Background

Our client, an online retailer with an instore presence within the Automotive Industry, was in need of a critical business management application which would assist them with their business operation decisions, market offerings, and metrics. My role in this extensive project was to measure various aspects of their business and provide our Software Development Team with the necessary parameters of the business so that they could implement within the program application.

The first measurement was to affirm their claim that the batteries which they offered, instore and online, lasted as long as the more expensive competitor's brand. Our client wanted to make sure that the claim that they were making was not false, at 5% level of significance. They tested the batteries continuously by mimicking normal use, and elapsed time until the batteries no longer functioned had been recorded. Based on the data that the firm provided, I ran a two-sample test assuming unequal variances (please see my results within the file titled "Automotive Retail.xls", first tab "Battery Life").

A: Client's Brand
B: Competitor's Brand

Null Hypothesis H0: mu_A = mu_B → mu_A – mu_B = 0
Alternate Hypothesis Ha: mu_A ≠ mu_B → mu_A – mu_B ≠ 0

Moreover, Tires were among the most commonly replaced parts on vehicles. Since they were not cheap and shopping for them was confusing, our client wanted to test their warranty offering to optimize on the best warranty that they were providing their customers. When customers looked for tires, many of them paid attention to the warranty that accompanied them as a way of getting reliable and long-lasting tires - sometimes paying a hefty premium for high performance. Our client, again tested their tires in house. Tires were inflated to the optimal pressure and then the tires spun on specially designed treadmills.

The thread-life warranties were presented in a mileage estimate. If a tire had been worn out evenly across the tread well before its estimated mileage limit, it might've qualified for replacement.

Of course, they wanted to minimize warranty replacements by making sure that their tires met the mileage standard. To this end, our client was testing their mileage warranty of 30,000 miles. To do the test, 60 tires had been put through the treadmill test. When the tires were worn down to the 2/32nd thread-wear indicators, the mileage was recorded. I was provided with this data (please see tab titled "Tire Mileage" within the same spreadsheet).

Based on the sample, I calculated the lower bound of the 95% confidence interval of the mean mileage for the tires to be 29,489.94 and the upper bound of the 95% confidence interval of the mean mileage for the tires to be 30,781.59. The probability of finding a sample with a mean of less than 30,000 miles was 0.337. For our client to be within 250 miles of the actual population mean, the required sample size needed to 401.

Based on what I observed here, I had recommended that our client changed their warranty to 25,000 miles instead of the 30,000-mile warranty they were offering. With a 25,000-mile warranty, I was 95% confident that they could mitigate the risk of having too many warranty replacements from customers who had worn their tires prior to reaching 25,000 miles. The benefits of this recommendation were fewer claims and replacement of tires; however, I did notify them that the drawback could mean a negative perception of the tire which in turn would impact the demand and, ultimately, affect price. If they kept the 30,000-mile warranty, they would need to improve the quality of the tire they were offering so that it met and/or exceeded the warranty window. This would add to offering a more expensive tire which would be more expensive to source.

Another concern for our client was estimating its shipping costs. Part of their business was providing parts and accessories on their website. They received orders from all around the world for both small and large parts. However, they did not have a method to estimate the shipping costs for the ordered parts and instead had to tell their customers "ETA TBD" (Estimated Time Arrival, To Be Determined), which had become less and less acceptable given the wide variation in shipping costs due to the size of the part and the shipping destination.

I was provided with a random sample of some of their recent orders (which had included only orders for the 48 contiguous states). Their data included the total dollar amount of the merchandise to be shipped (Merchandise Total), the weight of the order (Shipping Weight), and the actual cost of shipping incurred (Shipping Charge).

I attempted to use a simple linear regression for this case. With their data I arrived to the following regression models identified below:

Model A

SUMMARY OUTPUT

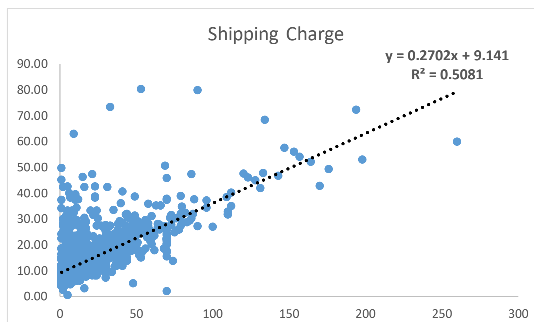
Regression Statistics	
Multiple R	0.712795017
R Square	0.508076736
Adjusted R Square	0.507966932
Standard Error	4.252761672
Observations	4482

Average	2.59
Std Dev	11.6739
Max	8.41
Min	-180.61
two Std Dev	23.3479

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	83685.85399	83685.85399	4627.111469	0
Residual	4480	81025.19864	18.08598184		
Total	4481	164711.0526			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	9.140977894	0.072835871	125.5010447	0	8.998183632	9.283772157
Shipping Weight	0.270172761	0.003971793	68.0228746	0	0.262386087	0.277959436



Model A: A regression model with only the Shipping Weight as the explanatory variable and the Shipping Charge as the dependent variable.

Model B

SUMMARY OUTPUT

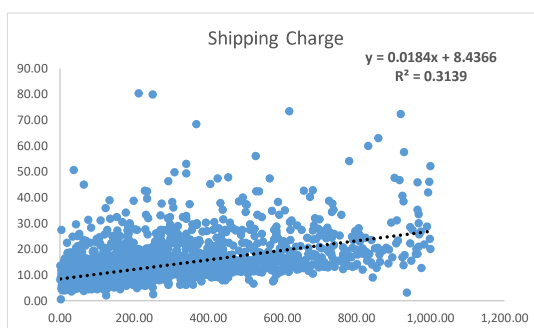
Regression Statistics	
Multiple R	0.560263088
R Square	0.313894728
Adjusted R Square	0.313741579
Standard Error	5.022474525
Observations	4482

Average	-158.06
Std Dev	180.7806
Max	7.16
Min	-973.12
two Std Dev	361.5611

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	51701.93103	51701.93103	2049.610224	0
Residual	4480	113009.1216	25.22525036		
Total	4481	164711.0526			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	8.436555925	0.101995424	82.7150438	0	8.236594545	8.636517306
Merchandise Total	0.018442923	0.000407375	45.27262113	0	0.017644267	0.019241578



Model B: A regression model with only the Merchandise Total as the explanatory variable and the Shipping Charge as the dependent variable.

Using the data as it stood (without cleaning/omitting the outliers), I discovered that both models were poor predictive models for providing a method to estimate the shipping costs for the ordered parts because neither model had an average error that sums up to zero, where we would've collectively overestimated and underestimated in the same way. For Model A I had an average error of 2.59 (overestimating) and for Model B I had an average error of -158.06 (underestimated). In addition, the R^2 for Model A was 0.51 of the variation and for Model B it was even lower at 0.31 of the variation which meant that since Model A was closer to 1, it was a stronger regression model where 51% of the variation in the response variable "Y" (shipping charges) could be explained by the independent variable "X" (Shipping Weight). The R^2 for Model B was even weaker at 0.31 which was closer to 0. The correlation of Model A was +0.7128 vs. Model B at +0.5602.

The Null Hypothesis for Model A was Shipping Weight and Shipping Charges were not correlated and the Null Hypothesis for Model B was Merchandise Total and Shipping Charges were not correlated. The Alternate Hypothesis was the opposite: for Model A it was Shipping Weight and Shipping Charges were correlated and for Model B it was Merchandise Total and Shipping Charges were correlated. Since the p-value in both models were less than the alpha of significance (0.05) then I rejected the Null Hypothesis and accepted the Alternate Hypothesis for each model.

When I analyzed and compared the standard deviation of the prediction errors for both models, Model A was a better model than Model B because, the measure of error (standard deviation) was smaller which had less amount of variability between prediction and actual Shipping Charge. This made Model A the more reliable forecasting model. In addition, the overprediction in Model A was much smaller when compared to Model B.

In conclusion, if the client wanted to choose between one model over the other, my recommendation was to choose to use Model A over Model B based on the comparisons outlined above.