

Coffee Beans Project Overview:

- Dataset link: <https://www.kaggle.com/datasets/saadharoon27/coffee-bean-sales-raw-dataset>
- Tools used: Microsoft Excel.
- Goal: Data manipulation and visualization.

Summary of cells, sheets, rows, columns, and chart manipulation:

Customer name cell:

- =XLOOKUP (lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])
- Where the lookup_value is the customer ID from the orders sheet, the lookup_array is the customer id from the customer's sheet, the return_array is the customer name from the customer's sheet, if_not_found is left empty with "" sign, match_mode is 0 which is known as the exact match and then at last close the bracket.

Email cell:

- =XLOOKUP (lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])
- Where the lookup_value is the customer ID from the orders sheet, the lookup_array is the customer id from the customer's sheet, the return_array is the email column from the customer's sheet, if_not_found is left empty with "" sign, match_mode is 0 which is known as the exact match and then at last close the bracket.
- Since some of the email cell values didn't have any values, we use IF (logical_test, [value_if_true], [value_if_false]) to find the missing emails.
- The IF (logical_test, [value_if_true], [value_if_false]) is used on the =XLOOKUP (lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode]).

- It is used like,

=IF(XLOOKUP(C2,customers!\$A\$1:\$A\$1001,customers!\$C\$1:\$C\$1001,,0)=0,"",XLOOKUP(C2,customers!\$A\$1:\$A\$1001,customers!\$C\$1:\$C\$1001,,0)).

Where the logical_test is

XLOOKUP(C2,customers!\$A\$1:\$A\$1001,customers!\$C\$1:\$C\$1001,,0)=0,

the value_if_true is "" and the value_if_false is

XLOOKUP(C2,customers!\$A\$1:\$A\$1001,customers!\$C\$1:\$C\$1001,,0).

Country cell:

- =XLOOKUP (lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode]).
- Where the lookup_value is the customer ID from the orders sheet, the lookup_array is the customer id from the customer's sheet, the return_array is the country's column from the customer's sheet, if_not_found is left empty with "" sign, match_mode is 0 which is known as the exact match and then at last close the bracket.

Coffee Type:

- INDEX (array, row_num, [column_num]) and MATCH(lookup_value, lookup_array, [match_type]).
- Where the array is the whole sheet of the products sheet, then use the match function the lookup_value is the product ID column from the orders sheet, the lookup_array is the product ID column from the products sheet, the match type is 0 which is known as the exact match(note: this match syntax was used for the row value), again use match syntax but this time for the column. For the second match syntax, the lookup_value is the coffee type column from the orders sheet, the lookup_array is only the entire header row columns of the products sheet, the exact match is 0 and then lastly close the bracket twice.
- =INDEX(products!\$A\$1:\$G\$49,MATCH(orders!\$D2,products!\$A\$1:\$A\$49,0),
MATCH(orders!\$1,products!\$A\$1:\$G\$1,0)) - this is the entire index and double match function used which I have explained.

Roast type, size and unit price column:

- Auto-populate these column cells by dragging the coffee-type cell to the right and the values will be filled.

	I	J	K	L	M
	Coffee Type	Roast Type	Size	Unit Price	Sales
s	Rob	M	1	9.95	
s	Exc				
s	Ara				

Sales column:

- To calculate the sales column, a simple multiplication procedure is needed where the unit price cell 2 * quantity cell 2 equals the sales total.

E	F	G	H	I	J	K	L	M
Quantity	Customer Name	Email	Country	CoffeeType	Roast Type	Size	Unit Price	Sales
2	Aloisia Allner	aallner0@lulu.com	United States	Rob	M	1	9.95	=L2 * E2
5	Piotr Bote	aallner0@lulu.com	United States	Exc	M	0.5	8.25	

M
Sales
19.9
41.3
13
27.5
55
38.9
21.9
4.76
17.0

Coffee type name cell:

- IF (logical_test, [value_if_true], [value_if_false])
- =IF(I2="Rob","Robusta",IF(I2="Exc","Excelsa",IF(I2="Ara","Arabica",IF(I2="Lib","Liberica",""))))
- the IF function is used to correct the short abbreviation of the coffee name to the full form. For E.g.: Rob to Robusta for a clearer understanding.
- The logical test was filled with if the cell I2 = "Rob" then in the true condition is filled with the full name of the coffee which is "Robusta", the full names of all the remaining coffee are stated using the IF functions again and at last in the value if false condition "" a blank value is left if the condition is false.

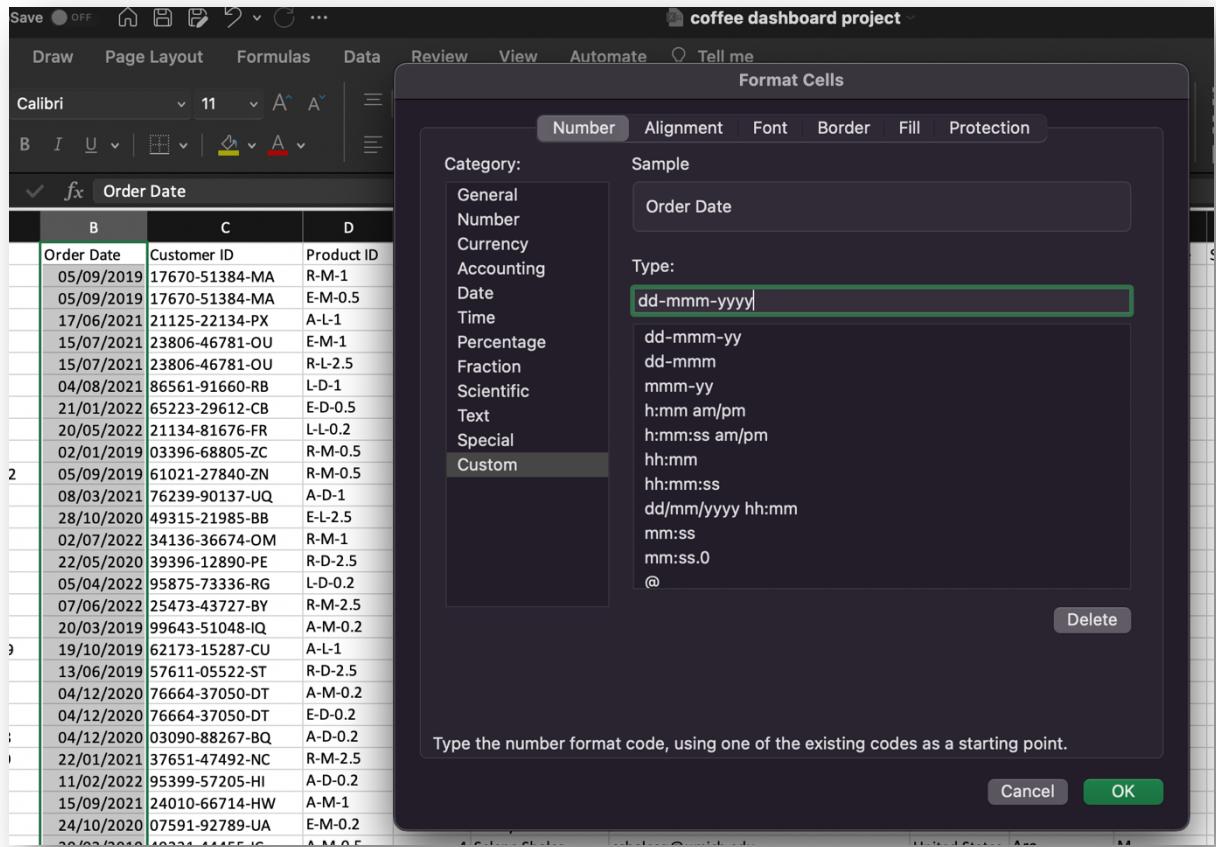
- Brackets are used to close each IF function stated.

Roast Type Name:

- IF (logical_test, [value_if_true], [value_if_false])
- =IF(J2="M","Medium",IF(J2="L","Large",IF(J2="D","Dark","")))
- Used the same IF function used for the cell coffee type name to rename the short abbreviations to full names for better understanding purposes.

Order date cell:

- Formatted the date to European date where the format is dd-mmm-yyyy.
- In this format the day comes first, the month is shown in abbreviation like sep and the year is shown in full.



Size cell:

- Formatted the size cell to show metrics. Previously the size cell didn't have any metrics and was represented in single digits which are then formatted to decimal points with metrics shown. For e.g.: 1 to 1.0 Kg.

The screenshot shows the 'Number' format dialog box in Excel. The 'Category' dropdown is set to 'Size'. The 'Sample' section shows the text 'Size'. In the 'Type:' field, the user has typed '0.0 "Kg"' (without quotes). Below this, there is a list of other number format codes. At the bottom of the dialog, there are 'Delete', 'Cancel', and 'OK' buttons. The background of the image shows a table with columns A, B, C, and D, and another table on the right with columns K, L, M, and N.

Date	Customer ID	Product ID
p-2019	17670-51384-MA	R-M-1
p-2019	17670-51384-MA	E-M-0.5
n-2021	21125-22134-PX	A-L-1
jl-2021	23806-46781-OU	E-M-1
jl-2021	23806-46781-OU	R-L-2.5
g-2021	86561-91660-RB	L-D-1
n-2022	65223-29612-CB	E-D-0.5
y-2022	21134-81676-FR	L-L-0.2
n-2019	03396-68805-ZC	R-M-0.5
p-2019	61021-27840-ZN	R-M-0.5
r-2021	76239-90137-UQ	A-D-1
t-2020	49315-21985-BB	E-L-2.5
jl-2022	34136-36674-OM	R-M-1
y-2020	39396-12890-PE	R-D-2.5
r-2022	95875-73336-RG	L-D-0.2
n-2022	25473-43727-BY	R-M-2.5
ir-2019	99643-51048-IQ	A-M-0.2
t-2019	62173-15287-CU	A-L-1
n-2019	57611-05522-ST	R-D-2.5
sc-2020	76664-37050-DT	A-M-0.2
sc-2020	76664-37050-DT	E-D-0.2
sc-2020	03090-88267-BQ	A-D-0.2
n-2021	37651-47492-NC	R-M-2.5
b-2022	95399-57205-HI	A-D-0.2
p-2021	24010-66714-HW	A-M-1
rt-2020	07591-92789-UA	E-M-0.2
b-2019	49231-44455-IC	A-M-0.5
zt-2019	50124-88608-EO	A-M-0.2
g-2022	00888-74814-UZ	A-D-0.5
b-2019	14158-30713-OB	A-D-1
p-2021	51427-89175-QJ	L-M-0.2
p-2021	51427-89175-QJ	A-D-0.5

	Unit Price	Sales	Coffee type	name
1	9.95	19.9	Arabica	Selene Shales
0.5	8.25	41.3	Excelsa	Vivie Danneil
1	12.95	13	Arabica	Theresita Newbury
1	13.75	27.5	Excelsa	Mozelle Calcutt
2.5	27.485	55	Robusta	Adrian Swaine
1	12.95	38.9	Liberica	
0.5	7.29	21.9	Excelsa	
0.2	4.755	4.76	Liberica	
0.5	5.97	17.9	Robusta	
0.5	5.97	5.97	Robusta	
1	9.95	39.8	Arabica	
2.5	34.155	171	Excelsa	
1	9.95	49.8	Robusta	
2.5	20.585	41.2	Robusta	
0.2	3.885	11.7	Liberica	
2.5	22.885	114	Robusta	
0.2	3.375	20.3	Arabica	
1	12.95	77.7	Arabica	
2.5	20.585	82.3	Robusta	
0.2	3.375	16.9	Arabica	
0.2	3.645	14.6	Excelsa	
0.2	2.985	17.9	Arabica	
2.5	22.885	91.5	Robusta	
0.2	2.985	11.9	Arabica	
1	11.25	11.3	Arabica	
0.2	4.125	12.4	Excelsa	
0.5	6.75	27	Arabica	
0.5	3.375	16.9	Arabica	
0.5	5.97	17.9	Arabica	
1	9.95	39.8	Arabica	
0.2	4.365	21.8	Liberica	
0.5	5.97	35.8	Arabica	

Unit price and sales cell:

- Updated the two cells with \$ currency values. Used the accounting format in the home tab and selected the \$ currency format to update both cells.

Duplicate values check:

- Selected all the sheets with data and then selected the data Tab which has a remove duplicate option, when clicked it removes all the duplicate values.

The screenshot shows a Microsoft Excel spreadsheet with the Data tab selected in the ribbon. A 'Remove Duplicates' dialog box is displayed, overlaid on the main table area. The dialog box includes a checkbox for 'My list has headers' which is checked, and a list of columns with checkboxes next to them: Select All, Order ID, Order Date, Customer ID, Product ID, and Customer Name. All checkboxes are checked. At the bottom of the dialog box are 'Cancel' and 'OK' buttons.

Order ID	Quantity	Customer Name	Email	Country	Coffee Type	Roast Type	Size	Unit Price	Sales	Coffee type name	Roast Type Name
1	2	Aloisia Allner	aallner0@lulu.com	United States	Ara	M	1.0 Kg				
0.5	5	Aloisia Allner	aallner0@lulu.com	United States	Exc	M	0.5 Kg				
1	1	Jami Redholes	jredholes2@tmall.com	United States	Ara	L	1.0 Kg				
1.5	2	Christoffer O' Shea		Ireland	Exc	M	1.0 Kg				
1.5	2	Christoffer O' Shea		Ireland	Rob	L	2.5 Kg				
1	3	Beryle Cottier		United States	Lib	D	1.0 Kg				
0.5	3	Shaylynn Lobe	slobe6@nifty.com	United States	Exc	D	0.5 Kg				
2	1	Melvin Wharfe		Ireland	Lib	L	0.2 Kg				
0.5	3	Guthrey Petracci	gpetracci8@livejournal.com	United States	Rob	M	0.5 Kg				
0.5	1	Rodger Raven	rraven9@ed.gov	United States	Rob	M	0.5 Kg				
1	4	Ferrell Ferber	fferbera@businesswire.com	United States	Ara	D	1.0 Kg				
0.5	5	Duky Phizackerly	dphizackerlyb@utexas.edu	United States	Exc	L	2.5 Kg				
1	5	Rosaleen Scholar	rscholarc@nyu.edu	United States	Rob	M	1.0 Kg				
2.5	2	Terence Vanyutin	tvanyutind@wix.com	United States	Rob	D	2.5 Kg				
0.2	3	Patrice Trobe	ptrobee@wunderground.com	United States	Lib	D	0.2 Kg				
-2.5	5	Llywelyn Oscroft	loscroftf@ebay.co.uk	United States	Rob	M	2.5 Kg				
0.2	6	Minni Alabaster	malabasters@boxxus.com	United States	Ara	M	0.2 Kg				

To create tables:

- Used keyboard shortcut: Command + T in Mac and for Windows use, control + T.
- Created tables for ease of pivot table creation and ease of updating pivot table data.

Table Name:		Summarise with Pivot Table												
Orders		Remove Duplicates		Insert Slicer		Refresh		Header Row		First Column		Filter Button		
								Total Row		Last Column				
A	B	C	D	E	F	G	H	I	J	K	L	M	N	
Order ID	Order Date	Customer ID	Product ID	Quantity	Customer Name	Email	Country	Coffee Type	Roast Type	Unit Price	Sales	Coffee type name	Roast Type	
QEV-37451-860	05-Sep-2019	17670-51384-MA	R-M-1	2	Aloisia Allner	aallner0@lulu.com	United States	Ara	M	1.0 Kg	\$ 9.5	19.30	Arabica	Medium
QEY-37451-860	05-Sep-2019	17670-51384-MA	E-M-0.5	5	Aloisia Allner	aallner1@lulu.com	United States	Exc	M	0.5 Kg	\$ 8.25	41.25	Excella	Medium
FAA-43335-268	17-Jun-2021	21125-22134-PA	A-L-1	1	Jami Redholes	jredholes2@tmall.com	United States	Ara	L	1.0 Kg	\$ 12.95	12.95	Arabica	Large
KAC-83089-793	15-Jul-2021	23806-46578-OU	E-M-1	2	Christoffer O' Shea	cosherr2@tmall.com	Ireland	Exc	M	1.0 Kg	\$ 13.75	27.75	Excella	Medium
KAC-83089-793	15-Jul-2021	23806-46578-OU	R-L-2.5	2	Christoffer O' Shea	cosherr3@tmall.com	Ireland	Rob	L	2.5 Kg	\$ 27.49	54.97	Robusta	Large
VCP-18956-553	04-Aug-2021	86561-91660-RB	L-D-1	3	Berlye Cottier	bcoettier1@tmall.com	United States	Lib	D	1.0 Kg	\$ 12.95	38.85	Umerica	Dark
IPR-31994-879	23-Jan-2022	65233-39612-CB	E-D-0.5	3	Shaylawn Lobe	slobe6@nifty.com	United States	Exc	D	0.5 Kg	\$ 7.29	21.87	Excella	Dark
SNZ-63540-705	20-May-2022	21134-81676-FR	L-L-0.2	1	Melvin Wharfe	mwharfe1@tmall.com	Ireland	Lib	L	0.2 Kg	\$ 4.76	4.76	Umerica	Large
ZCT-46571-659	02-Jan-2019	03996-68805-ZC	R-M-0.5	3	Guthrie Petracci	gpetracci8@livejournal.com	United States	Rob	M	0.5 Kg	\$ 5.97	17.91	Robusta	Medium
NWQ-70061-912	05-Sep-2019	61021-27840-ZN	R-M-0.5	1	Rodger Raven	rraven9@ed.gov	United States	Rob	M	0.5 Kg	\$ 5.97	5.97	Robusta	Medium
BKX-47233-845	08-Mar-2021	76239-90137-UQ	A-D-1	4	Ferril Ferber	fferber@businesswire.com	United States	Ara	D	1.0 Kg	\$ 9.95	39.80	Arabica	Dark
VGR-00102-970	28-Oct-2020	49515-21985-BB	E-L-2.5	5	Dukly Phizackerley	dphizackerlyb@utexas.edu	United States	Exc	L	2.5 Kg	\$ 34.16	170.78	Excella	Large
SZW-48378-399	02-Jul-2022	34136-36674-RG	R-M-1	5	Roseline Scholar	rscolar@nyu.edu	United States	Rob	M	1.0 Kg	\$ 9.95	45.97	Robusta	Medium
ITI-87418-783	22-May-2022	39396-12890-PF	R-D-2.5	2	Terrence Vanyutin	tvanuytin@wix.com	United States	Rob	D	2.5 Kg	\$ 20.59	41.17	Robusta	Dark
GNZ-46006-507	05-Apr-2022	95875-73336-RG	L-D-0.2	3	Patrice Trobe	ptrobee@wunderground.com	United States	Lib	D	0.2 Kg	\$ 3.89	11.66	Umerica	Dark
FYQ-78248-319	07-Jun-2022	25474-33727-YR	R-M-2.5	5	Lijwwellyn Oscroft	loscroft@ebay.co.uk	United States	Rob	M	2.5 Kg	\$ 22.89	114.43	Robusta	Medium
VAU-44387-224	20-Mar-2019	99945-51048-IQ	A-M-0.2	6	Minni Alabaster	malabaster9@hexun.com	United States	Ara	M	0.2 Kg	\$ 3.38	20.25	Arabica	Medium
RDW-33315-159	19-Oct-2019	62173-15827-CU	A-L-1	6	Rhanion Broxup	rbroxup9@jmdio.com	United States	Ara	L	1.0 Kg	\$ 12.95	77.70	Arabica	Large
TDZ-59011-211	13-Jun-2019	57611-05522-ST	R-D-2.5	4	Pall Redford	predfor1@ow.ly	Ireland	Rob	D	2.5 Kg	\$ 20.59	82.34	Robusta	Dark
IDU-25739-399	04-Dec-2020	76664-37050-DT	A-M-0.2	5	Aureo Corradino	accorradino9@harvard.edu	United States	Ara	M	0.2 Kg	\$ 3.38	16.88	Arabica	Medium
IDU-25739-399	04-Dec-2020	76664-37050-E	D-E-0.2	4	Aureo Corradino	accorradino9@harvard.edu	United States	Exc	D	0.2 Kg	\$ 3.65	14.58	Excella	Dark
NUO-20013-488	04-Dec-2020	03090-88267-QD	A-D-0.2	6	Avril Davidowsky	adavidowsky@netvibes.com	United States	Ara	D	0.2 Kg	\$ 2.99	17.91	Arabica	Dark
UGU-65630-479	22-Jan-2021	37651-14749-NC	R-M-2.5	4	Anabel Antuk	antuk9@kickstarter.com	United States	Rob	M	2.5 Kg	\$ 22.89	91.54	Robusta	Medium
IAU-13844-332	11-Feb-2022	95395-57205-HI	A-D-0.2	4	lorge Kleinten	lkleinten1@timesonline.co.uk	United States	Ara	D	0.2 Kg	\$ 2.99	11.94	Arabica	Dark
TKY-71558-096	15-Sep-2021	24010-66714-HW	A-M-1	1	Chrisy Blofield	cblofield0@amazon.co.uk	United States	Ara	M	1.0 Kg	\$ 11.25	11.25	Arabica	Medium
OKY-65322-253	24-Oct-2020	07591-29789-U	E-M-0.2	3	Colley Farris	colleyfarris1@tmall.com	United States	Exc	M	0.2 Kg	\$ 4.13	12.38	Excella	Medium
EVP-44500-491	20-Feb-2019	49214-44455-IC	A-M-0.5	4	Seleine Shales	sshales9@umich.edu	United States	Ara	M	0.5 Kg	\$ 6.75	27.00	Arabica	Medium
WAG-26945-689	08-Oct-2019	50124-88608-EQ	A-M-0.2	5	Vivie Dannell	v.dannell9@mtv.com	Ireland	Ara	M	0.2 Kg	\$ 3.38	16.88	Arabica	Medium
CHE-7895-767	02-Aug-2022	20088-74814-UZ	A-D-0.5	3	Theresita Newbury	tnewbury@usda.gov	Ireland	Ara	D	0.5 Kg	\$ 5.97	17.91	Arabica	Dark
RYZ-14633-602	20-Feb-2019	14158-30713-OH	A-D-1	4	Moellez Calcurati	mcalcutt9@baidu.com	Ireland	Ara	D	1.0 Kg	\$ 9.95	39.80	Arabica	Dark
WOO-36015-429	25-Sep-2021	51427-89175-QJ	L-M-0.2	5	Adrian Swaine	adswayne1@tmall.com	United States	Lib	M	0.2 Kg	\$ 4.37	21.83	Umerica	Medium
WOO-36015-429	25-Sep-2021	51427-89175-QJ	A-D-0.5	6	Adrian Swaine	adswayne2@tmall.com	United States	Ara	D	0.5 Kg	\$ 5.97	35.82	Arabica	Dark

Pivot table:

- Use of pivot table by selecting the pivot table from the insert tab.
- Renamed the pivot table name and sheet to TotalSales.
- In the pivot table Fields, Dragged the order date into the row's fields.
- Right-click on the year rows and select the grouping option to select month and year as those are the options wanted.
- Go to the design tab then click on the reports layout and select Show in tabular form, click grand totals select off for rows and columns and lastly click on subtotals then select do not show subtotals.
- Drag the coffee type names to the column fields and drag the sum to the values field.
- Format the sum cell which is dragged into the values fields. Formatted the sum cell by selecting the format of the number to have no decimal places and clicked on the 1000 comma separator option. This makes the sheet much cleaner.
- 2D line chart visualization- click on the pivot chart sheet go to the insert tab then click on the line chart option and select 2D line chart.
- Formatting the 2D line chart- double click on the chart and then fill the gradient with solid fill. Used the custom RGB colour option and used light purple as a gradient fill.
- Formatted the axis label and chart label by selecting the design tab and clicking on the add chart element option.
- Timeline- to insert timeline go to pivot chart analyse and click on insert timeline.

Screenshot of Microsoft Excel showing the "Create Pivot Table" dialog box.

Choose the data that you want to analyse.

- Select a table or range
- Table/Range: Orders
- Use an external data source

Choose where to place the Pivot Table.

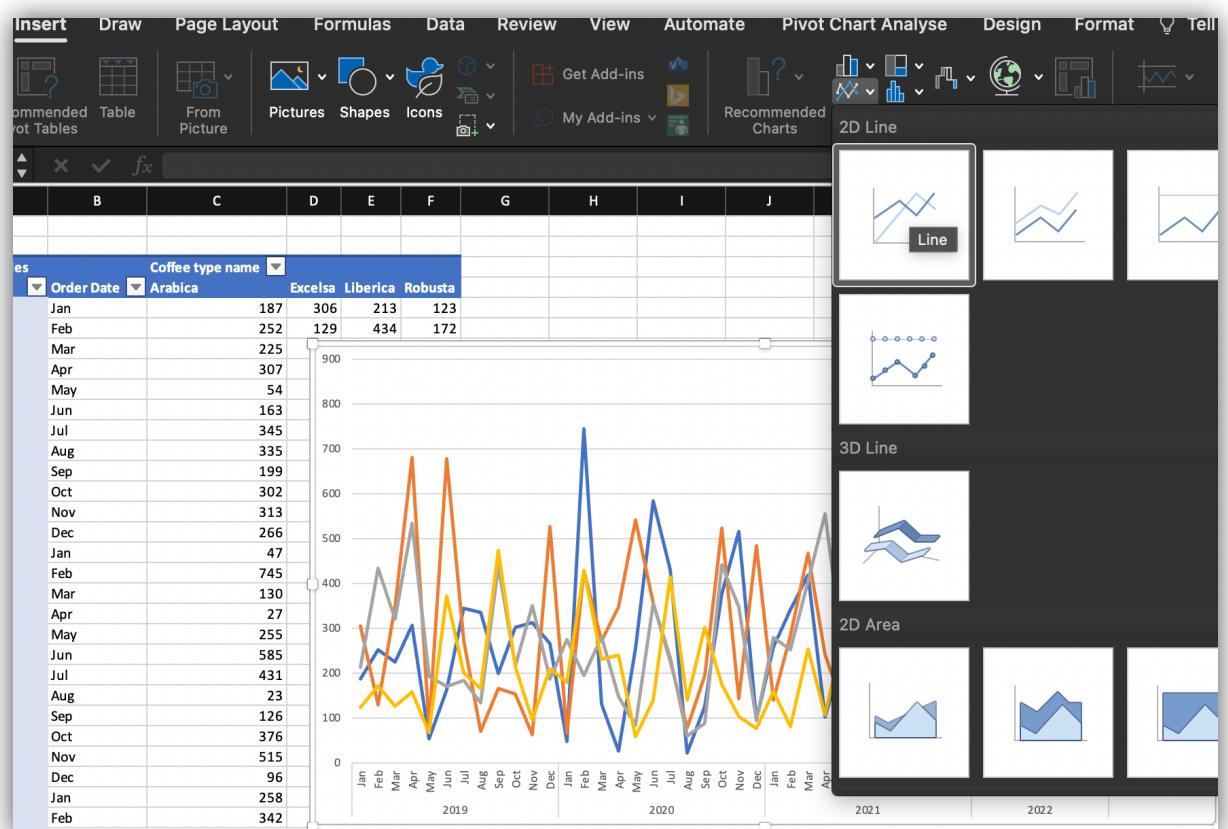
- New worksheet
- Existing worksheet

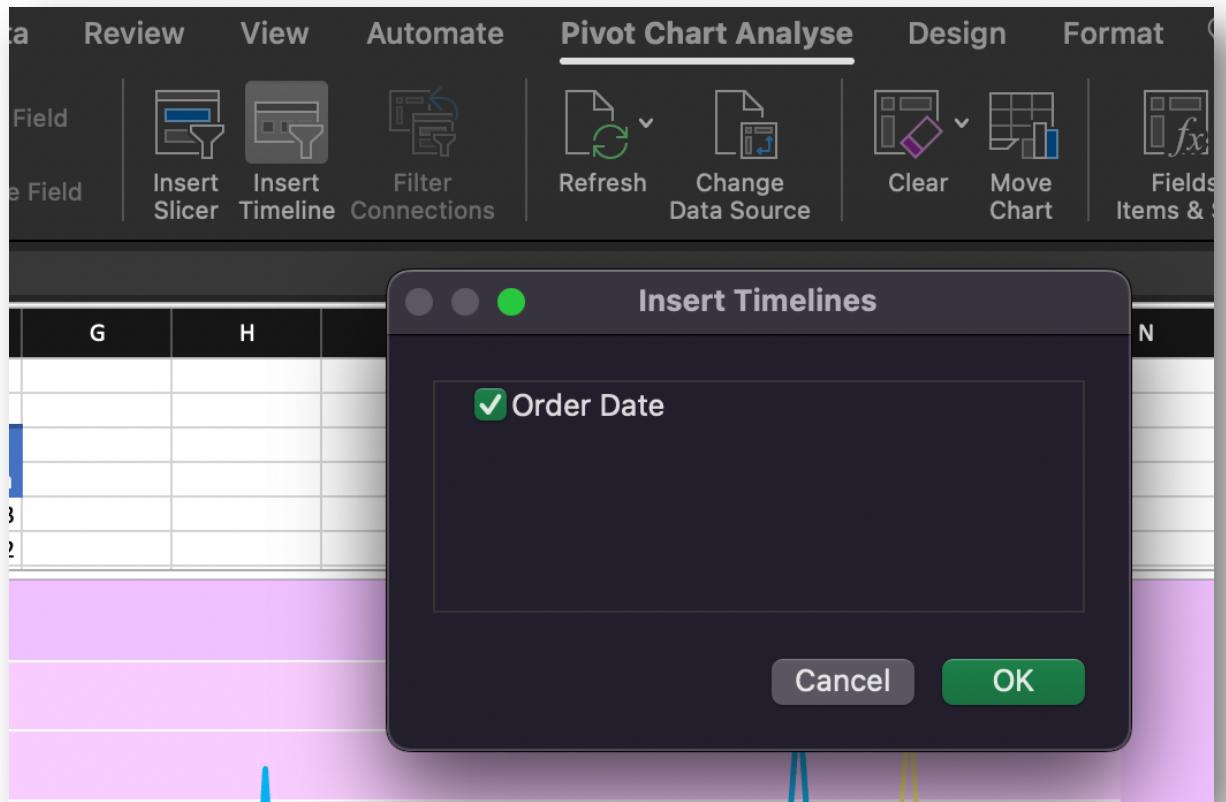
Table/Range: []

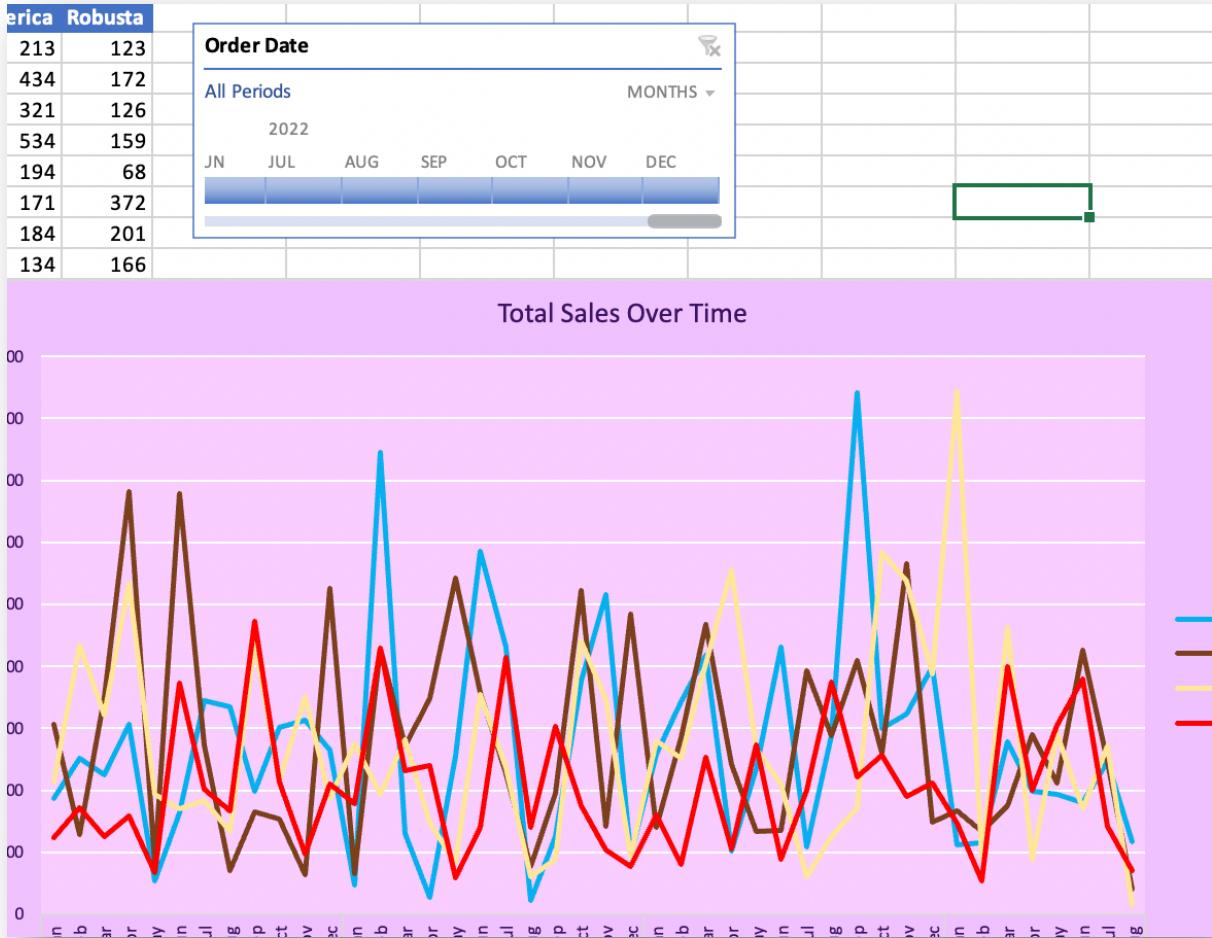
Cancel OK

Below the dialog, a portion of the Excel ribbon and a table are visible:

ID	Order Date	Customer ID	Product ID	Quantity
7451-860	05-Sep-2019	17670-51384-MA	R-M-1	
7451-860	05-Sep-2019	17670-51384-MA	E-M-0.5	
3335-268	17-Jun-2021	21125-22134-PX	A-L-1	
3089-793	15-Jul-2021	23806-46781-OU	E-M-1	
3089-793	15-Jul-2021	23806-46781-OU	R-L-2.5	
8956-553	04-Aug-2021	86561-91660-RB	L-D-1	
1994-879	21-Jan-2022	65223-29612-CB	E-D-0.5	
5340-705	20-May-2022	21134-81676-FR	L-L-0.2	
5571-659	02-Jan-2019	03396-68805-ZC	R-M-0.5	
70061-912	05-Sep-2019	61021-27840-ZN	R-M-0.5	
7233-845	08-Mar-2021	76239-90137-UQ	A-D-1	
1002-970	28-Oct-2020	49315-21985-BB	E-L-2.5	
18378-399	02-Jul-2022	34136-36674-OM	R-M-1	
1418-783	22-May-2020	39396-12890-PE	R-D-2.5	
6006-527	05-Apr-2022	95875-73336-RG	L-D-0.2	
8248-319	07-Jun-2022	25473-43727-BY	R-M-2.5	
4387-624	20-Mar-2019	99643-51048-IQ	A-M-0.2	







Insert slicer:

- From the pivot chart analyze tab select Insert slicer.
- Create slicers for size, coffee roast type and loyalty card.

Loyalty card cell creation:

- =XLOOKUP (lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])
- Use of the xlookup function.
- Look up value is the customer ID, look up array customer ID from the customer's sheet, return is the Loyalty card, lastly 0 for an exact match and then close the bracket.

Duplicate total sales pivot table:

- Rename it to country bar chart.

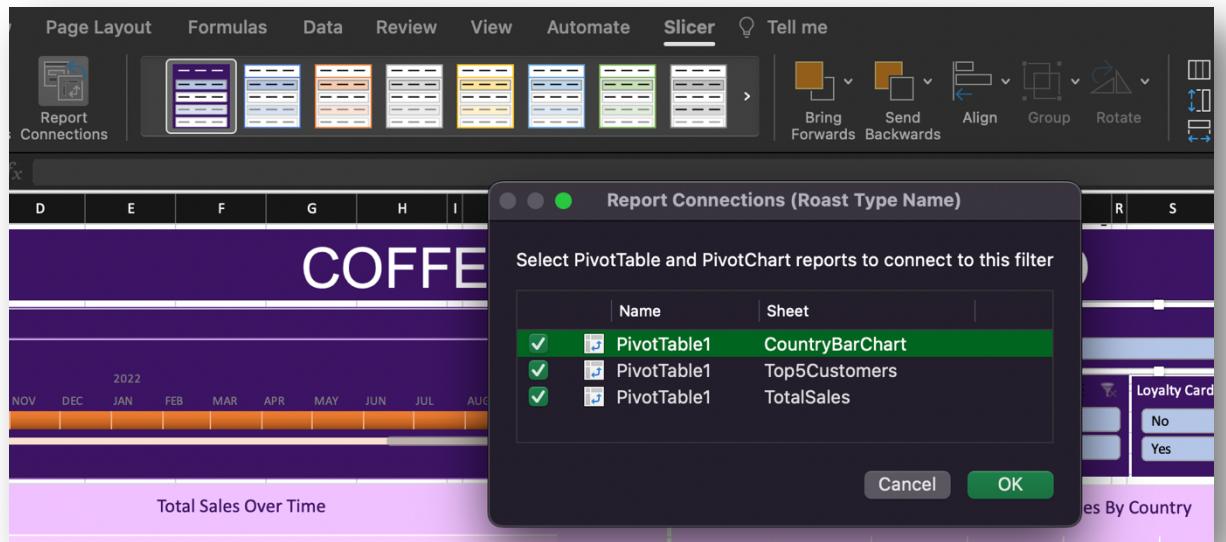
- Remove all the pivot table fields except for sum and drag the country to the rows field.
- Go to the insert tab and insert a bar chart.
- Format the pivot table sheet of the country in ascending order and use the sum of sales option to rearrange the bar charts.
- Format the bar chart's colour, font size, font colour and axis colour.
- Go to the design tab and insert data labels.
- Format the currency of the data labels to \$.

Duplicate country bar chart pivot table sheet:

- Rename it to Top5Customers.
- Change the country axis to the customer's name axis which is done by first dropping the country from the axis field and then dragging the customer's name into the axis field of the pivot table fields.
- Then go into the customer's name column of the pivot table sheet click on value filters and then select top 10 and then modify it to top 5 and order by the sum of sales.

Create a new worksheet:

- Name it Dashboard.
- Insert a rectangle shape (press the option key and hold it while dragging the shape) fill it will colour and name it the coffee sales dashboard. This is the header of the visualisation.
- Go to the timeline tab click on report connections and then select all the pivot tables which will then allow for the slicer to visualize all the charts.
- Do the same for the slicers and select all the pivot tables by clicking on report connections.
- Remove the gridlines by selecting the view tab.



Final visualisation result:

