DATA ANALYSIS PROJECT IN EXCEL

Sales of a Coffee Shop

Postolache Calin

STEP ONE OF THE PROJECT

Getting and cleaning the Data

01 Search for a database on Github/Kaggle.

02 Download the database and load it in excel.

03 Explore the data and determine the next steps.

04 Clean and transform the data for the analysis process.

05 Preparing for the data analysis process.

STEP TWO OF THE PROJECT

Analyzing the Data

1. I've created a main sheet called ,,orders'' where i'll do the main analysis.

	A	B	C	D	E	F	G	H
	Order ID	Order Date	Month	Customer ID	Product ID	Quantity	Customer Name	Email
2	QEV-37451-860	05-Sep-2019	9	17670-51384-MA	R-M-1	2	Aloisia Allner	aallner0@lulu.com
3	QEV-37451-860	05-Sep-2019	9	17670-51384-MA	E-M-0.5	5	Aloisia Allner	aallner0@lulu.com
4	FAA-43335-268	17-Jun-2021	6	21125-22134-PX	A-L-1	1	Jami Redholes	jredholes2@tmall.com
5	KAC-83089-793	15-Jul-2021	7	23806-46781-OU	E-M-1	2	Christoffer O' Shea	
6	KAC-83089-793	15-Jul-2021	7	23806-46781-OU	R-L-2.5	2	Christoffer O' Shea	
7	CVP-18956-553	04-Aug-2021	8	86561-91660-RB	L-D-1	3	Beryle Cottier	
8	IPP-31994-879	21-Jan-2022	1	65223-29612-CB	E-D-0.5	3	Shaylynn Lobe	slobe6@nifty.com
9	SNZ-65340-705	20-May-2022	5	21134-81676-FR	L-L-0.2	1	Melvin Wharfe	
10	EZT-46571-659	02-Jan-2019	1	03396-68805-ZC	R-M-0.5	3	Guthrey Petracci	gpetracci8@livejournal.com
11	NWQ-70061-912	05-Sep-2019	9	61021-27840-ZN	R-M-0.5	1	Rodger Raven	rraven9@ed.gov
12	BKK-47233-845	08-Mar-2021	3	76239-90137-UQ	A-D-1	4	Ferrell Ferber	fferbera@businesswire.com
13	VQR-01002-970	28-Oct-2020	10	49315-21985-BB	E-L-2.5	5	Duky Phizackerly	dphizackerlyb@utexas.edu
14	SZW-48378-399	02-Jul-2022	7	34136-36674-OM	R-M-1	5	Rosaleen Scholar	rscholarc@nyu.edu
15	ITA-87418-783	22-May-2020	5	39396-12890-PE	R-D-2.5	2	Terence Vanyutin	tvanyutind@wix.com
16	GNZ-46006-527	05-Apr-2022	4	95875-73336-RG	L-D-0.2	3	Patrice Trobe	ptrobee@wunderground.com
17	FYQ-78248-319	07-Jun-2022	6	25473-43727-BY	R-M-2.5	5	Llywellyn Oscroft	loscroftf@ebay.co.uk
18	VAU-44387-624	20-Mar-2019	3	99643-51048-IQ	A-M-0.2	6	Minni Alabaster	malabasterg@hexun.com
19	RDW-33155-159	19-Oct-2019	10	62173-15287-CU	A-L-1	6	Rhianon Broxup	rbroxuph@jimdo.com
20	TDZ-59011-211	13-Jun-2019	6	57611-05522-ST	R-D-2.5	4	Pall Redford	predfordi@ow.ly
21	IDU-25793-399	04-Dec-2020	12	76664-37050-DT	A-M-0.2	5	Aurea Corradino	acorradinoj@harvard.edu
22	IDU-25793-399	04-Dec-2020	12	76664-37050-DT	E-D-0.2	4	Aurea Corradino	acorradinoj@harvard.edu
23	NUO-20013-488	04-Dec-2020	12	03090-88267-BQ	A-D-0.2	6	Avrit Davidowsky	adavidowskyl@netvibes.com
24	UQU-65630-479	22-Jan-2021	1	37651-47492-NC	R-M-2.5	4	Annabel Antuk	aantukm@kickstarter.com
25	FEO-11834-332	11-Feb-2022	2	95399-57205-HI	A-D-0.2	4	Iorgo Kleinert	ikleinertn@timesonline.co.uk
26	TKY-71558-096	15-Sep-2021	9	24010-66714-HW	A-M-1	1	Chrisy Blofeld	cblofeldo@amazon.co.uk
27	OXY-65322-253	24-Oct-2020	10	07591-92789-UA	E-M-0.2	3	Culley Farris	
28	EVP-43500-491	20-Feb-2019	2	49231-44455-IC	A-M-0.5	4	Selene Shales	sshalesq@umich.edu

I	J	K	L	M	N	O	P	Q
Country	Coffee Type	Coffee Full Name	Roast Type	Roast Full Name	Size	Unit Price	Sales	Loyalty Card
United States	Rob	Robusta	M	Medium	1.0 kg	$ 9.95	$ 19.90	Yes
United States	Exc	Excelsa	M	Medium	0.5 kg	$ 8.25	$ 41.25	Yes
United States	Ara	Arabica	L	Light	1.0 kg	$ 12.95	$ 12.95	Yes
Ireland	Exc	Excelsa	M	Medium	1.0 kg	$ 13.75	$ 27.50	No
Ireland	Rob	Robusta	L	Light	2.5 kg	$ 27.49	$ 54.97	No
United States	Lib	Liberica	D	Dark	1.0 kg	$ 12.95	$ 38.85	No
United States	Exc	Excelsa	D	Dark	0.5 kg	$ 7.29	$ 21.87	Yes
Ireland	Lib	Liberica	L	Light	0.2 kg	$ 4.76	$ 4.76	Yes
United States	Rob	Robusta	M	Medium	0.5 kg	$ 5.97	$ 17.91	No
United States	Rob	Robusta	M	Medium	0.5 kg	$ 5.97	$ 5.97	No
United States	Ara	Arabica	D	Dark	1.0 kg	$ 9.95	$ 39.80	No
United States	Exc	Excelsa	L	Light	2.5 kg	$ 34.16	$ 170.78	Yes
United States	Rob	Robusta	M	Medium	1.0 kg	$ 9.95	$ 49.75	No
United States	Rob	Robusta	D	Dark	2.5 kg	$ 20.59	$ 41.17	No
United States	Lib	Liberica	D	Dark	0.2 kg	$ 3.89	$ 11.66	Yes
United States	Rob	Robusta	M	Medium	2.5 kg	$ 22.89	$ 114.43	No
United States	Ara	Arabica	M	Medium	0.2 kg	$ 3.38	$ 20.25	No
United States	Ara	Arabica	L	Light	1.0 kg	$ 12.95	$ 77.70	No
Ireland	Rob	Robusta	D	Dark	2.5 kg	$ 20.59	$ 82.34	Yes
United States	Ara	Arabica	M	Medium	0.2 kg	$ 3.38	$ 16.88	Yes
United States	Exc	Excelsa	D	Dark	0.2 kg	$ 3.65	$ 14.58	Yes
United States	Ara	Arabica	D	Dark	0.2 kg	$ 2.99	$ 17.91	No
United States	Rob	Robusta	M	Medium	2.5 kg	$ 22.89	$ 91.54	Yes
United States	Ara	Arabica	D	Dark	0.2 kg	$ 2.99	$ 11.94	Yes
United States	Ara	Arabica	M	Medium	1.0 kg	$ 11.25	$ 11.25	No
United States	Exc	Excelsa	M	Medium	0.2 kg	$ 4.13	$ 12.38	Yes

2. In this sheet i've added new columns from other sheets that will help me run the analysis.

3. In this sheet called ,,orders'' i've added new columns from other sheets that will help me run the analysis.
To do this i've used functions like ,,MONTH, XLOOKUP, INDEX, MATCH and also IF AND NESTED IF functions to get the data from all sheets into the ,,orders'' one.

ps: i intentionally left the blank sections in the email column.

Exemples of functions used:

```
=XLOOKUP(D2,customers!$A:$A,customers!$G:$G,,0)

=INDEX(products!$A$1:$G$49,MATCH(orders!$E2,products!$A$1:$A$49,0),MATCH(orders!J$1,products!$A$1:$G$1,0))

=IF([@[Coffee Type]]="Rob","Robusta",IF([@[Coffee Type]]="Exc","Excelsa",IF([@[Coffee Type]]="Ara","Arabica",IF([@[Coffee Type]]="Lib","Liberica",""))))
```

-Found in the ,,orders'' sheet

STEP THREE OF THE PROJECT

Pivot tables

First pivot table created is related to Total Sales Over Time.

Sum of Sales		Coffee Full Name			
Years	Order Date	Arabica	Excelsa	Liberica	Robusta
2019	Jan	187	306	213	123
	Feb	252	129	434	172
	Mar	225	349	321	126
	Apr	307	681	534	159
	May	54	83	194	68
	Jun	163	678	171	372
	Jul	345	274	184	201
	Aug	335	71	134	166
	Sep	179	166	439	493
	Oct	302	154	216	214
	Nov	313	63	351	96
	Dec	266	527	187	211
2020	Jan	47	66	275	179
	Feb	745	429	194	430
	Mar	130	271	281	232
	Apr	27	347	148	240
	May	255	542	83	59
	Jun	585	357	355	141
	Jul	431	227	236	415
	Aug	23	78	61	140
	Sep	126	195	89	303
	Oct	376	523	441	174
	Nov	515	143	347	104
	Dec	96	485	94	77
2021	Jan	258	140	280	160
	Feb	342	284	252	81
	Mar	418	468	405	253
	Apr	102	242	555	106
	May	235	133	267	273
	Jun	430	136	210	88
	Jul	109	394	61	199
	Aug	288	289	126	374
	Sep	841	410	171	221
	Oct	299	260	585	256
	Nov	323	566	538	189
	Dec	399	148	388	212
2022	Jan	113	166	844	147
	Feb	115	134	91	54
	Mar	278	175	463	400
	Apr	198	290	89	200
	May	193	212	292	304
	Jun	180	426	170	379
	Jul	247	247	271	142
	Aug	116	41	16	71

This will be useful later on for the vizualization and the dashboard.

Country	Sum of Sales
United Kingdom	$2,799
Ireland	$6,697
United States	$35,639

Sales by country so we can see where is the most revenue coming from which is clearly the United States.

Top 5 customers which also indicates that the biggest revenue comes from customers that live in the United States.

Customer Name	Country	Sum of Sales
Don Flintiff	United Kingdom	$278
Nealson Cuttler	United States	$282
Terri Farra	United States	$289
Brenn Dundredge	United States	$307
Allis Wilmore	United States	$317

Revenue based on loyalty card. This shows that the most revenue comes from customers with no loyalty card.

Sum of Sales	Loyalty Car	
Years	No	Yes
2019	$6,292	$5,896
2020	$6,419	$5,699
2021	$7,926	$5,840

Coffee Full Nam	Sum of Sales
Robusta	$9,005
Arabica	$11,768
Liberica	$12,054
Excelsa	$12,306

Sales based on Coffee type. Excelsa is the best selling product.

Sales based on roast type. Light roast is the best selling roast type.

Roast Full Name	Sum of Sales
Dark	$13,179
Medium	$14,600
Light	$17,354

Sum of Sales	Years			
Month	2019	2020	2021	2022
1	$829	$567	$838	$1,269
2	$987	$1,798	$959	$394
3	$1,021	$915	$1,545	$1,315
4	$1,681	$762	$1,006	$776
5	$399	$939	$908	$1,002
6	$1,385	$1,438	$865	$1,155
7	$1,004	$1,309	$763	$907
8	$706	$300	$1,076	$244
9	$1,277	$713	$1,644	
10	$885	$1,515	$1,400	
11	$823	$1,109	$1,616	
12	$1,190	$752	$1,148	

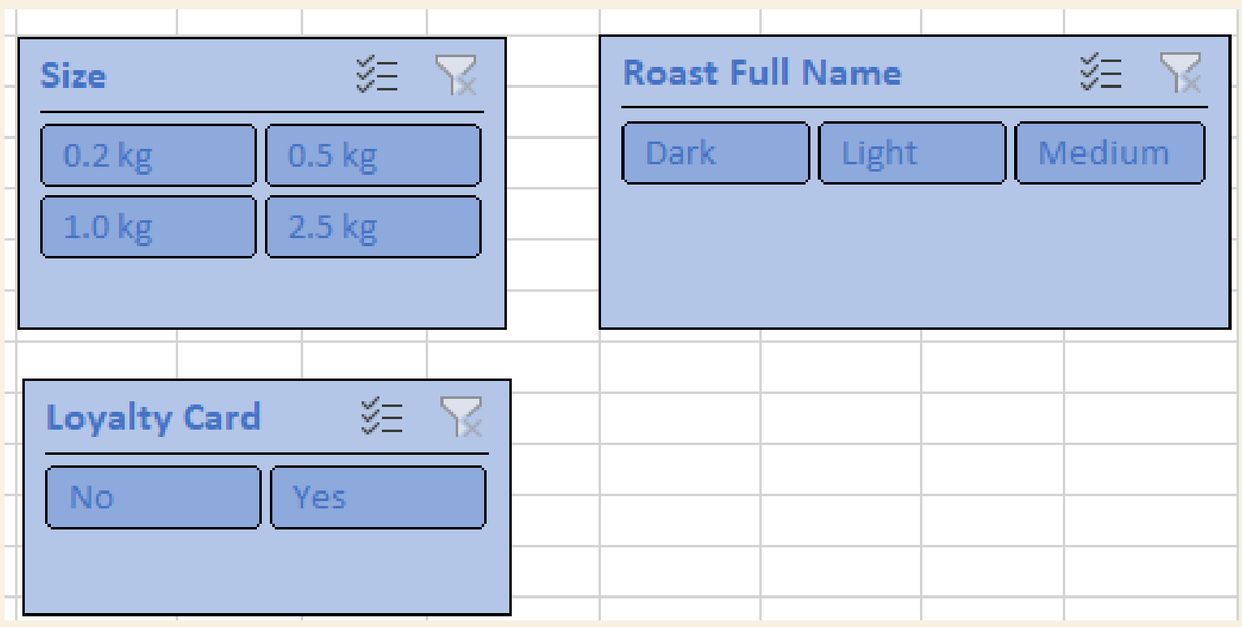
Sales performance over the year by month. This will help see which months perform the best.

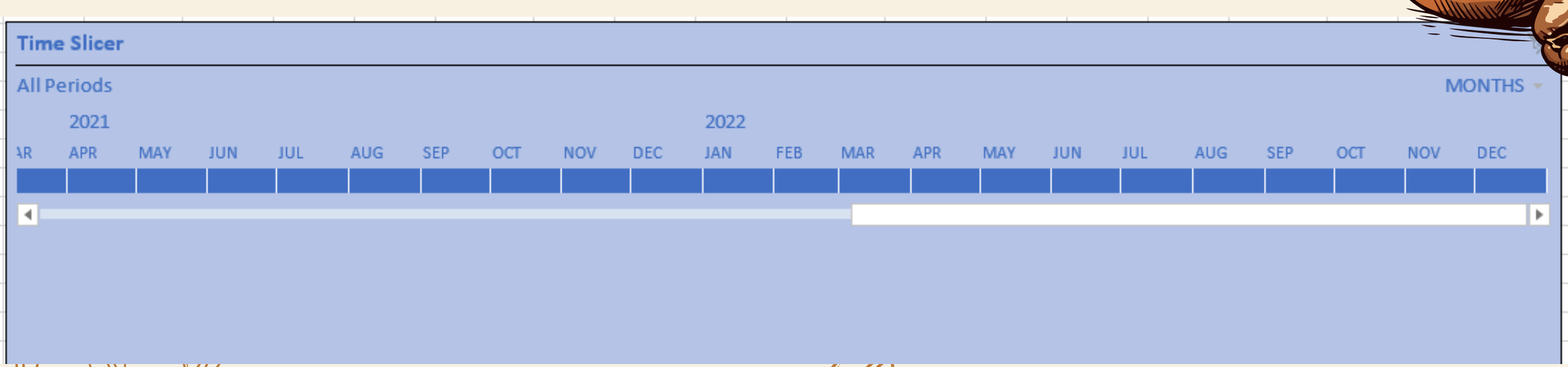
Years	Average of Sales
2019	$47.05
2020	$46.07
2021	$44.12
2022	$42.55

Average order value which is clearly decreasing every year.

Also added some slicers that can help vizualize the „Total Sales over time" chart for better insights.

This will be found in the Dashboard

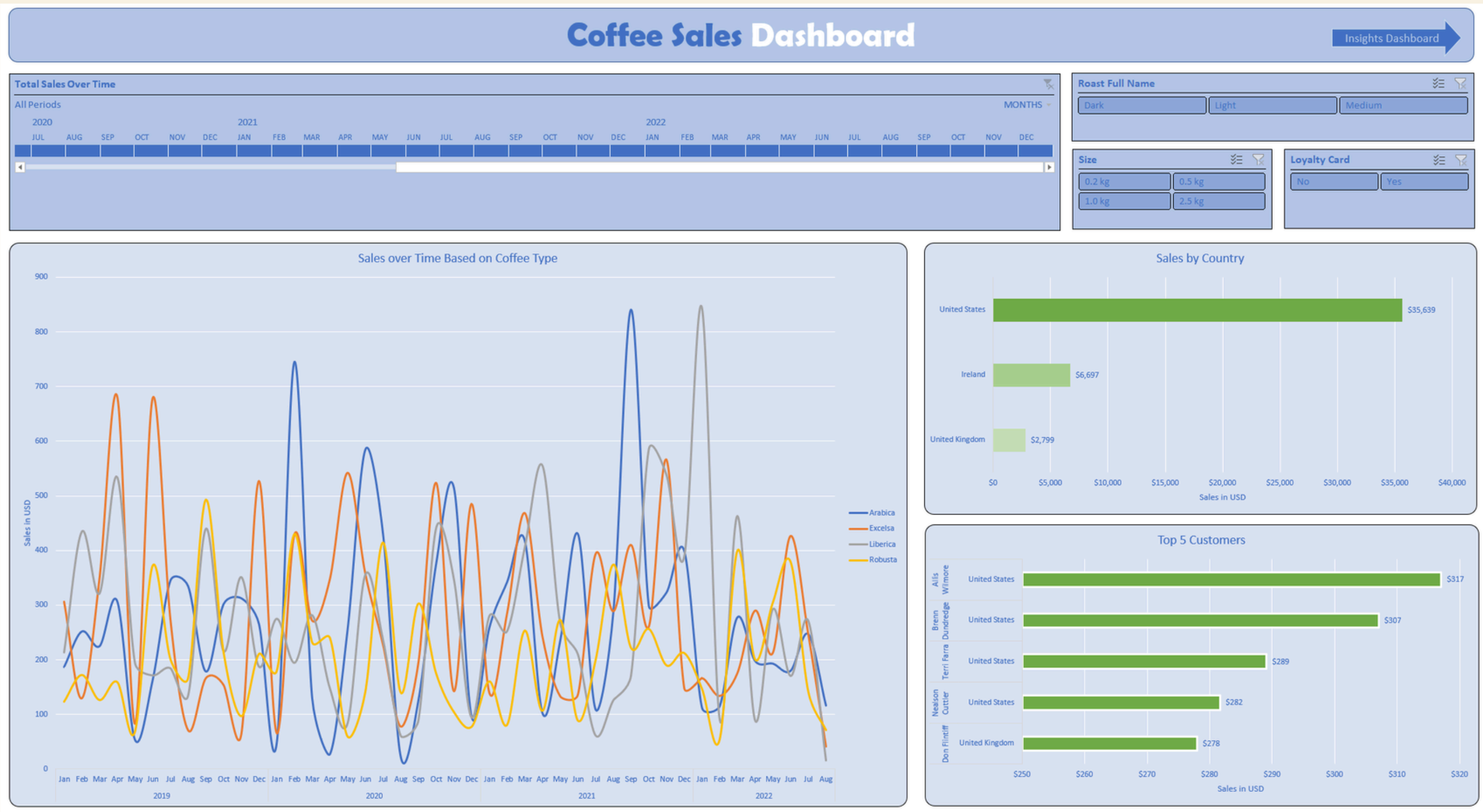




STEP FOUR OF THE PROJECT

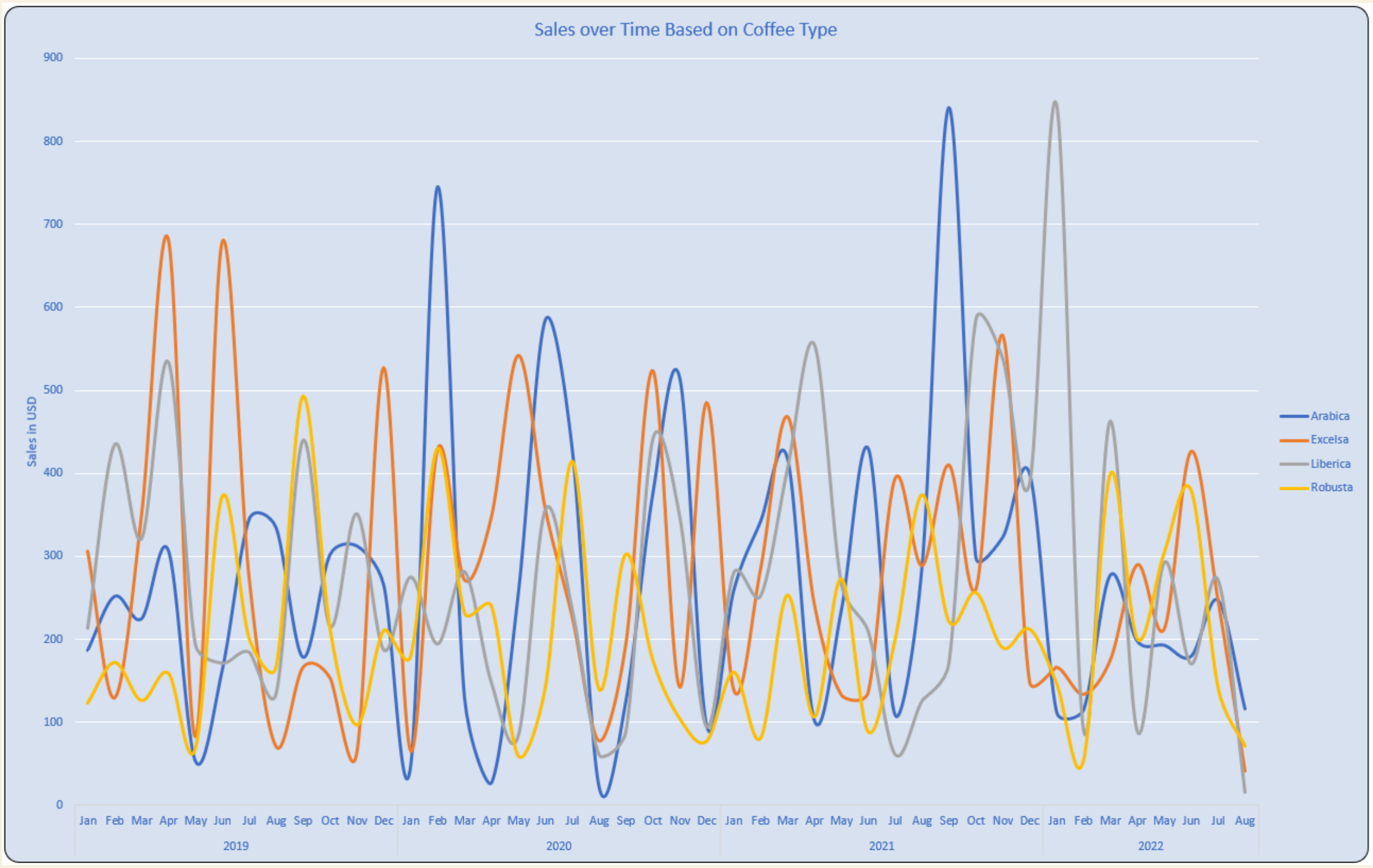
Creating the Dashboards

This is the main dashboard:

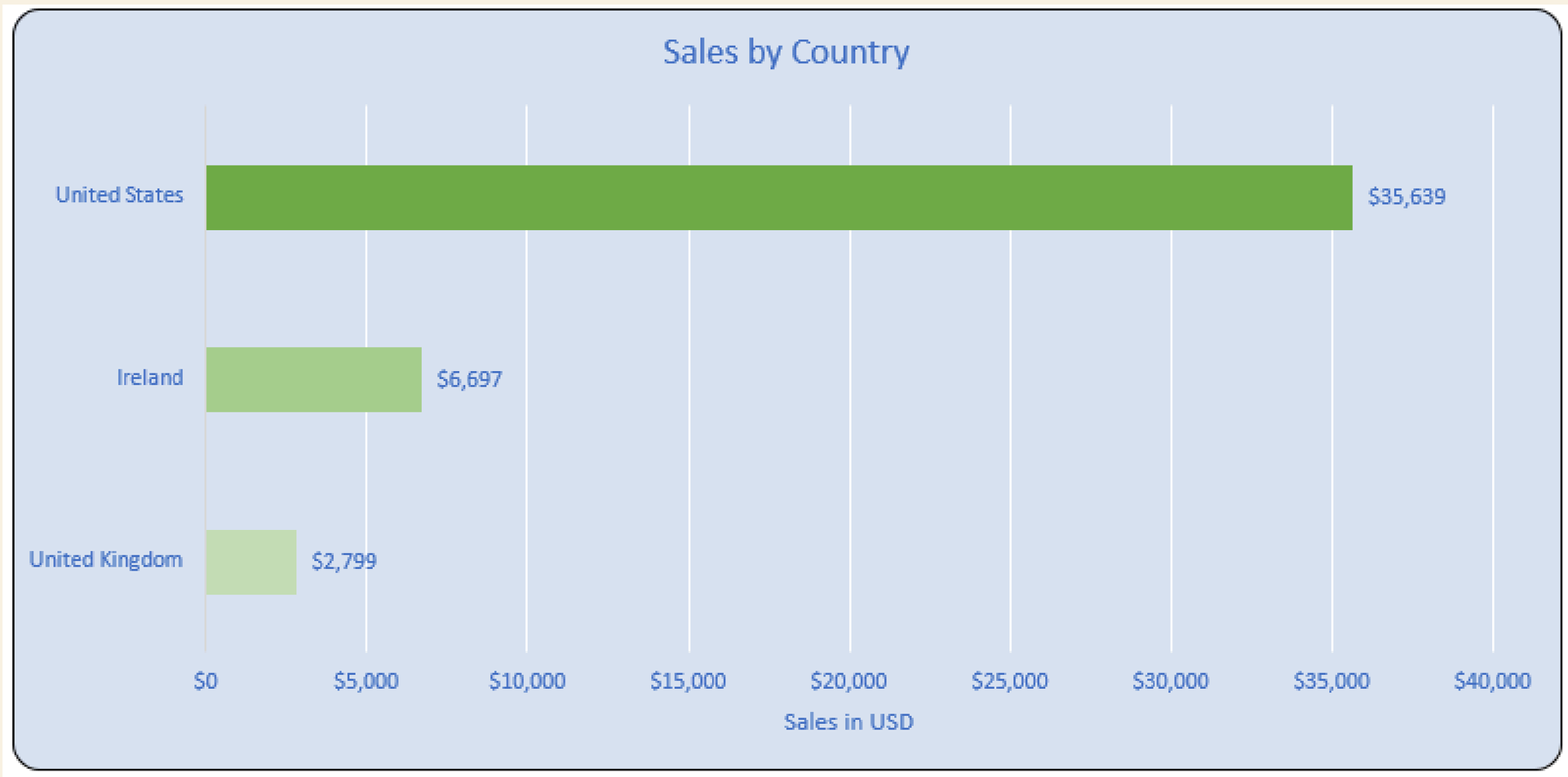


For charts, we can find the Sales over time based on Coffee Type, Sales by Country and Top 5 Customers.

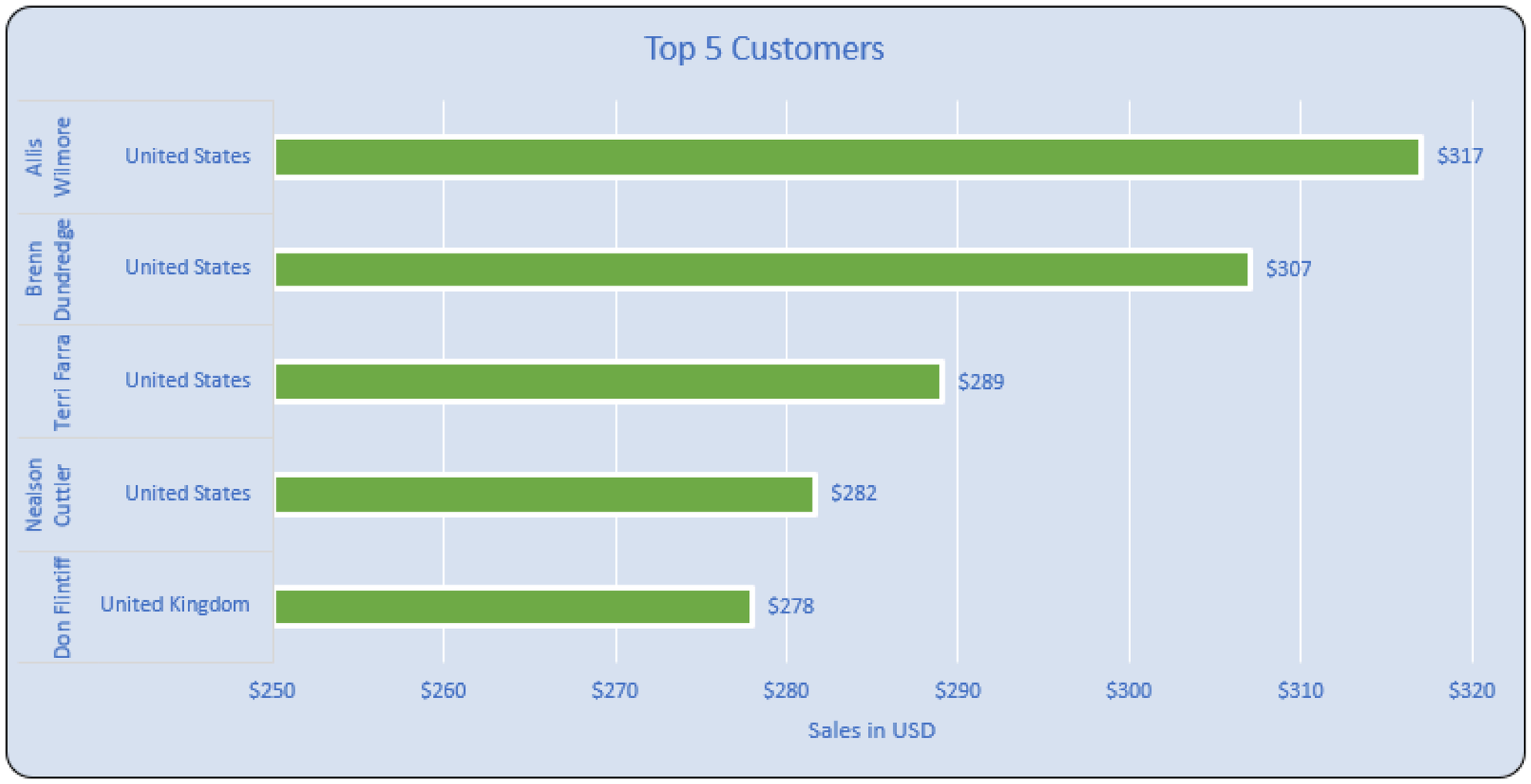
For slicers, we can find the Time, Roast type, Size and Loyalty card slicers for the Sales over time chart.



In this chart we can see that sales by coffee type fluctuate but Excelsa tends to stay high and constant the most when it comes to sales.



When it comes to sales, United States generates the most sales from all three countries.



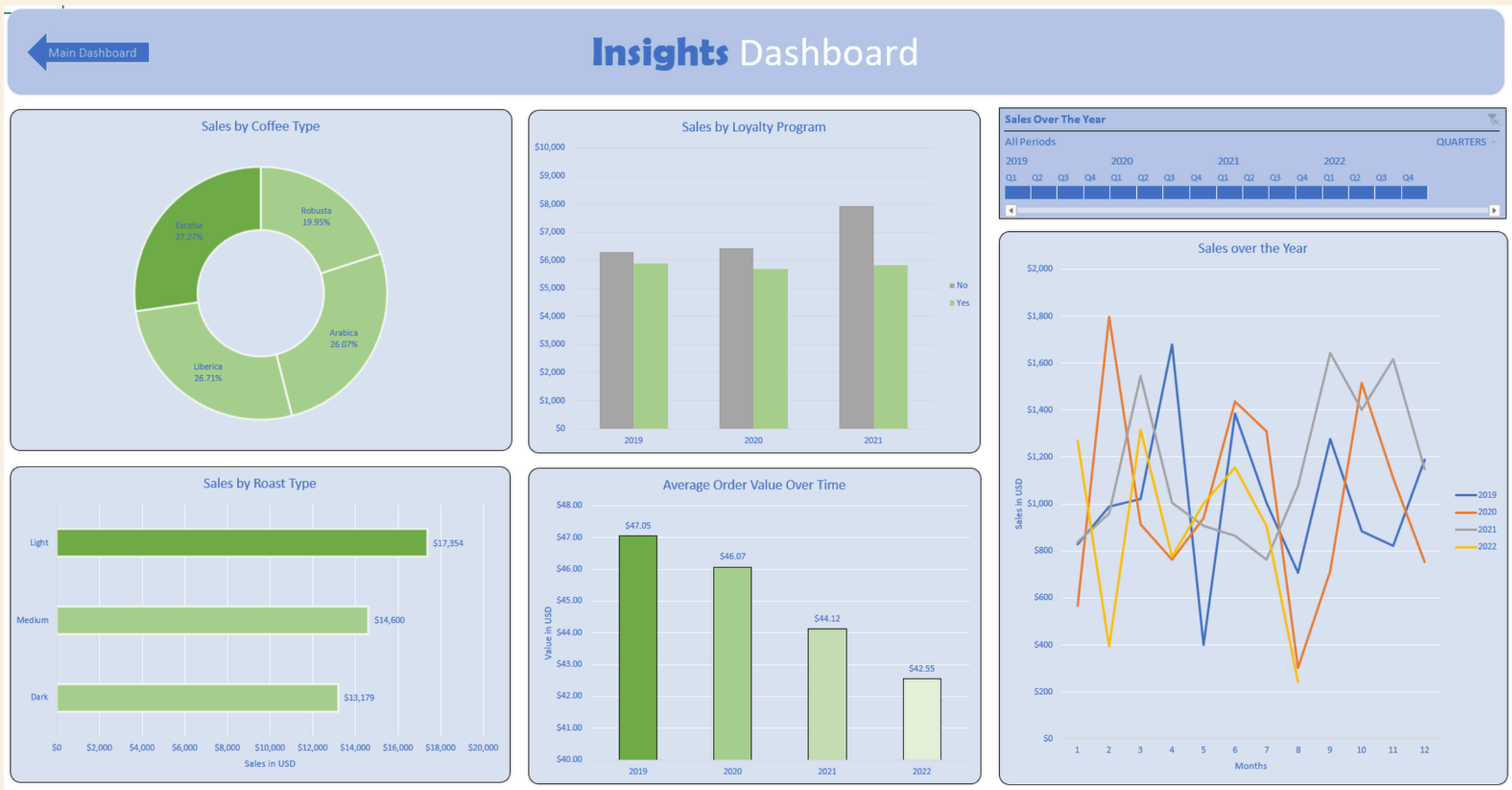
Now to the Top 5 Customers chart, we can see that four out of five customers are also from the United States which confirms even more that the US is the best market here.

From this Dashboard we can understand the following things:

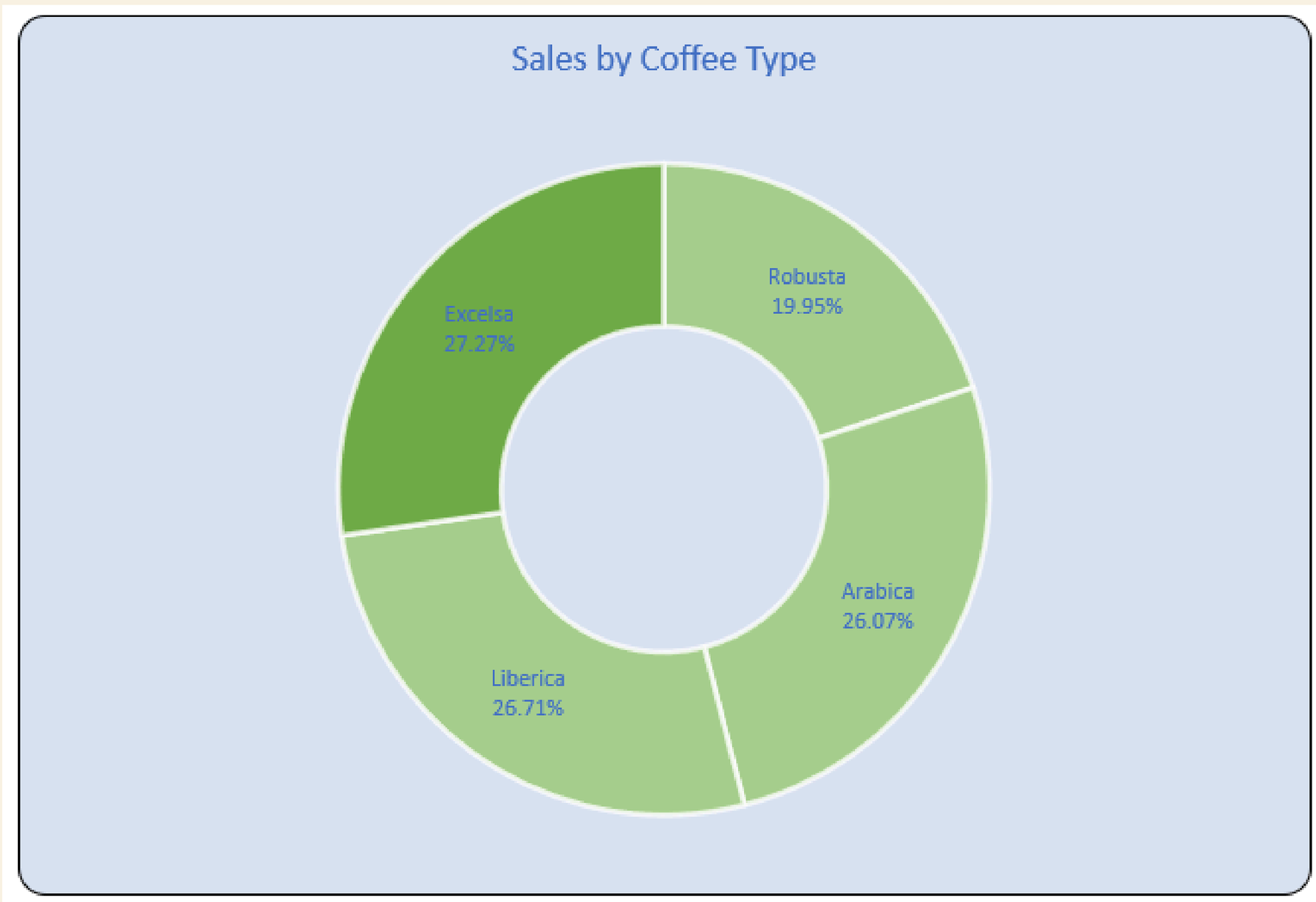
1.Coffee type sales vary based on seasons but Excelsa performs the best overall.

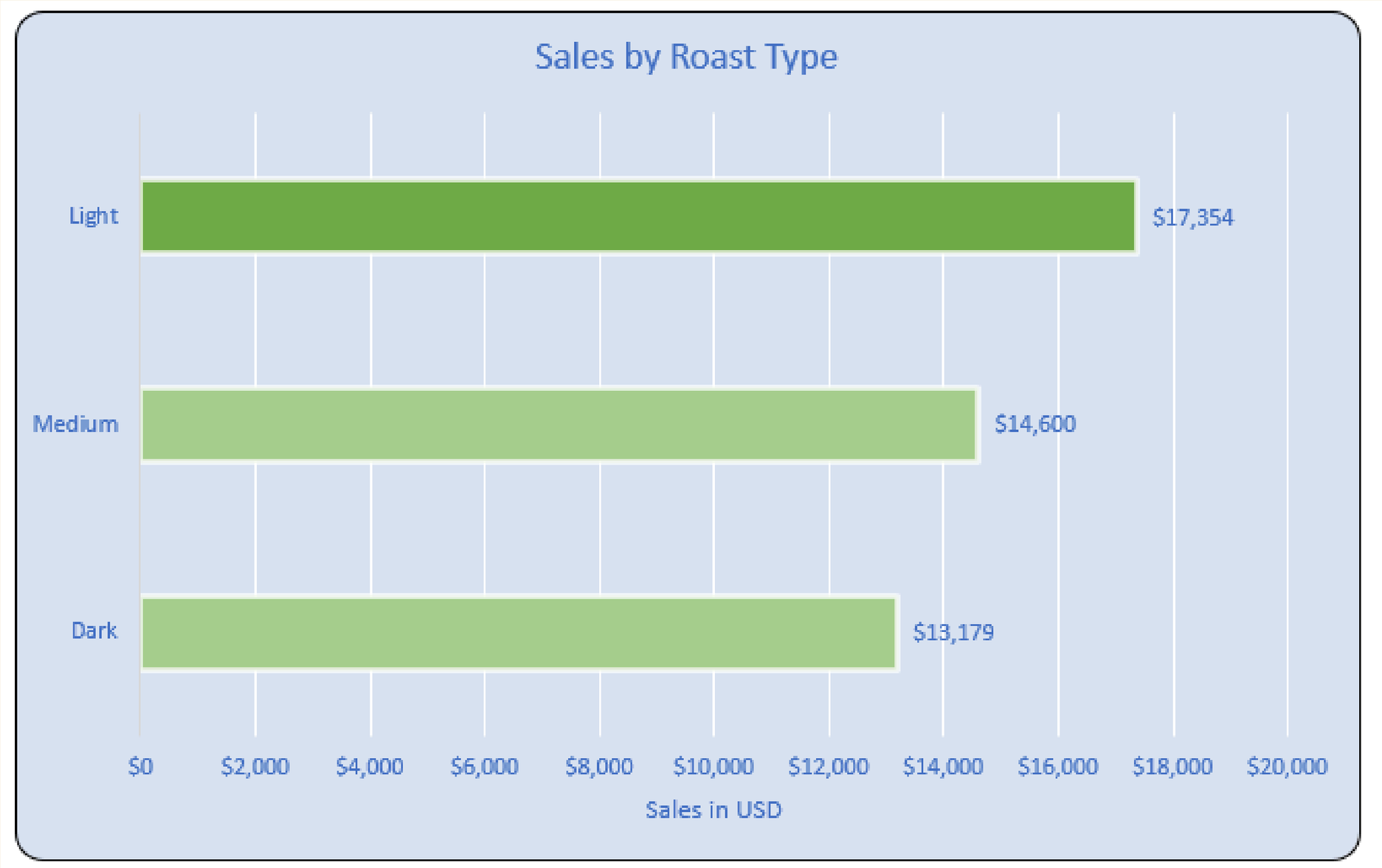
2.United States generates the most amount of Sales by a big difference compared to the other two countries.

This is the insights dashboard:

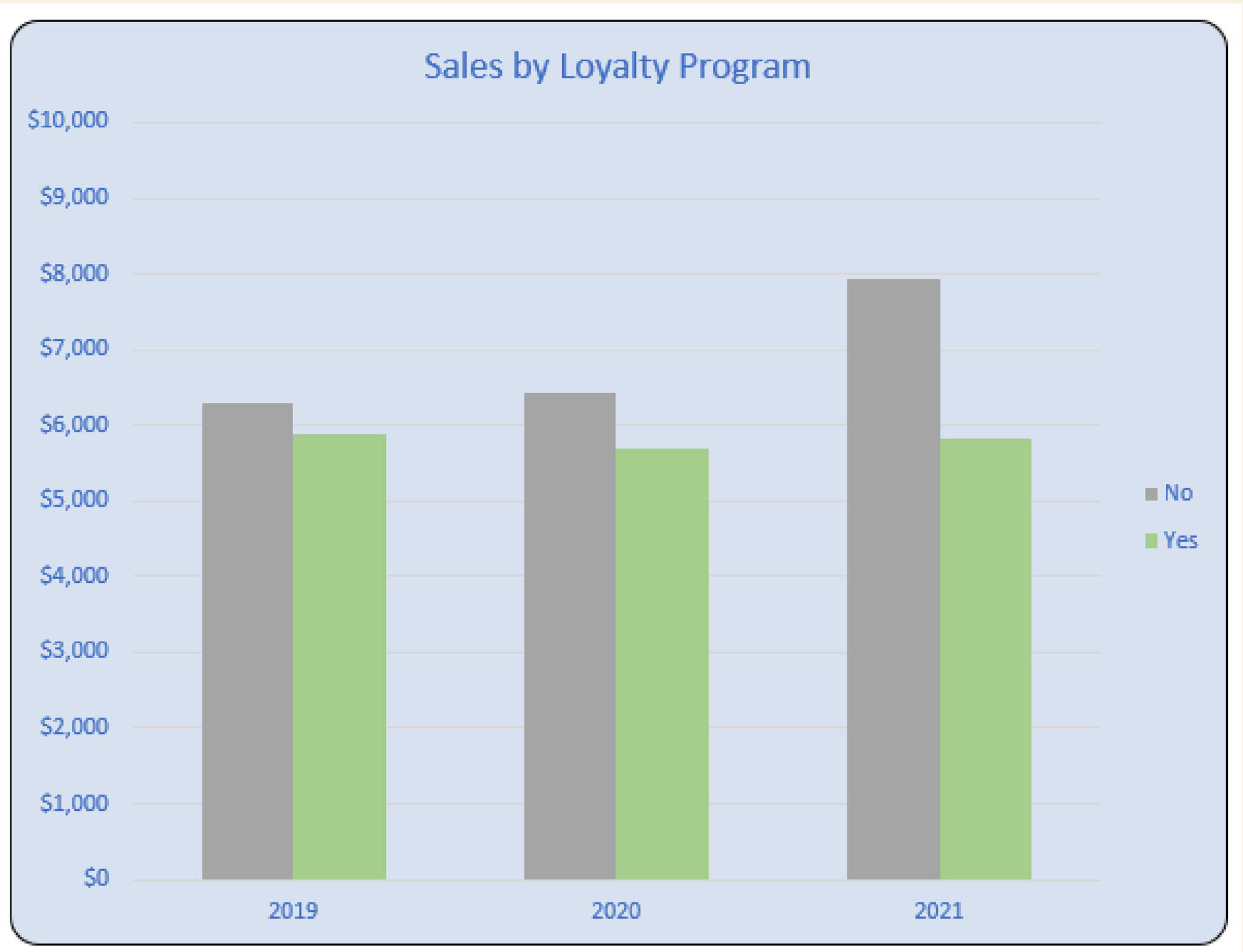


Here we can see and confirm that Excelsa is the best selling coffee type.



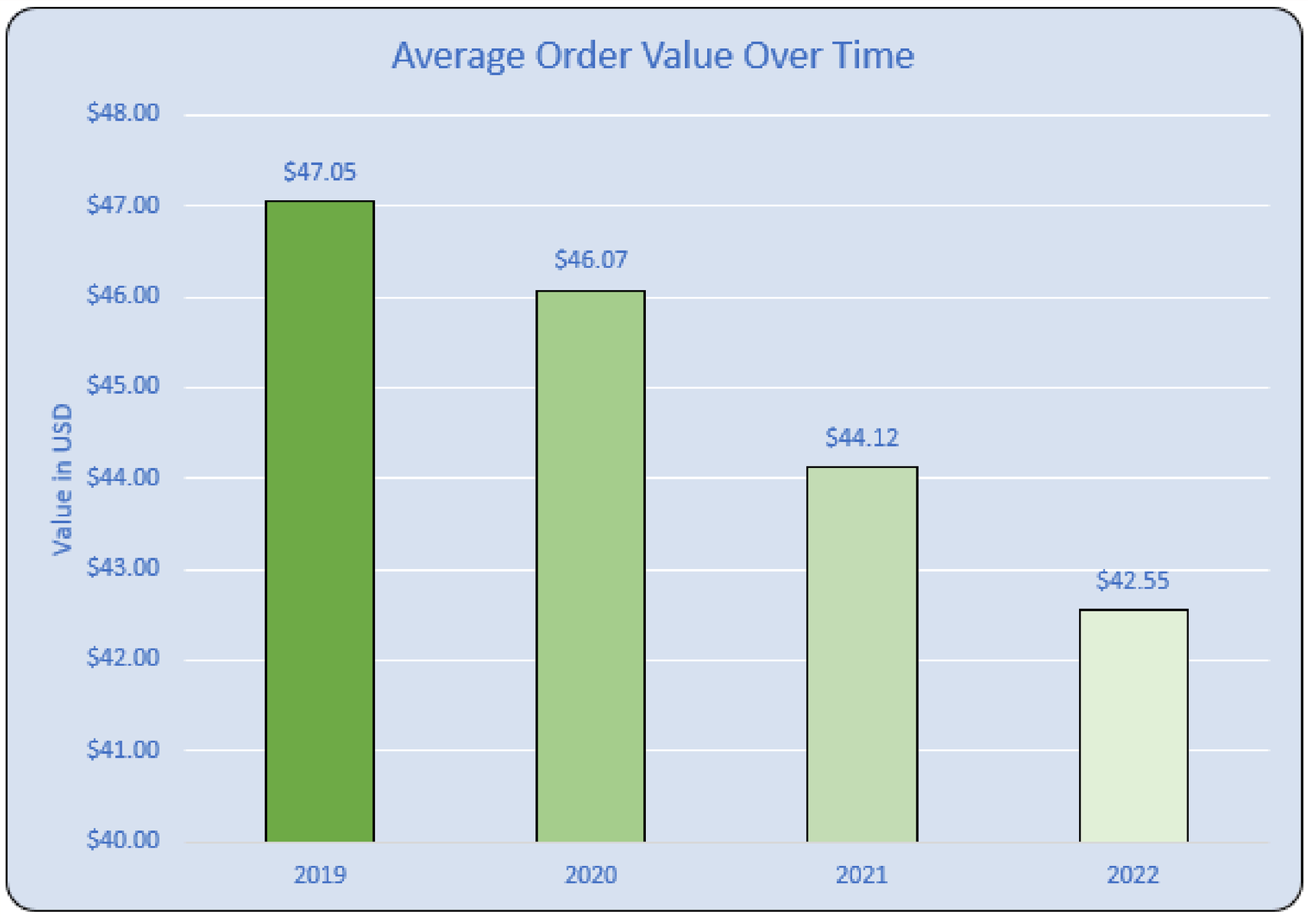


When it comes to Roast Type, Light Roast is the best selling type with a pretty significant difference.



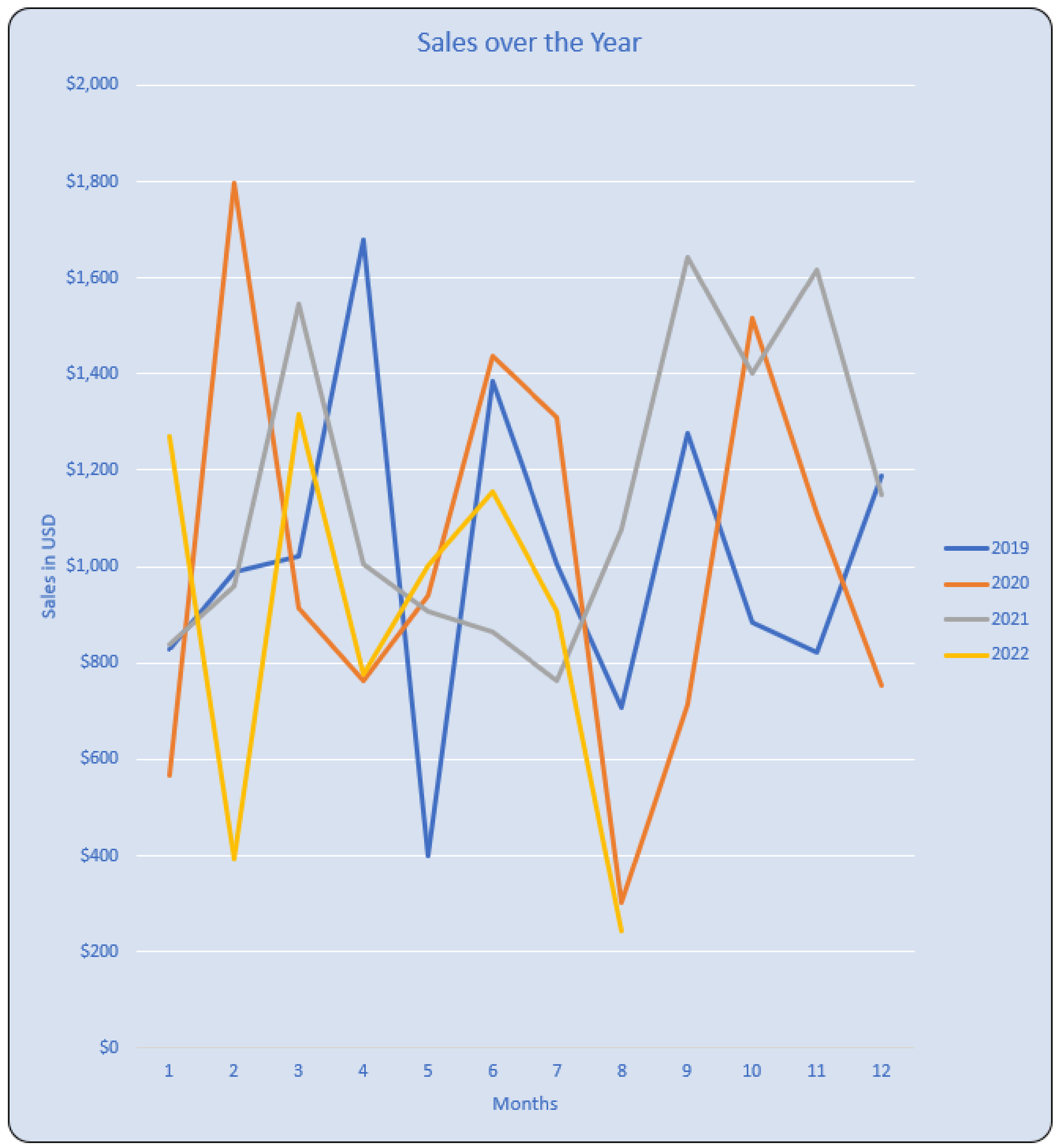
In the Loyalty Program chart we can see that most revenue comes from people that are not part of it.

This is an interesting insight as it could mean that the Loyalty Program isn't attractive enough for more people to join.



Average order value over time is decreasing from one year to another which is a real problem.

This could happen because of competition offering lower prices, way too bigger discounts or a problem with the product.



Sales over the Year shows that first half of the year, especially march and april generate the most income with a significant drop to the end of the summer in August, which is also the worst month.

This is followed by a rapid rise when the cold weather comes.

Key takeaways:

1. **United States brings the most sales.**
2. **Non-Loyalty Program members generate the most sales.**
3. **Excelsa and Ligh Roast generate the most sales when it comes to coffee and roast type.**
4. **March and April are the best months.**
5. **Average Order Value(AOV) is decreasing which needs to be more investigated.**

The end...