**To Vice President of Operations, Portugal:**

Opening: We are honored to assist you in the Portuguese market and look forward to sharing our insights through well-curated data. We understand the importance of evaluating key factors for your company's future structure and strive to provide the right data at the right time to serve as a blueprint for the next 2-5 years. We hope this information will help build a stable model to be repeated as you see fit during the restructuring of your company. An appendix section can be found at the conclusion of this analysis to explain the Extraction, Transformation, and loading of provided data.

As desired, we have taken a deeper dive into the pharmaceutical department and with the available data given to us, we can help pull out key insights. **Our first question**, starts by looking at the most profitable and counter profitable stores to identify unique points of interest. We are asking the question, which pharmacy stores have the highest and lowest sales?
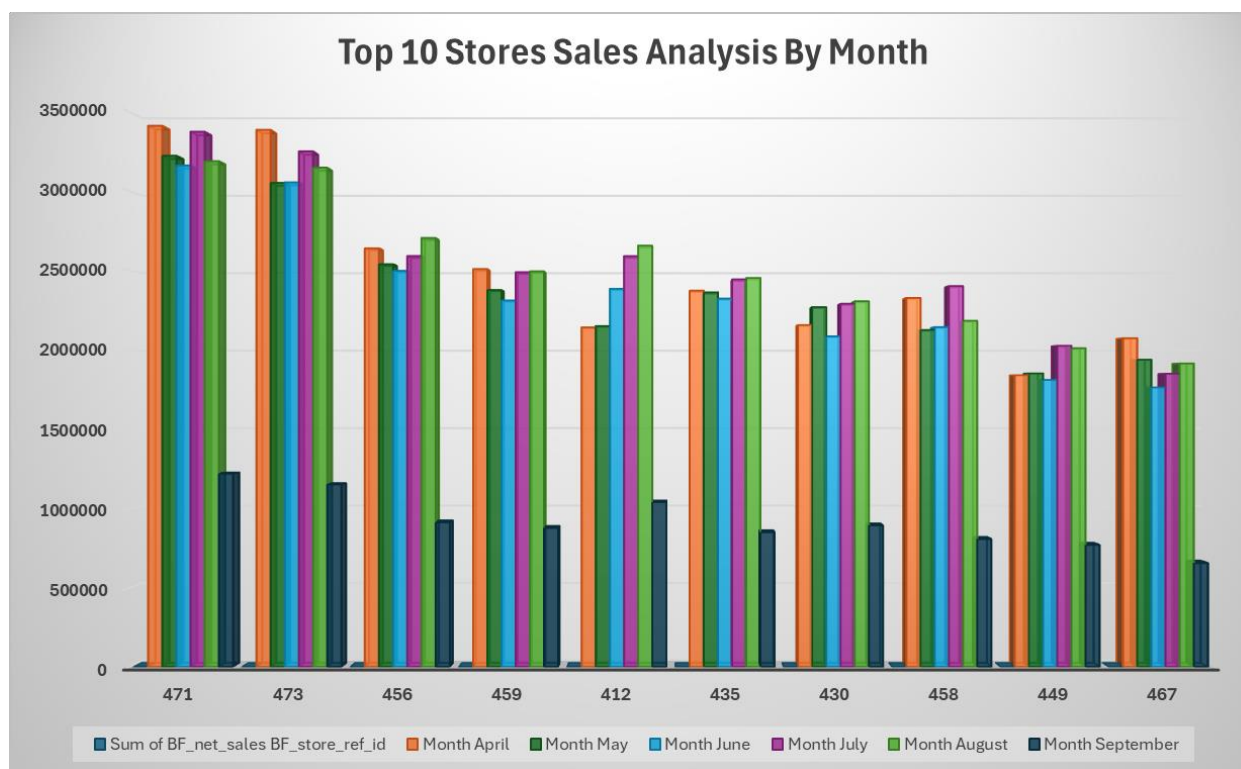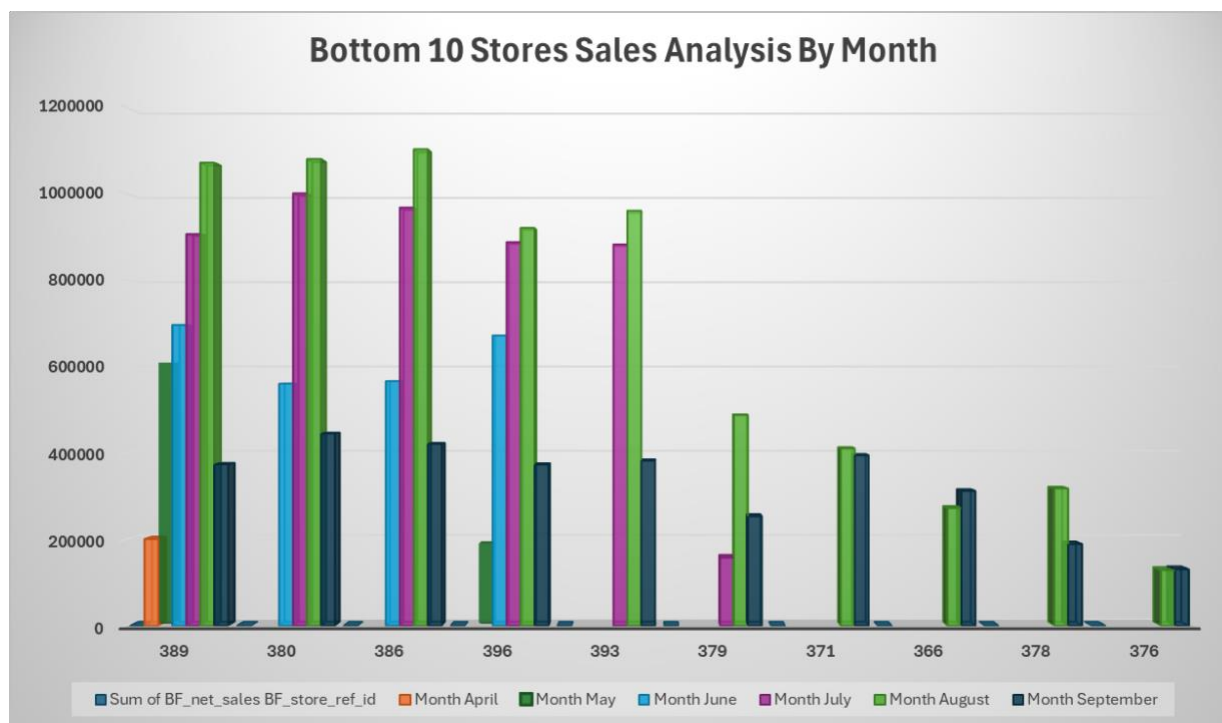

Top 10 Stores Sales Analysis By Month

Store Ref ID	Number of Drugs Sold	Total Qty Bill	Number of Customers Visited	Net Sales After Sales Return	Sales Discount (MRP-Spend Bill)
471	23	275,240	11,729	$17,576,557.63	$11,430,356.63
473	29	252,091	9,312	$17,041,460.27	$10,089,218.39
456	27	208,107	8,739	$13,883,667.76	$8,587,228.19
459	28	196,322	10,393	$13,055,650.02	$8,270,698.31
412	29	187,602	10,716	$12,959,373.28	$7,802,118.00
435	24	180,808	9,761	$12,802,546.09	$7,578,688.39
430	34	175,387	8,264	$11,997,495.79	$6,905,694.87
458	26	175,775	8,442	$11,983,768.07	$7,746,476.44
449	24	146,538	7,387	$10,305,204.59	$5,286,568.15
467	24	157,277	10,463	$10,190,852.36	$6,276,629.33

According to the analysis, the chart above shows the **top 10** stores' sales analysis by month, and the table shows a deeper analysis of the number of drugs sold, total quantity in the bill, number of customers who visited the stores, net sales after the sales return, and the sales discount derived from MRP, subtracting the total spend bill for the top 10 stores. The sales discount can be considered as the customer insurance coverage.

We have the sales records from 4/1/2019 to 9/13/2019, five and a half months of sales in our analysis, shown in the DateID column. Store ID 471, 473, 456, 412, 435, 430, 458, 449, and 467 are the top ten stores with the highest net sales

among 41 stores. Store ID 471 has the highest net sales of $17.6 million and sales discounts of $11.4 million, with the highest number of customers visiting the store of 11,729 and total items sold of 275,240. However, it has the lowest variety of drugs sold of 23 products compared to other stores. In contrast, Store ID 430 had the highest variety, offering 34 different products for sale. Although it has a variety of drug products, it is not a popular store, it is 7th place for the store sales comparison.

From the chart, we found that the stores' sales fluctuated from April to August 2019. But store ID 412 showed a sales improvement. We exclude the analysis for September 2019 for the comparison because it is only a half-month record.

Bottom 10 Stores Sales Analysis By Month

Store Ref ID	Number of Drugs Sold	Total Qty Bill	Number of Customers Visited	Net Sales After Sales Return	Sales Discount (MRP-Spend Bill)
389	20	57,399	5,095	$3,858,361.51	$2,329,002.93
380	27	48,782	3,762	$3,086,283.65	$2,287,465.76
386	24	49,650	4,535	$3,057,640.02	$2,160,709.68
396	26	49,227	4,673	$3,043,442.43	$2,146,494.38
393	20	32,552	3,042	$2,225,119.60	$1,206,904.36
379	15	14,400	1,563	$898,723.73	$607,943.91
371	17	11,210	1,418	$805,237.23	$467,580.92
366	17	8,203	1,188	$584,475.09	$387,288.80
378	14	8,699	1,284	$507,059.98	$470,704.14
376	12	4,123	581	$259,676.58	$233,760.82

The chart above shows the **bottom 10** stores' sales analysis by month, and the table shows a deeper analysis of the number of drugs sold, total quantity in the bill, number of customers who visited the stores, net sales after the sales return, and the sales discount derived from MRP subtract the total spend bill for bottom 10 stores. We assume the sales discount is considered as the customer insurance coverage.

From the data analysis, we found that store ID 396 was newly opened in May 2019, store IDs 380 and 386 were newly opened in June 2019, store IDs 393 and 379 were newly opened in July 2019, and store IDs 371, 366, 378, and 376 were newly opened in August 2019. Therefore, it will be inappropriate to compare the total sales from April to September 2019.

For example, store ID 378 was newly opened on 8/16/2019, it only has half-month August sales. Store ID 376 was newly opened on 8/19/2019, store ID 371 was newly opened on 8/24/2019, and lastly, store ID 366 was newly opened on 8/27/2019, which only has five days of August sales. Hence, we concluded that store ID 389, which has the complete period (April to September) has the lowest store sales compared to other stores. We excluded the newly opened stores mentioned above for the sales comparison because they need time to build sales growth.

An investigation is required regarding the lower sales for the new store ID 376, business started on 8/19/2019 as compared to new store ID 378, business started on 8/16/2019. Only four days difference, the net sales have a discrepancy of $247,883.40 ($507,059.98 - $259,676.58). Why does store ID 376 only have 581 customers visiting the store? Is it a location problem, or does a new store need advertisements to promote the new store and enhance sales growth?

Store ID ⤓	Total Qty Bill	Sales Return	% of Sales Return	Number of Customers Visited
459	196,322	$553,494.66	6.41%	10,393
471	275,240	$547,409.84	6.34%	11,729
473	252,091	$503,193.32	5.83%	9,312
467	157,277	$327,376.25	3.79%	10,463
439	128,721	$321,248.92	3.72%	6,897
435	180,808	$305,982.34	3.54%	9,761
455	163,758	$301,269.05	3.49%	9,839
458	175,775	$290,835.33	3.37%	8,442
424	169,656	$271,962.91	3.15%	8,038
412	187,602	$256,721.71	2.97%	10,716

The table above shows the top 10 stores' sales return by value and number of customers. Store IDs 459 and 471 have the top two highest sales returns with about 6% of sales returns among all stores. Followed by store ID 473 with 5.83%. We used a count distinct for the customers coming for sales return. Store IDs 459 and 471 have several customers for the sales return of 10,393 and 11,729, respectively. An investigation is required by the management to identify the reason for frequent returns, whether due to customer dissatisfaction with pricing issues, or product quality concerns.

Our second question: We then wanted to take a deeper approach to focusing on which months have the highest sales revenue? This approach seeks to capture the data over a time series by using DateID as an anchor for the analysis.

One of the obstacles when seeking sales revenue comparisons amongst stores is that not every retail location has reported data. One noteworthy finding in the data is that those stores who have the most missing values seem to perform the worst for those years of which they do report; this includes stores 366, 371, 376, & 379 all who only report two of the six-month available timelines. Meanwhile those stores in the top ten ranking all share the common theme of having reported data for six out of the six months- the remainder of this answer will focus on those top ten retail locations.

It is important to note September data only accounts for the first thirteen days of the month and therefore there is not a fair comparison between April through August. For the remainder of answering this question only the months of April through August will be assessed. A full set of monthly sales can be found as an exhibit in the pivot table excerpt below. The top ten highest revenue retail locations are stores 417, 473, 456, 459, 412, 430, 458, 449, 467- in ascending order. The chart below confirms what was explored initially in question#1 above while showing the data within a matrix structure.

Sum of net_sales, per month, per store with distinct counts of the number of days activity in each month.

Column Labels

Store	April — Apr Net Sales, Sum	# of Days	May — May Net Sales, Sum	# of Days	June — Jun Net Sales, Sum	# of Days	July — Jul Net Sales, Sum	# of Days	August — Aug Net Sales, Sum	# of Days	September — Sep Net Sales, Sum	# of Days	Total Net Sales, Sum
366									$272,498.40	5	$311,976.69	13	$584,475.09
371									$410,631.20	8	$394,606.04	13	$805,237.23
376									$128,905.63	13	$130,770.95	13	$259,676.58
378									$317,899.71	16	$189,160.27	13	$507,059.98
379							$157,731.37	10	$487,627.52	31	$253,364.84	13	$898,723.73
380					$560,213.69	21	$1,000,878.79	31	$1,080,276.53	31	$444,914.64	13	$3,086,283.65
386					$566,093.15	21	$967,714.46	31	$1,102,932.99	31	$420,899.43	13	$3,057,640.02
389	$200,724.21	15	$609,268.45	31	$696,361.02	30	$906,627.12	31	$1,071,732.52	31	$373,648.18	13	$3,858,361.51
392	$579,330.32	30	$620,886.48	31	$640,144.82	30	$746,741.37	31	$951,888.24	31	$363,448.74	13	$3,902,439.97
393							$882,567.73	28	$960,382.44	31	$382,169.42	13	$2,225,119.60
396			$189,879.35	15	$671,890.36	30	$887,923.11	31	$920,794.40	31	$372,955.22	13	$3,043,442.43
399	$470,867.98	30	$765,455.27	31	$912,342.73	30	$1,088,156.66	31	$1,223,001.12	31	$528,490.26	13	$4,988,314.02
402	$942,814.60	30	$905,504.66	31	$875,366.49	30	$1,122,238.51	31	$1,155,678.29	31	$507,269.19	13	$5,508,871.75
405	$978,853.68	30	$1,044,862.96	31	$949,492.59	30	$1,254,061.65	31	$1,167,524.69	31	$453,630.11	13	$5,848,425.68
408	$1,756,690.97	30	$1,663,280.90	31	$1,732,465.14	30	$1,889,706.39	31	$2,055,829.80	31	$705,506.92	13	$9,803,480.12
411	$1,275,573.56	30	$1,420,304.72	31	$1,471,248.83	30	$1,602,363.48	31	$1,702,568.08	31	$631,185.11	13	$8,103,243.78
412	$2,140,683.27	30	$2,149,161.81	31	$2,384,180.47	30	$2,589,559.51	31	$2,656,319.88	31	$1,039,468.33	13	$12,959,373.28
414	$1,046,694.19	30	$1,138,475.81	31	$1,057,021.78	30	$1,041,670.18	31	$1,079,798.00	31	$396,142.32	13	$5,759,802.28
418	$1,482,697.31	30	$1,269,873.15	31	$1,374,771.38	30	$1,429,891.46	31	$1,662,472.40	31	$574,267.69	13	$7,793,973.40
421	$1,869,476.64	30	$1,809,142.28	31	$1,789,390.74	30	$1,795,835.77	31	$1,950,940.79	31	$717,127.52	13	$9,931,913.74
424	$1,690,850.41	30	$1,817,285.97	31	$1,685,083.71	30	$1,995,064.56	31	$1,815,782.58	31	$694,226.89	13	$9,698,294.11
430	$2,155,661.63	30	$2,268,778.91	31	$2,084,188.39	30	$2,289,834.59	31	$2,305,563.21	31	$893,469.07	13	$11,997,495.79
432	$1,790,530.98	30	$1,689,315.32	31	$1,721,587.13	30	$1,872,622.84	31	$1,868,844.17	31	$758,173.40	13	$9,701,073.83
435	$2,372,574.64	30	$2,361,385.71	31	$2,322,499.17	30	$2,443,748.21	31	$2,452,556.26	31	$849,782.10	13	$12,802,546.09
436	$1,216,632.64	30	$1,241,673.83	31	$1,035,098.82	30	$1,273,014.41	31	$1,303,906.09	31	$544,366.07	13	$6,614,691.85
439	$1,577,561.08	30	$1,517,402.03	31	$1,424,962.74	30	$1,374,970.64	31	$1,455,398.13	31	$497,550.34	13	$7,847,844.97
441	$1,658,331.77	30	$1,569,075.50	31	$1,590,809.73	30	$911,619.68	20	$1,333,435.26	31	$469,744.64	13	$7,533,016.56
444	$1,740,696.35	30	$1,837,705.85	31	$1,923,919.77	30	$1,956,297.60	31	$1,988,447.37	31	$698,606.12	13	$10,145,673.06
447	$1,865,909.43	30	$1,821,051.17	31	$1,609,577.32	30	$1,785,654.78	31	$1,832,210.92	31	$710,252.34	13	$9,624,655.97
449	$1,839,460.20	30	$1,850,970.39	31	$1,810,328.97	30	$2,025,128.45	31	$2,009,827.71	31	$769,488.87	13	$10,305,204.59
452	$1,682,434.97	30	$1,710,237.48	31	$1,538,996.02	30	$1,678,313.79	31	$1,589,339.86	31	$612,162.96	13	$8,811,485.08
454	$1,453,134.03	30	$1,457,772.86	31	$1,392,756.86	30	$1,437,042.46	31	$1,367,901.34	31	$510,811.97	13	$7,619,419.53
455	$1,888,497.08	30	$1,813,344.27	31	$1,685,627.23	30	$2,006,958.69	31	$1,942,550.75	31	$672,026.20	13	$10,009,004.22
456	$2,639,452.55	30	$2,538,935.55	31	$2,497,597.34	30	$2,590,572.06	31	$2,704,622.77	31	$912,487.49	13	$13,883,667.76
458	$2,326,445.45	30	$2,124,299.63	31	$2,142,486.03	30	$2,401,110.82	31	$2,182,590.33	31	$806,835.81	13	$11,983,768.07
459	$2,508,561.40	30	$2,376,815.86	31	$2,310,623.05	30	$2,488,735.34	31	$2,494,054.23	31	$876,860.14	13	$13,055,650.02
462	$1,382,432.96	30	$1,533,840.91	31	$1,344,640.36	30	$1,387,487.22	31	$1,285,271.84	31	$502,836.11	13	$7,436,509.39
467	$2,072,964.03	30	$1,937,710.66	31	$1,762,502.76	30	$1,848,935.78	31	$1,913,547.02	31	$655,192.12	13	$10,190,852.36
468	$1,545,165.47	30	$1,470,692.45	31	$1,423,236.57	30	$1,452,404.87	31	$1,621,572.12	31	$531,568.48	13	$8,044,639.96
471	$3,411,166.29	30	$3,222,435.25	31	$3,161,430.80	30	$3,373,988.62	31	$3,188,654.94	31	$1,218,881.73	13	$17,576,557.63
473	$3,386,074.03	30	$3,051,853.65	31	$3,055,011.00	30	$3,250,168.66	31	$3,147,448.27	31	$1,150,904.66	13	$17,041,460.27
~	$54,948,944.14	30	$54,798,679.08	31	$55,203,946.94	31	$61,207,341.65	30	$64,163,227.78	31	$24,527,229.38	13	$314,849,368.96

The highest retail revenue locations for the months April through August are listed below and in this subset of data shows in red the fifteen lowest performing months per location along with the three least performing stores overall.

Sales Analysis by Month ~Top/Bottom 15 Months & Top/Bottom 3 Retail Locations

| Sum of net_sales Store_id	Month April	May	June	July	August	Grand Total
412	$2,140,683.27	$2,149,161.81	$2,384,180.47	$2,589,559.51	$2,656,319.88	$11,919,904.95
430	$2,155,661.63	$2,268,778.91	$2,084,188.39	$2,289,834.59	$2,305,563.21	$11,104,026.72
435	$2,372,574.64	$2,361,385.71	$2,322,499.17	$2,443,748.21	$2,452,556.26	$11,952,763.99
449	$1,839,460.20	$1,850,970.39	$1,810,328.97	$2,025,128.45	$2,009,827.71	$9,535,715.71
456	$2,639,452.55	$2,538,935.55	$2,497,597.34	$2,590,572.06	$2,704,622.77	$12,971,180.27
458	$2,326,445.45	$2,124,299.63	$2,142,486.03	$2,401,110.82	$2,182,590.33	$11,176,932.26
459	$2,508,561.40	$2,376,815.86	$2,310,623.05	$2,488,735.34	$2,494,054.23	$12,178,789.88
467	$2,072,964.03	$1,937,710.66	$1,762,502.76	$1,848,935.78	$1,913,547.02	$9,535,660.25
471	$3,411,166.29	$3,222,435.25	$3,161,430.80	$3,373,988.62	$3,188,654.94	$16,357,675.90
473	$3,386,074.03	$3,051,853.65	$3,055,011.00	$3,250,168.66	$3,147,448.27	$15,890,555.61
Grand Total	$24,853,043.50	$23,882,347.41	$23,530,847.97	$25,301,782.04	$25,055,184.62	$122,623,205.55

Our third question looks at what are the top 10 pharmacy stores with the highest sales returns? A high percentage of returns for a specific drug type, indicated by the return_value_bill, can negatively affect profitability? This question was not entirely possible to answer to the fullest extent we would have liked to answer due to limited data on hand. We did, however, observe the top ten retail locations listed below are the ten months with the greatest dollar amount of returned merchandise, along with the top three highest revenue earning locations among total returns for the months spanning April to August.

Returns, 10 Highest Months & Top 3 Locations

| Sum of return_value_bill	Column Labels					
Store ID#	April	May	June	July	August	~
412	$1,647.88	$5,001.17	$20,242.93	$98,729.72	$98,602.37	$224,224.07
430	$5,881.95	$2,720.37	$6,327.89	$111,012.77	$84,297.34	$210,240.32
435	$1,332.79	$5,653.05	$40,688.75	$124,545.12	$105,768.13	$277,987.84
449	$2,983.27	$1,834.13	$9,027.54	$80,211.24	$79,978.59	$174,034.77
456	$4,544.46	$7,245.47	$15,006.56	$84,747.28	$82,415.92	$193,959.68
458	$1,692.36	$3,256.63	$15,175.36	$140,545.59	$95,389.76	$256,059.70
459	$15,639.56	$121,185.15	$78,743.65	$141,156.57	$150,920.02	$507,644.96
467	$7,275.96	$70,102.30	$55,523.67	$75,194.74	$96,841.57	$304,938.23
471	$17,397.61	$97,909.43	$130,899.49	$138,583.06	$124,682.03	$509,471.62
473	$18,304.26	$95,245.94	$129,909.87	$138,121.60	$84,929.58	$466,511.24
~	$76,700.09	$410,153.64	$501,545.72	$1,132,847.69	$1,003,825.30	$3,125,072.44

We then showcase among the ten lowest preforming locations, the ten highest months of returns and the top three stores for returns by monthly dollar amount. The matrix style data table below also demonstrates a constraint we ran into of missing data points in our analysis.

Returns, 10 Highest Months & Top 3 Locations ~ Lowest Preforming Stores

Sum of return_value_bill	Column Labels					
Store ID#	April	May	June	July	August	~
366					$30,422.75	$30,422.75
371					$34,553.42	$34,553.42
376					$9,491.60	$9,491.60
378					$24,010.52	$24,010.52
380			$38,432.29	$45,395.26	$36,141.84	$119,969.39
386			$41,229.86	$44,263.20	$68,748.10	$154,241.16
389	$214.32	$662.39	$4,021.65	$44,483.20	$69,070.52	$118,452.09
393				$47,620.61	$55,854.07	$103,474.67
396		$5,284.73	$42,590.76	$55,740.19	$34,157.06	$137,772.73
~	$214.32	$5,947.12	$126,274.56	$237,502.47	$362,449.88	$732,388.34

None-the-less, we are then able to take the three stores identified above and compare their return rate against total sales for each of the three locations as identified below. It is noteworthy to mention this April and May are largely incomplete data sources we will only focus on June, July, and September. *This purpose of this model is to identify the weakest location with the highest amount of returns in hopes to either A) target a remedy to lower returns and thus increase revenue margin B) close the location altogether.* We can see below store# 386 would be the ideal candidate to focus on first- this analysis model can then be repeated as needed to identify further locations.
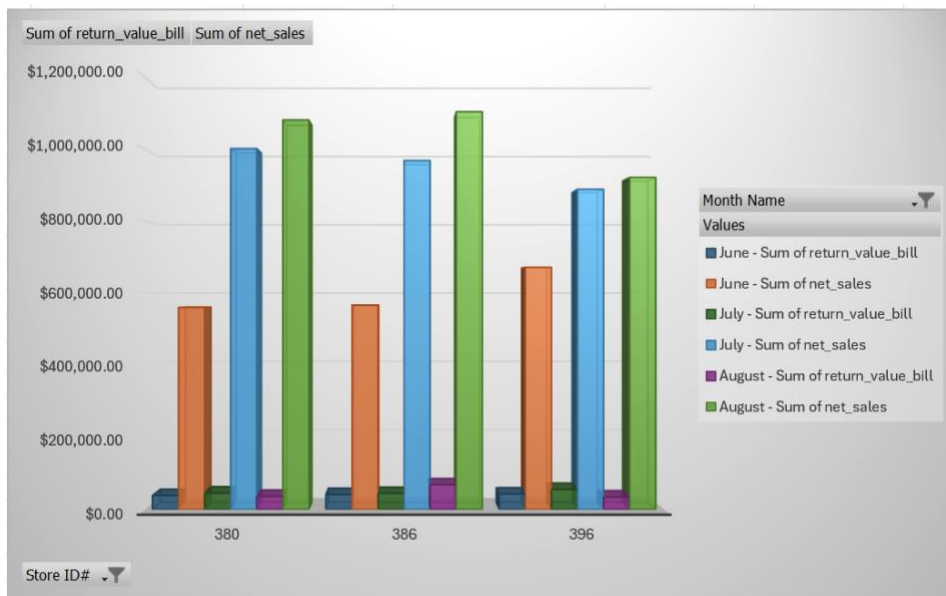
Going beyond revenue purely at the store level we next take a look at the individual consumer level. **Our fourth main question** then becomes there are null values in the payment_ref_id column, accounting for a total of $3.2M across all stores, with no specified payment type (cash, check, or card). Does this indicate outstanding payment? In our opinion, any such null values in payments become a red flag and be a cause for further investigation below.
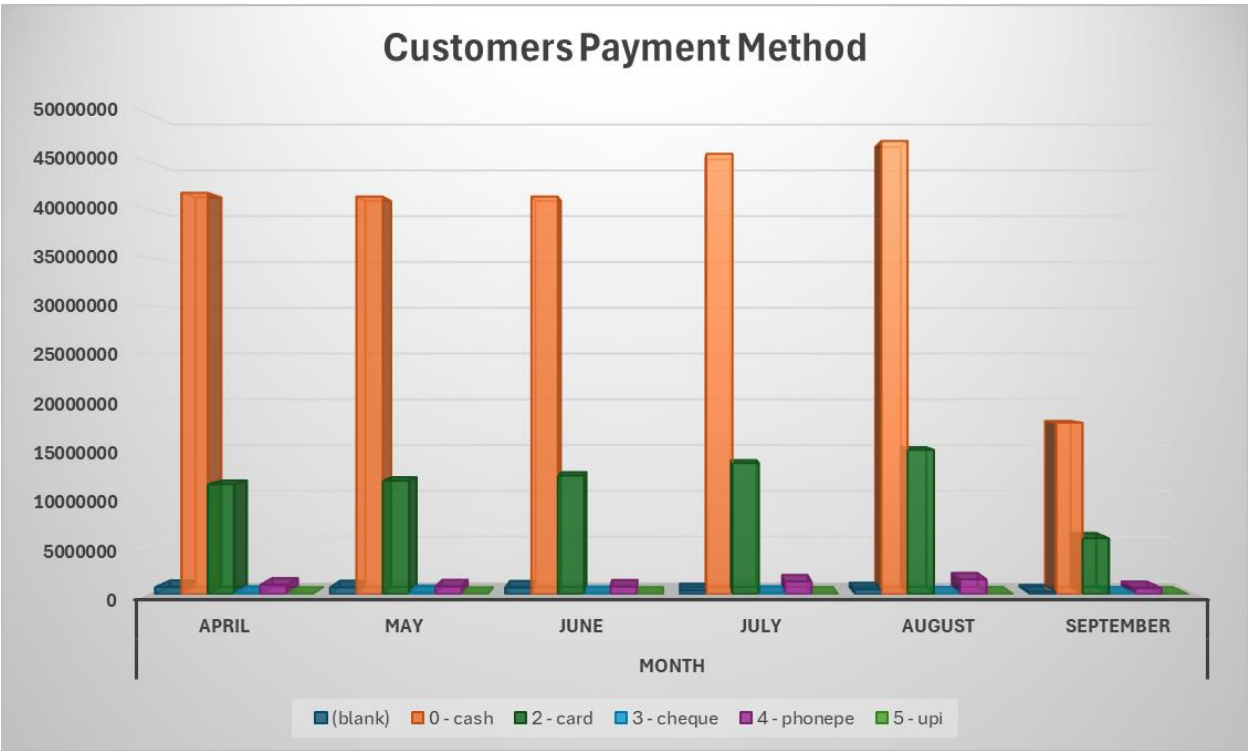

From the customers' payment method chart above, we found that customers' preferred payment method is cash (75%), followed by card (22%), and by phone (1.92%).

Payment Ref	April	May	June	July	August	September	Grand Total	Percentage
(blank)	733,467.83	705,702.76	676,279.71	386,163.55	485,746.53	252,838.77	3,240,199.16	1.03%
0 - cash	41,695,771.98	41,311,018.70	41,309,694.12	45,739,589.25	47,084,628.41	17,809,221.52	234,949,923.98	74.62%
2 - card	11,427,041.45	11,784,249.11	12,323,241.08	13,615,126.02	14,987,377.47	5,805,267.83	69,942,302.95	22.21%
3 - cheque	110,225.77	189,881.73	89,910.85	124,082.74	43,691.92	34,763.72	592,556.72	0.19%
4 - phonepe	976,530.08	802,579.58	790,849.71	1,317,708.62	1,536,448.42	615,695.97	6,039,812.40	1.92%
5 - upi	5,907.02	5,247.21	13,971.47	24,671.47	25,335.02	9,441.57	84,573.75	0.03%
						Total:	314,849,368.96	100.00%

On the following page we can observe the top ten blank payment methods by net sales. This can help your company identify missing data points which if filled can be leveraged for further analysis.

Consultant Team: Jacelyn Lim, Michael ORourke, Anchal Sharda

Top 10 Stores - Blank Payment Method by Net Sales
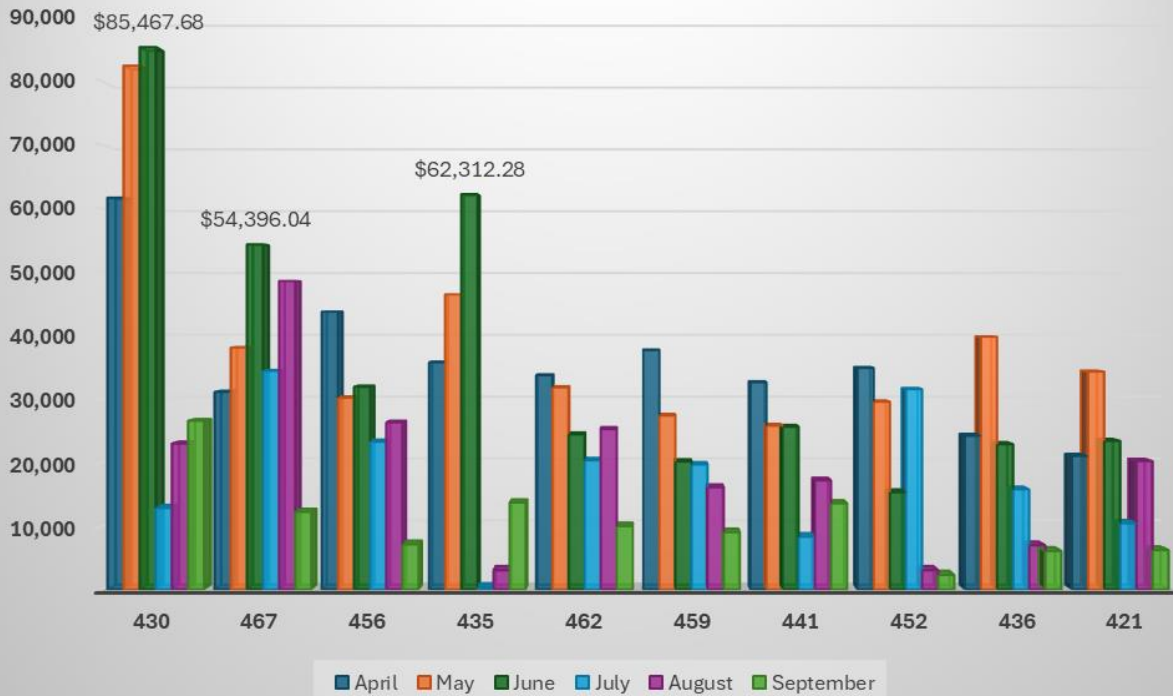
Store ID	April	May	June	July	August	September	Grand Total
430	$61,757.91	$82,526.23	$85,467.68	$12,857.35	$22,994.28	$26,585.07	$292,188.52
467	$31,131.91	$38,119.65	$54,396.04	$34,455.43	$48,622.02	$12,221.89	$218,946.94
456	$43,782.55	$30,211.44	$31,943.00	$23,298.50	$26,326.60	$7,112.59	$162,674.67
435	$35,787.59	$46,504.22	$62,312.28	$310.79	$3,052.55	$13,716.57	$161,683.99
462	$33,753.30	$31,851.33	$24,445.22	$20,374.36	$25,289.11	$9,993.61	$145,706.93
459	$37,717.28	$27,412.72	$20,156.52	$19,696.91	$16,112.79	$9,039.73	$130,135.95
441	$32,647.56	$25,818.94	$25,643.63	$8,318.09	$17,188.71	$13,589.68	$123,206.61
452	$34,916.82	$29,576.92	$15,332.79	$31,574.86	$3,046.83	$2,296.14	$116,744.35
436	$24,229.76	$39,815.75	$22,893.27	$15,784.47	$6,970.52	$6,065.52	$115,759.29
421	$21,019.32	$34,333.84	$23,415.10	$10,432.23	$20,233.89	$6,210.34	$115,644.73

From the data analysis, we noticed that there was a 1% blank payment method, a total of $314 million in net sales. It appeared in each month from April to September 2019. The top 10 stores on blank payment method are store ID 430, 467, 456, 435, 462, 459, 441, 452, 436, and 421. It doesn't look like a system error because it appears each month in each store. Therefore, we assume there were the customers' outstanding amounts.

The chart above shows that Jun 2019 has the highest blank payment method, especially for store IDs 430, 467, and 435. Followed by May 2019. Based on our assumption, if the blank payment is the amount outstanding. The top 10 stores that have the highest outstanding amount are store ID 430, 467, 456, 435, 462, 459, 441, 452, 436, and 421. An investigation is required for store IDs 430 and 467 for more than $200k outstanding from April to September 2019. An action needs to be followed up by the accounts receivables department on the outstanding payment.

Top 10 Stores - Blank Payment Method by Net Sales
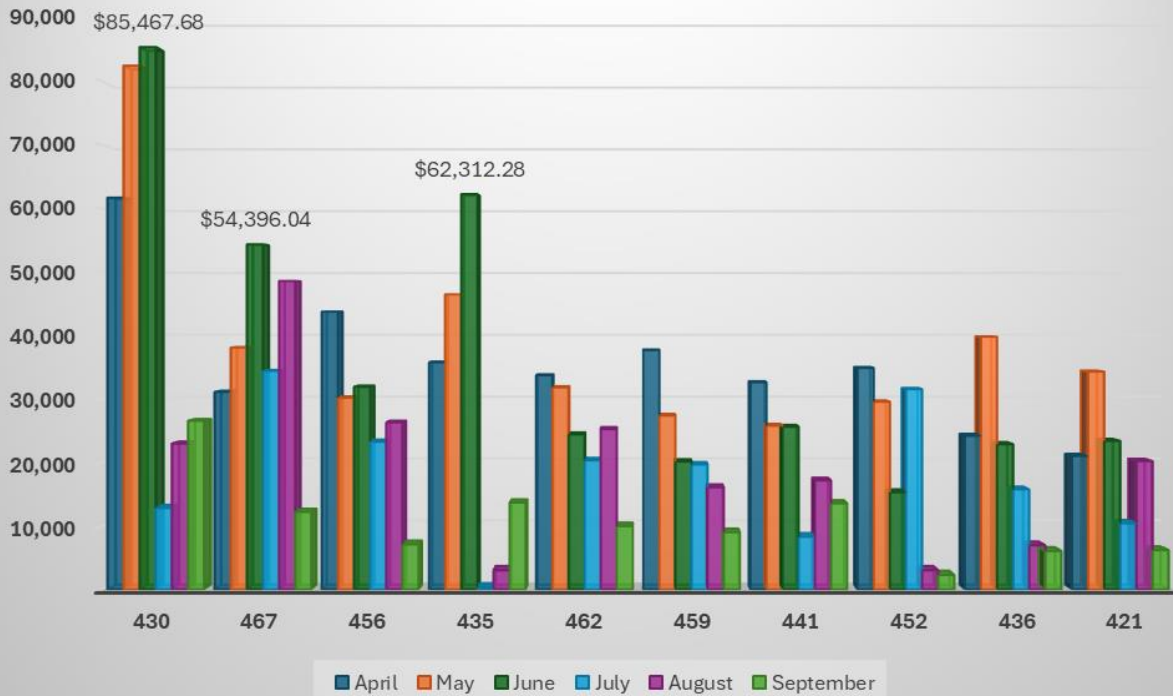
We further analyzed the top 10 stores that have a blank payment method by number of bills. We applied a count distinct for the number of bills. The trends will not be the same as analyzed by net sales amount. This represents some bills having a larger outstanding amount, but fewer outstanding bills. For example, store IDs 467 and 435 have the number of outstanding bills of 162 and 144, respectively. However, their outstanding amounts are reversed pattern (store ID 467: $54k, 435: $62k). The chart above shows that store ID 430 has the highest number of outstanding bills, with 202 in May and 213 in June 2019.
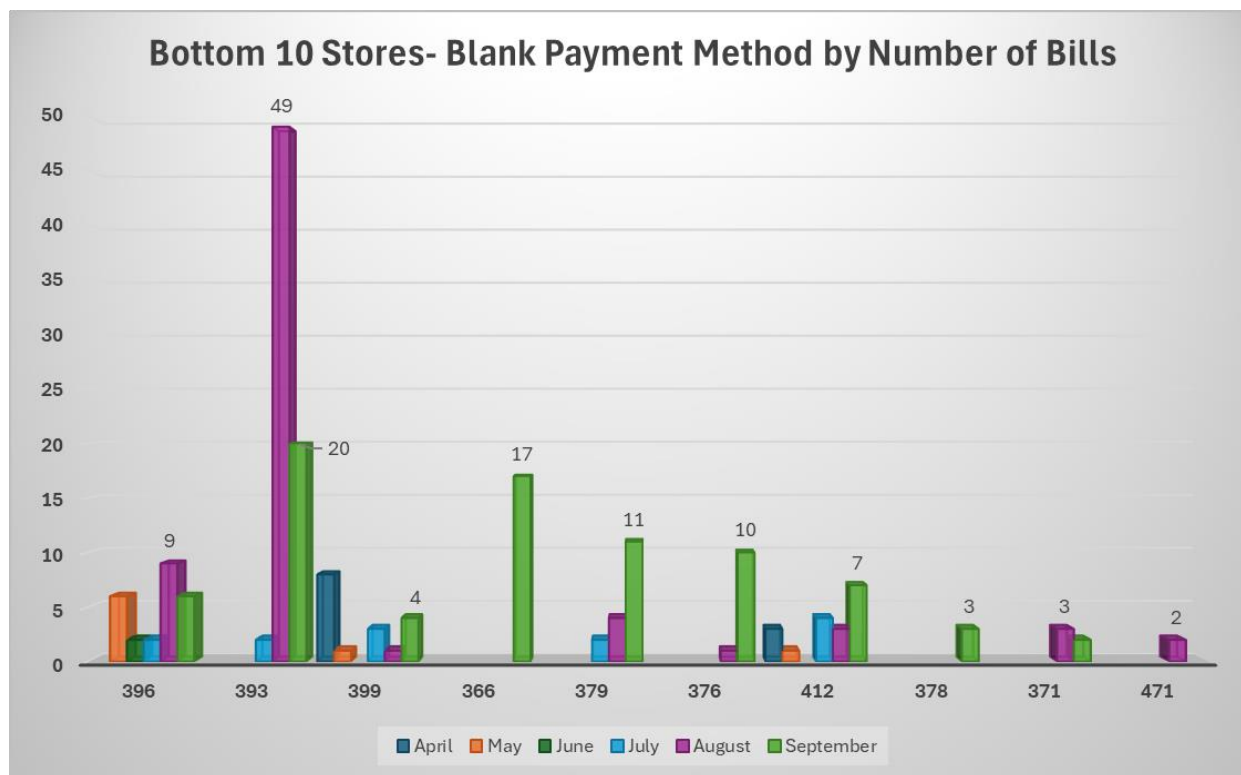

Bottom 10 Stores - Blank Payment Method By Net Sales
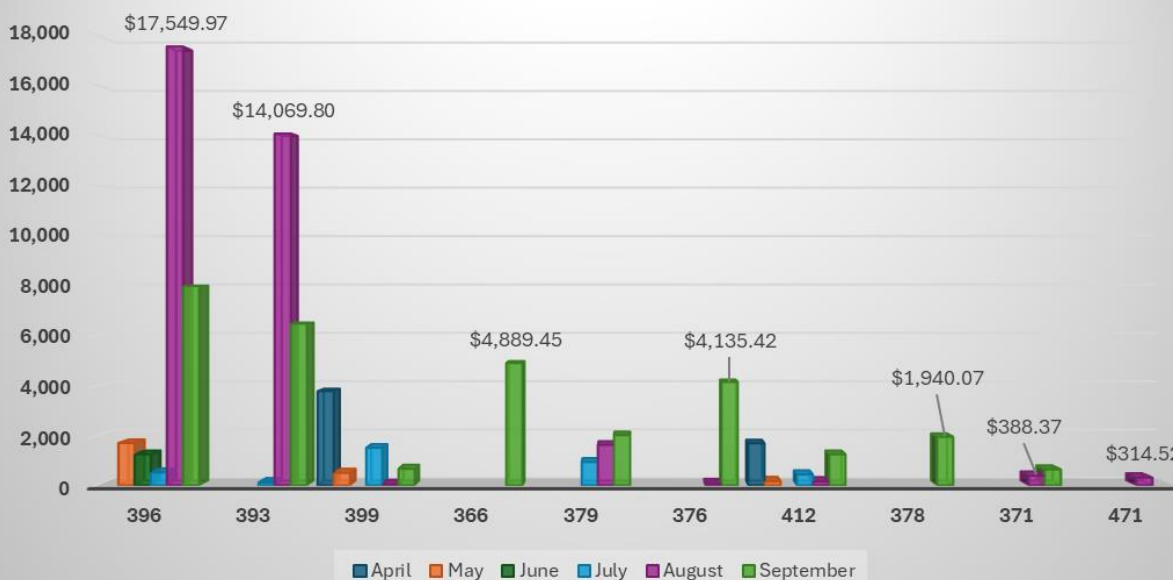
Store ID	April	May	June	July	August	September	Grand Total
396		$1,694.31	$1,242.91	$513.73	$17,549.97	$7,971.23	$28,972.15
393				$117.17	$14,069.80	$6,463.69	$20,650.66
399	$3,762.42	$496.32		$1,504.73	$48.95	$661.22	$6,473.64
366						$4,889.45	$4,889.45
379				$934.81	$1,620.51	$2,010.25	$4,565.57
376					$84.15	$4,135.42	$4,219.57
412	$1,683.05	$166.86		$432.66	$155.54	$1,243.99	$3,682.10
378						$1,940.07	$1,940.07
371					$388.37	$638.58	$1,026.95
471					$314.52		$314.52

We assume that the blank payment method is the customer's outstanding amount. Based on our assumption above, the bottom 10 stores that have the lowest outstanding amount are store ID 396, 393, 399, 366, 379, 376, 412, 378, 371, and 421. The outstanding amount does not happen every month.

The chart above shows that store IDs 471 and 371 have the least customer amount outstanding of $314.52 and $1,026.95, respectively. However, store IDs 396 and 393 had a higher number of outstanding amounts in August 2019.

Bottom 10 Stores- Blank Payment Method by Number of Bills

We further analyzed the bottom 10 stores that have a blank payment method by number of bills, which is considered a better performance with fewer outstanding. We applied a count distinct for the number of bills. The trends were not the same compared to the analysis by net sales figures. This represents some bills having a larger outstanding amount, but fewer outstanding bills. For example, store IDs 393 and 396 have the number of outstanding bills of 49 and 9, respectively. However, their outstanding amounts are reversed pattern (store ID 393: $14k, 396: $18k). The chart above shows that store ID 471 and 371 had the lowest number of outstanding bills of 2 and 3, respectively in August 2019.

In conclusion, store ID 471 has the highest net sales performance at $17.6 million, and the number of customers visiting 11,729 during the period, with a lower amount outstanding of $314.52 for only two bills. We assume that the blank

payment is the customer's outstanding amount. Conversely, store ID 389 has the lowest net sales performance at $3.9 million among all 41 stores with complete sales data (April to September) excluding nine new opening stores. Store ID 389 has a total outstanding amount of $36k for 135 bills, ranking 12th for the least amount outstanding by bills. We have excluded nine newly opened stores: store IDs 380, 386, 396, 393, 379, 371, 366, 378, and 376 for the sales comparison due to incomplete sales data from new openings. Finally, a follow-up is required by the management to identify the reason for lower sales, whether due to location, sales price, services, etc. The number of outstanding issues need to be handled closely by the accounts receivable department to increase the company's liquidity.

On the other hand, store IDs 459 and 471 have the top two highest sales returns with about 6% of sales returns among all stores and have several customers for the sales return of 10,393 and 11,729, respectively. Therefore, a follow-up is required by the management to identify the reason for frequent returns, whether due to customer dissatisfaction regarding pricing issues, or product quality concerns and also the outstanding amount issues.

Lastly we focus on the product itself by taking a closer look at the drugs sold leading us to our **fifth main question,** what is the percentage cost for each drug to calculate profitability? We asked this question since your needs were specific on how to either contract or expand locations into the future, so naturally a deeper dive into the main inventory seemed fitting to help solve your business problem.

Initial our approach was to examine the Ethical, Generic, Chronic, and Acute medications as the top 4 drugs, with sales quantities ranging from 1.95 million to 2.57 million among the 9 medications in the provided data. However, it was discovered that we needed to be careful due to the classification structure in our data dictionary listing classification 1 & 2 which could have run the risk of double counting within our analysis.

The provided data gives an overview of the sales performance of four main drug categories – Ethical, Generic, Chronic and Acute medications, with sales quantities ranging from 1.95 million to 2.57 million among the 9 medications across different stores.
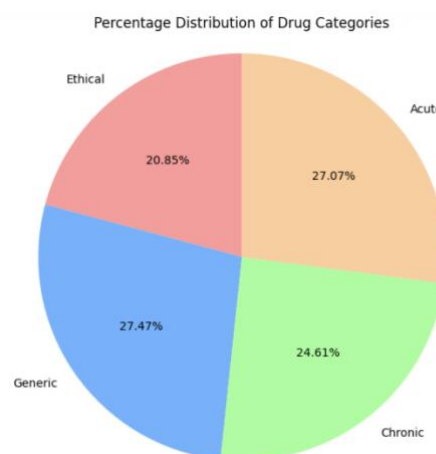
store_id	Sum of total_spend_bill	Sum of Quantity Ethical	Sum of quantity Generic	Sum of Quantity Chronic	Sum of Quantity Acute	Total quantity Units
Grand Total	323,481,688	1,954,534	2,572,985	2,304,355	2,537,716	9,369,590

Generic drugs are the highest in sales quantity measuring 2.57 million units making up a large portion of total drug sales across all stores. Although generics are typically sold at a lower price their higher volume compensates for this.

Acute medications are second highest in sales quantity at 2.54 million units sold. These drugs are typically sold based on urgent medical needs leading to high demand during peak illness seasons.

Chronic medications are in high demand due to their long-term usage selling at 2.30 million units. Although their total quantity sold is less than acute and generic drugs they still represent a large portion of overall sales.

Ethical medication has the lowest quantity sold at 1.95 million units but they are often high in price contributing to a significant portion of overall revenue.

Percentage Distribution of Drug Categories

Ethical 20.85%
Acute 27.07%
Generic 27.47%
Chronic 24.61%

Here we have identified generic and acute drugs are the highest selling quantity while ethical drugs sell the lowest quantity yet contribute the largest total spend. This analysis can help identify the drugs with the greatest demand levels for prioritizing inventory negotiations with vendors.

In in brief the analysis shows that generic and acute drugs have high sales volume while ethical drugs contribute significantly to the total spend due to their higher pricing. To capitalize on these insights we first look at inventory management by prioritizing high volume drugs for timely stock replenishment and maintain safety stock for ethical drugs.

Next we can observe vendor negotiations by recommending leveraging high volume sales to negotiate discounts for generics and secure stable pricing for ethical drugs. Leading to pricing strategies using competitive and tiered pricing for generics and ethical drugs to balance cost and demand. Finally, regular monitoring by continuously tracking sales and market trends to adjust strategies accordingly. The sum of which might favor your business in a balanced approach to managing both high volume and high-cost drugs categories effectively.

In our analysis we identified four drug categories ethical generic chronic and acute, as the most popular. We found those stores with IDs. 456, 471, and 473 have the highest demand for these drugs resulting in higher sales for each category.

These stores are the leaders in generating the highest profits of all the stores listed in the data. Further analysis into qualitative factors for these high preforming stores is recommended to better understand how to apply specific management strategies to the lowest preforming locations.
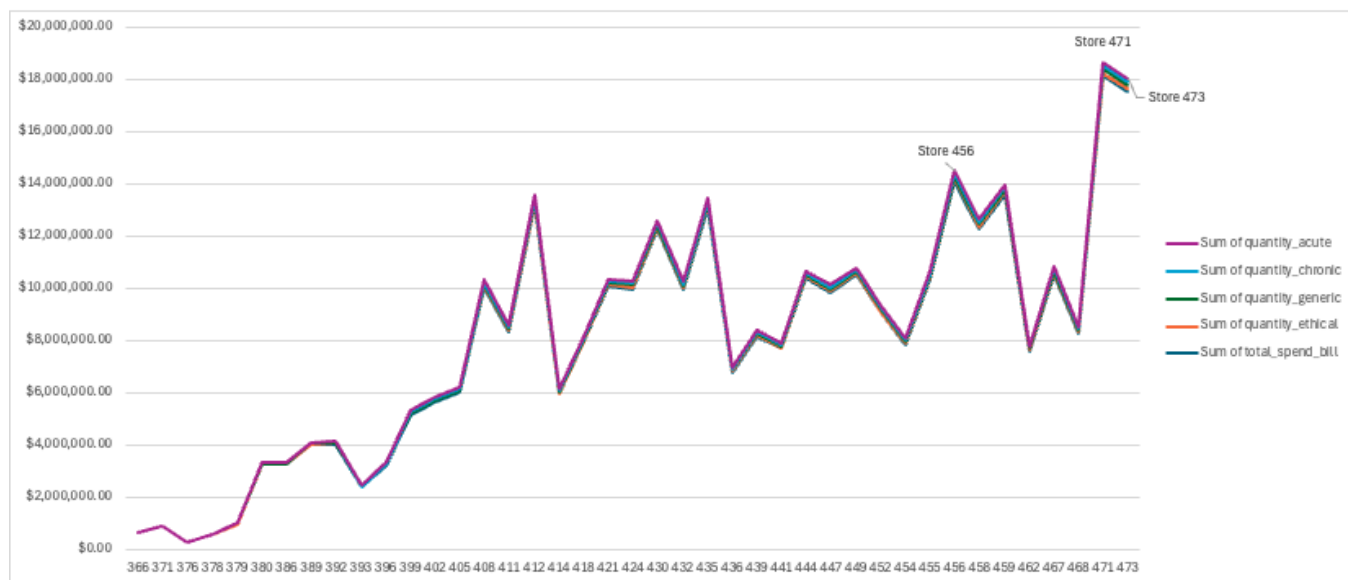
store_id	Sum of total_spend_bill	Sum of Quantity Ethical	Sum of quantity Generic	Sum of Quantity Chronic	Sum of Quantity Acute
456	14,104,719	81,154	116,448	116,339	91,768
471	18,123,967	116,070	146,380	146,628	128,612
473	17,544,654	114,926	128,472	145,347	106,744
Grand Total	49,773,340	312,150	391,300	408,314	327,124

In provided data we are comparing the top regions performances across 3 different seasons: Spring, Summer and Fall, highlighting the top performing regions. Below is a breakdown of the data showcasing key metrics for each region and providing insights into seasonal sales trends.

Highlights
Summer: Region 471 shows the highest sales and total spend **Spring:** Region 471 still leads but region 473 and region 456 perform well. **Early fall:** Region 471 maintains its position as the top region in terms of revenue followed by a region 473 and region 459

Summer:

Region 471 is top for performer with the highest sales in total plan significant quantity sold across Ethical, Generic, Chronic and Acute drug categories It has a total spent bill of $10,118,239 and relatively low number of returns (4508).

Region 473 follows closely behind with total spent bill of $9,805,588. It has slightly lower returns than region 471 with 3806 units returned.

Region 459 also shows strong sales particularly in the chronic and acute drug categories with the total spent bill of $7,664,232.

Spring:

Region 471 again leads with the highest total spend bill of $6,748,908 and strong sales across ethical generic chronic and acute drug categories.

Region 473 is a close second with the total spend of 6,551,477 continuing its strong performance in chronic and acute drugs.

Region 456 performs well both ethical and generic drug categories with a total bill of $5,190,178 although it has significantly fewer returns than other regions.

Early fall:

Region 471 remains a leader contributing $ 1,256,817 in total spent maintaining strong sales in all categories.

Region 473 and region 459 performed well in early fall as well with region 459 having a notable share of market in all categories generating 922709 in total sales.
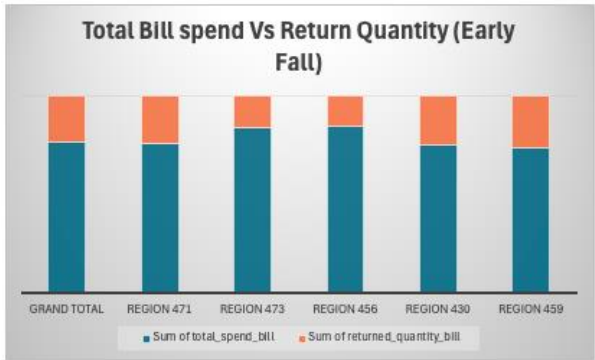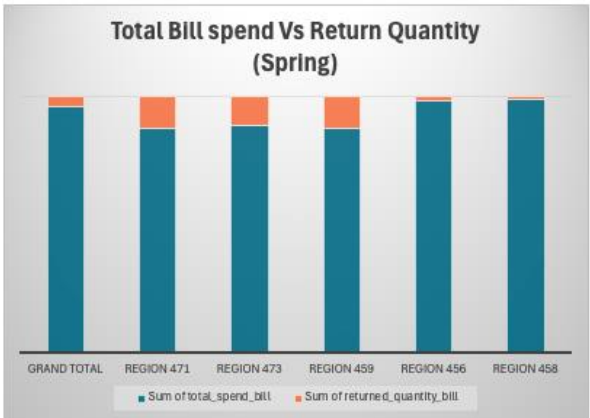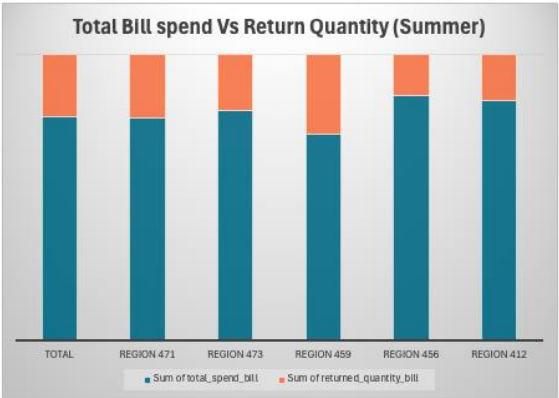
Our conclusion tells us that region 471 is the consistent leader across all the seasons having the highest sales volume in ethical genetic chronic and acute drugs as well as the highest total spent.

Region 473 is another strong performer particularly in the chronic and acute categories ranking just below region 471 in terms of revenue generation. Region 459 stands out in both spring and summer particularly in the chronic and acute categories though it falls behind region 471 and 473 in total revenue. Region 456 shows strong sales especially in the spring with the total spend off over $5 million in that season.

This analysis can help target regions for your company to focus on in regard to high sales potential while improving supply chain inventory and vendor negotiations to further boost performance in high demand regions.

In conclusion, in this analysis, we have built a model which can be replicated as many times as necessary throughout the life cycle of your company's restructuring. We have observed both the highest and lowest preforming stores by revenue as an aggregate in addition to monthly sales figures. We have also touched on the payment process and null values found within the data, new customer activity, the impacts of health insurance, the impacts of returns and their effects on revenue in relation to profit margin within low preforming retail locations.

Further, we have observed highlights from the main inventory being highest and lowest drugs sold throughout different season of the year to help build a historical database for your company records which we are more than happy to continue with further analysis partnerships with your firm. Thank you for your time and please feel free to reach out with any questions or concerns- thank you again from Team two.

Appendix Section

ETL Cleaning Process:

First we extracted the data by using Power Query to extract the data from multiple sources, the five separate data files, i.e., bill_fact, dim_store, dim_customer, dim_doctor, and dim_payment_method. These files contain information related to retail locations, sales, customers, doctors, and payment methods. Each file includes unique identifiers such as store_ref_id, customer_ref_id, doctor_ref_id, and payment_ref_id.

Next in the transform stage, we use Power Query to review, clean, and consolidate the data. Next, we address data inconsistencies by standardizing values and formats. Finally, there is no adjustment made to the null value in the payment method column. After the analysis, it happened to all stores and every month, so it doesn't look like errors. We assume that there are outstanding amounts.

We then convert the DateID column to date formats and add a new column for monthly analysis of sales revenue using the "Month" format. We also convert all the date related columns in dim_customer, dim_doctor, and dim_store to the date format .

In our flat file, we add the abbreviations in front of each field name for easy differentiation derived from which data file, i.e., BF is for bill_fact, DP is for dim_payment, DC is for dim_customer, DD is for dim_doctor, and DS is for dim_store.

We updated the drug categories and separated them into two categories. Drugs included in category one are Cat1_quantity_ethical, Cat1_quantity_generic, Cat1_quantity_surgical, Cat1_quantity_ayurvedic, Cat1_quantity_general, Cat1_quantity_otc. Meanwhile, category two are Cat2_quantity_chronic, Cat2_quantity_acute, and Cat2_quantity_h1.

We merge the five data files, as mentioned above, using unique identifiers (such as store_ref_id, customer_ref_id, doctor_ref_id, and payment_ref_id) to form a unified dataset for easier analysis. This merging process also includes transposing the "flat_file_bill_fact_combine" file with the help of a data dictionary to accurately review and map item descriptions.

The final step was to load the transformed and cleaned data is into Excel for final analysis. We leverage pivot tables and conditional formatting to visually identify and highlight the lowest and highest-performing retail locations and used the pivot tables to construct the chart making it easier to draw actionable insights.

We calculate net sales by subtracting return_value_bill from total_spend_bill. Additionally, we determine the sales discount by subtracting total_spend_bill from mrp_bill. This analysis provides insights into the sales performance and growth potential of each store. We assume that the sales discount is the customer insurance coverage.