

Content

[Scrub](#)

[Explore](#)

[Visualize](#)

Scrub

You are being provided with data about sales at a store. You can access this data by [clicking here \[+ Activity Dataset \]](#) and clicking on Use Template in the upper right corner.

This data has issues that need to be fixed before it can be used for your data analysis project. some issues may include:

- Duplicate records
- Missing values
- Obviously wrong values

Using Google Sheets, review the data, find the issues, and clean the dataset for your use in this project.

What are some errors you identified in this dataset? Identify the column or row where possible. For example: "Duplicate transaction data on rows 45 and 46." You do not need to include every error, but include at least three.

Missing Values in Columns : Product Name, Customer ID, Size, Product Category, Product Line
Duplicate Customer ID.
Duplicate Order Numbers.

How did you fix the errors that you identified in the previous question? For example: "Deleted the duplicate transaction data on row 46."

I replaced the missing values with N/A , the duplicate values that i found in customer id and order numbers were not deleted or made any changes since i came to the conclusion that a single user/customer can order different items and a single order number can consists of various items.

Copy of Data Analytics Course 2 Activity Dataset										
File Edit View Insert Format Data Tools Extensions Help										
100% 123 Default...										
C100 b20d6c81-2021-4eb2-bbdd-2263721d652f										
1	A	B	C	D	E	F	G	H	I	J
1	Date	Order_Number	Customer_ID	Product_Name	SKU	Price	Size	Quantity	Product_Category	Product_Line
98	1/5/2021	LDV-9304-2273-1933	b269404c-5903-4795-9153-d3d6c719ba28	Reddy Beddy	DJWE1V9LZK	\$23.07	large	3	N/A	dog
99	2/16/2020	3N3-3660-1039-1510	b22c9c00-9900-4bd1-9e44-4c006f75e827	Purrrfect Puree	28LQOI0LSK	\$19.96	N/A	2	N/A	cat
100	4/24/2020	LAQ-1426-7637-9037	b20d6c81-2021-4eb2-bbdd-2263721d652f	Chewie Dental	C3Q531YP08	\$38.17	N/A	3	N/A	dog
101	10/7/2019	JET-0967-2500-3073	b2021c96-ca3d-4d16-9e89-6949706e8729	Chomp-a Plush	3HDXSH4WTM	\$39.24	N/A	1	toy	dog
102	3/2/2021	PON-0395-5741-4360	aef74a79-4c4d-436d-9bba-ee01c3939e7	Kitty Climber	W86BRJ9SSG	\$39.32	N/A	3	toy	cat
103	6/30/2019	KGA-9818-3253-4157	ae46598d-caab-4f3b-bc5a-e3a30014bd94	Fetch Blaster	M291KH4JLW	\$29.47	N/A	2	N/A	dog
104	11/16/2019	MAS-0296-0725-6519	ae46598d-caab-4f3b-bc5a-e3a30014bd94	Scratchy Post	MPH6SCD7UT	\$26.95	N/A	3	N/A	cat
105	1/20/2021	T8D-8382-7701-8222	ae1366f5-96ec-4487-a135-31e85fe63195	Reddy Beddy	I2GQUNYDXW	\$36.30	medium	3	bedding	dog
106	1/20/2021	T8D-8382-7701-8222	ae1366f5-96ec-4487-a135-31e85fe63195	N/A	05FYJLBEDH	\$24.53	N/A	3	N/A	cat
107	4/15/2021	XSD-2338-9958-0974	ABC	Reddy Beddy	I2GQUNYDXW	\$36.30	medium	3	bedding	dog
108	4/21/2021	KEC-4821-3488-4229	ABC	Ball and String	1W0NSO3XOD	\$16.71	N/A	3	N/A	cat
109	6/6/2019	269-9595-0109-8368	abb12d96-c95b-45dc-816c-ecaaa39603e8	Snoozer Hammock	V4B4RNS3ZP	\$34.38	N/A	3	N/A	cat
110	4/6/2020	376-9941-7035-3941	ab84d135-5c2b-459a-805a-b985880c8cb8	Chomp-a Plush	3HDXSH4WTM	\$39.24	N/A	1	toy	dog
111	10/14/2020	QXJ-3013-1038-0974	aac8f226-43d2-4b06-987b-5c4224c11b9d	N/A	8PYSMLYINS	\$10.80	N/A	2	N/A	cat
112	1/29/2020	TEM-6615-3245-7573	ad9b601a-f558-48e0-94d5-6888d5890f87	Cat Cave	058G0P7V60	\$29.66	N/A	2	N/A	cat
113	11/18/2020	AOR-8548-3628-4086	ad9b601a-f558-48e0-94d5-6888d5890f87	Whole Chemistry Recipe	6K4AUUS730	\$18.53	small	3	N/A	dog
114	5/14/2021	HAO-7903-2679-8356	ae043696-512d-451b-b21f-7d76e81f17bc	Yum Fish-Dish	GZCJZ3ET04	\$30.63	N/A	2	food	cat
115	3/18/2020	AY0-1789-8103-7174	ae69cca0-92e6-462e-bf58-b6fe9ec2c2f2	Tuna Tasties	A8SU9C4GLF	\$14.53	N/A	2	N/A	cat
116	7/25/2021	STC-6536-1945-1455	ae7b0a06-4c02-4162-b46d-192e8e29dc03	Purrrfect Puree	28LQOI0LSK	\$19.96	N/A	3	N/A	cat
117	3/18/2019	Ti9-8009-3658-7140	a52b4ba6-c632-4ae7-85ae-75c67150da7	Chomp-a Plush	3HDXSH4WTM	\$39.24	N/A	3	N/A	dog
118	7/7/2020	E3W-1000-2722-6137	a4596e54-ab68-4ce5-ac70-6c4c7fde7a29	All Veggie Yummies	OWFPW3WZHG	\$25.48	N/A	1	treat	dog
119	7/9/2020	CH6-5719-2716-5430	a4583689-3ee9-4b2b-8363-0218493e75fb	Snoozer Hammock	V4B4RNS3ZP	\$34.38	N/A	3	N/A	cat
120	10/20/2020	SM2-3854-6886-7834	ae1366f5-96ec-4487-a135-31e85fe63195	Foofy Mouse	8PYSMLYINS	\$10.80	N/A	3	N/A	cat
121	8/24/2021	HMP-2151-3946-4748	ae1366f5-96ec-4487-a135-31e85fe63195	Foofy Mouse	8PYSMLYINS	\$10.80	N/A	2	N/A	cat

Copy of Data Analytics Course 2 Activity Dataset										
File Edit View Insert Format Data Tools Extensions Help										
100% 123 Default...										
B297 YUS-3737-0554-8330										
1	A	B	C	D	E	F	G	H	I	J
1	Date	Order_Number	Customer_ID	Product_Name	SKU	Price	Size	Quantity	Product_Category	Product_Line
2	9/10/2019	01D-9110-4462-5421	525e2268-c032-4dd2-8811-7061d72c0e0	Whole Chemistry Recipe	6K4AUUS730	\$18.53	small	3	food	dog
3	12/10/2019	0H1-4183-3360-8151	bc664ccb-a551-4d5f-a9f9-5572c464cb30	Tuna Tasties	A8SU9C4GLF	\$14.53	N/A	3	treat	cat
4	2/2/2020	0NH-1764-6245-7955	e18b862b-ceba-4438-985f-74b97b90c9f7	Cat Cave	058G0P7V60	\$29.66	N/A	2	N/A	cat
5	4/7/2019	0NV-9141-7747-8049	29cec00b-b9ef-45de-af94-f5231986ea36	Scratchy Post	MPH6SCD7UT	\$26.95	N/A	2	toy	cat
6	7/27/2019	0RO-5522-3689-6676	d63ce3da-22dd-46b6-b3e2-5281cdf6c0f2	Whole Chemistry Recipe	6K4AUUS730	\$18.53	small	3	N/A	dog
7	7/21/2019	148-7275-8810-6218	58bdc8c-2038-4cb6-bcb0-0df4bbcb99759	Yum Fish-Dish	GZCJZ3ET04	\$30.63	N/A	3	food	cat
8	10/9/2020	15T-3023-4253-7686	4825bcac-8271-463d-a098-4d3705f72de7	Reddy Beddy	KDTPMPSB2KZ	\$13.84	small	1	N/A	dog
9	10/9/2020	15T-3023-4253-7686	4825bcac-8271-463d-a098-4d3705f72de7	Purrrfect Puree	28LQOI0LSK	\$19.96	N/A	2	treat	cat
10	8/22/2020	17G-5632-1455-7199	dfb86128-c62a-4234-b628-4312d2d17219	Tug-a-Back	IZBHF5KR79	\$21.82	N/A	2	N/A	cat
11	9/4/2020	17W-0107-3788-4067	c89289f5-c3cc-409c-ac21-a8f6ad8b1ae2	N/A	GZCJZ3ET04	\$30.63	N/A	1	N/A	cat
12	2/1/2020	1D7-1383-8416-8120	8ecbad18-b446-4131-beaa-13459c42b2aa	Reddy Beddy	KDTPMPSB2KZ	\$13.84	small	1	bedding	dog
13	12/12/2020	1FA-9299-1314-8439	b06aaf60-186a-431d-882e-2ed49097e1d9	Reddy Beddy	KDTPMPSB2KZ	\$13.84	small	1	bedding	dog
14	12/12/2020	1FA-9299-1314-8439	b06aaf60-186a-431d-882e-2ed49097e1d9	Feline Fix Mix	RKAPY311TP	\$39.55	N/A	2	treat	cat
15	5/6/2019	1L9-7976-4352-2432	543a9164-7b05-4195-a5fd-b5741f0efe91	Foofy Mouse	8PYSMLYINS	\$10.80	N/A	3	toy	cat
16	7/14/2019	1LD-8694-7766-7200	4d9d52c2-a50b-4bb0-af40-806b119a28e8	Foofy Mouse	8PYSMLYINS	\$10.80	N/A	1	toy	cat
17	10/18/2019	1MK-5419-1276-8899	ee6d3390-4969-43d2-afcf-3601877ac29d	Cat Cave	058G0P7V60	\$29.66	N/A	1	N/A	cat
18	6/20/2019	1NH-9182-3837-7124	49d51a2d-83d1-43eb-a856-37ba73cf8f52	Reddy Beddy	I2GQUNYDXW	\$36.30	medium	3	N/A	dog
19	6/20/2019	1NH-9182-3837-7124	49d51a2d-83d1-43eb-a856-37ba73cf8f52	Yum Fish-Dish	GZCJZ3ET04	\$30.63	N/A	2	N/A	cat
20	3/12/2020	1YR-3971-6042-1854	9d983d64-cd8f-4d91-b1f9-419e044cb2c0	Tug-a-Back	IZBHF5KR79	\$21.82	N/A	1	N/A	cat
21	6/6/2019	269-9595-0109-8368	abb12d96-c95b-45dc-816c-ecaaa39603e8	Snoozer Hammock	V4B4RNS3ZP	\$34.38	N/A	3	N/A	cat
22	1/22/2020	2CT-5208-4608-4140	50802cca-d159-4002-84aa-44bf970badf3	Scratchy Post	MPH6SCD7UT	\$26.95	N/A	2	toy	cat
23	7/19/2020	2OI-1191-6735-8296	98395aed-edab-450a-a45b-df382fbaecb3	N/A	V4B4RNS3ZP	\$34.38	N/A	2	N/A	cat
24	7/19/2020	2OI-1191-6735-8296	N/A	Scratchy Post	MPH6SCD7UT	\$26.95	N/A	1	toy	cat
25	7/19/2020	2OI-1191-6735-8296	N/A	Endless Fix Mix	DKVADY31TD	\$39.55	N/A	2	treat	cat

Explore

Using the spreadsheet tools in Google Sheets, explore the data. You are encouraged to use spreadsheet functions like AVERAGE and CORREL as well as SQL queries like ORDER BY and LIMIT.

When you have used these tools, create a chart that highlights a relationship you discovered in the data. For instance, you might create a bar chart that shows sales of a particular item in different months of the year to showcase how well it sells in warmer months.

What spreadsheet functions did you use and what results did you get? For example: “The AVERAGE of the “sales” column was \$35.55.” You do not need to include every function you used, but include at least three.

```
=AVERAGE('transactions-pet_store-small'!F:F)
=MIN('transactions-pet_store-small'!F:F)
=MAX('transactions-pet_store-small'!F:F)
=MEDIAN('transactions-pet_store-small'!F:F)
=MODE('transactions-pet_store-small'!F:F)
=CORREL('transactions-pet_store-small'!F:F,'transactions-pet_store-small'!H:H)
OUTPUT:
```

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Average Price	Min Price	Max Price	Median Price	Mode Price	Correlation								
2	\$26.03	\$10.80	\$39.55	\$25.48	15.34	-0.00951544446								
3														
4	QUERY EXECUTED:	QUERY('transactions-pet_store-small'!A:J,'SELECT * WHERE F>20 AND F<30 AND I='bedding' ORDER BY H')												
5														
6	Date	Order_Number	Customer_ID	Product_Name	SKU	Price	Size	Quantity	Product_Catego	Product_Line				
7	10/18/2019	6GK-1195-1412	bc064ccb-a551-	Reddy Beddy	DJWE1V9LZK	\$23.07	large	1	bedding	dog				
8	3/18/2021	S11-5756-9839-f	c62dbd98-4e48-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	1	bedding	dog				
9	4/2/2021	T40-2709-4968-	3f444056-2da3-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	1	bedding	dog				
10	6/3/2019	5JP-6492-2648-	00c30446-d919-	Cat Cave	058G0P7V60	\$29.66	N/A	1	bedding	cat				
11	7/19/2020	UIX-8608-7364-	9a14f33e-7caf-4	Reddy Beddy	DJWE1V9LZK	\$23.07	large	2	bedding	dog				
12	4/16/2019	W7D-8546-4966	d3007de5-e5c1-	Reddy Beddy	DJWE1V9LZK	\$23.07	large	2	bedding	dog				
13	3/8/2021	JQQ-6840-1731	6aa6c58-681d-	Cat Cave	058G0P7V60	\$29.66	N/A	2	bedding	cat				
14	8/11/2020	GX9-0927-4423-	cf02289-035d-	Reddy Beddy	DJWE1V9LZK	\$23.07	large	3	bedding	dog				
15	11/12/2020	3FW-2274-2877	934bd721-4fab-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	3	bedding	dog				
16	8/18/2019	4PN-9464-4192-	f3c97e19-9a88-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	3	bedding	dog				
17														
18														
19														
20														
21														
22														
23														
24														
25														

What SQL queries did you use and why? For example: “Used the AND clause to see sales that were for a certain item and above a certain price” You do not need to include every query you used, but include at least three.

I used the following query :

```
=QUERY('transactions-pet_store-small'!A:J,"SELECT* WHERE F>20 AND F<30 AND I='bedding' ORDER BY H")
```

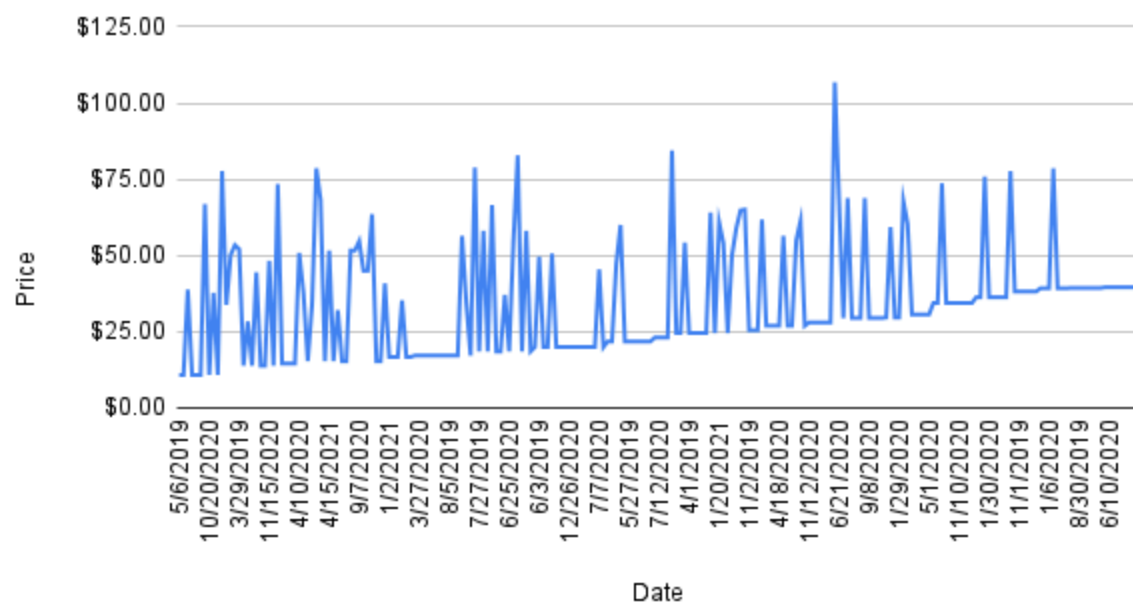
Just for the sake of practicing i used all the clauses and functions i knew in my excel sheet.

The screenshot shows a Google Sheets interface with a spreadsheet titled "Copy of Data Analytics Course 2 Activity Dataset". The spreadsheet has columns A through N. The data is organized into two main sections. The first section, starting from row 1, contains summary statistics: Average Price (\$26.03), Min Price (\$10.80), Max Price (\$39.55), Median Price (\$25.48), Mode Price (\$15.34), and Correlation (-0.00951544446). The second section, starting from row 6, contains a table of transaction data with columns: Date, Order_Number, Customer_ID, Product_Name, SKU, Price, Size, Quantity, Product_Catego, and Product_Line. The data rows show transactions for various products, including "Reddy Beddy" and "Snoozer Essenti". A query bar at the bottom of the spreadsheet shows the query: "QUERY('transactions-pet_store-small'!A:J,"SELECT* WHERE F>20 AND F<30 AND I='bedding' ORDER BY H")".

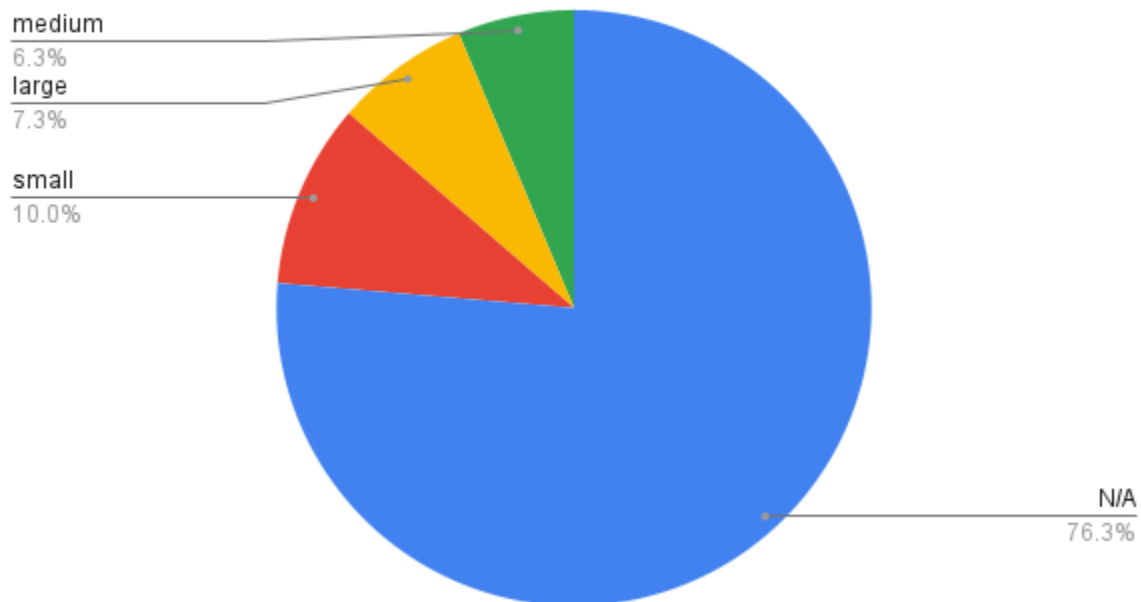
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Average Price	Min Price	Max Price	Median Price	Mode Price	Correlation								
2	\$26.03	\$10.80	\$39.55	\$25.48	\$15.34	-0.00951544446								
3														
4	QUERY EXECUTED:	QUERY('transactions-pet_store-small'!A:J,"SELECT* WHERE F>20 AND F<30 AND I='bedding' ORDER BY H")												
5														
6	Date	Order_Number	Customer_ID	Product_Name	SKU	Price	Size	Quantity	Product_Catego	Product_Line				
7	10/18/2019	6GK-1195-1412	bc664ccb-a551-	Reddy Beddy	DJWE1V9LZK	\$23.07	large	1	bedding	dog				
8	3/18/2021	S11-5756-9839-f	c62dbd98-4e48-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	1	bedding	dog				
9	4/2/2021	T40-2709-4868-	3f444056-2d33-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	1	bedding	dog				
10	6/3/2019	5JP-6492-2648-	00c30446-d919-	Cat Cave	058G0P7V60	\$29.66	N/A	1	bedding	cat				
11	7/19/2020	UIX-8608-7364-f	9a14f33e-7caf-4	Reddy Beddy	DJWE1V9LZK	\$23.07	large	2	bedding	dog				
12	4/16/2019	W7D-8546-4966	d3007de5-e5c1-	Reddy Beddy	DJWE1V9LZK	\$23.07	large	2	bedding	dog				
13	3/8/2021	JQJ-6840-1731	6aa6c58-681d-	Cat Cave	058G0P7V60	\$29.66	N/A	2	bedding	cat				
14	8/11/2020	GX9-0927-4423-	cf02289-035d-	Reddy Beddy	DJWE1V9LZK	\$23.07	large	3	bedding	dog				
15	11/12/2020	3FW-2274-2877	934bd721-4fab-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	3	bedding	dog				
16	8/18/2019	4PN-9464-4192-	f3c97e19-9a88-	Snoozer Essenti	GABWVMEL2R	\$28.04	N/A	3	bedding	dog				
17														
18														
19														
20														
21														
22														
23														
24														
25														

Copy and paste at least one chart into this document that was created from the dataset.

Price vs Date



Count of Size



Visualize

Using the dataset you have scrubbed and explored, create a dashboard with at least two charts and at least one interaction.

Copy and paste the URL for your published Tableau Public dashboard

https://public.tableau.com/views/TableauWork_17177079264840/Dashboard1?:language=en-US&publish=yes&:sid=&:display_count=n&:origin=viz_share_link

Copy and paste an image of the dashboard downloaded from Tableau Public

