Date: May 31, 2021
To: Pharmacy Managers
From: Charlie Evert
RE: **Pharmacy Rankings Based on PreSales Data**

Pharmadata has provided a data set named PreSales of prescription drug sales that can assist in identifying revenue drivers that can be used to maximize revenues. Average income of surrounding customers, floor space (along with the percentage used for prescription departments), parking spaces and whether the store is from a shopping center were key metrics that were tracked in PreSales dataset. Sales were also tracked, and were driven by the aforementioned variables.

The managerial problem with this data set is that there are no defined rankings for each pharmacy. The following analysis will rank each pharmacy in terms of relative performance with given resources in producing sales. There are many variables with which to rank this dataset, but the following will explain why two variables rise above the rest. Variables that demonstrate the best fit are kept in the model, and those that do not fit as well are removed.

The following will explain the data, including estimates and recommendations for managers to maximize information derived from PreSales. Managers can utilize this predictive model by increasing or decreasing variables that they have control over, which for the most part could not be covered in this analysis (since the primary forces behind the rankings regarded floor space and the pharmacy as a percentage of the store. Since managers have finite resources, this analysis should help managers allocate these resources to achieve the best possible outcome. This will help to explain what the most successful stores are doing to maximize pharmacy revenues, and how stores themselves can improve performance.

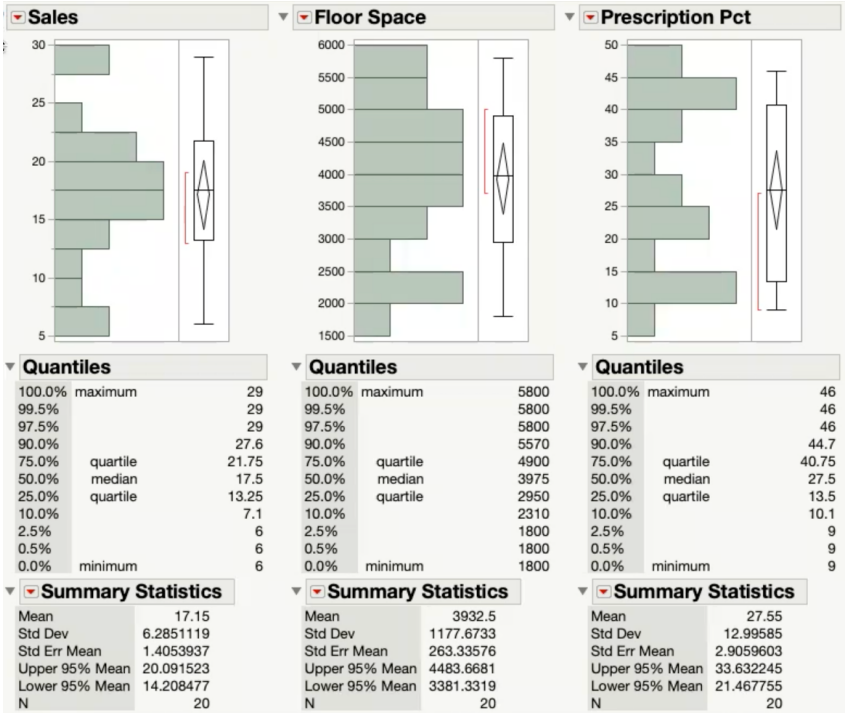
Data Summary

The data indicates that some variables are slightly more relevant than others in pinpointing store performance (by driving increased income). These variables are sales (in hundreds), floor space (in square feet), and the percentage of the store dedicated to the pharmacy (prescription percentage).

I designated stores within shopping centers as 1, and those not in shopping centers as 0. Additionally, I omitted the shopping center attribute, since it only serves as a primary/surrogate key. Finally, I omitted parking spaces, average income of the surrounding area and the aforementioned dummy variable due to their weak correlation with the multivariate model created by JMP. These attributes are not present in the following statistics nor the subsequent multivariate model. However, they are located within Appendix A.

Descriptive Statistics

The following displays histograms with descriptive statistics for each relevant variable:



The average sales are slightly less than median sales. The range of sales is between 600 and 2900 weekly. There is a slight leftward skew, and the data follows a bell curve pattern, meaning that most data occurs towards the median. This normal distribution suggests little outside influence (regulations, quotas, etc.), in that sales are generally between first and third quartile measures of 13.25 and 21.75, respectively.

The average floor space is slightly less than the median floor space, indicating a slight leftward skew. This data generally observes a bell curve pattern, with more data occurring towards the median than peripherally. This suggests that this data occurred naturally without outside interference, in that commercial real estate tends to vary quite a bit in terms of the floor space available but typically hosts subject stores at properties between 3500-5000 square feet.

The average prescription percentage is slightly greater than the mean, indicating a very slight rightward skew. That data has several peaks occurring towards the periphery, and does not appear to occur as a normal distribution. Despite this, these peaks occur just before the minimum and maximum values, indicating that there is a typical pattern for stores in terms of the allocation of their pharmacy in overall floor space, and this may be mandated by upper management, regulations, or some other unknown factor.

Correlation Statistics

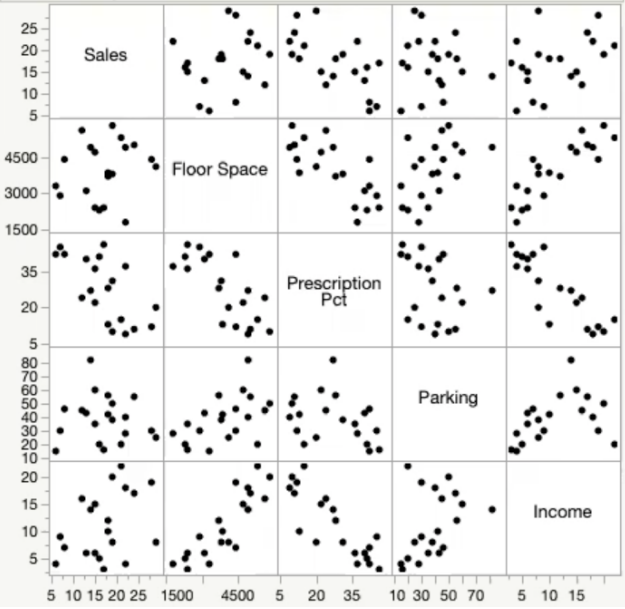
The following displays correlations of each variable:

Correlations

	Sales	Floor Space	Prescription Pct	Parking	Income
Sales	1.0000	0.1831	-0.6628	-0.0693	0.3847
Floor Space	0.1831	1.0000	-0.7506	0.5041	0.8630
Prescription Pct	-0.6628	-0.7506	1.0000	-0.3279	-0.8454
Parking	-0.0693	0.5041	-0.3279	1.0000	0.3928
Income	0.3847	0.8630	-0.8454	0.3928	1.0000

The correlations are estimated by Row-wise method.

Scatterplot Matrix



An interesting correlation to note here are that average incomes highly correlate with floor space. This means that richer communities tend to have larger locations, which makes sense given more discretionary income being available to those with higher incomes. This may also explain why prescription departments are proportionally smaller in richer communities; since pharmacies tend to have inelastic demand, and richer people tend to spend more, it makes sense that stores should allocate as much floor space as possible in these richer stores to goods with elastic demands (meaning everything that isn't the pharmacy).

Additionally, we see that floor space varies very minutely with sales, indicating that floor space has little to do with increasing sales.

Multivariate Model Information

The following is the multivariate model that is statistically significant for the variables of floor space and prescription percentage in predicting sales:

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	48.290855	6.890435	7.01	<.0001*
Floor Space	-0.003842	0.001133	-3.39	0.0035*
Prescription Pct	-0.58189	0.102637	-5.67	<.0001*

This model has a Y intercept of 48.291, a floor space coefficient of -.004 and a prescription percentage coefficient of -.582. Essentially, as floor space and the allocation of the store to the pharmacy increase, predicted sales decrease. The Y Intercept means that sales should equal $4829 when inputs are equal to zero. More floor space decreases this figure, as does an increasing allocation of the store to its pharmacy.

All P values are less than the level of significance (.005). This means that the model is statistically significant, and can be relied upon for its estimates.

Estimates

The figures below rank the differences between actual and predicted sales. This figure was arrived at by subtracting the multivariate predicted sales from actual sales figures.

Larger values indicate outperforming the aforementioned multivariate model, whereas smaller values indicate underperforming this model. Pharmacy 7 (at the top) outperformed all other pharmacies, and pharmacy 5 (at the bottom) underperformed all other pharmacies. Thus, the pharmacies are ranked in descending order, with the best performing pharmacies being higher up.

Pharmacy	DifferenceSales
7	8.1003102841
20	4.6975786095
4	3.597872365
12	3.3484192194
14	2.1551959789
6	1.8015977541
3	1.3213516081
17	1.0545825275
16	0.4038986521
10	0.2473687645
11	0.2185199395
15	-0.104165577
9	-0.808704406
2	-1.186712621
1	-2.226657333
8	-2.430539457
13	-3.121324778
19	-3.963170411
18	-5.171928373
5	-7.933492746

Recommendations

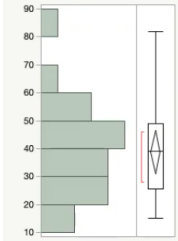
Stores must allocate less to pharmacies and rent less floor space as they are in order to reap the greatest number of sales.

Stores should be able to put all floorspace to good use, and this model indicates that stores may have more floor space than they need. To make up for this, stores should hire better personnel to occupy this floor space or improve the floor space's appearance. Additionally, it is probable that real estate with less floor space costs less. While data is not available in this set for this, data collection regarding the rent costs per sift of floor space would be helpful in determining the degree to which stores should reallocate their resources.

Further data should be collected regarding the allocation of the pharmacy departments in each store; based upon observable data, there is "sweet spot" of allocation.

Appendix A

Parking



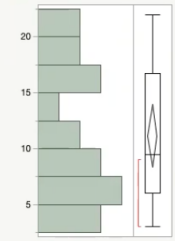
Quantiles

100.0%	maximum	82
99.5%		82
97.5%		82
90.0%		59.6
75.0%	quartile	49
50.0%	median	39
25.0%	quartile	25.75
10.0%		16.4
2.5%		15
0.5%		15
0.0%	minimum	15

Summary Statistics

Mean	38.8
Std Dev	16.841678
Std Err Mean	3.7659136
Upper 95% Mean	46.682148
Lower 95% Mean	30.917852
N	20

Income



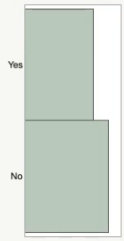
Quantiles

100.0%	maximum	22
99.5%		22
97.5%		22
90.0%		19.9
75.0%	quartile	16.75
50.0%	median	9.5
25.0%	quartile	6
10.0%		4
2.5%		3
0.5%		3
0.0%	minimum	3

Summary Statistics

Mean	11.15
Std Dev	6.0199231
Std Err Mean	1.3460957
Upper 95% Mean	13.967411
Lower 95% Mean	8.3325893
N	20

Shopping Center



Frequencies

Level	Count	Prob
No	11	0.55000
Yes	9	0.45000
Total	20	1.00000

N Missing 0

2 Levels