



*Value Through Professional
Asset Management*

2017 National Education Seminar

YOUR EDUCATIONAL TRACK TO SUCCESS

July 31 - August 3

JW MARRIOTT DESERT RIDGE RESORT & SPA | PHOENIX, AZ





Excel Skills for Fleet Managers

Part 2:

Pivot Tables, Outliers, and Data Filtering and
Substitution

Presented by:

Scott Conlon, Mercury Associates

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



2017 National Education Seminar

YOUR EDUCATIONAL TRACK TO SUCCESS

July 31 - August 3

JW MARRIOTT DESERT RIDGE RESORT & SPA | PHOENIX, AZ



Value Through Professional
Asset Management

What does a pivot table do?

- If you just joined us, it creates a summary table that is derived from a larger table of data
- If you were here for Part 1 this morning, it makes the same table that was shown in that session, but with no formulas and in a fraction of the time.



FILE HOME **INSERT** PAGE LAYOUT FORMULAS DATA REVIEW

PivotTable Recommended Table Pictures Online Shapes SmartArt Screenshot
PivotTables PivotTables Pictures Pictures Illustrations

A1 X ✓ fx Stock Number

	A	B	C	D
1	Stock Numb =	Account Name =	Transaction Da =	Vehicle Description
2	X375562	Wisconsin	9/11/2015	
3	X375562	Wisconsin	9/29/2015	
4	X375561	Wisconsin	9/9/2015	
5	X375561	Wisconsin	9/9/2015	

Select your data, then click **Insert > PivotTable**

Create PivotTable

Create PivotTable

Choose the data that you want to analyze

☒ Select a table or range

Table/Range: Sheet1!\$A:\$G

☐ Use an external data source

Choose where you want the PivotTable report to be placed

☒ New Worksheet

☐ Existing Worksheet

Location:

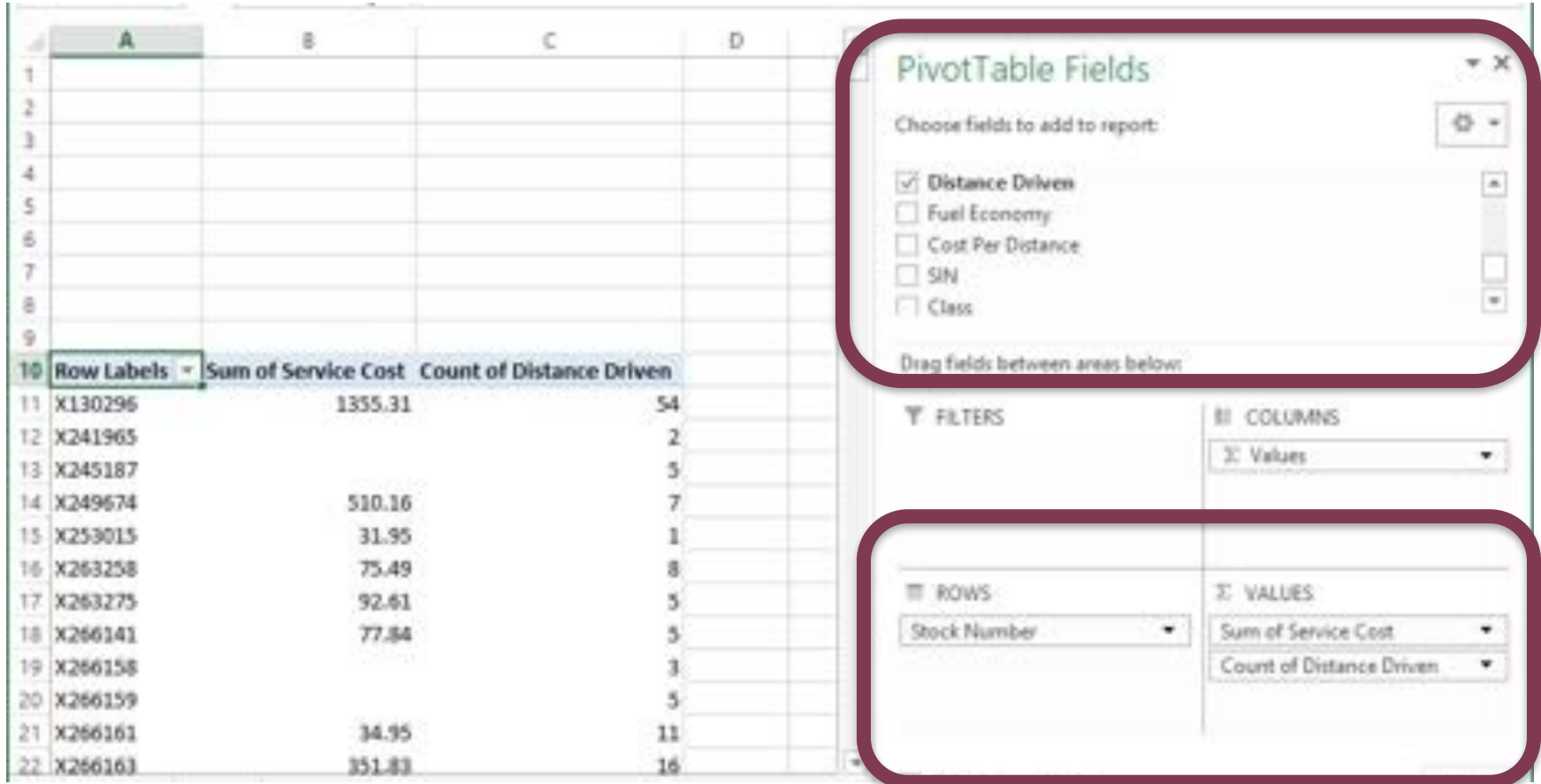
Choose whether you want to analyze multiple tables

☐ Add this data to the Data Model

OK Cancel

Copyright © 2017 Mer

Select PivotTable Fields by dragging and dropping:
Stock Number into the **Rows** area
Service Cost & Distance Driven into the **Values** area



The screenshot shows an Excel PivotTable and the PivotTable Fields task pane. The PivotTable is located in the range A10:D22. The task pane is on the right side of the screen.

PivotTable Fields Task Pane:

- Choose fields to add to report:**
 - ☒ Distance Driven
 - ☐ Fuel Economy
 - ☐ Cost Per Distance
 - ☐ SIN
 - ☐ Class
- Drag fields between areas below:**
 - FILTERS:** (Empty)
 - COLUMNS:** (Empty)
 - ROWS:** Stock Number
 - VALUES:** Sum of Service Cost, Count of Distance Driven

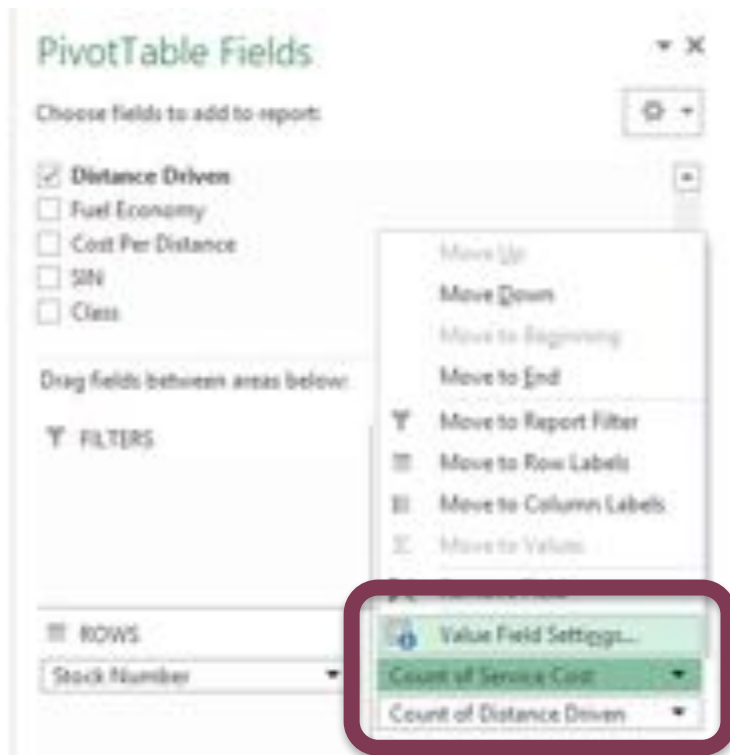
PivotTable Data:

Row Labels	Sum of Service Cost	Count of Distance Driven
X130296	1355.31	54
X241965		2
X245187		5
X249674	510.16	7
X253015	31.95	1
X263258	75.49	8
X263275	92.61	5
X266141	77.84	5
X266158		3
X266159		5
X266161	34.95	11
X266163	351.83	16

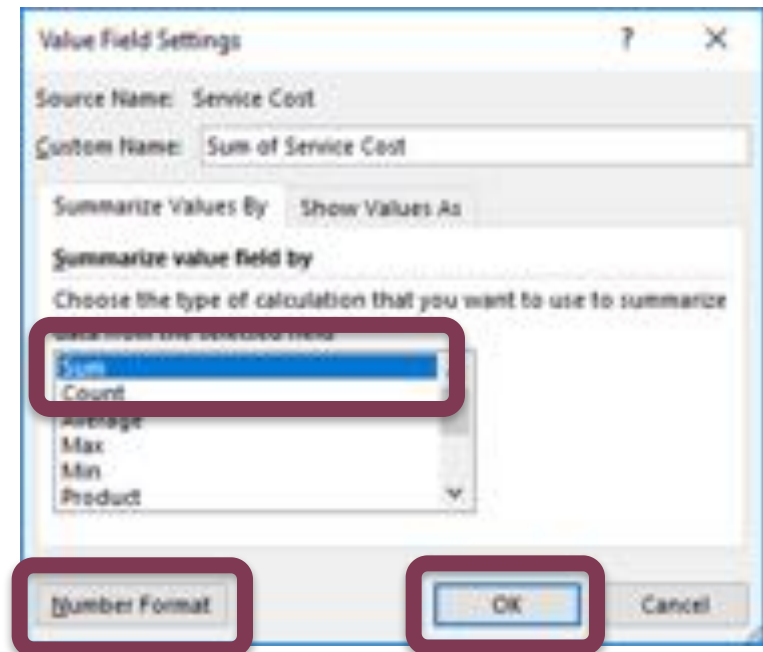
Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Change Calculation from Count to Sum

Select Value Field Settings



Change from Count to Sum, then change Number Format



Copyright © 2017 Mercury Associates, Inc. All rights reserved.



Pivot Table Result

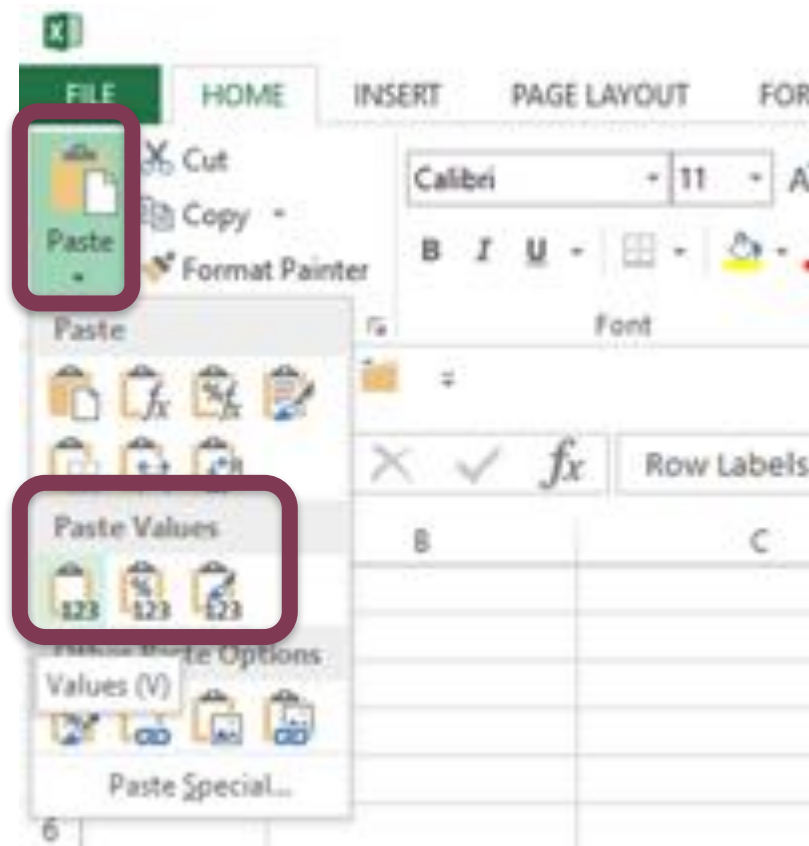
The table below now shows the **SUM** of the distance driven (12,503 for the first vehicle) rather than the **COUNT** (54 mileage entries)

8			
9			
10	Row Labels	Sum of Service Cost	Sum of Distance Driven
11	X130296	\$ 1,355.31	12503
12	X241965		327
13	X245187		726
14	X249674	\$ 510.16	1719
15	X253015	\$ 31.95	164642
16	X263258	\$ 75.49	2118
17	X263275	\$ 92.61	1140
18	X266141	\$ 77.84	1217
19	X266158		696

Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Explore the Data

Copy your pivot table and paste it as values into a new spreadsheet tab



Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Calculate Cost per Mile

D16 $\times$ $\checkmark$ f_x **=B16/C16**

	A	B	C	D	E
15	Row Label	Sum of Service Co	Sum of Distance Drive	CPM	
16	X130296	\$ 1,355.31	12503	\$ 0.11	
17	X241965		327	\$	
18	X245187		726	\$ -	
19	X249678	\$ 510.16	1719	\$ 0.30	
20	X253015	\$ 31.95	164642	\$ 0.00	
21	X263258	\$ 75.49	2118	\$ 0.04	
22	X263275	\$ 92.61	1140	\$ 0.08	
23	X266141	\$ 77.84	1217	\$ 0.06	

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



Assess Data Integrity

This data contains some obvious errors that are likely caused by improper meter entry; additional reported mileage will cause CPM to appear lower than it should

While this error is easy to spot, there are some commonly accepted techniques to identifying these types of outliers in the data

16	X130296	\$	1,355.31	12503	\$	0.11
17	X241965			327	\$	-
18	X245187			726	\$	-
19	X249678	\$	510.16	1719	\$	0.30
20	X253015	\$	31.95	164642	\$	0.00
21	X263258	\$	75.49	2118	\$	0.04
22	X263275	\$	92.61	1140	\$	0.08
23	X266141	\$	77.84	1217	\$	0.06

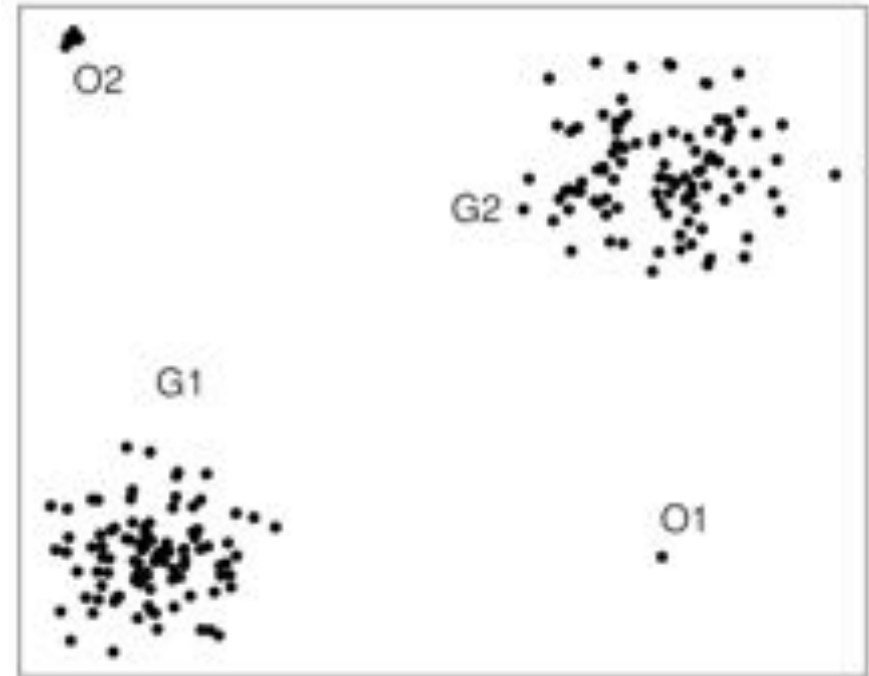
Copyright © 2017 Mercury Associates, Inc. All rights reserved.

What is an outlier?

Outlier:

Statistical data which is extremely different from the others in the same sample.

There is no rigid definition of what constitutes an outlier, but there are some simple methods of determination.



Sources: <http://www.businessdictionary.com/definition/outlier.html>

https://en.wikipedia.org/wiki/Outlier#/media/File:Two-dimensional_Outliers_Example.png

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



Outlier Identification

- Calculate the “Five Number Summary”
 - Maximum value
 - Third quartile
 - Median value
 - First quartile
 - Minimum value
- Calculate interquartile range and fences



Copyright © 2017 Mercury Associates, Inc. All rights reserved.
Source: <http://mvpprograms.com/help/mvstats/graphics/WhatAreBoxAndWhiskerPlots>

Five Number Summary

The formulas for maximum, median, and minimums are shown below

Note: the dollar signs (\$) that are embedded in the formulas are called **Reference Locks**. These are optional in this case, but will make it more convenient to fill the formulas to adjacent columns

C1					
			X	✓	f_x
					=MAX(C\$16:C\$8701)
	A				B
1	Max Value				=MAX(B\$16:B\$8701)
2	Median				=MEDIAN(B\$16:B\$8701)
3	Min Value				=MIN(B\$16:B\$8701)
4					
5	Third Quartile				=QUARTILE.EXC(B\$16:B\$8701,3)
6	First Quartile				=QUARTILE.EXC(B\$16:B\$8701,1)

Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Five Number Summary

The calculated values are shown below

The median values seem reasonable, but the maximum values (\$11,666 and 173M miles for the sums of Service Cost and Distance Driven, respectively) seem unreasonable

A quarter of the vehicles travel over 9,246 miles per year, but also that a quarter of the vehicles travel fewer than 3,607 miles

	A	B	C	D
1	Max Value	\$ 11,666.18	172845580	\$ 444.44
2	Median	\$ 139.85	6022.5	\$ 0.02
3	Min Value	\$ -	1	\$ -
4				
5	Third Quartile	\$ 492.90	9246.25	\$ 0.08
6	First Quartile	\$ 45.18	3607	\$ 0.01
7				

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



2017 National Education Seminar
YOUR EDUCATIONAL TRACK TO SUCCESS



Calculate “Fences”

The “fences” are the boundaries around data that is likely accurate
A “weak” outlier is one that is more than 1.5 times the IQR above/below
the third/first quartile

A “strong” outlier is one that is more than 3 times the IQR above/below
the third/first quartile

Negative mileages (regardless of fence) should also be reviewed

5	Third Quartile	\$	492.90		9246.25
6	First Quartile	\$	45.18		3607
7					
8	Interquartile Range (row 5 - row 6)	\$	447.71		5639.25
9	3 x IQR (row 8 * 3)	\$	1,343.14	\$	16,917.75
10					
11	Upper Fence (row 5 + row 9)	\$	1,836.03	\$	26,164.00
12	Lower Fence (row 6 - row 9)	\$	(1,297.96)	\$	(13,310.75)

Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Apply to Detailed Data

	A	N	AB
1	Max Value	\$ 8,840.99	55555550
2	Median	\$ 46.45	279
3	Min Value	\$ (1,905.25)	1
4			
5	Third Quartile	\$ 118.22	357
6	First Quartile	\$ 24.00	197
7			

Once the Five Number Summary is applied to the transactional data, it is obvious that the maximum mileage entry is 55,555,550

This resulted from an employee not taking mileage entry seriously
These inaccuracies will need to be taken into account when trying to derive insights from the transactional data

Copy

Filtering and Substitution

What do I do with the bad data now that I have found it?

- For ALD/VLD, locate, research, and correct errors in agency FMIS
- For fleet analysis:
 - For averages: filter data out and move on
 - For totals: substitute data with median value

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



2017 National Education Seminar
YOUR EDUCATIONAL TRACK TO SUCCESS



Locate Spurious Entries

The screenshot shows an Excel spreadsheet with a 'Number Filters' dialog box open. The dialog box has two red boxes: one around 'Number Filters' and another around 'Greater Than...'. The spreadsheet shows statistical calculations in rows 1-14 and data in rows 15-17.

	A		AC	AD	
1	Max Value				
2	Median				
3	Min Value				
4					
5	Third Quartile				
6	First Quartile				
7					
8	Interquartile Range (row 5				
9	1.5 x IQR (row 8 * 1.5)				
10					
11	Upper Fence (row 5 + row 9				
12	Lower Fence (row 6 - row 9				
13					
14					
15	Stock Number	Account Name	Distance Drive	Fuel Economy	Cost Per Dista
16	X375562	Wisconsin	8	0.29	8
17	X375562	Wisconsin	415	18.15	0

Copyright

Flag and Filter Erroneous Data



The **IF** function checks whether a condition is met, and returns one value if true and another value if false. Here, the condition is the distance driven being greater than the fence

=IF(logical_test,[value_if_true],[value_if_false])

In plain language, this could be written as “if the distance driven is unreasonable, say ‘discard’, if not, say ‘use’”

9	3 x IQR (row 8 * 3)	480	
10			
11	Upper Fence (row 5 + row 9)	837	
12	Lower Fence (row 6 - row 9)	-283	
13			
14			
15	Stock Number	Distance Driven	Distance Driven Flag
16	X1375562		=IF(AB16>\$AB\$11,"Discard","Use")

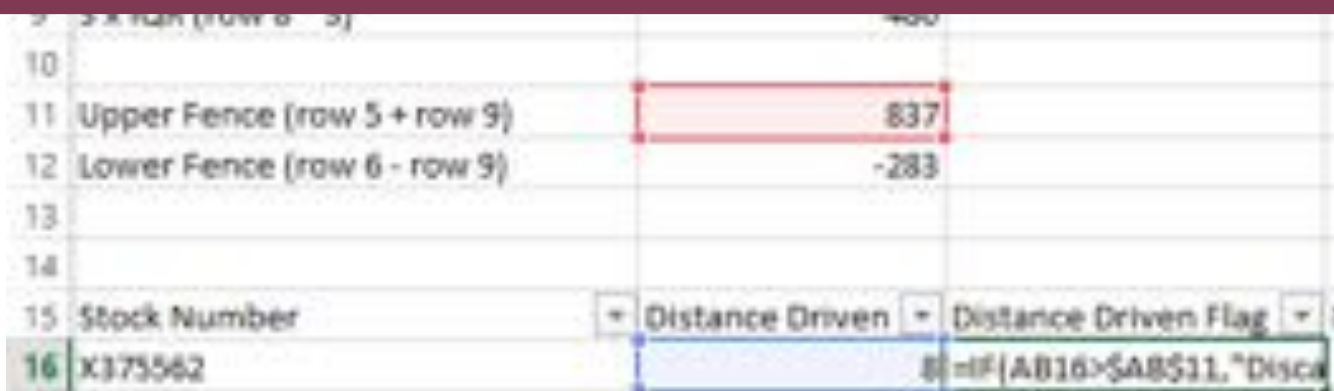
Copyright © 2017 / Mercury Associates, Inc. All rights reserved.

Flag and Filter Erroneous Data



	A	AB	AC
1	Max Value	55555550	
2	Median	320	

Reference Locks change a cell reference from relative to absolute. By changing **AB11** to **\$AB\$11**, the formula can be filled to the rows below without the reference to AB11 changing to AB12 and so on. Dollar signs that precede the letters lock the column reference, while letters that precede the numbers lock the row reference.



9	3 x 12in (row 8 - 5)		400
10			
11	Upper Fence (row 5 + row 9)	837	
12	Lower Fence (row 6 - row 9)	-283	
13			
14			
15	Stock Number	Distance Driven	Distance Driven Flag
16	X375562		=IF(AB16>\$AB\$11,\"Discard\",\"Use\")

Copyright © 2017 / Mercury Associates, Inc. All rights reserved.

Filter Pivot Table Data

The screenshot displays an Excel PivotTable and the PivotTable Fields task pane. The PivotTable is filtered by 'Distance Driven Flag' (All). The task pane shows the following configuration:

- PivotTable Fields:** Choose fields to add to report.
- Distance Driven:** ☒ Distance Driven, ☒ Distance Driven Flag
- Drag fields between areas below:**
- FILTERS:** Distance Driven Flag
- ROWS:** Stock Number
- VALUES:** Sum of Service Cost, Sum of Distance Driven

Row Labels	Sum of Service Cost	Sum of Distance Driven
X299590	\$ 1,291.37	172845580
X352545		23031009
X281939	\$ 144.56	15949316
X360929	\$ 1,213.42	9448322
X360615	\$ 332.49	8211527
X335952	\$ 50.78	6177728
X361090	\$ 83.95	4877129
X359641		3648207
X373338		2868338
X335766	\$ 144.75	2501201
X353225	\$ 549.48	1723405
X346206	\$ 240.91	1706183
X335423	\$ 113.96	1588428
X352650		1430369
X338141	\$ 389.11	1419131
X352136	\$ 282.86	1416473
X330651	\$ 6.00	1406078

Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Filtered Results

C11			46744
	A	B	C
7			
8	Distance Driven Flag	Use	.T
9			
10	Row Labels	Sum of Service Cost	Sum of Distance Driven
11	X365378	\$ 3,995.74	46744
12	X369587	\$ 2,378.61	40674
13	X368518	\$ 3,009.60	36942
14	X359132	\$ 1,685.34	33997
15	X357297	\$ 1,330.87	31689
16	X319112	\$ 1,251.04	30119
17	X368515	\$ 1,339.68	29413
18	X352409	\$ 978.35	29216
19	X321215	\$ 834.33	28745
20	X359873	\$ 972.49	28020
21	X369030	\$ 7.00	27919
22	X368041	\$ 806.23	27840
23	X344681	\$ 395.09	27517
24	X336394	\$ 1,350.19	27509
25	X368521	\$ 786.15	27350
26	X352727	\$ 2,639.02	27149

The **Sum of Distance Driven** now shows that the maximum mileage per year for an individual vehicle is **46,744**
This is still a high number, but is more feasible than before

FILTERS
Distance Driven Flag

COLUMNS
Σ Values

ROWS
Stock Number

VALUES
Sum of Service Cost
Sum of Distance Driven

Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Substitute Median Value

If I need to estimate the total distance driven for the fleet, I will need to substitute “good” data in for the “bad” data that was previously excluded

Using the same IF function, but instead of “**Discard**” and “**Use**”, use the median (**\$AB\$2**) and the actual (**AB16**)

Then fill the formula down

=IF(AB16>\$AB\$11,\$AB\$2,AB16)			
Max Value	55555550	837	
Median	279	276	
Min Value	1	0	
Third Quartile	357	352	
First Quartile	197	190	
Interquartile Range (row 5 - row 6)	160	162	
3 x IQR (row 8 * 3)	480	486	
Upper Fence (row 5 + row 9)	837	838	
Lower Fence (row 6 - row 9)	-283	-296	
Stock Number	Distance Driven	Normalized Distance Driven	
X375562	8	=IF(AB16>\$AB\$11,\$AB\$2,AB16)	
X375562	415	415	
X375561	250	250	
X375561	250	250	
X375561	3	3	
X375561	335	335	
X375560	14	14	

Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Substituted Results

Excel screenshot showing a PivotTable and the PivotTable Fields task pane.

PivotTable Fields Task Pane:

- Choose fields to add to report:
 - ☐ Distance Driven Flag
 - ☒ Normalized Distance Driven
 - ☐ Fuel Economy
 - ☐ Cost Per Distance
- FILTERS:** (Empty)
- ROWS:** Stock Number
- COLUMNS:** Σ Values
- VALUES:**
 - Σ Sum of Service Cost
 - Σ Sum of Normalized Distance...

PivotTable Data:

Row Labels	Sum of Service Cost	Sum of Normalized Distance Driven
X130296	\$ 1,355.31	12503
X241965		327
X245187		726
X249674	\$ 510.16	1719
X253015	\$ 31.95	279
X263258	\$ 75.49	2118
X263275	\$ 92.61	1140
X266141	\$ 77.84	1217
X266158		696
X266159		1221
X266161	\$ 34.95	3742
X266163	\$ 351.83	5087
X266164	\$ 37.40	2947
X267107	\$ 1,508.99	8343
X267108		1201

Copyright © 2017 Mercury Associates, Inc. All rights reserved.

Listing of Skills Covered

Excel Functions Used in this Presentation

- Filters – sift through data to find outliers
- Reference Locks – drag and fill formulas
- Pivot Tables – easily sum or average large data by vehicle or other characteristic
- Five Number Summary – MAX, MEDIAN, MIN, QUARTILE.EXC
- Conditional Logic – IF statements

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



2017 National Education Seminar
YOUR EDUCATIONAL TRACK TO SUCCESS



Homework

- Pull in data from another sheet – INDEX & MATCH or VLOOKUP

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



2017 National Education Seminar
YOUR EDUCATIONAL TRACK TO SUCCESS



For More Information

Scott Conlon

Mercury Associates, Inc.

sconlon@mercury-assoc.com

301 519 0535,1027

www.mercury-assoc.com

Copyright © 2017 Mercury Associates, Inc. All rights reserved.



2017 National Education Seminar
YOUR EDUCATIONAL TRACK TO SUCCESS