSE	659170	1182850	20171015
1976	Alabama	AL	1	63	41	US President	Ford, Gerald	republican	FALSE	504070	1182850	20171015
1976	Alabama	AL	1	63	41	US President	Maddox, Lester	american independent party	FALSE	9198	1182850	20171015
1976	Alabama	AL	1	63	41	US President	Bubar, Benjamin "Ben"	prohibition	FALSE	6669	1182850	20171015
1976	Alabama	AL	1	63	41	US President	Hall, Gus	communist party use	FALSE	1954	1182850	20171015
1976	Alabama	AL	1	63	41	US President	Macbride, Roger	libertarian	FALSE	1481	1182850	20171015
1976	Alabama	AL	1	63	41	US President			TRUE	308	1182850	20171015
1976	Alaska	AK	2	94	81	US President	Ford, Gerald	republican	FALSE	71555	123574	20171015
1976	Alaska	AK	2	94	81	US President	Carter, Jimmy	democrat	FALSE	44058	123574	20171015
1976	Alaska	AK	2	94	81	US President	Macbride, Roger	libertarian	FALSE	6785	123574	20171015
1976	Alaska	AK	2	94	81	US President			TRUE	1176	123574	20171015
1976	Arizona	AZ	4	86	61	US President	Ford, Gerald	republican	FALSE	418642	742719	20171015
1976	Arizona	AZ	4	86	61	US President	Carter, Jimmy	democrat	FALSE	295602	742719	20171015



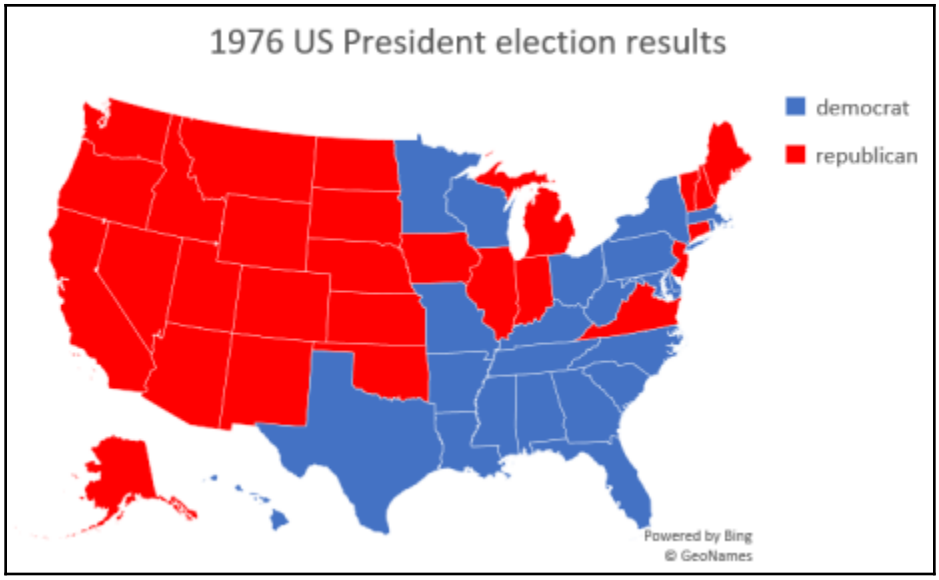
	ear	state	state_po	state_fips	state_cen	state_ic	office	candidate
1	1976	Alabama	AL	1	63	41	US President	Carter, Jimmy
2	1976	Alabama	AL	1	63	41	US President	Ford, Gerald
3	1976	Alabama	AL	1	63	41	US President	Maddox, Lester
4	1976	Alabama	AL	1	63	41	US President	Bubar, Benjamin "Ben"
5	1976	Alabama	AL	1	63	41	US President	Hall, Gus
6	1976	Alabama	AL	1	63	41	US President	Macbride, Roger
7	1976	Alabama	AL	1	63	41	US President	null


Group By
Basic
Advanced
Specify the columns to group by and one or more outputs.
Group by
state
Add grouping
New column name
Operation
Column
Winner
Sum
party
Votes
Max
candidatevotes
Add aggregation
OK
Cancel


= Table.Group(#"Changed Type", {"state"}, {{"Winner", each List.Sum([party]), type text}, {"Votes", each List.Max
Queries
state
Winner
Votes
1 Alabama Error 659170
2 Alaska Error 71555
3 Arizona Error 418642
4 Arkansas Error 498604
5 California Error 3882244
6 Colorado Error 584278
7 Connecticut Error 719261
8 Delaware Error 122461
9 District of Columbia Error 137818
10 Florida Error 1636000
11 Georgia Error 979409

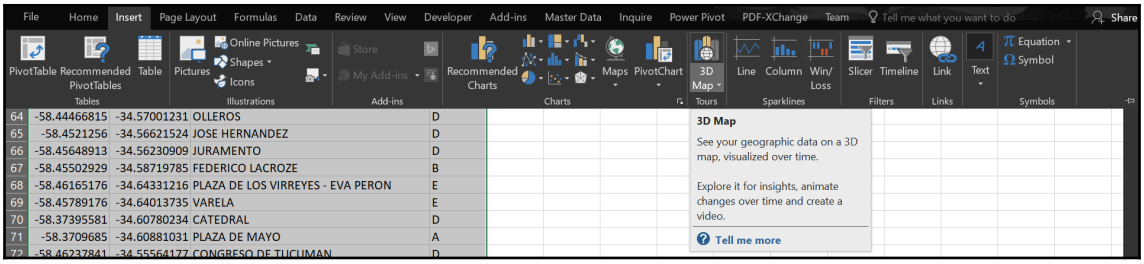


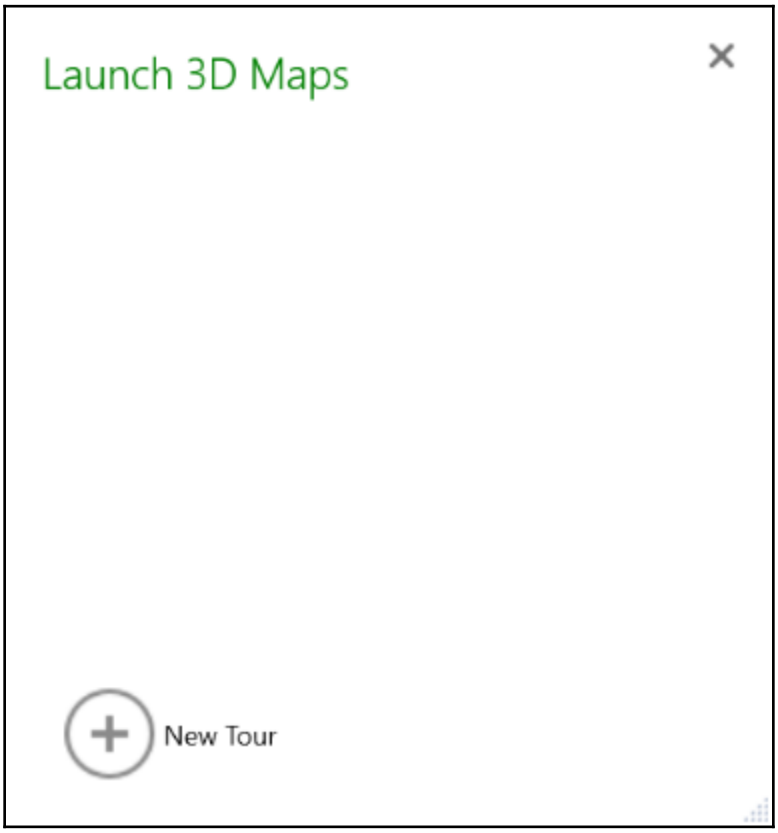
	state	Winner	Votes
1	Alabama	democrat	659170
2	Alaska	republican	71555
3	Arizona	republican	418642
4	Arkansas	democrat	498604
5	California	republican	3882244
6	Colorado	republican	584278
7	Connecticut	republican	719261
8	Delaware	democrat	122461

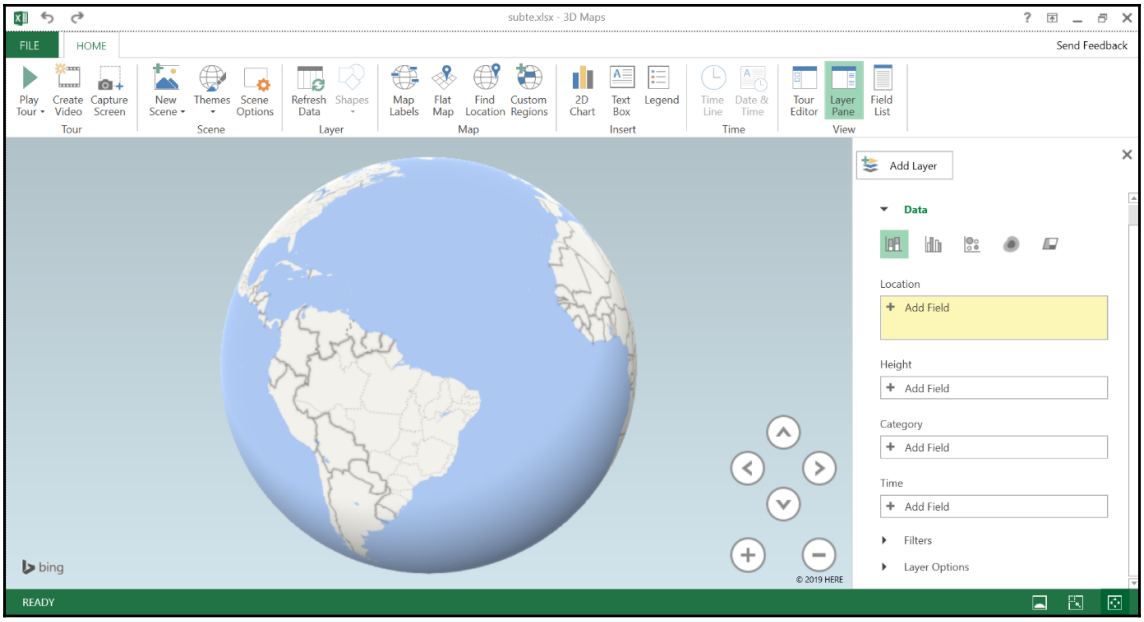


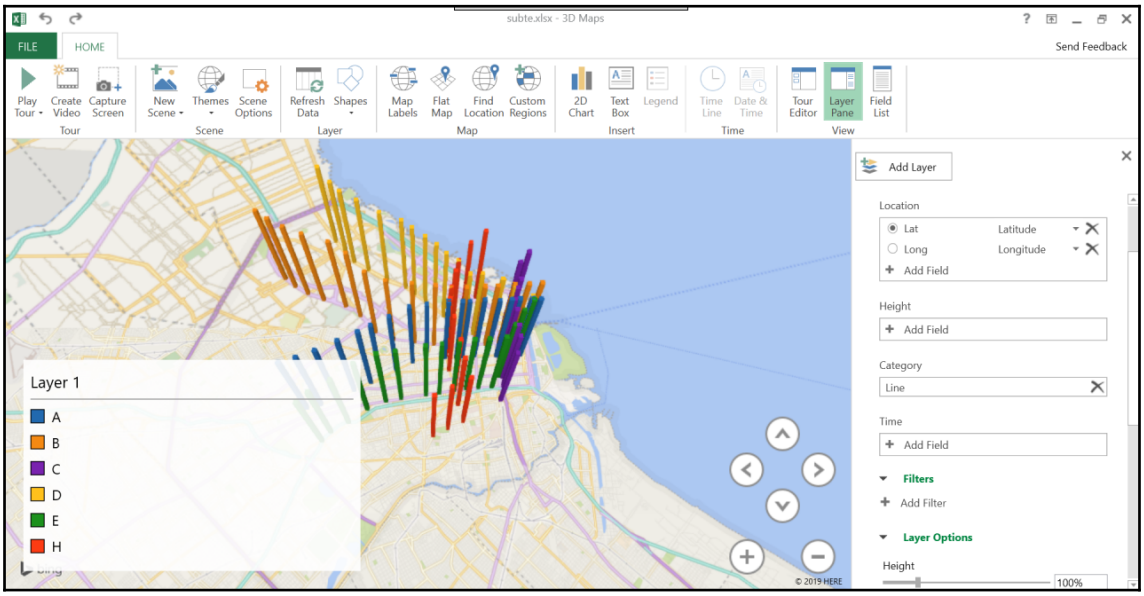
	A	B	C	D
1	Long	Lat	Station	Line
2	-58.39892759	-34.63575018	CASEROS	H
3	-58.40096956	-34.62937566	INCLAN - MEZQUITA AL AHMAD	H
4	-58.40232273	-34.62309232	HUMBERTO 1°	H
5	-58.40473172	-34.61524215	VENEZUELA	H
6	-58.40603638	-34.60893524	ONCE - 30 DE DICIEMBRE	H
7	-58.38057434	-34.6042452	9 DE JULIO	D
8	-58.39792376	-34.59975708	FACULTAD DE MEDICINA	D
9	-58.38514236	-34.60158717	TRIBUNALES - TEATRO COLÓN	D
10	-58.40716132	-34.59162784	AGÜERO	D
11	-58.41595542	-34.58515594	R.SCALABRINI ORTIZ	D
12	-58.42119601	-34.58141119	PLAZA ITALIA	D
13	-58.42571144	-34.57842202	PALERMO	D
14	-58.37401822	-34.59119381	RETIRO	C
15	-58.37815578	-34.60176992	LAVALLE	C
16	-58.37952998	-34.60484374	DIAGONAL NORTE	C
17	-58.38061072	-34.60898331	AV. DE MAYO	C
18	-58.38044447	-34.61261728	MORENO	C
19	-58.38017361	-34.6181256	INDEPENDENCIA	C
20	-58.38143443	-34.62761945	CONSTITUCION	C
21	-58.37507152	-34.60329729	FLORIDA	B
22	-58.38071485	-34.60363711	C. PELLEGRINI	B
23	-58.38729613	-34.60409355	URUGUAY	B
24	-58.39231424	-34.60441954	CALLAO - MAESTRO ALFREDO BRAVO	B
25	-58.39947426	-34.60464297	PASTEUR - AMIA	B
26	-58.40539944	-34.60458106	PUEYRREDON	B
27	-58.4117626	-34.60407952	CARLOS GARDEL	B

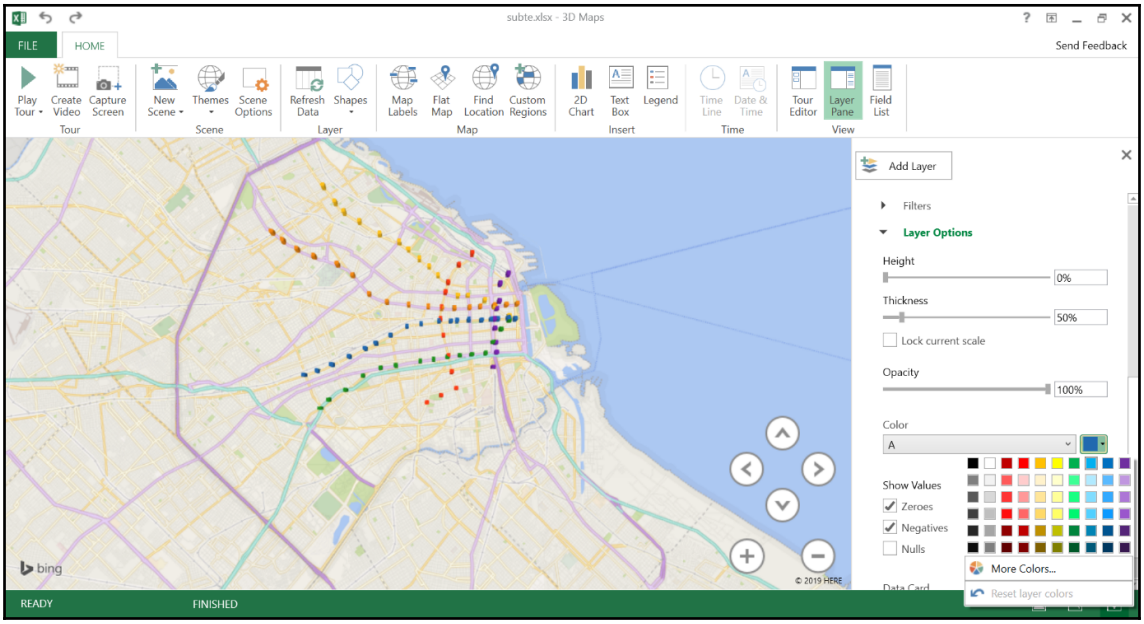
Underground Stations

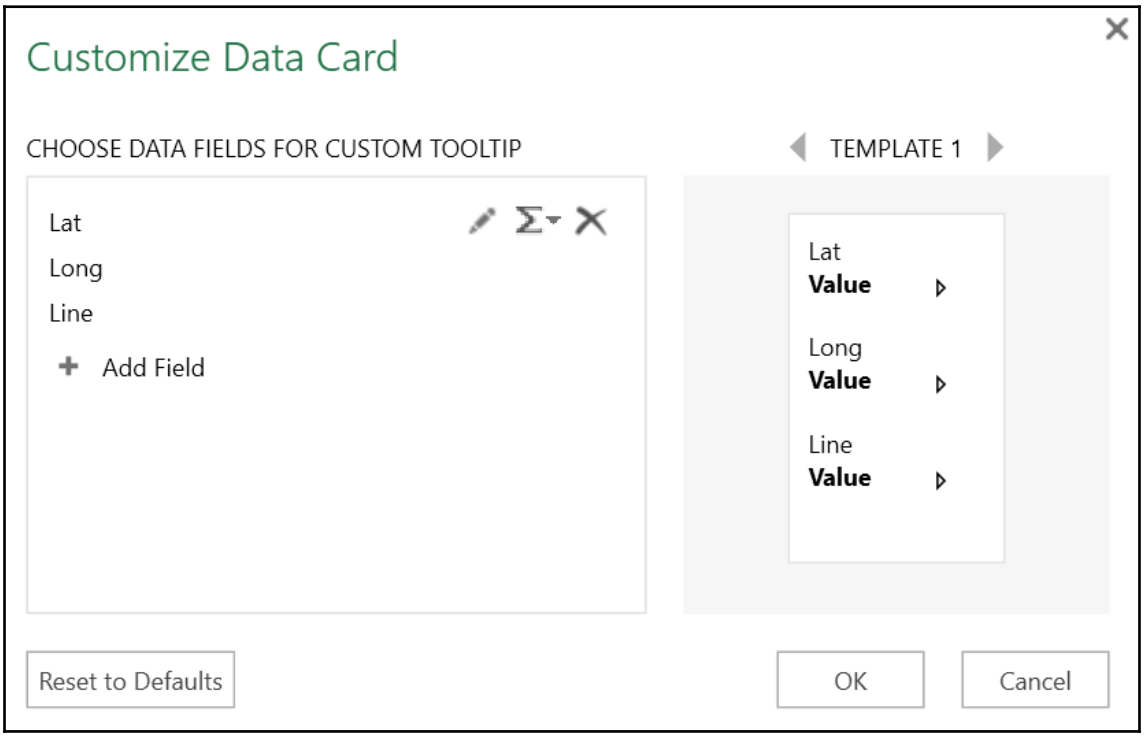

File
Home
Insert
Page Layout
Formulas
Data
Review
View
Developer
Add-ins
Master Data
Inquire
Power Pivot
PDF-XChange
Team
Tell me what you want to do
Share
PivotTable
Recommended PivotTables
Table
Tables
Pictures
Online Pictures
Shapes
Icons
Illustrations
Store
My Add-ins
Add-ins
Recommended Charts
Maps
PivotChart
Charts
3D Map
Tours
Line
Column
Win/ Loss
Sparklines
Slicer
Timeline
Filters
Link
Links
Text
Equation
Symbol
Symbols
64 -58.44466815 -34.57001231 OLLEROS D
65 -58.4521256 -34.56621524 JOSE HERNANDEZ D
66 -58.45648913 -34.56230909 JURAMENTO D
67 -58.45502929 -34.58719785 FEDERICO LACROZE B
68 -58.46165176 -34.64331216 PLAZA DE LOS VIRREYES - EVA PERON E
69 -58.45789176 -34.64013735 VARELA E
70 -58.37395581 -34.60780234 CATEDRAL D
71 -58.3709685 -34.60881031 PLAZA DE MAYO A
72 -58.46237841 -34.55564177 CONGRESO DE TUCUMAN D
3D Map
See your geographic data on a 3D map, visualized over time.
Explore it for insights, animate changes over time and create a video.
Tell me more

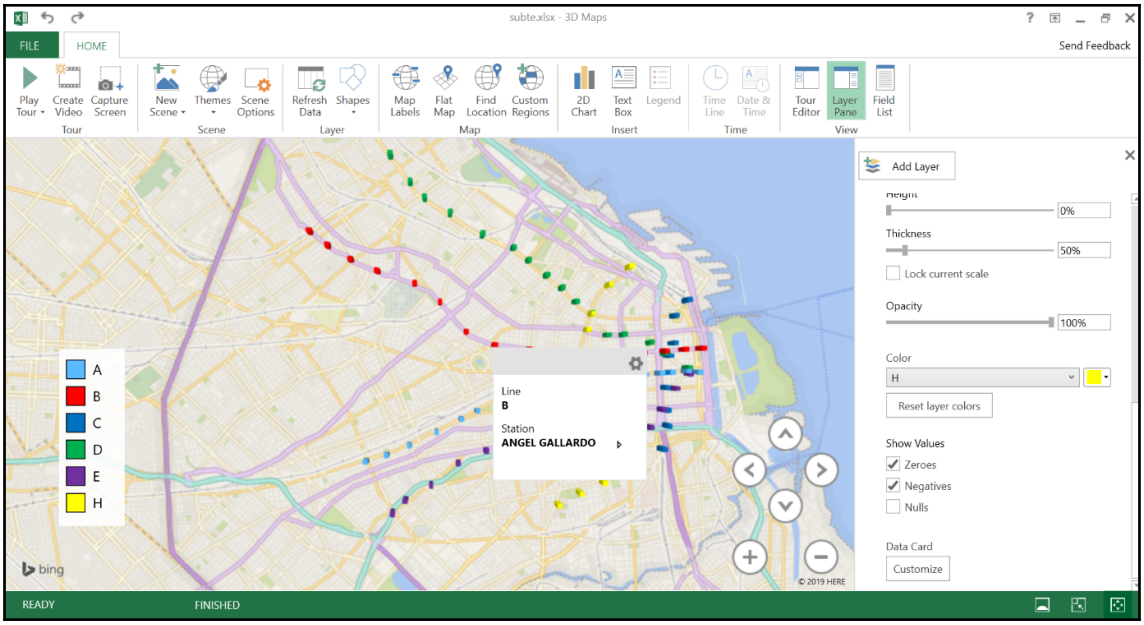

Launch 3D Maps
New Tour

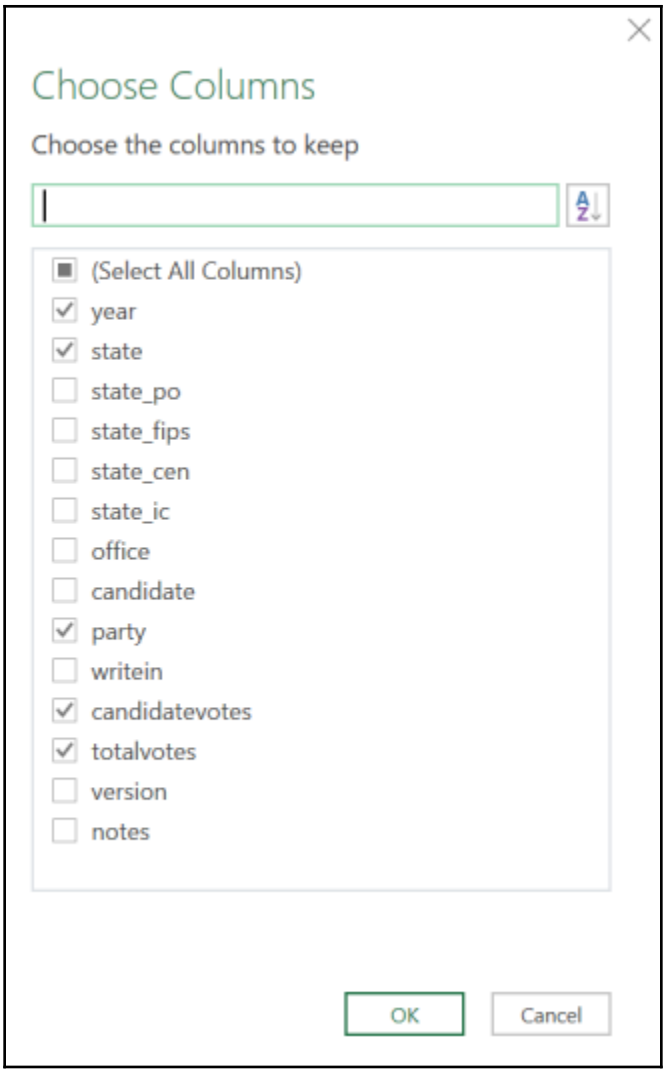

subte.xlsx - 3D Maps
FILE
HOME
Send Feedback
Play Tour
Create Video
Capture Screen
Tour
New Scene
Themes
Scene Options
Scene
Refresh Data
Shapes
Layer
Map Labels
Flat Map
Find Location
Custom Regions
Map
2D Chart
Text Box
Legend
Insert
Time Line
Date & Time
Time
Tour Editor
Layer Pane
Field List
View
Add Layer
Data
Location
Add Field
Height
Add Field
Category
Add Field
Time
Add Field
Filters
Layer Options
bing
© 2019 HERE
READY

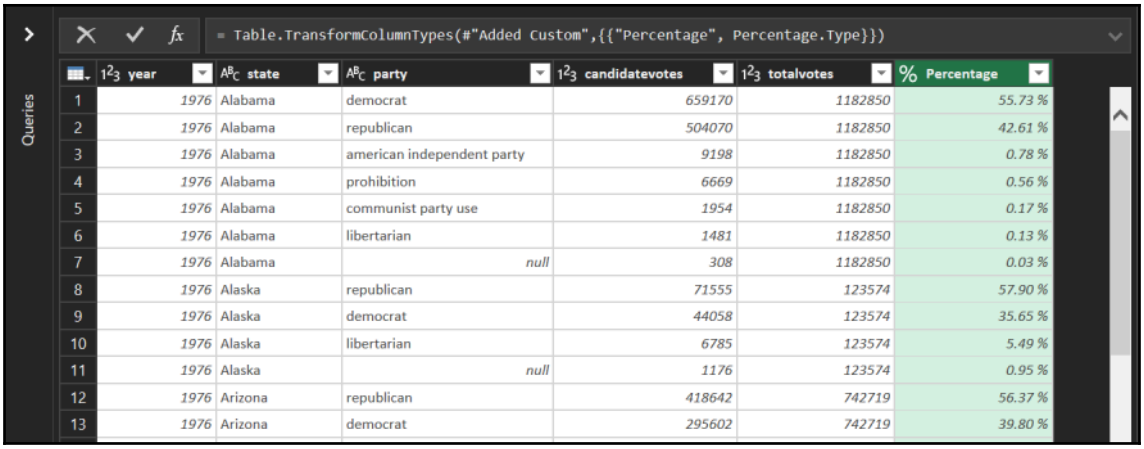

subte.xlsx - 3D Maps
FILE
HOME
Send Feedback
Play Tour
Create Video
Capture Screen
Tour
New Scene
Themes
Scene Options
Scene
Refresh Data
Shapes
Layer
Map Labels
Flat Map
Find Location
Custom Regions
Map
2D Chart
Text Box
Legend
Insert
Time Line
Date & Time
Time
Tour Editor
Layer Pane
Field List
View
Add Layer
Location
Lat
Latitude
Long
Longitude
Add Field
Height
Add Field
Category
Line
Time
Add Field
Filters
Add Filter
Layer Options
Height
100%
Layer 1
A
B
C
D
E
H
© 2019 HERE


subte.xlsx - 3D Maps
FILE
HOME
Send Feedback
Play Tour
Create Video
Capture Screen
Tour
New Scene
Themes
Scene Options
Scene
Refresh Data
Shapes
Layer
Map Labels
Flat Map
Find Location
Custom Regions
Map
2D Chart
Text Box
Legend
Insert
Time Line
Date & Time
Time
Tour Editor
Layer Pane
Field List
View
Add Layer
Filters
Layer Options
Height
0%
Thickness
50%
Lock current scale
Opacity
100%
Color
Show Values
Zeroes
Negatives
Nulls
More Colors...
Reset layer colors
Data Card
bing
© 2019 HERE
READY
FINISHED

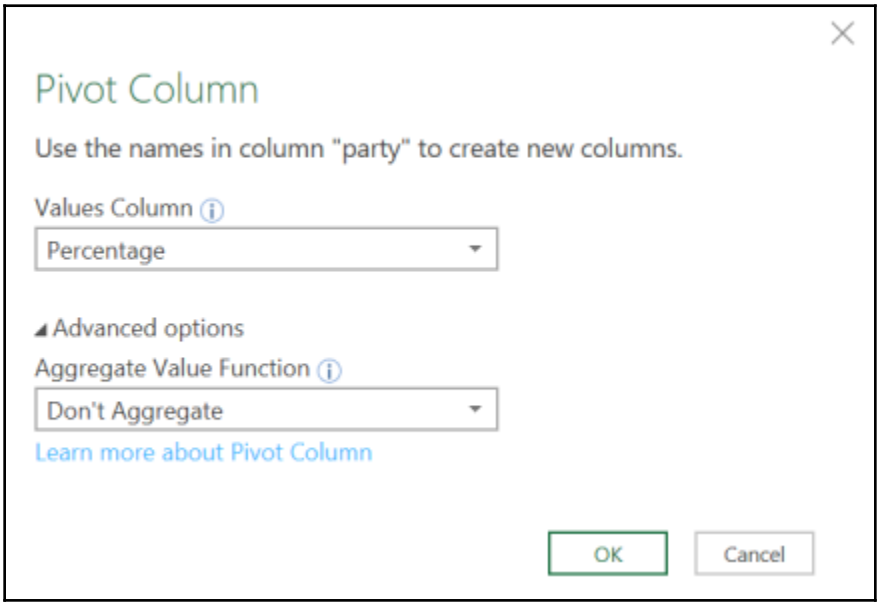

Customize Data Card
CHOOSE DATA FIELDS FOR CUSTOM TOOLTIP
TEMPLATE 1
Lat
Long
Line
Add Field
Lat
Value
Long
Value
Line
Value
Reset to Defaults
OK
Cancel

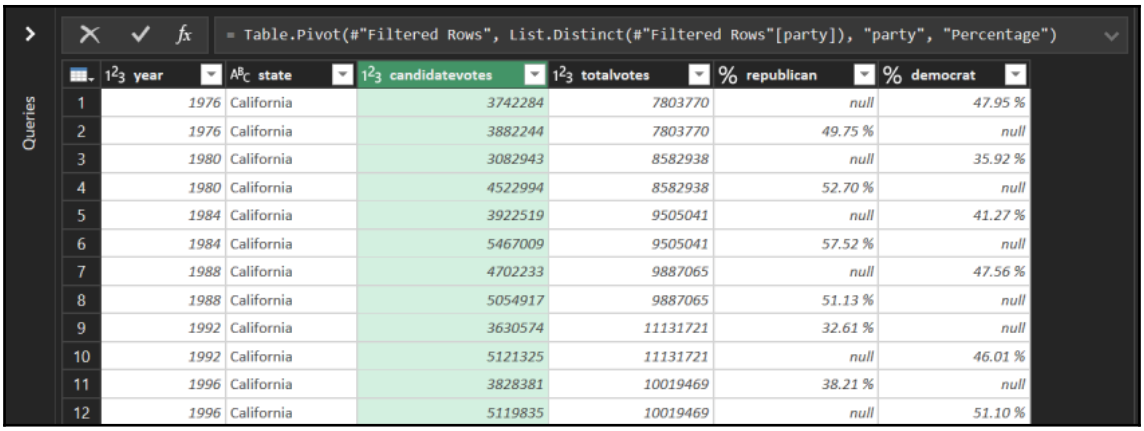

subte.xlsx - 3D Maps
FILE
HOME
Send Feedback
Play Tour
Create Video
Capture Screen
Tour
New Scene
Themes
Scene Options
Scene
Refresh Data
Shapes
Layer
Map Labels
Flat Map
Find Location
Custom Regions
Map
2D Chart
Text Box
Legend
Insert
Time Line
Date & Time
Time
Tour Editor
Layer Pane
Field List
View
A
B
C
D
E
H
Line
B
Station
ANGEL GALLARDO
bing
© 2019 HERE
Add Layer
0%
Thickness
50%
Lock current scale
Opacity
100%
Color
H
Reset layer colors
Show Values
Zeroes
Negatives
Nulls
Data Card
Customize
READY
FINISHED


Choose Columns
Choose the columns to keep
(Select All Columns)
year
state
state_po
state_fips
state_cen
state_ic
office
candidate
party
writein
candidatevotes
totalvotes
version
notes
OK
Cancel



	year	state	party	candidatevotes	totalvotes	Percentage
1	1976	Alabama	democrat	659170	1182850	55.73 %
2	1976	Alabama	republican	504070	1182850	42.61 %
3	1976	Alabama	american independent party	9198	1182850	0.78 %
4	1976	Alabama	prohibition	6669	1182850	0.56 %
5	1976	Alabama	communist party use	1954	1182850	0.17 %
6	1976	Alabama	libertarian	1481	1182850	0.13 %
7	1976	Alabama	null	308	1182850	0.03 %
8	1976	Alaska	republican	71555	123574	57.90 %
9	1976	Alaska	democrat	44058	123574	35.65 %
10	1976	Alaska	libertarian	6785	123574	5.49 %
11	1976	Alaska	null	1176	123574	0.95 %
12	1976	Arizona	republican	418642	742719	56.37 %
13	1976	Arizona	democrat	295602	742719	39.80 %



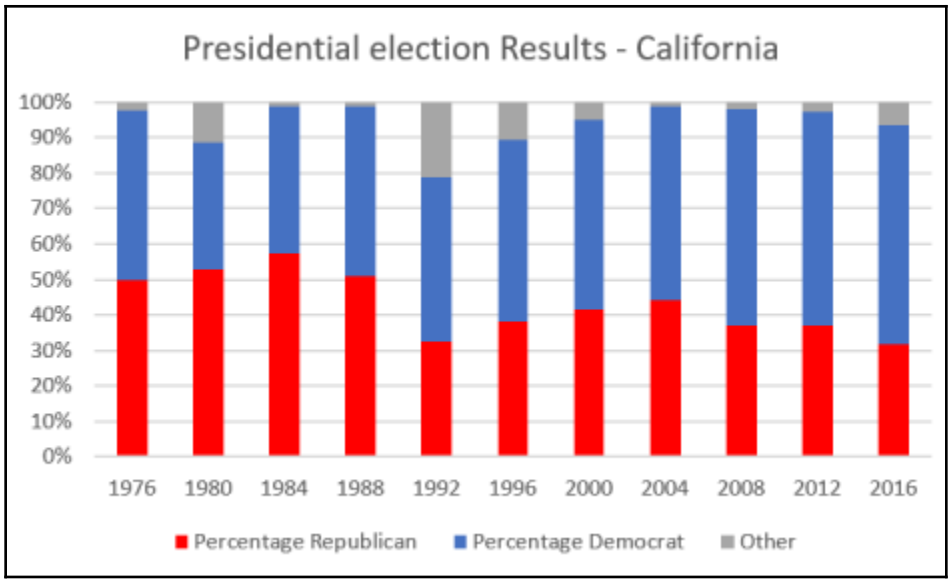


Queries

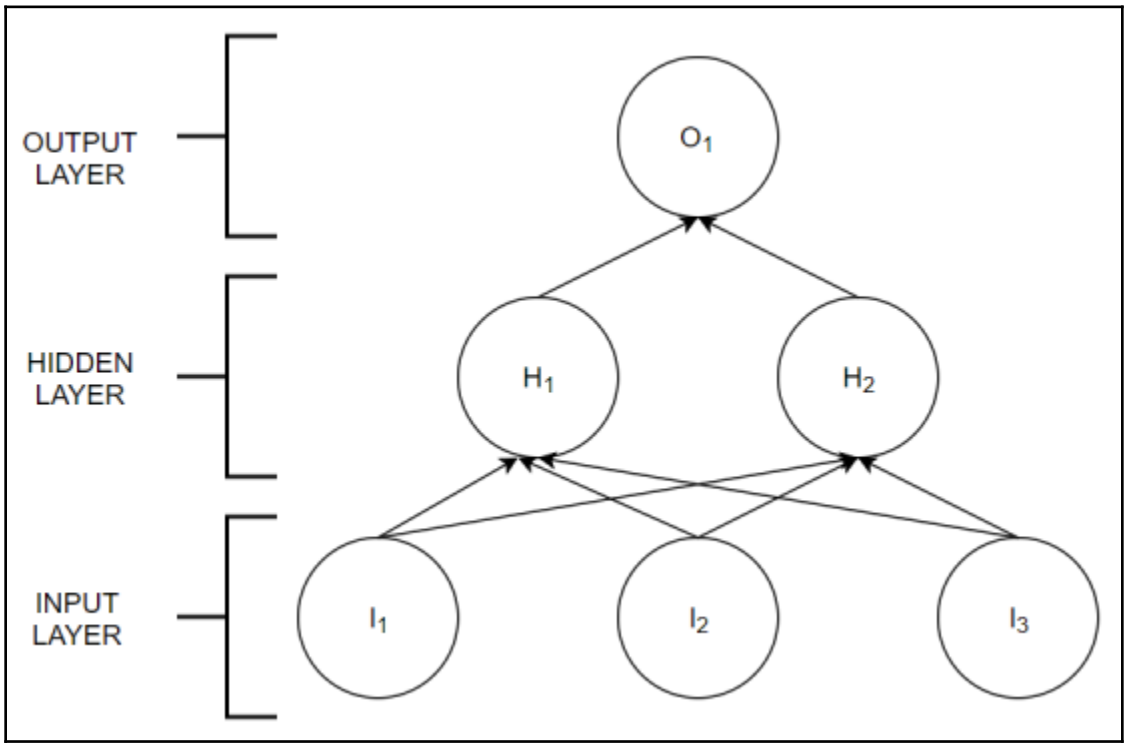
```
= Table.Pivot(#"Filtered Rows", List.Distinct(#"Filtered Rows"[party]), "party", "Percentage")
```

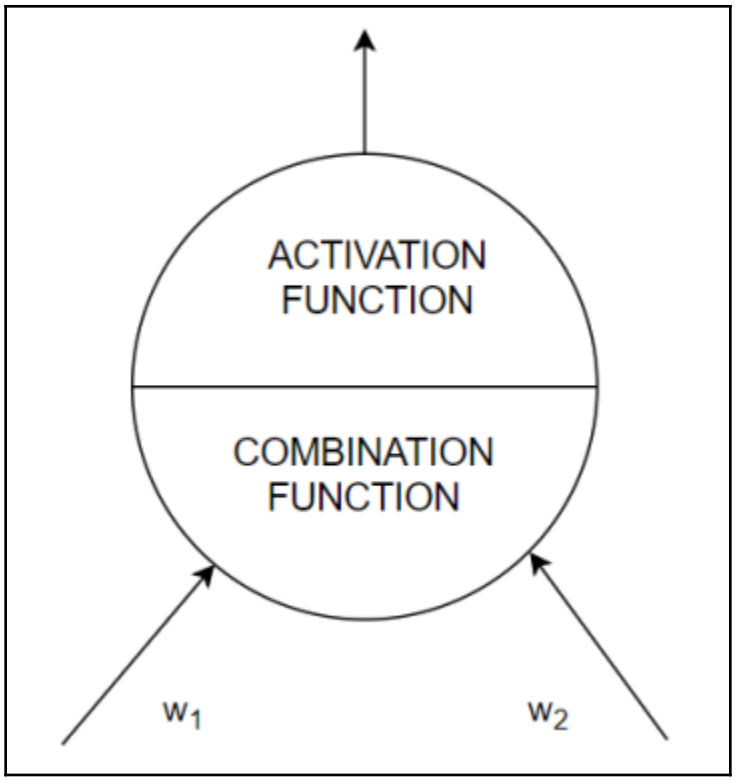
	year	state	candidatevotes	totalvotes	republican	democrat
1	1976	California	3742284	7803770	null	47.95 %
2	1976	California	3882244	7803770	49.75 %	null
3	1980	California	3082943	8582938	null	35.92 %
4	1980	California	4522994	8582938	52.70 %	null
5	1984	California	3922519	9505041	null	41.27 %
6	1984	California	5467009	9505041	57.52 %	null
7	1988	California	4702233	9887065	null	47.56 %
8	1988	California	5054917	9887065	51.13 %	null
9	1992	California	3630574	11131721	32.61 %	null
10	1992	California	5121325	11131721	null	46.01 %
11	1996	California	3828381	10019469	38.21 %	null
12	1996	California	5119835	10019469	null	51.10 %

	A	B	C	D
1	year	Percentage Republican	Percentage Democrat	Other
2	1976	49.75%	47.95%	2.30%
3	1980	52.70%	35.92%	11.38%
4	1984	57.52%	41.27%	1.22%
5	1988	51.13%	47.56%	1.31%
6	1992	32.61%	46.01%	21.38%
7	1996	38.21%	51.10%	10.69%
8	2000	41.65%	53.45%	4.90%
9	2004	44.36%	54.31%	1.34%
10	2008	36.95%	61.01%	2.03%
11	2012	37.12%	60.24%	2.64%
12	2016	31.62%	61.73%	6.66%


Presidential election Results - California
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1976
1980
1984
1988
1992
1996
2000
2004
2008
2012
2016
Percentage Republican
Percentage Democrat
Other

Chapter 9: Artificial Neural Networks

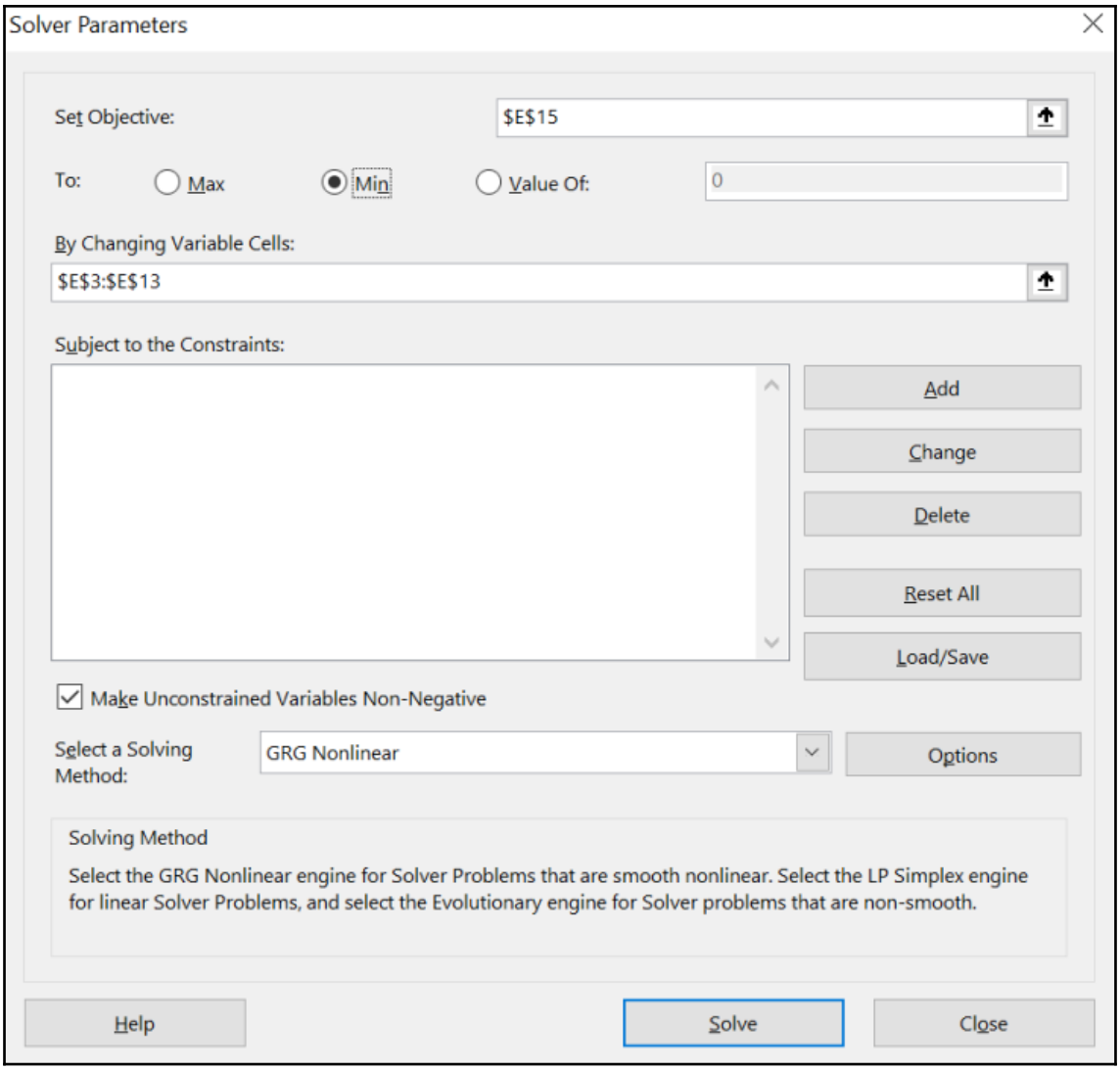


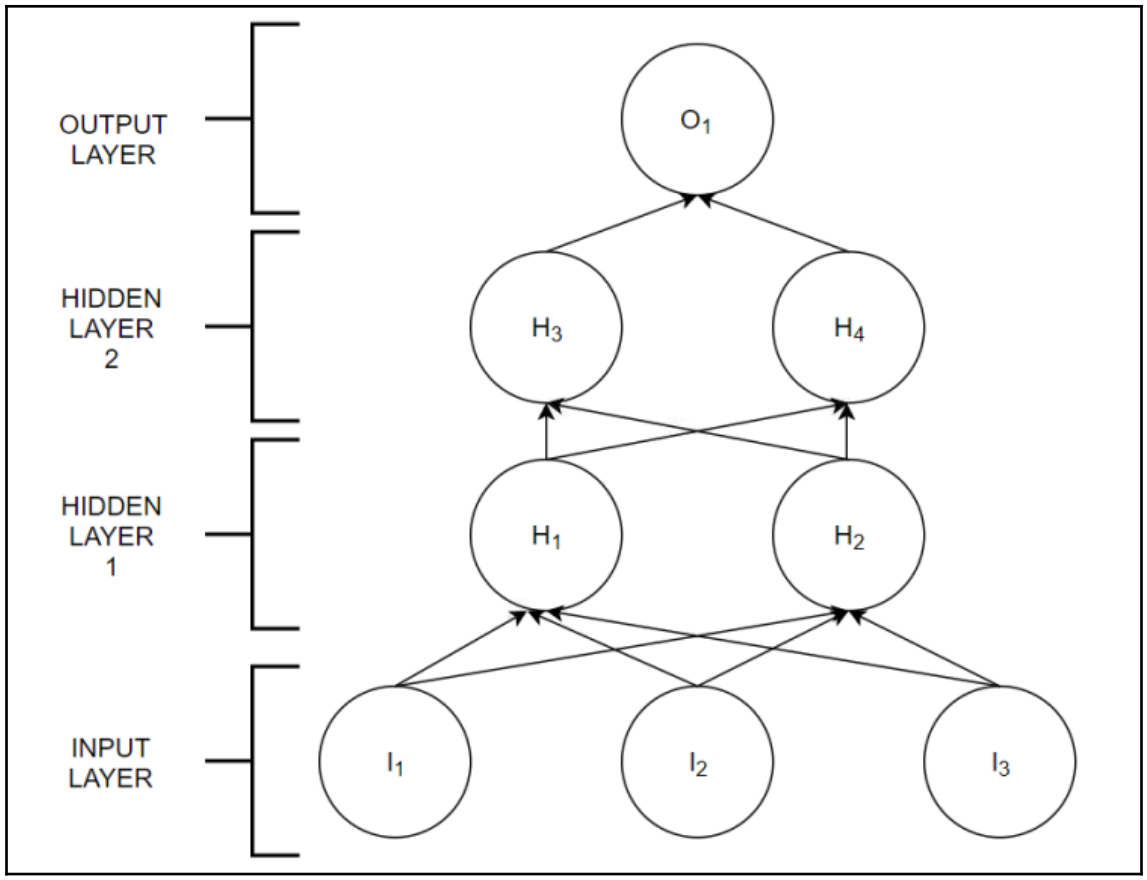

ACTIVATION
FUNCTION
COMBINATION
FUNCTION
w1
w2

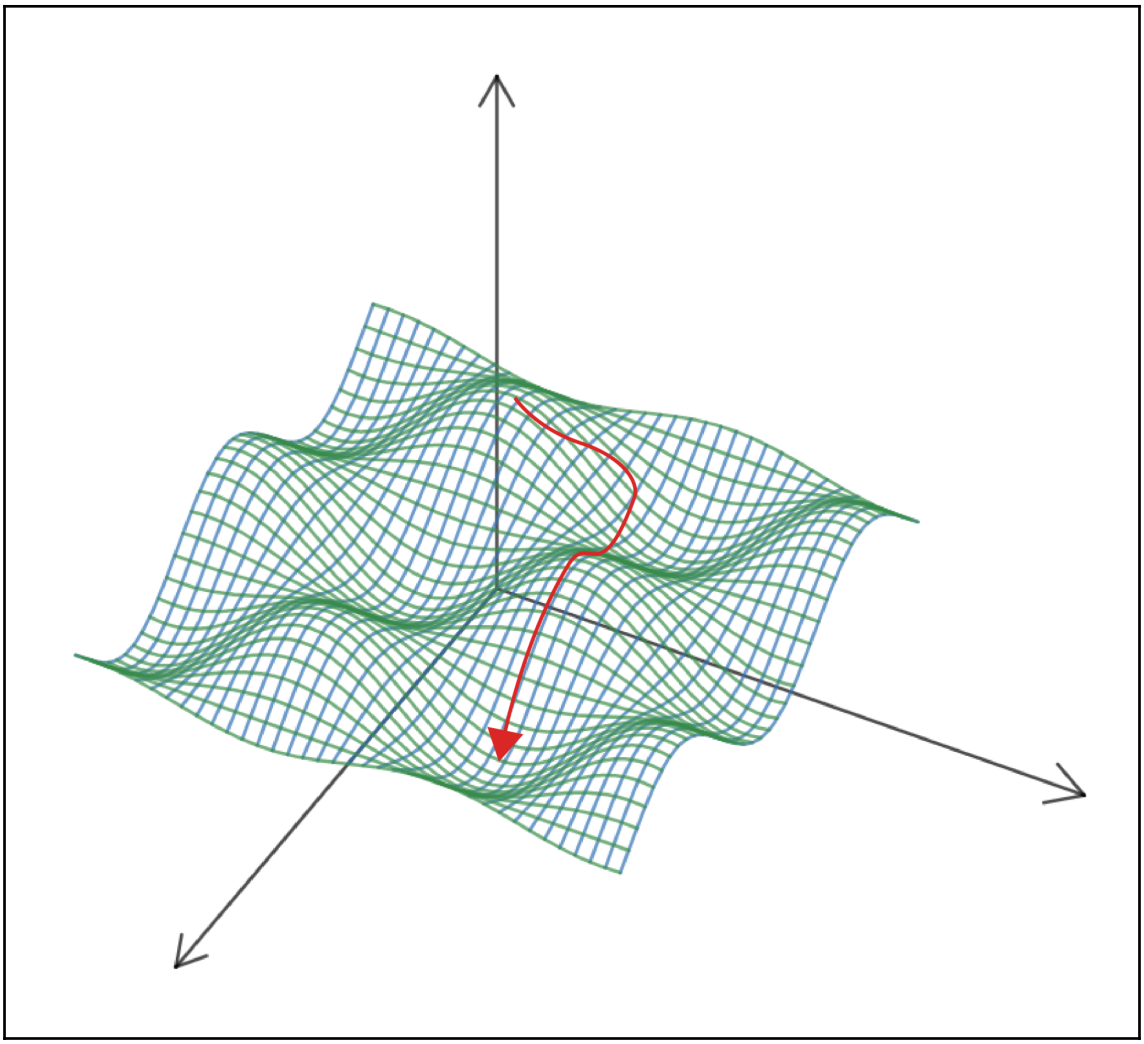
	A	B	C	D	E	F
1	Recency (months)	Frequency (times)	Monetary (c.c. blood)	Time (months)	whether he/she donated blood in March 2007	
2	2	50	12500	98	1	
3	0	13	3250	28	1	
4	1	16	4000	35	1	
5	2	20	5000	45	1	
6	1	24	6000	77	0	
7	4	4	1000	4	0	
8	2	7	1750	14	1	
9	1	12	3000	35	0	
10	2	9	2250	22	1	
11	5	46	11500	98	1	
12	4	23	5750	58	0	
13	0	3	750	4	0	
14	2	10	2500	28	1	
15	1	13	3250	47	0	
16	2	6	1500	15	1	
17	2	5	1250	11	1	
18	2	14	3500	48	1	
19	2	15	3750	49	1	
20	2	6	1500	15	1	
21	2	3	750	4	1	
22	2	3	750	4	1	
23	4	11	2750	28	0	
24	2	6	1500	16	1	
25	2	6	1500	16	1	

Reference | transfusion | training | test

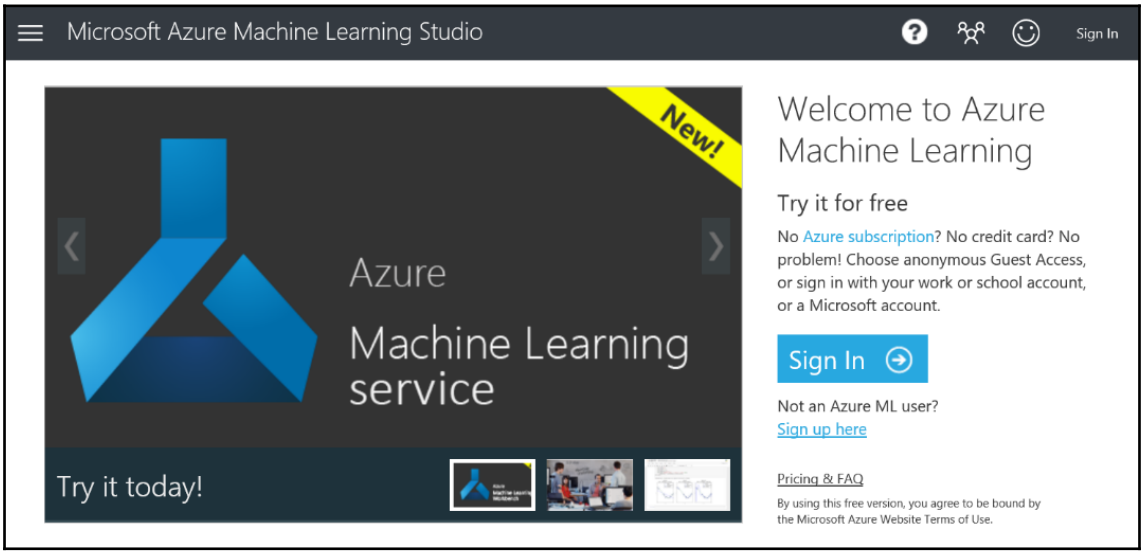
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Neural network with 4 inputs, one hidden layer with 2 neurons														
2															
3				w11											
4				w12											
5				w13											
6				w14											
7				w21											
8				w22											
9				w23											
10				w24											
11				theta0											
12				theta1											
13				theta2											
14															
15				Sq error											
16															
17															
18															
19	Train data														
20															
21	#	x1	x2	x3	x4	y	wsum1	wsum2	at1	at2	Output	Error	Sq error	round_out	
22															
23															

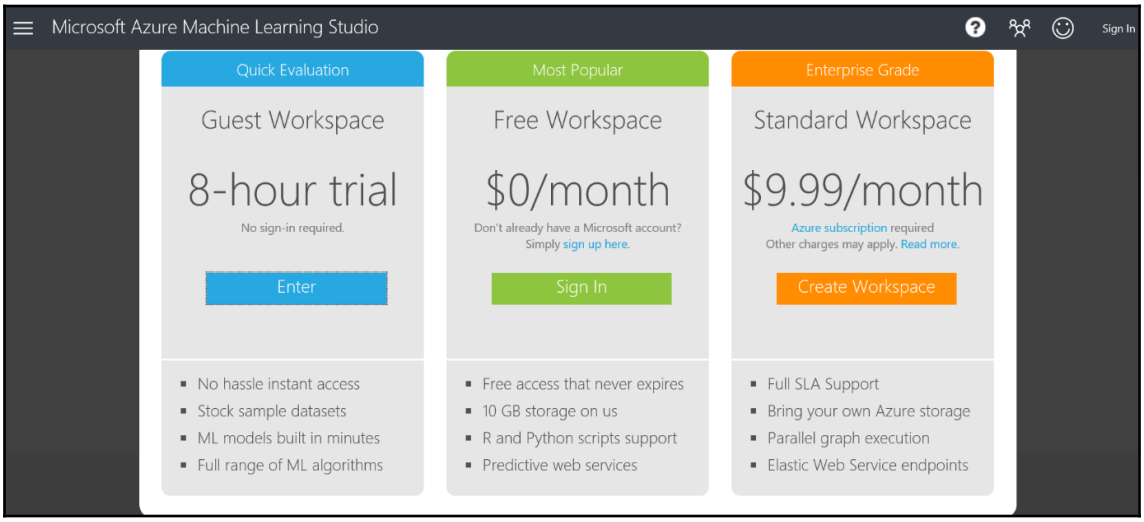

Solver Parameters
Set Objective:
$E$15
To:
Max
Min
Value Of:
0
By Changing Variable Cells:
$E$3:$E$13
Subject to the Constraints:
Add
Change
Delete
Reset All
Load/Save
Make Unconstrained Variables Non-Negative
Select a Solving Method:
GRG Nonlinear
Options
Solving Method
Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.
Help
Solve
Close

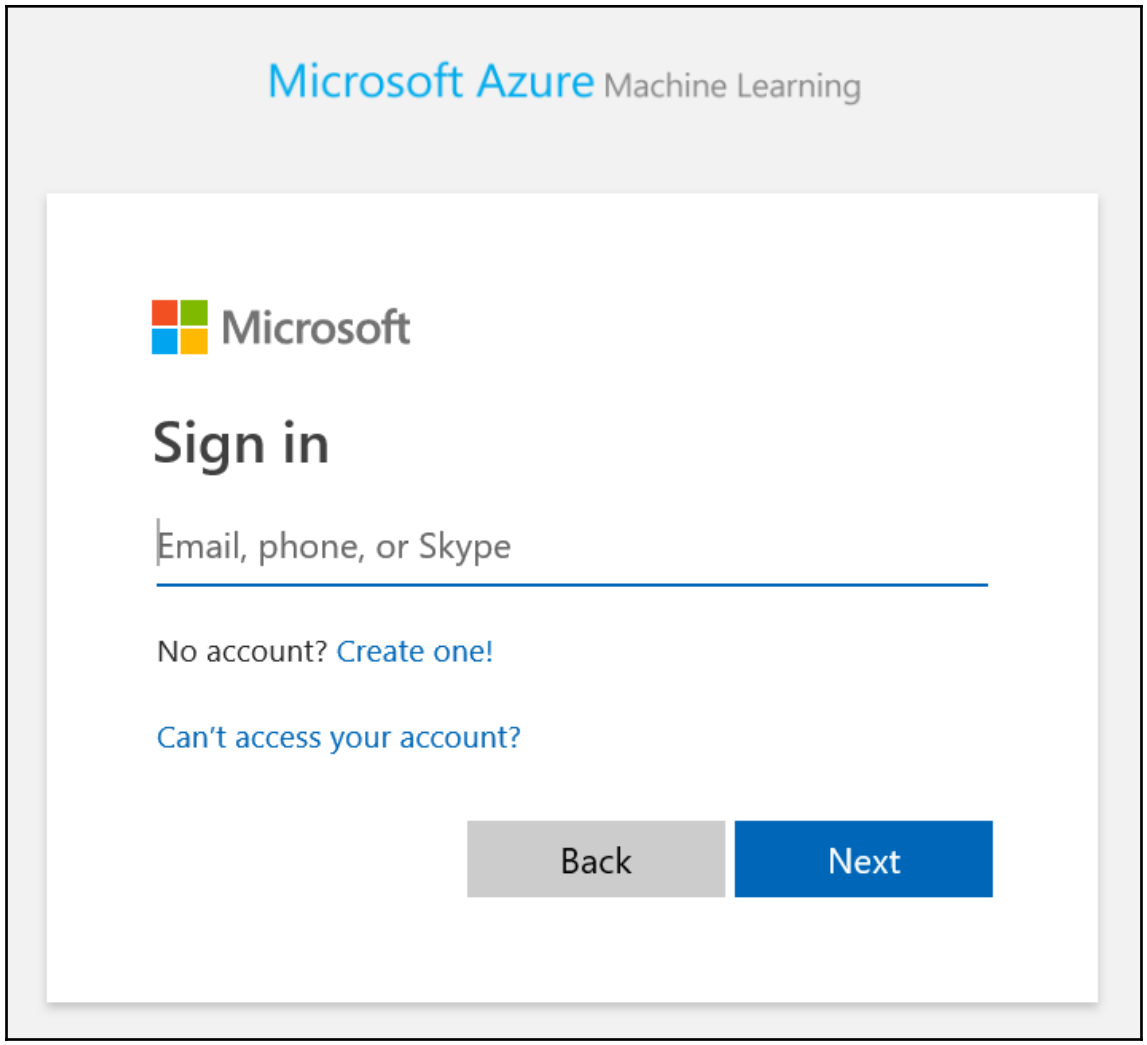

OUTPUT LAYER
HIDDEN LAYER 2
HIDDEN LAYER 1
INPUT LAYER
O1
H3
H4
H1
H2
I1
I2
I3

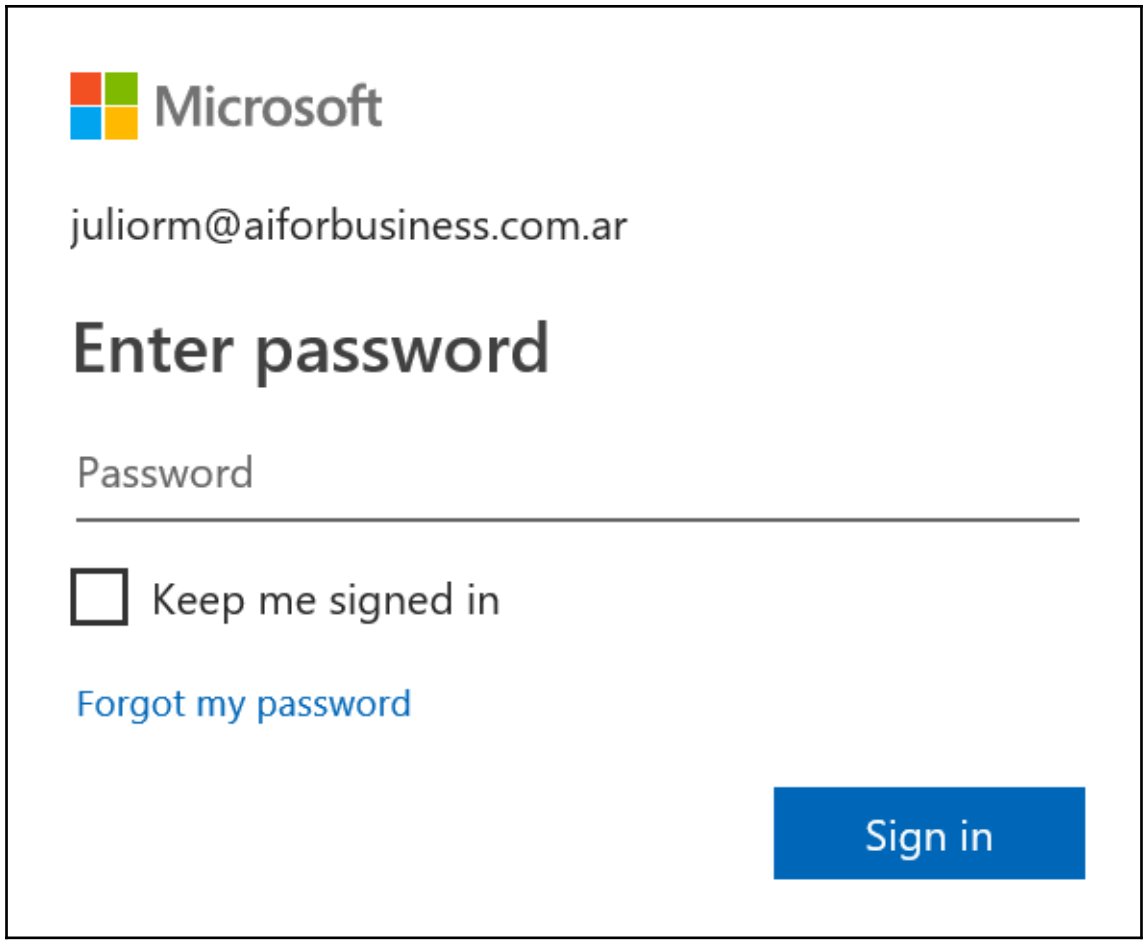


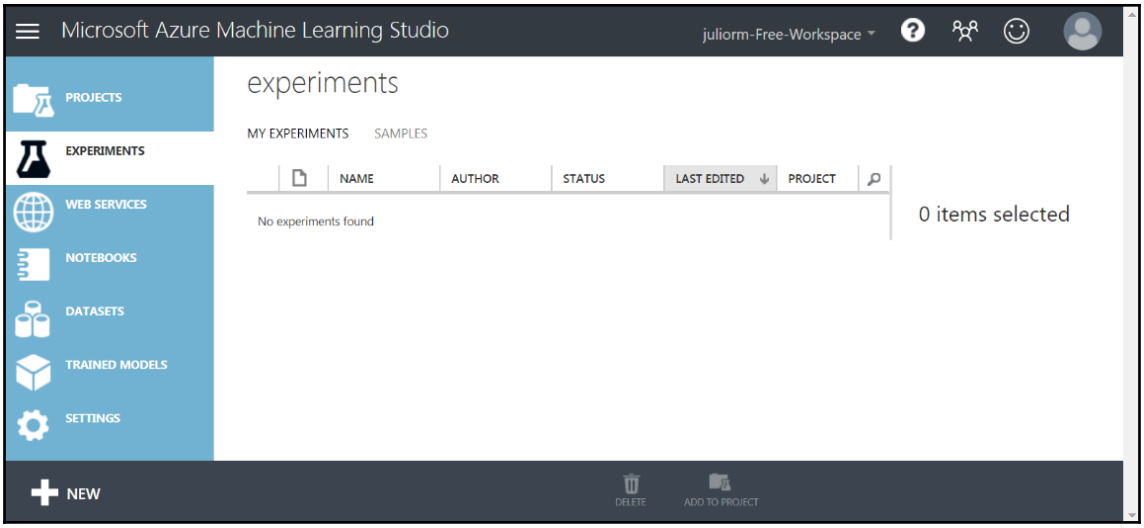
Chapter 10: Azure and Excel - Machine Learning in the Cloud

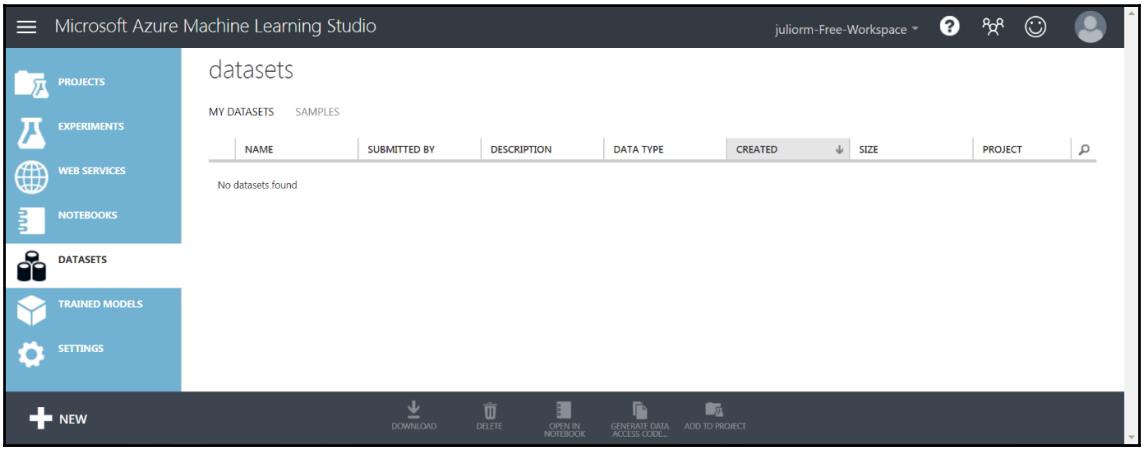


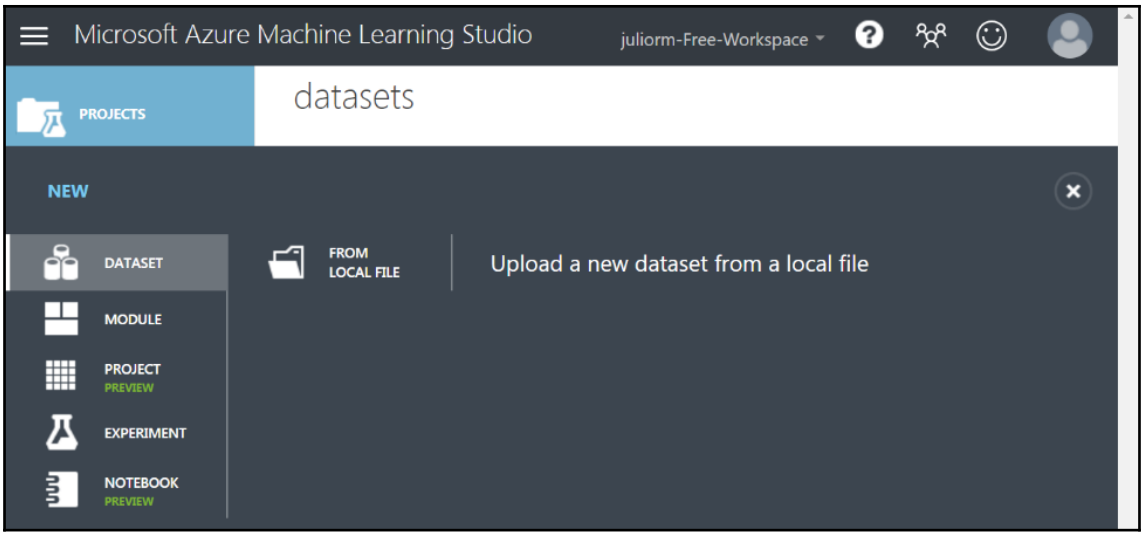


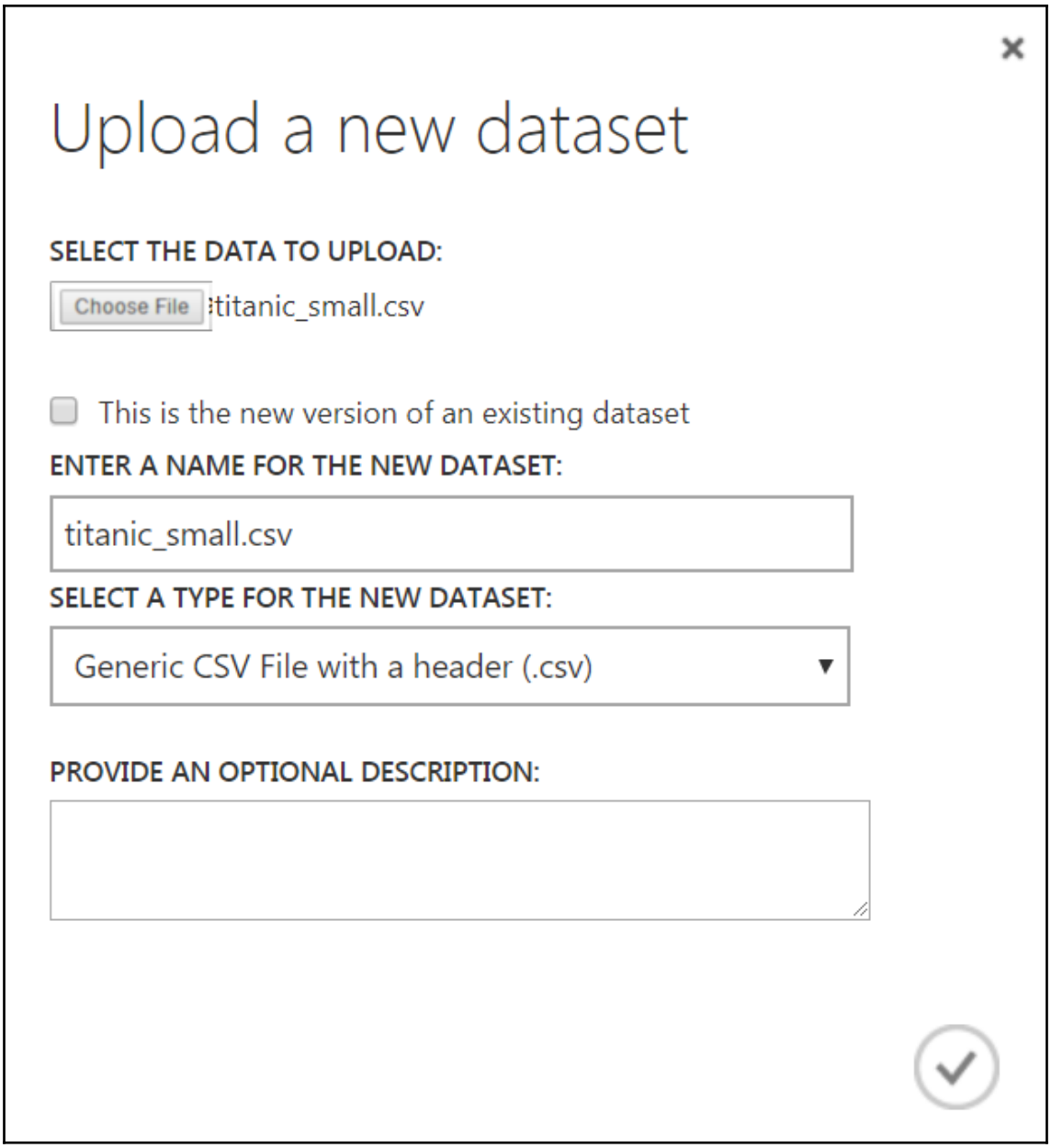

Microsoft Azure Machine Learning
Microsoft
Sign in
Email, phone, or Skype
No account? Create one!
Can't access your account?
Back
Next

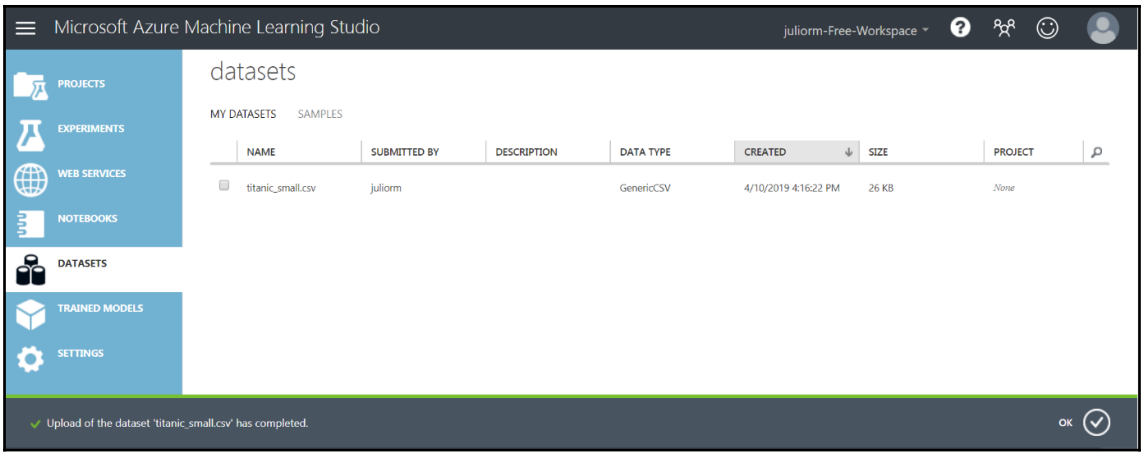

Microsoft
juliorm@aiforbusiness.com.ar
Enter password
Password
Keep me signed in
Forgot my password
Sign in

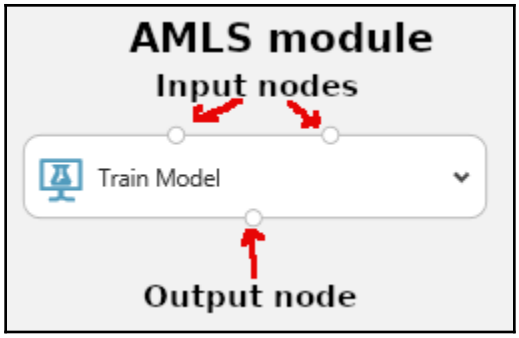

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
PROJECTS
EXPERIMENTS
WEB SERVICES
NOTEBOOKS
DATASETS
TRAINED MODELS
SETTINGS
experiments
MY EXPERIMENTS
SAMPLES
NAME
AUTHOR
STATUS
LAST EDITED
PROJECT
No experiments found
0 items selected
NEW
DELETE
ADD TO PROJECT

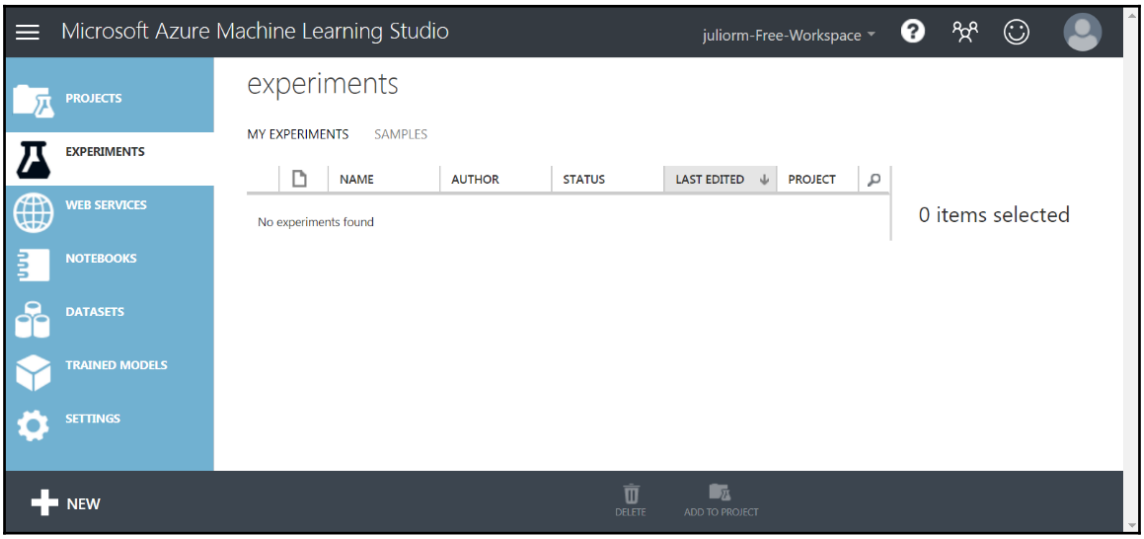

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
PROJECTS
EXPERIMENTS
WEB SERVICES
NOTEBOOKS
DATASETS
TRAINED MODELS
SETTINGS
datasets
MY DATASETS
SAMPLES
NAME
SUBMITTED BY
DESCRIPTION
DATA TYPE
CREATED
SIZE
PROJECT
No datasets found
NEW
DOWNLOAD
DELETE
OPEN IN NOTEBOOK
GENERATE DATA ACCESS CODE...
ADD TO PROJECT

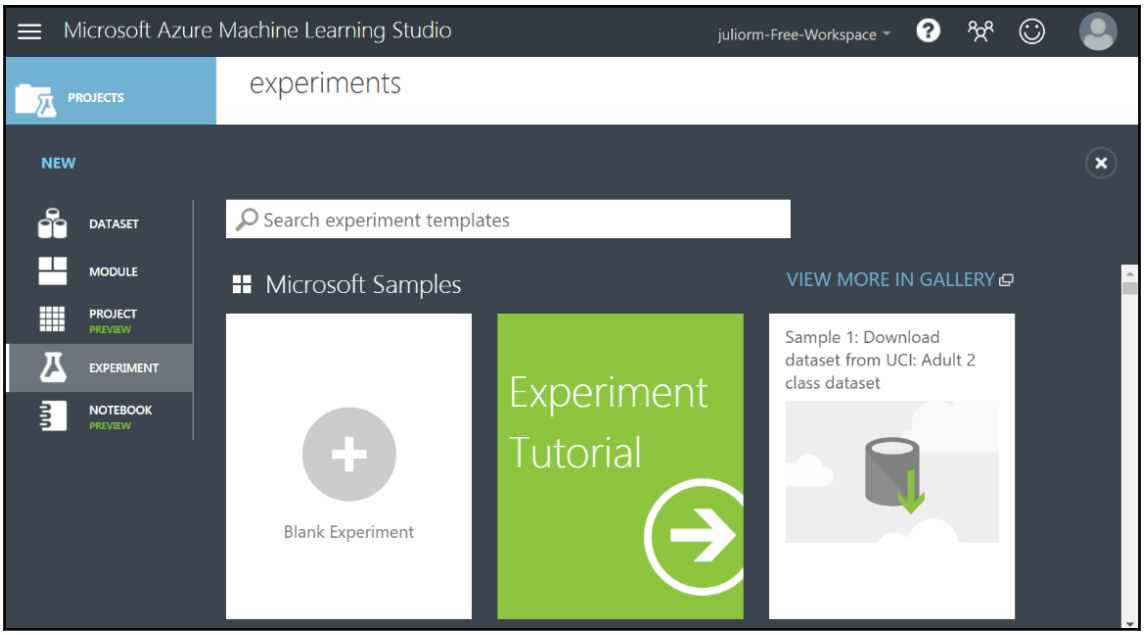

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
PROJECTS
datasets
NEW
DATASET
MODULE
PROJECT
PREVIEW
EXPERIMENT
NOTEBOOK
PREVIEW
FROM
LOCAL FILE
Upload a new dataset from a local file

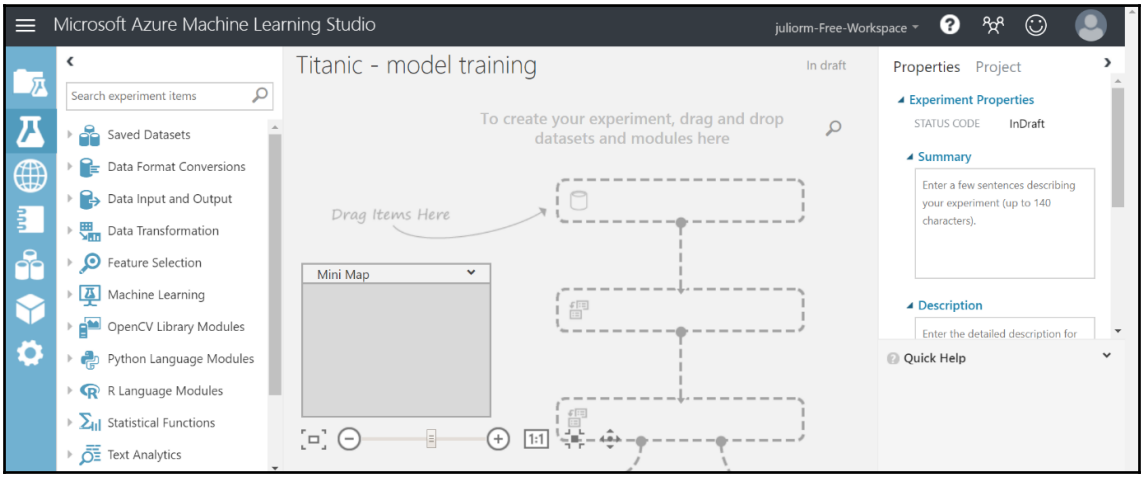

Upload a new dataset
SELECT THE DATA TO UPLOAD:
Choose File
titanic_small.csv
This is the new version of an existing dataset
ENTER A NAME FOR THE NEW DATASET:
titanic_small.csv
SELECT A TYPE FOR THE NEW DATASET:
Generic CSV File with a header (.csv)
PROVIDE AN OPTIONAL DESCRIPTION:

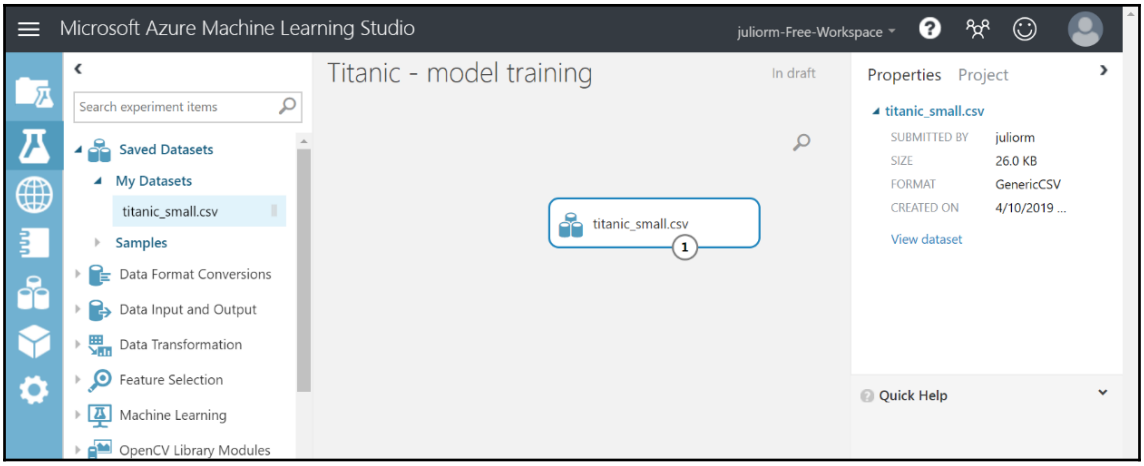

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
PROJECTS
EXPERIMENTS
WEB SERVICES
NOTEBOOKS
DATASETS
TRAINED MODELS
SETTINGS
datasets
MY DATASETS
SAMPLES
NAME
SUBMITTED BY
DESCRIPTION
DATA TYPE
CREATED
SIZE
PROJECT
titanic_small.csv
juliorm
GenericCSV
4/10/2019 4:16:22 PM
26 KB
None
Upload of the dataset 'titanic_small.csv' has completed.
OK

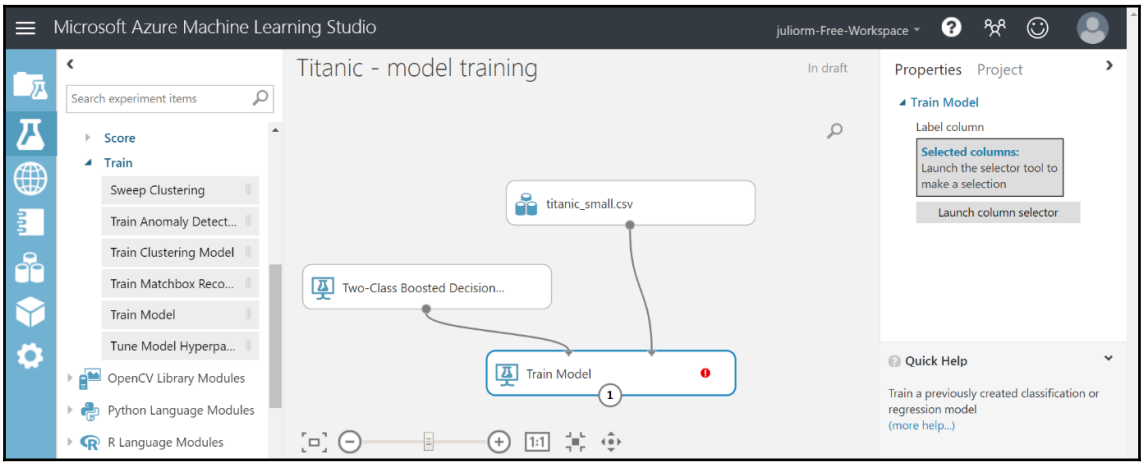

AMLS module
Input nodes
Train Model
Output node

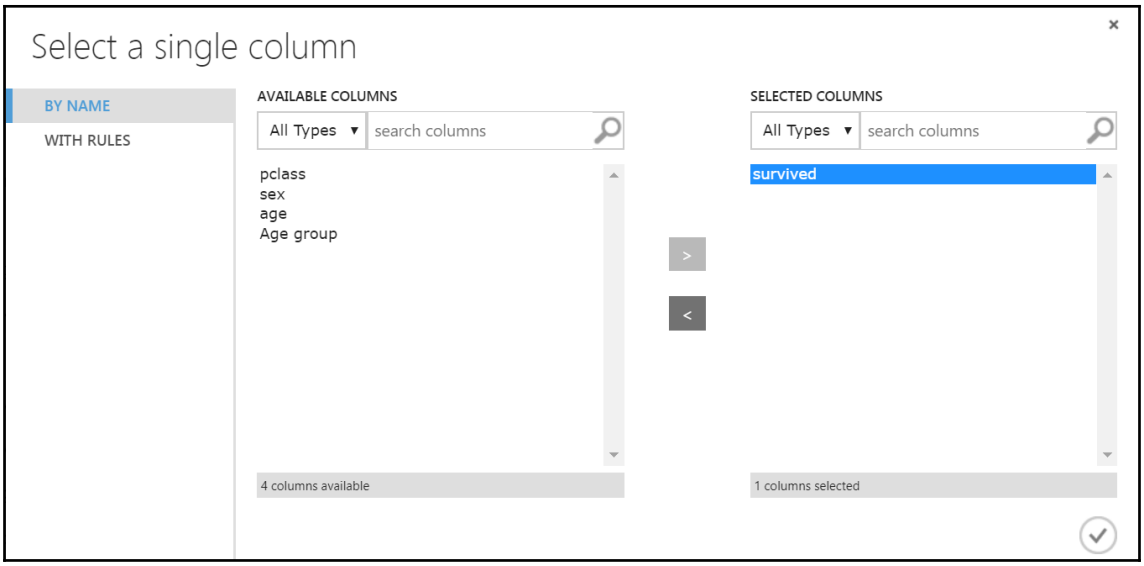

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
PROJECTS
EXPERIMENTS
WEB SERVICES
NOTEBOOKS
DATASETS
TRAINED MODELS
SETTINGS
experiments
MY EXPERIMENTS
SAMPLES
NAME
AUTHOR
STATUS
LAST EDITED
PROJECT
No experiments found
0 items selected
NEW
DELETE
ADD TO PROJECT


Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
PROJECTS
experiments
NEW
DATASET
MODULE
PROJECT
PREVIEW
EXPERIMENT
NOTEBOOK
PREVIEW
Search experiment templates
Microsoft Samples
VIEW MORE IN GALLERY
Blank Experiment
Experiment Tutorial
Sample 1: Download dataset from UCI: Adult 2 class dataset

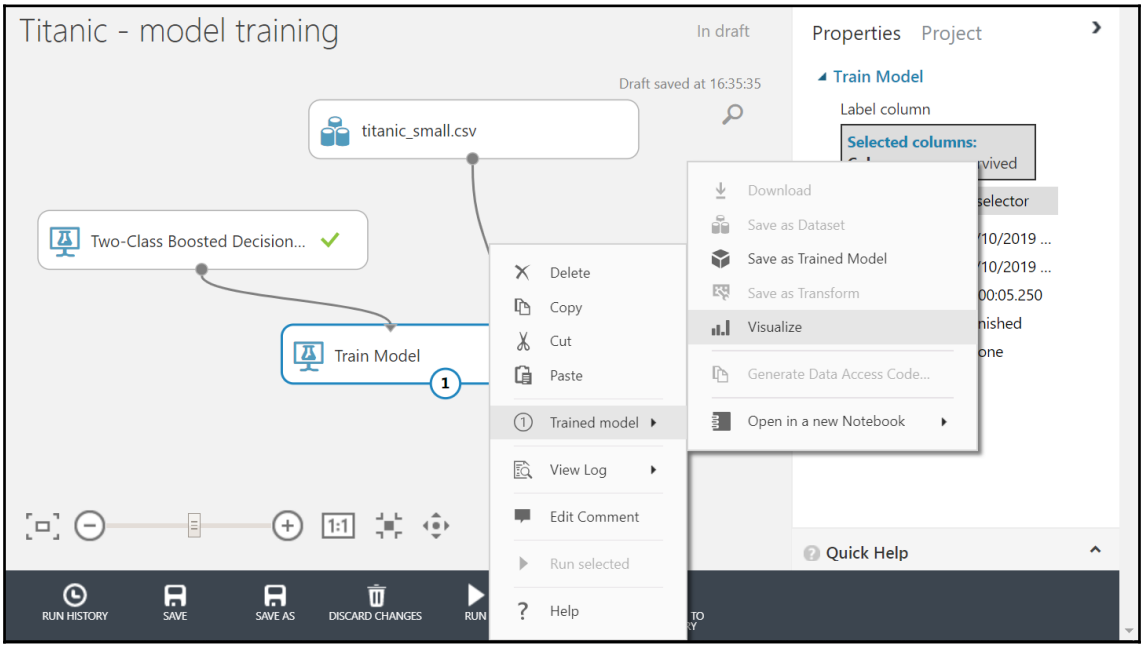

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
Search experiment items
Saved Datasets
Data Format Conversions
Data Input and Output
Data Transformation
Feature Selection
Machine Learning
OpenCV Library Modules
Python Language Modules
R Language Modules
Statistical Functions
Text Analytics
Titanic - model training
In draft
To create your experiment, drag and drop datasets and modules here
Drag Items Here
Mini Map
Properties
Project
Experiment Properties
STATUS CODE
InDraft
Summary
Enter a few sentences describing your experiment (up to 140 characters).
Description
Quick Help

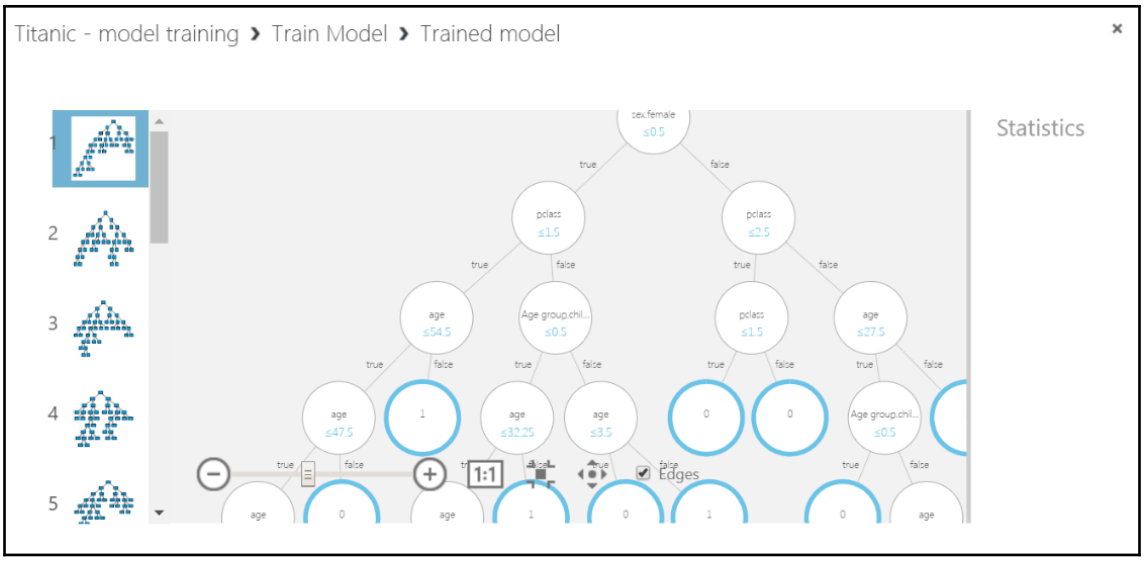

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
Titanic - model training
In draft
Search experiment items
Saved Datasets
My Datasets
titanic_small.csv
Samples
Data Format Conversions
Data Input and Output
Data Transformation
Feature Selection
Machine Learning
OpenCV Library Modules
titanic_small.csv
1
Properties
Project
titanic_small.csv
SUBMITTED BY
juliorm
SIZE
26.0 KB
FORMAT
GenericCSV
CREATED ON
4/10/2019 ...
View dataset
Quick Help

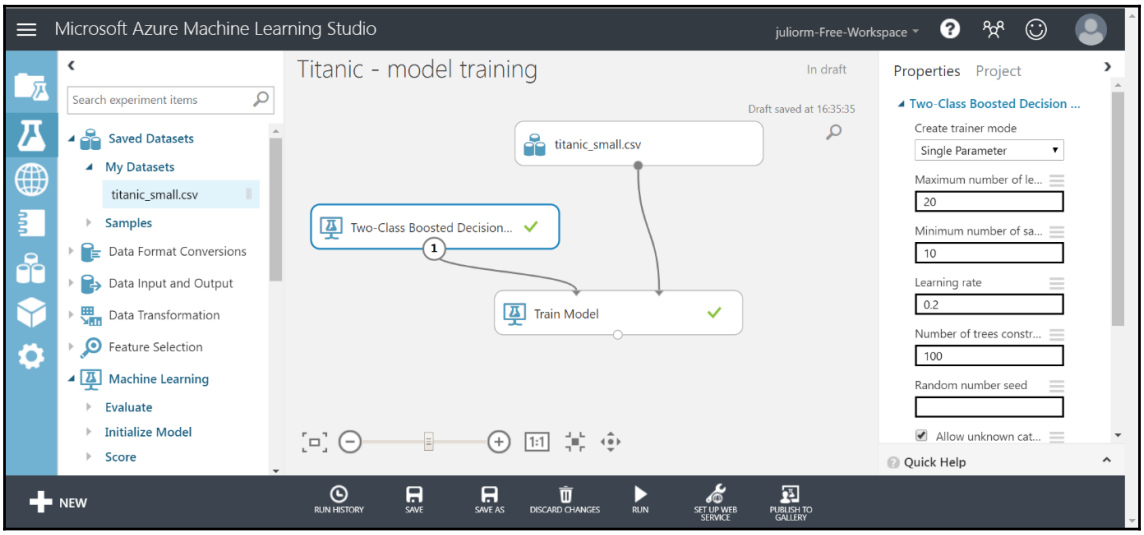

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
Titanic - model training
In draft
Search experiment items
Score
Train
Sweep Clustering
Train Anomaly Detect...
Train Clustering Model
Train Matchbox Reco...
Train Model
Tune Model Hyperpa...
OpenCV Library Modules
Python Language Modules
R Language Modules
titanic_small.csv
Two-Class Boosted Decision...
Train Model
1
Properties
Project
Train Model
Label column
Selected columns:
Launch the selector tool to make a selection
Launch column selector
Quick Help
Train a previously created classification or regression model
(more help...)

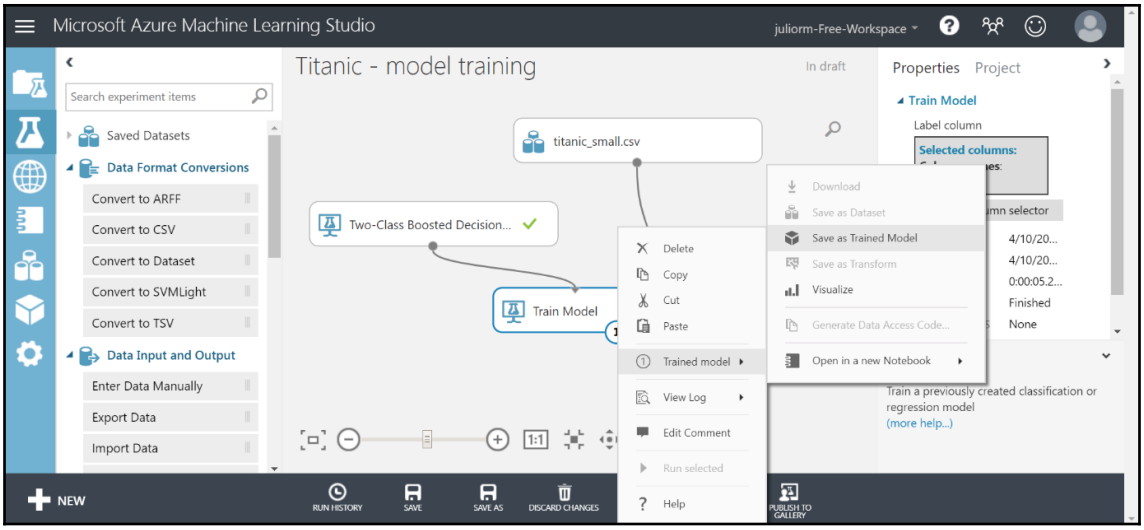

Select a single column
BY NAME
WITH RULES
AVAILABLE COLUMNS
All Types
search columns
pclass
sex
age
Age group
4 columns available
SELECTED COLUMNS
All Types
search columns
survived
1 columns selected

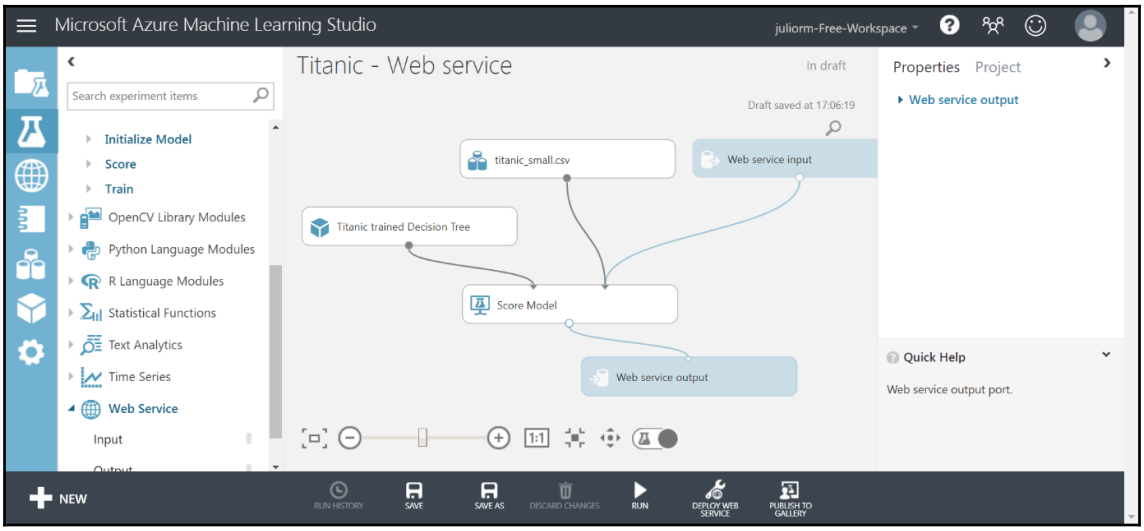

RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY

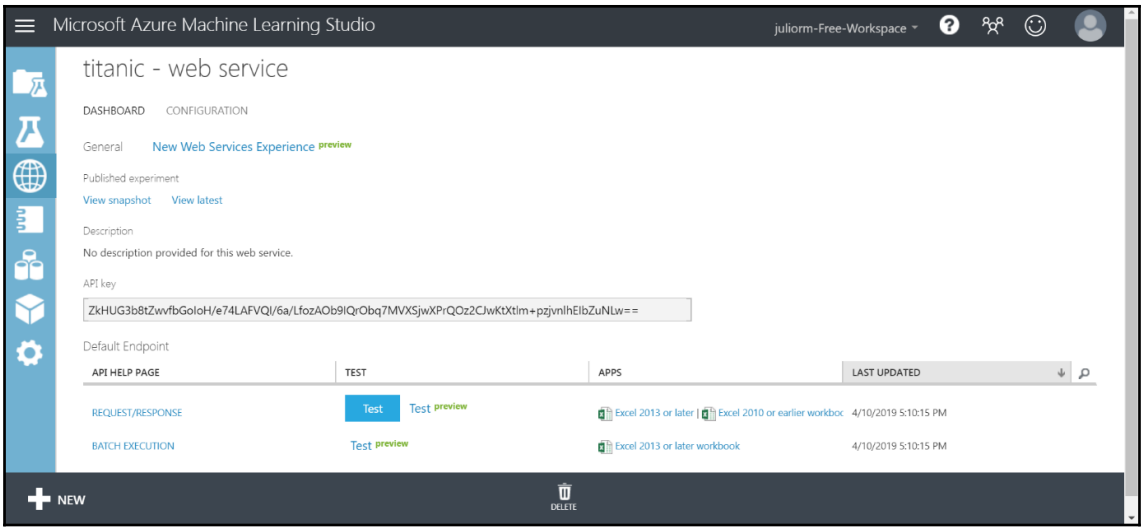

Titanic - model training
In draft
Draft saved at 16:35:35
titanic_small.csv
Two-Class Boosted Decision...
Train Model
Delete
Copy
Cut
Paste
Trained model
View Log
Edit Comment
Run selected
Help
Download
Save as Dataset
Save as Trained Model
Save as Transform
Visualize
Generate Data Access Code...
Open in a new Notebook
Properties
Project
Train Model
Label column
Selected columns:
Quick Help
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN

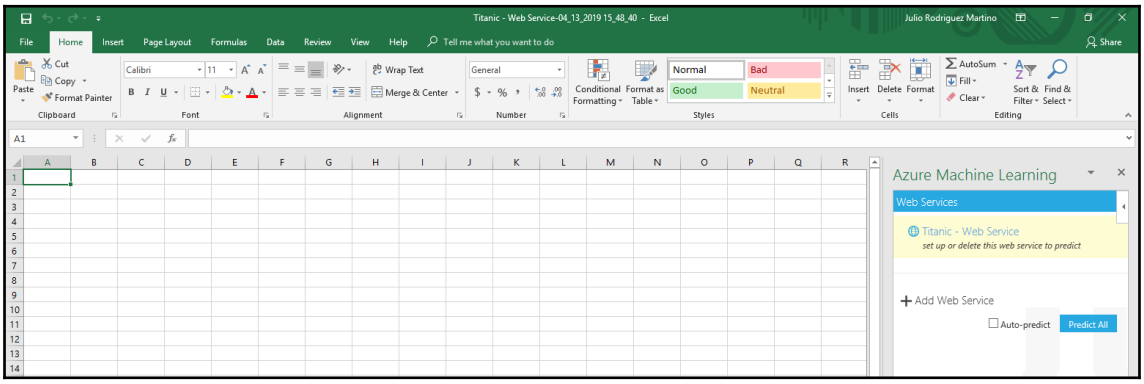

Titanic - model training › Train Model › Trained model
Statistics
Edges

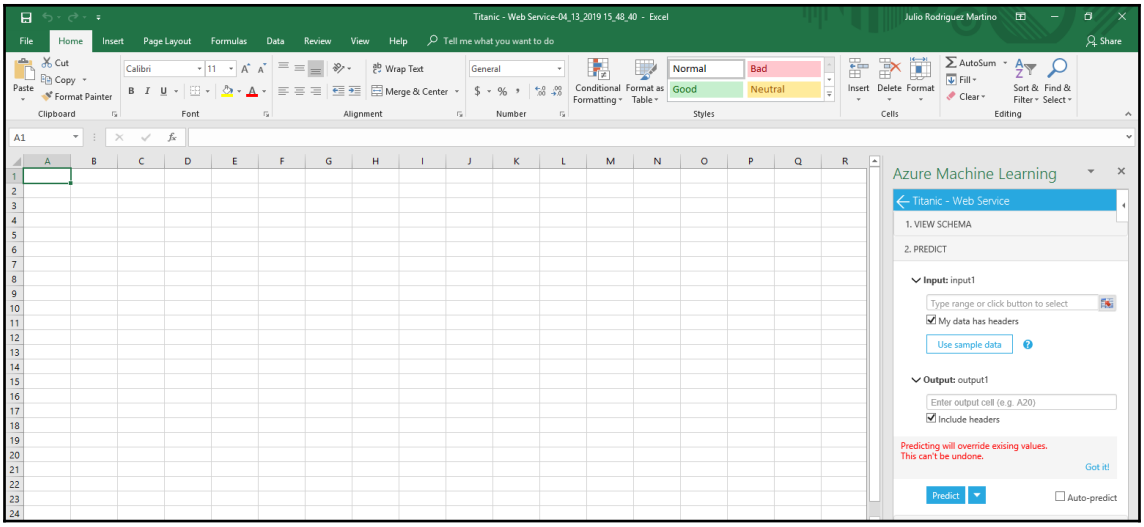

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
Titanic - model training
In draft
Draft saved at 16:35:35
Search experiment items
Saved Datasets
My Datasets
titanic_small.csv
Samples
Data Format Conversions
Data Input and Output
Data Transformation
Feature Selection
Machine Learning
Evaluate
Initialize Model
Score
titanic_small.csv
Two-Class Boosted Decision...
Train Model
Properties
Project
Two-Class Boosted Decision ...
Create trainer mode
Single Parameter
Maximum number of le...
20
Minimum number of sa...
10
Learning rate
0.2
Number of trees constr...
100
Random number seed
Allow unknown cat...
Quick Help
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY

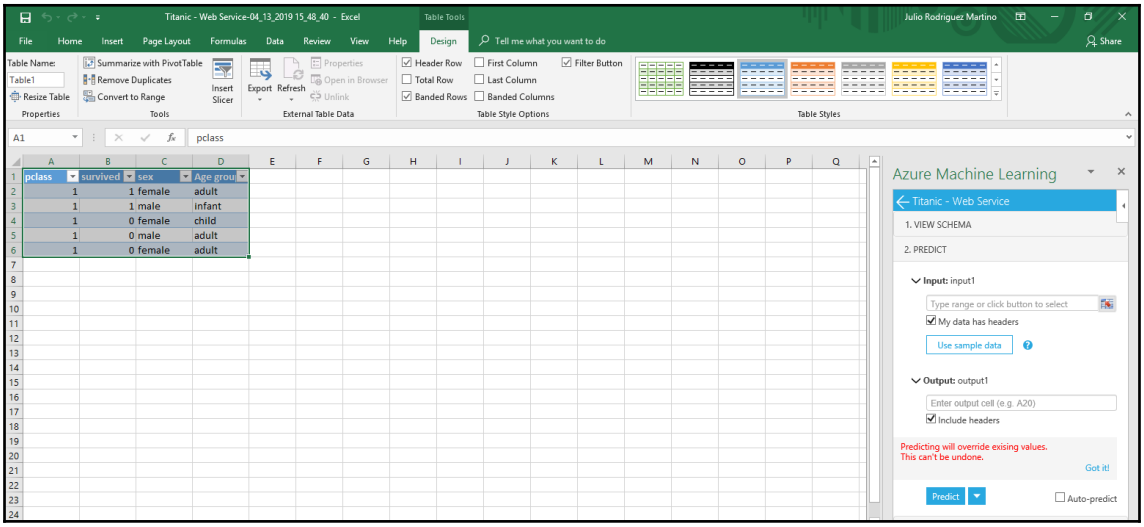

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
Titanic - model training
In draft
Search experiment items
Saved Datasets
Data Format Conversions
Convert to ARFF
Convert to CSV
Convert to Dataset
Convert to SVMLight
Convert to TSV
Data Input and Output
Enter Data Manually
Export Data
Import Data
titanic_small.csv
Two-Class Boosted Decision...
Train Model
Delete
Copy
Cut
Paste
Trained model
View Log
Edit Comment
Run selected
Help
Download
Save as Dataset
Save as Trained Model
Save as Transform
Visualize
Generate Data Access Code...
Open in a new Notebook
Properties
Project
Train Model
Label column
Selected columns:
Train a previously created classification or regression model
(more help...)
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
PUBLISH TO GALLERY

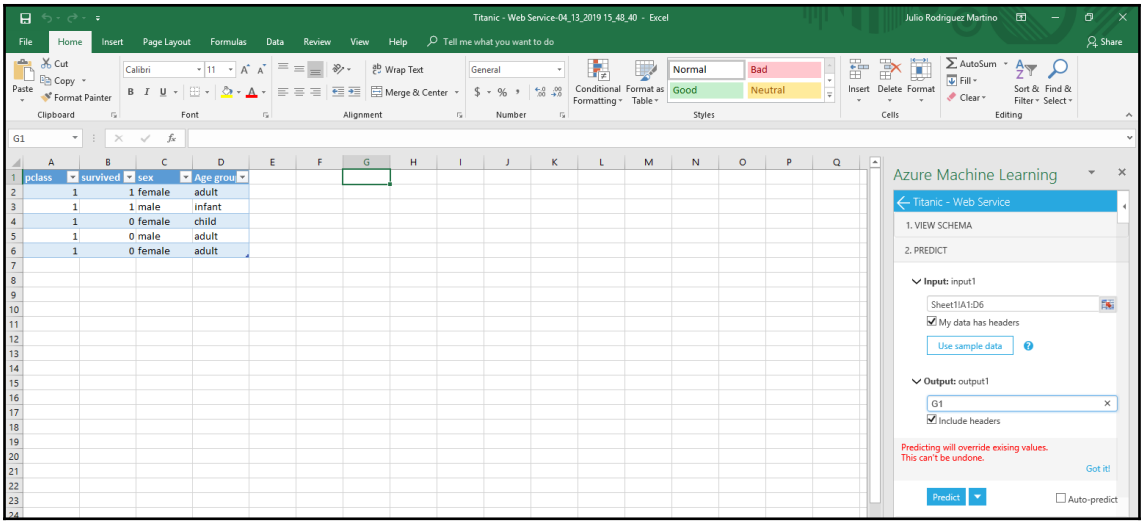

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
Titanic - Web service
In draft
Draft saved at 17:06:19
Search experiment items
Initialize Model
Score
Train
OpenCV Library Modules
Python Language Modules
R Language Modules
Statistical Functions
Text Analytics
Time Series
Web Service
Input
titanic_small.csv
Web service input
Titanic trained Decision Tree
Score Model
Web service output
Properties
Project
Web service output
Quick Help
Web service output port.
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
DEPLOY WEB SERVICE
PUBLISH TO GALLERY

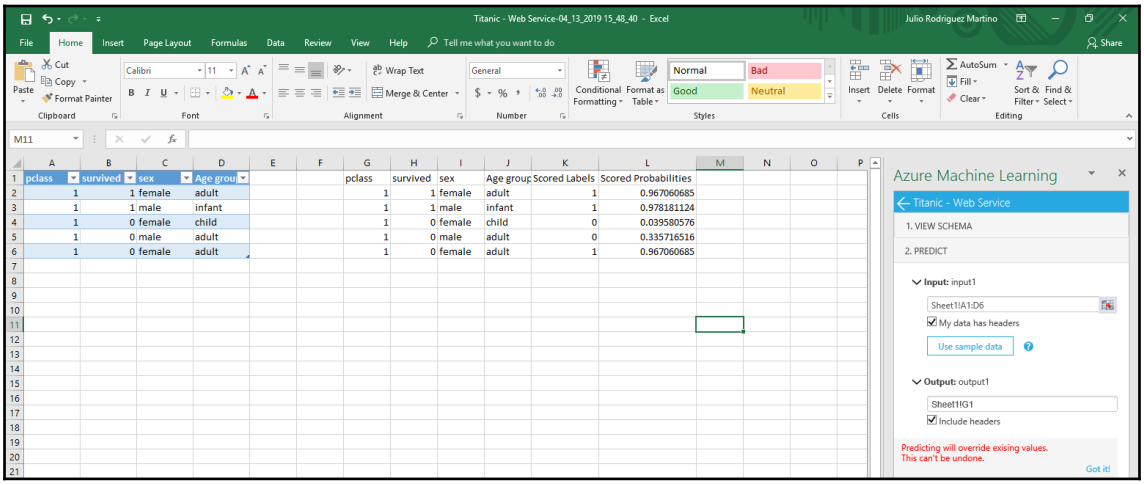

Microsoft Azure Machine Learning Studio
juliorm-Free-Workspace
titanic - web service
DASHBOARD
CONFIGURATION
General
New Web Services Experience preview
Published experiment
View snapshot
View latest
Description
No description provided for this web service.
API key
ZkHUG3b8tZwvfbGoloH/e74LAFVQl/6a/LfozAOb9lQrObq7MVXSjwXPrQOz2CJwKtXtlm+pzjvnlhElbZuNLw==
Default Endpoint
API HELP PAGE
TEST
APPS
LAST UPDATED
REQUEST/RESPONSE
Test
Test preview
Excel 2013 or later | Excel 2010 or earlier workbook
4/10/2019 5:10:15 PM
BATCH EXECUTION
Test preview
Excel 2013 or later workbook
4/10/2019 5:10:15 PM
NEW
DELETE


Titanic - Web Service-04_13_2019 15_48_40 - Excel
Julio Rodriguez Martino
File
Home
Insert
Page Layout
Formulas
Data
Review
View
Help
Tell me what you want to do
Share
Paste
Cut
Copy
Format Painter
Clipboard
Calibri
11
Font
Wrap Text
Merge & Center
Alignment
General
Number
Conditional Formatting
Format as Table
Normal
Bad
Good
Neutral
Styles
Insert
Delete
Format
Cells
AutoSum
Fill
Clear
Sort & Filter
Find & Select
Editing
A1
Azure Machine Learning
Web Services
Titanic - Web Service
set up or delete this web service to predict
Add Web Service
Auto-predict
Predict All

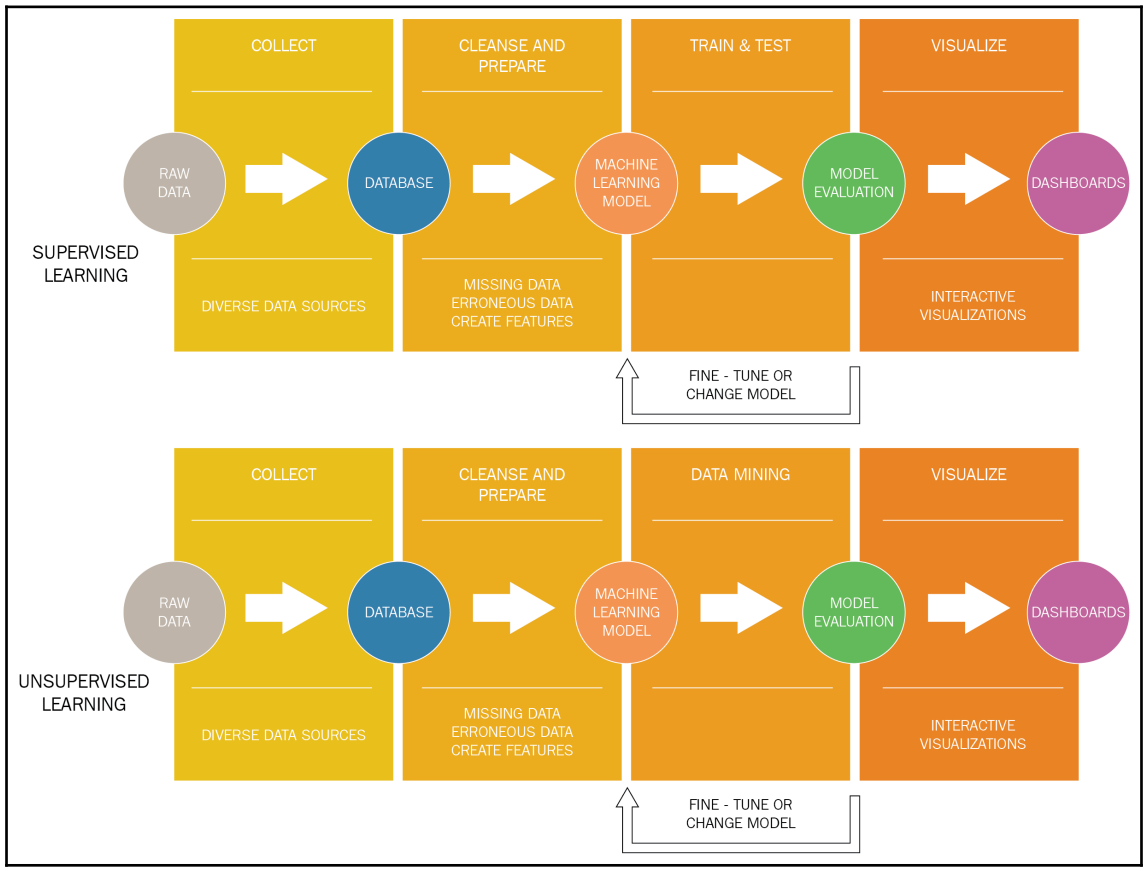

Titanic - Web Service-04_13_2019 15_48_40 - Excel
Julio Rodriguez Martino
Azure Machine Learning
Titanic - Web Service
1. VIEW SCHEMA
2. PREDICT
Input: input1
Type range or click button to select
My data has headers
Use sample data
Output: output1
Enter output cell (e.g. A20)
Include headers
Predicting will override exising values.
This can't be undone.
Got it!
Predict
Auto-predict

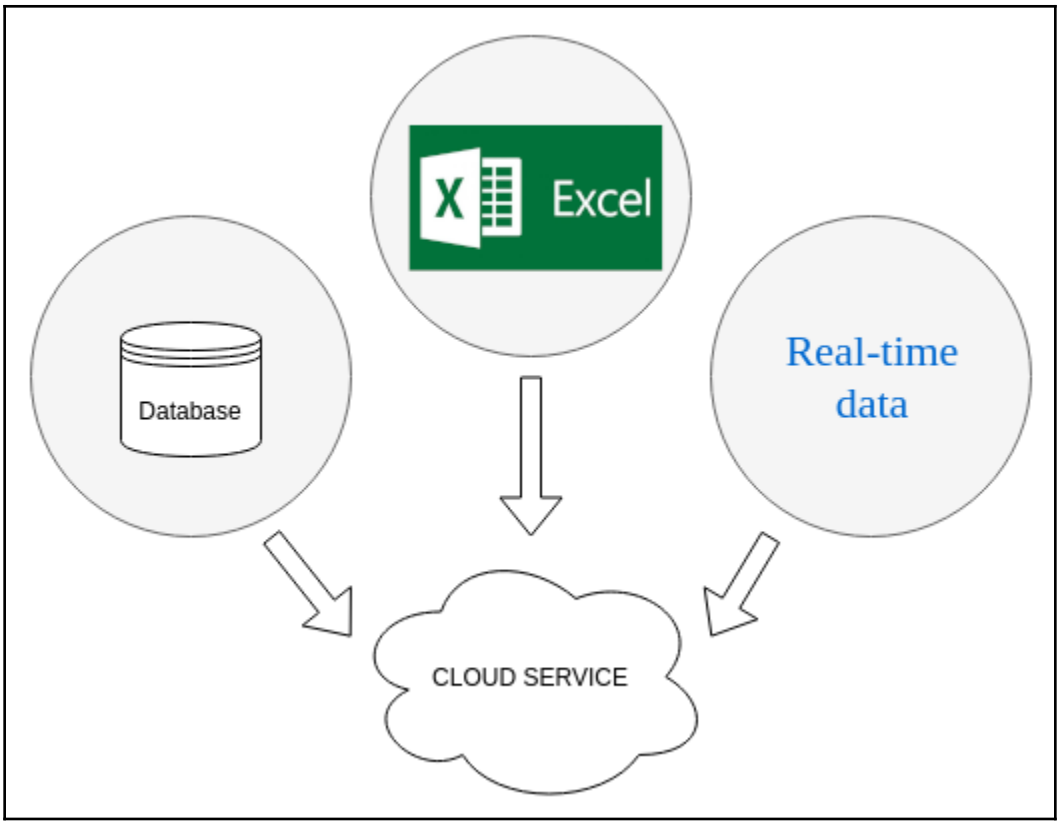

Titanic - Web Service-04_13_2019 15_48_40 - Excel
Julio Rodriguez Martino
pclass survived sex Age group
1 1 female adult
1 1 male infant
1 0 female child
1 0 male adult
1 0 female adult
Azure Machine Learning
Titanic - Web Service
1. VIEW SCHEMA
2. PREDICT
Input: input1
Type range or click button to select
My data has headers
Use sample data
Output: output1
Enter output cell (e.g. A20)
Include headers
Predicting will override exising values.
This can't be undone.
Got it!
Predict
Auto-predict

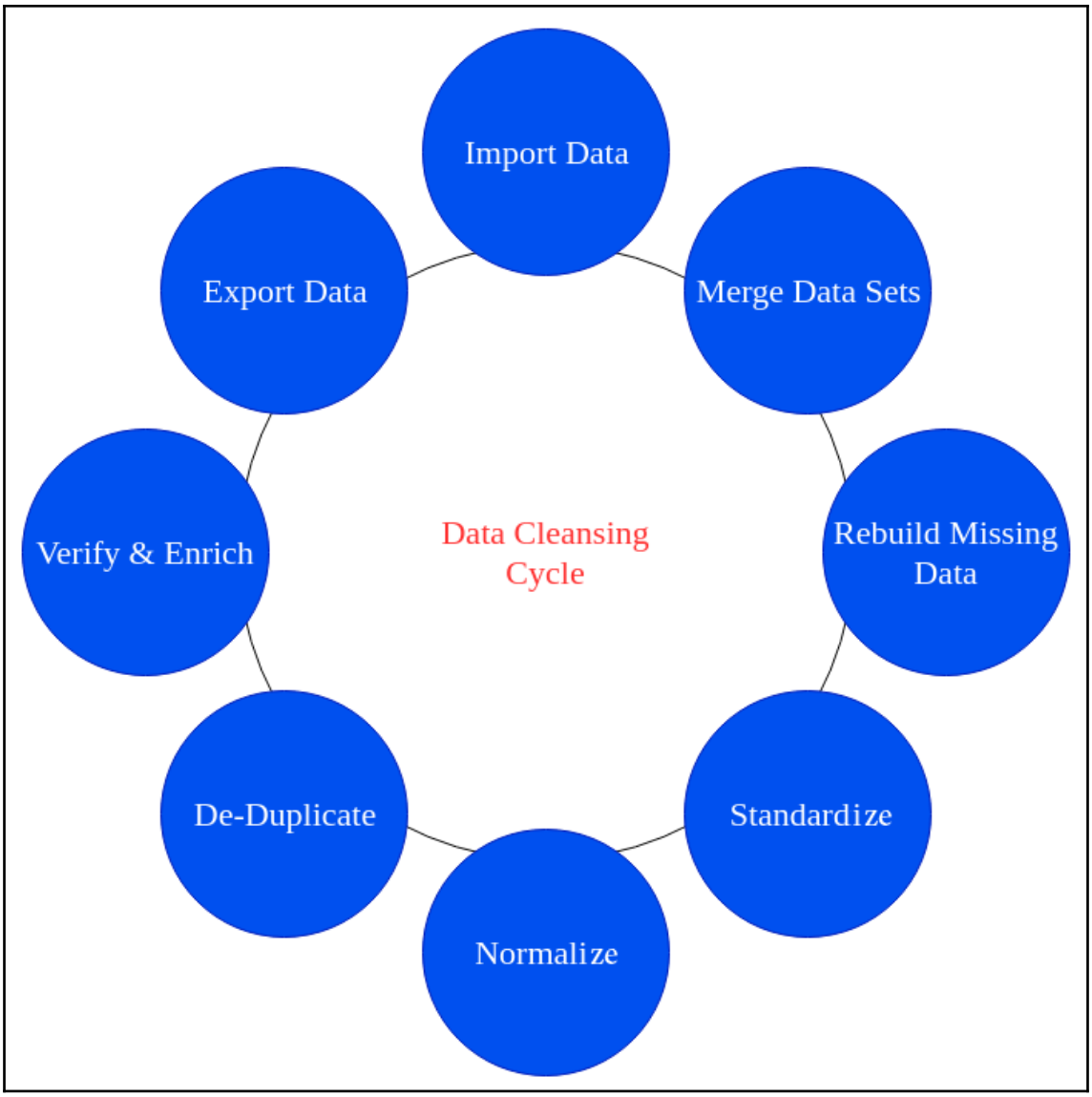

Titanic - Web Service-04_13_2019 15_48_40 - Excel
Julio Rodriguez Martino
Azure Machine Learning
Titanic - Web Service
1. VIEW SCHEMA
2. PREDICT
Input: input1
Sheet1!A1:D6
My data has headers
Use sample data
Output: output1
G1
Include headers
Predicting will override exising values.
This can't be undone.
Got it!
Predict
Auto-predict

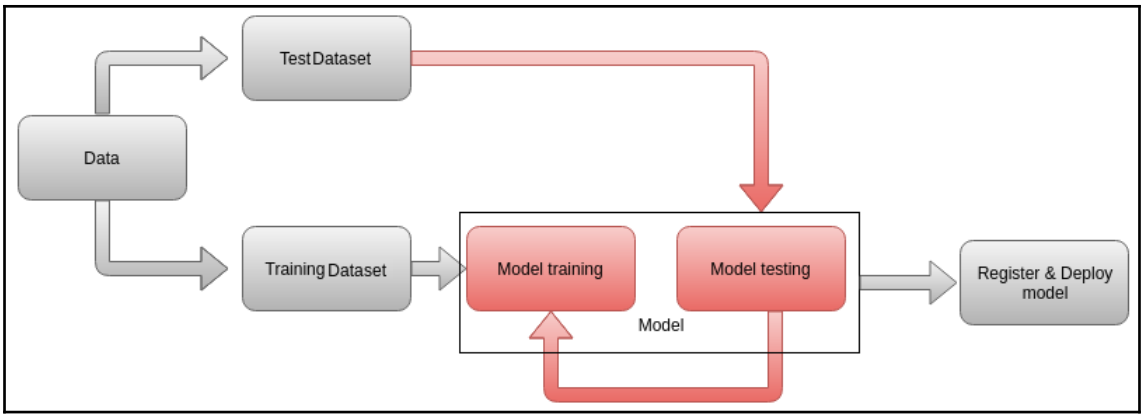

Titanic - Web Service-04_13_2019 15_48_40 - Excel
Julio Rodriguez Martino
Azure Machine Learning
Titanic - Web Service
1. VIEW SCHEMA
2. PREDICT
Input: input1
Sheet1!A1:D6
My data has headers
Use sample data
Output: output1
Sheet1!G1
Include headers
Predicting will override exising values.
This can't be undone.
Got it!

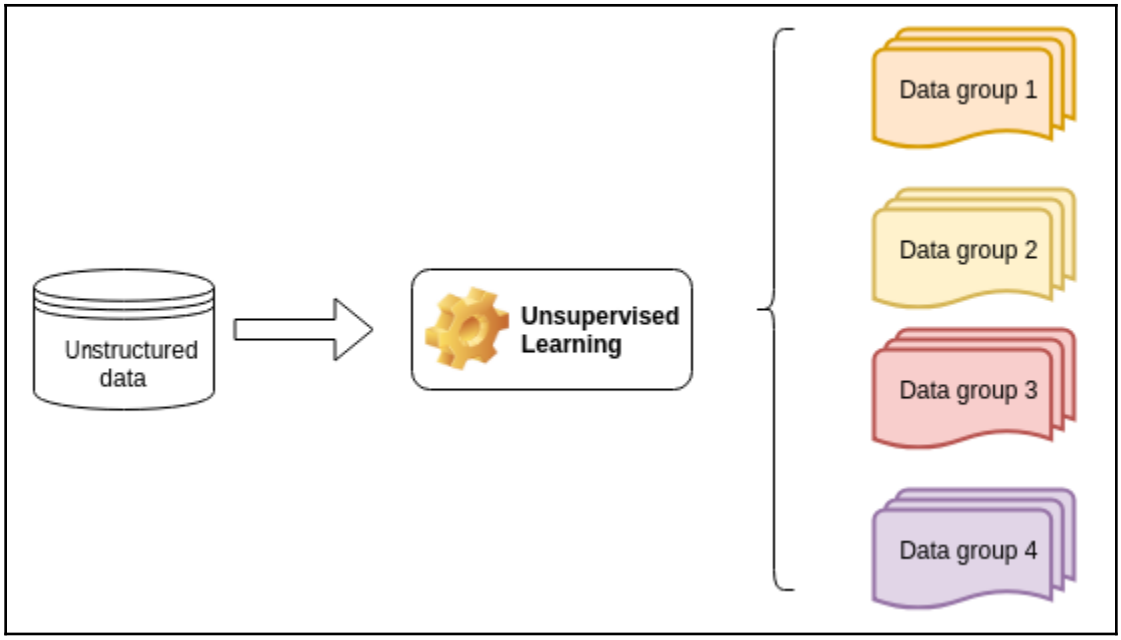
Chapter 11: The Future of Machine Learning

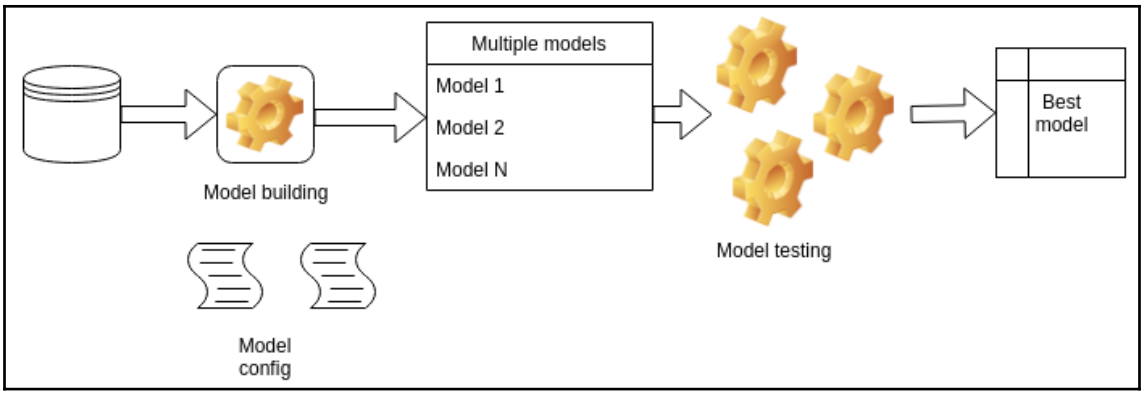



Excel
Database
Real-time
data
CLOUD SERVICE


Import Data
Merge Data Sets
Rebuild Missing Data
Standardize
Normalize
De-Duplicate
Verify & Enrich
Export Data
Data Cleansing Cycle


Test Dataset
Data
Training Dataset
Model training
Model testing
Model
Register & Deploy model


Unstructured data
Unsupervised Learning
Data group 1
Data group 2
Data group 3
Data group 4


Multiple models
Model 1
Model 2
Model N
Best model
Model building
Model testing
Model config