

DATA CLEANING & PIVOT TABLE WITH EXCEL

By Hafis

Excel for data analysis



Project Goal

- To produce clean dataset using excel.
- Use Pivot Table function in excel.



Information

Consist of two excel file :

- **Montgomery_Fleet_Equipment_Inventory_FA_PART_1_START**
(4 Column, 63 rows)
- **Montgomery_Fleet_Equipment_Inventory_FA_PART_2_START**
(3 Column, 50 rows



Data Cleaning

- Column widths:

Sort out the widths of all columns so that the data is clearly visible in all cells.



The image shows two side-by-side screenshots of Microsoft Excel. The left screenshot, titled 'Montgomery_Fleet_Equipment_Inventory_FA_PART_1_START - Excel', displays a spreadsheet with several columns of data. The columns are relatively narrow, causing some data to be truncated or less readable. The right screenshot, titled 'Montgomery_Fleet_Equipment_Inventory_FA_PART_1_END CLEANED - Excel', shows the same data after cleaning. The columns are wider, making the data clearly visible in all cells. A blue arrow points from the start state to the end state, indicating the process of adjusting column widths.

Department	Equipment Class	Equipment Count
Board of Elections	Van	1
Board of Elections	Off Road Vehicle Equipment	2
Circuit Court	SUV	1
Community Engagement Cluster	Pick Up Trucks	8
Community Engagement Cluster	Off Road Van	7
Community Engagement Cluster	SUV	2
Community Facilities	Sedan	1
Community Facilities	Sedan	1
Consumer Protection	Sedan	1
Correction and Rehabilitation	Off Road Van	3
Correction and Rehabilitation	Public Safety Sedan	1
Correction and Rehabilitation	Public Safety SUV	1
Correction and Rehabilitation	SUV	1
Correction and Rehabilitation	Pick Up Trucks	1
Correction and Rehabilitation	Van	1
Correction and Rehabilitation	Sedan	1
Correction and Rehabilitation	CUV	1
County Executives Office	Sedan	1
County Executives Office	Public Safety SUV	1
Economic Development	SUV	1
Environmental Protection	SUV	1
Environmental Protection	Sedan	1

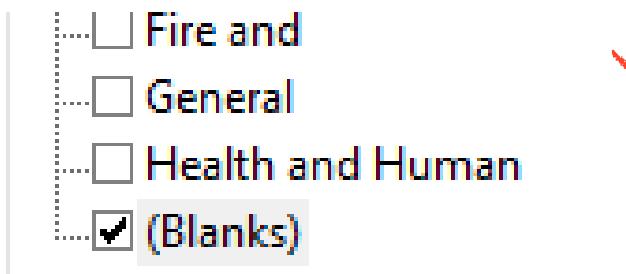
Data Cleaning

- Empty rows:

Use the Filter feature to look for blanks and remove all empty rows from the data.

Apply to the department column.

In real situation we need confirm with the other staff if there any missing value rather than delete it .



	A	B
1	Department	Department Equipment
2	Board of	Elections Van
3	Board of	Elections Off Ro
4	Circuit	Court SUV
5	Community Engagement	Cluster Pick Up
6	Community Engagement	Cluster Off Ro
7	Community Engagement	Cluster SUV
8		
9	Community Use of Public	Facilities Sedan
10	Community Use of Public	Facilities Sedan

	A	B
1	Department	Equipment
2	Board of Elections	Van
3	Board of Elections	Off Road V
4	Circuit Court	SUV
5	Community Engagement Cluster	Pick Up Tr
6	Community Engagement Cluster	Off Road V
7	Community Engagement Cluster	SUV
8	Community Use of Public Facilities	Sedan
9	Consumer Protection	Sedan
10	Correction and Rehabilitation	Off Road V

Data Cleaning

- Duplicate records:

Use either the Conditional Formatting or Remove Duplicates feature to look for and remove any duplicated records from the data.



A screenshot of a spreadsheet application showing rows 54 through 58. The first column contains row numbers (54, 55, 56, 57, 58) and the second column contains the text "Health and Human". Row 58 is entirely blank. A red arrow points from the "Remove Duplicates" button to the bottom of row 58, and a blue arrow points down to the bottom of the slide.

54	Health and Human
55	Health and Human
56	Health and Human
57	Health and Human
58	

A screenshot of a spreadsheet application showing rows 50 through 54. The first column contains row numbers (50, 51, 52, 53, 54) and the second column contains the text "Health and Human". Row 54 is entirely blank. A red arrow points from the "Remove Duplicates" button to the bottom of row 54, and a blue arrow points down to the bottom of the slide.

50	Health and Human
51	Health and Human
52	Health and Human
53	Health and Human
54	



Data Cleaning

- Spelling:

The original source file data has not been checked for errors in the spelling. Check for spelling mistakes in the data and fix them.

abc
✓
Spelling

21	Envirommental
22	Envirommental
23	Envirommental
24	Envirommental
25	Envirommental
26	Envirommental



21	Environmental
22	Environmental
23	Environmental
24	Environmental
25	Environmental
26	Environmental

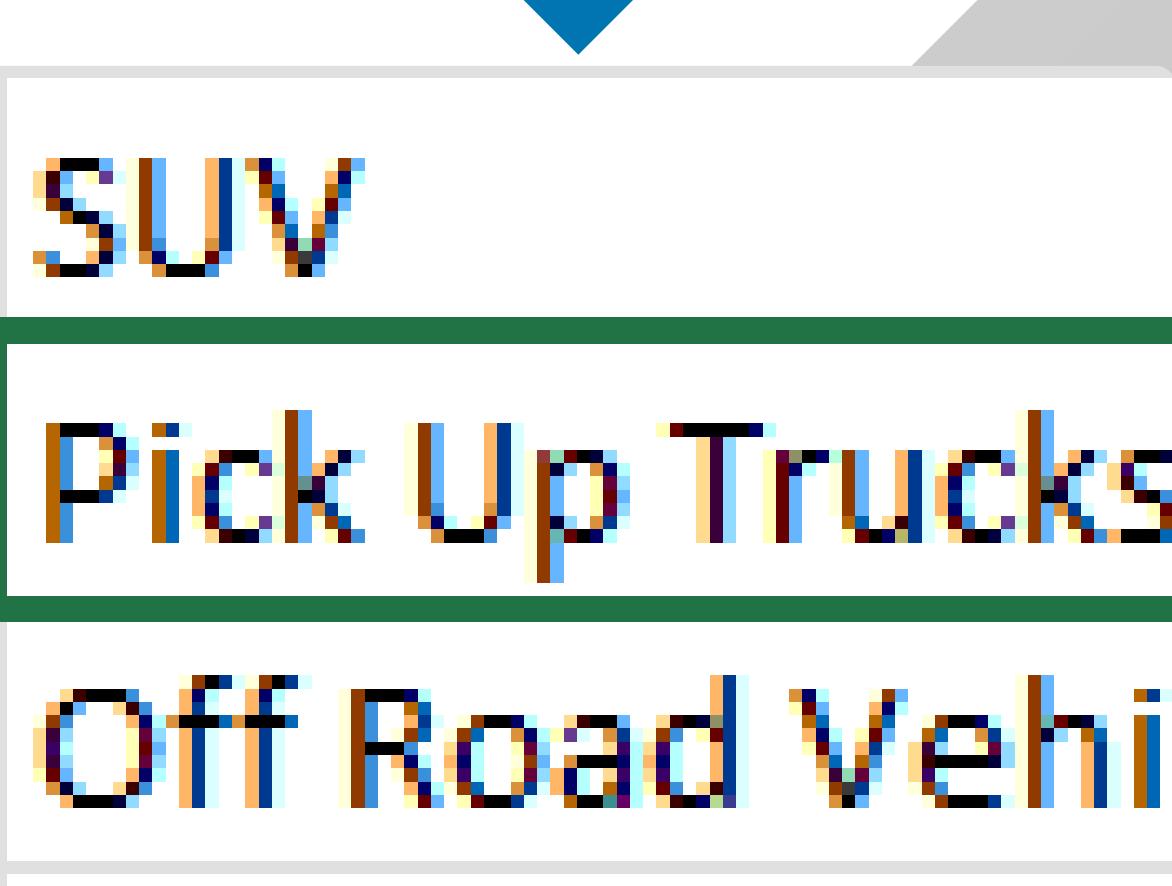
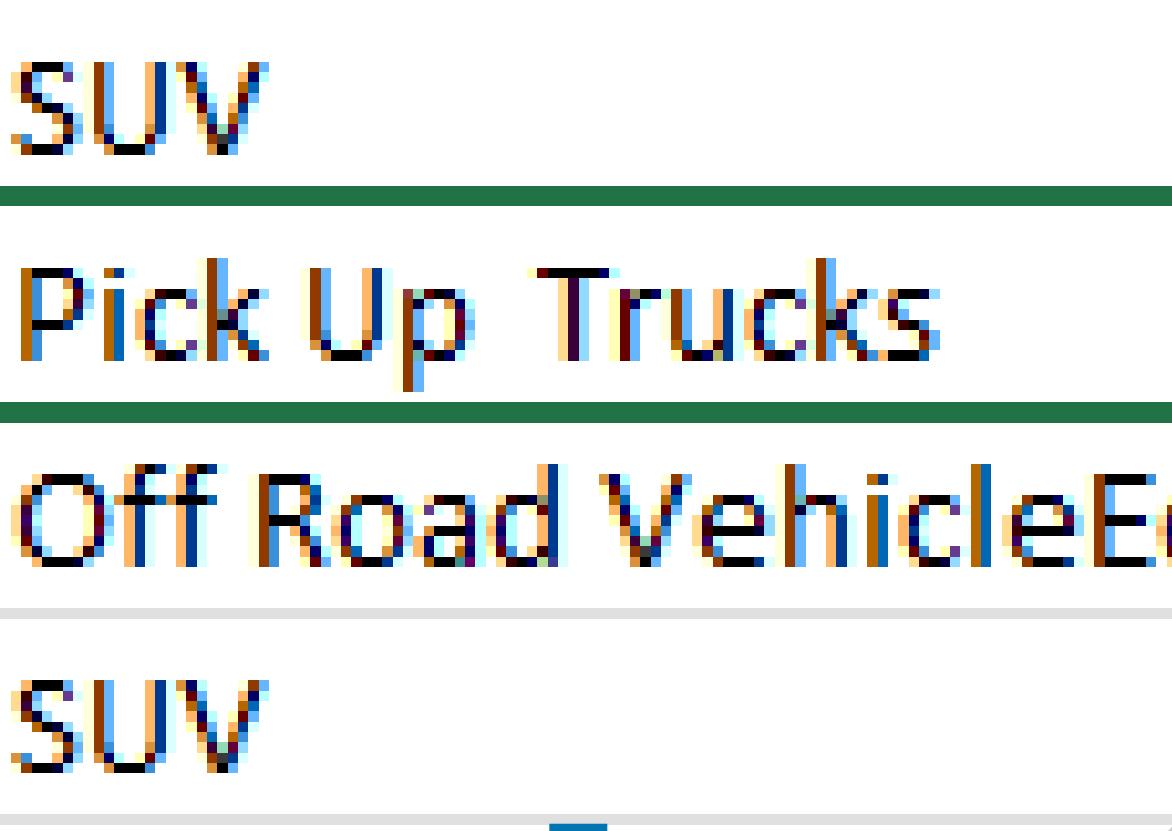
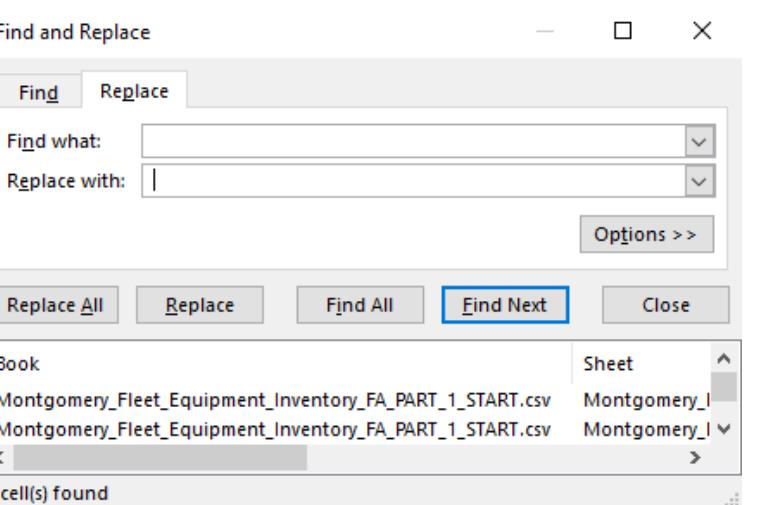


Data Cleaning

- Whitespace:

Use the Find and Replace feature to remove all double-spaces from the data.

2 Space replace with 1 space



Data Cleaning

- Department names:

When the data was converted from its data source, the department names (see correct list below) didn't import correctly and they are now split over two columns in the data. Use Flash Fill to reduce the department names to just one column, and then remove any unnecessary columns.

Insert new column and fill the 2 column department into 1.

1	Department	Department
2	Board of	Elections
3	Board of	Elections
4	Circuit	Court
5	Community Engagement	Cluster



1	Department
2	Board of Elections
3	Board of Elections
4	Circuit Court
5	Community Engagement Cluster



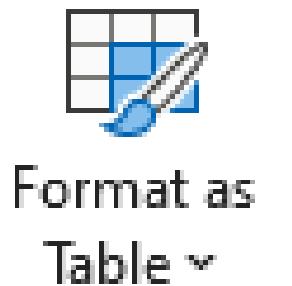
PIVOT TABLE



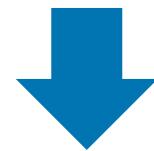
Pivot Table

- Format the data as a table:

Use the Format as Table option to format the data as a table.



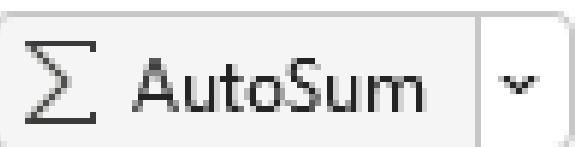
1	Department	Equipment Class
2	Housing and Community Affairs	Pick Up Trucks
3	Housing and Community Affairs	SUV
4	Housing and Community Affairs	Sedan
5	Human Rights	Sedan



1	Department	Equipment Class
2	Housing and Community Affairs	Pick Up Trucks
3	Housing and Community Affairs	SUV
4	Housing and Community Affairs	Sedan
5	Human Rights	Sedan



Pivot Table



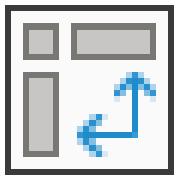
- Use AutoSum to calculate values:

- SUM
- AVERAGE
- MIN
- MAX
- COUNT

SUM Equipment	1582
Average Equipment	32.2857143
Min Equipment	1
Max Equipment	379
Count Equipment	49



Pivot Table



PivotTable

- Create a Pivot Table:

Use the PivotTable feature to create a pivot table that displays the Department field in the Rows section, and the Equipment Count in the Values section, so that the pivot table displays the sum of equipment count by department.

- Sort the pivot table data:

Use the Sort By Value setting on the pivot table to sort it in descending order by the sum of equipment count.



Row Labels	Sum of Equipment Count
Transportation	1221
Permitting Services	109
Sheriffs Office	85
Liquor Control	56
Housing and Community Affairs	45
Recreation	35
Technology Services	16
Libraries	6
State Attorneys Office	5
Human Rights	2
Office Of Homeland Security	1
Public Information Office	1

PivotTable Fields

Choose fields to add to report:

Search

Department

Equipment Class

Equipment Count

More Tables...

Drag fields between areas below:

Filters Columns

Rows Values

Department Sum of Equipment Co...

Pivot Table

Analyze data in the pivot table:

- In pivot table 2 add the Equipment Class field below the Department field so that the different vehicle types appear under each department with their respective counts.
- Collapse all fields except the top one - Transportation
- In pivot table 3 add the Equipment Class field above the Department field so that the different vehicle types appear first, with the different departments listed underneath each vehicle type with their respective counts.
- Collapse all fields except the top one - CUV

Row Labels	Sum of Equipment Count
Transportation	1221
CUV	5
Heavy Duty	248
Medium Duty	98
Off Road Vehicle Equipment	276
Pick Up Trucks	93
Sedan	37
SUV	53
Transit Bus	379
Van	32
Permitting Services	109
Sheriffs Office	85
Liquor Control	56
Housing and Community Affairs	45
Recreation	35
Technology Services	16
Libraries	6
State Attorneys Office	5
Human Rights	2
Office Of Homeland Security	1
Public Information Office	1

The screenshot shows a Pivot Table setup. The 'Rows' section contains 'Department' and 'Equipment Class'. The 'Values' section contains 'Sum of Equipment Co...' (partially visible). The main area displays data grouped by Department, with Equipment Class details shown as sub-items.

Row Labels	Sum of Equipment Count
cuv	15
Permitting Services	9
Transportation	5
Technology Services	1
Heavy Duty	290
Medium Duty	100
Off Road Vehicle Equipment	283
Pick Up Trucks	150
Public Safety CUV	4
Public Safety Pick Up Trucks	1
Public Safety Sedan	47
Public Safety SUV	20
Public Safety Van	8
Sedan	130
SUV	90
Transit Bus	379
Van	65

The screenshot shows a Pivot Table setup. The 'Rows' section contains 'Equipment Class' and 'Department'. The main area displays data grouped by Equipment Class, with Department details shown as sub-items. A single value 'Sum of Equipment Co...' is shown in the Values section.



Conclusion

- Excel can be used for data cleaning but have limited dataset limit.
- Pivot table can be used for finding an insight of the dataset.



THANK YOU

