

**Stock Prices of Las Vegas Sands
From 11/1/2021 to 10/31/2022
Compared to Other Stocks in the
Consumer Discretionary
Classification of S&P500**

By: Joshua Winchester

Executive Summary

Our purpose is to examine Las Vegas Sands (LVS) daily stock closing prices with respect to the other daily stock closing prices of the S&P500 consumer discretionary category, of which LVS is a component. The data is for the period 11/01/2021 to 10/31/2022 and includes 56 companies total. We want to get a multiple regression model that bests predicts the LVS stock price using the other stocks available. Throughout our analysis we discovered that half of those companies can be included in the final multiple regression model. While it was hard to disentangle signal from noise, a pattern appeared where stocks who offered “luxury” consumer discretionary spending (like Pulte Group or Starbucks) often rose in tandem with LVS, while stocks who offered “standard” consumer discretionary spending (like Ford or BorgWarner) fell with respect to a rising LVS stock price.

Data Overview & Background

The data source we are working with is a S&P500 of daily closing stock prices of the Consumer Discretionary classification from the period 11/01/2021 to 10/31/2022. There are 252 days recorded in this period, that is, 5 days for each week (Monday-Friday, when the stock exchanges are open). There are 56 companies listed, and given they are in a consumer discretionary spending category, they reflect people spending money on things that they might not directly need to survive, but are nice to have. Travel, entertainment, vehicles, fast food, and clothing are common areas that these businesses focus on.

In particular, we are going to focus on Las Vegas Sands (LVS) with respect to all these other stocks. LVS is classified as Casino's and Gaming, and within that category it is safe to place it in the luxury category. Its founding began in 1989 when Sheldon Adelson and a group of investors bought the Sands Hotel on the Las Vegas Strip. Adelson had been involved in the business conventions in Las Vegas, and soon added on a convention space to the Sands Hotel. Eventually the historic Sands Hotel was demolished in the early 00's to make way its replacement, the Venetian (and eventually Palazzo), which might be best thought of as the "crown jewels" of LVS. Reflecting the roots of its founding, LVS was most profitable at its convention business. LVS began expanding into Asia (especially Macau) in the mid 00's.

Throughout this whole period, Sheldon Adelson maintained his shareholder control of the company. In 2021, Adelson passed away. While the Adelson family today still maintains 56% of shareholder control, the "crown jewels" were sold immediately after Adelson passed: today, LVS owns 7 mega casino resorts, with 6 in Macau and 1 in Singapore and has no direct exposure in the USA (including Las Vegas). Macau might be best thought of as the "Chinese Las Vegas", and LVS is particularly well suited to capitalize on a solidified upper middle class that demands the same level of entertainment as many Westerners have access to as China cements its position as a global economic powerhouse.

Exploratory Data Analysis

R1 shows the reported end of day price for Las Vegas Sands (from here on out referred to as LVS). R2 shows the mean of all the other stocks in this data set for that given day. We can see that the general pattern is reflected in both of these, but its important to notice that the scales are different. From the lowest to the highest point, LVS experienced a difference of nearly 67%, whereas all the others only reached a difference of 43%. Although the peak was higher, the troughs were fairly similar. Overall, LVS seems to experience much more dramatic differences in price relative to all the others combined. In R3, we can see that LVS has a much smaller inter-quartile range, with numerous outliers at the peak compared to R4, which is a relatively balanced box-and-whisker plot.

Correlation Analysis

Here, we started by using Proc Corr to do a correlation analysis with LVS as the dependent variable and all the other stocks as independent variables. S6 is the Proc Corr output; only 2 stocks were not correlated with a significance level of $\alpha=.05$, and they are GPC (Genuine Parts Company) and ORLY (O'Reilly Auto Parts), two auto part companies. Of the other similar gaming and entertainment stocks, Wynn has the highest R value, which makes sense as they are very similar in their business model. MGM's R value is at .59, and Caesar's is .46: this is unique because there are quite a few other stocks that have R values near .50, but few above .55. This suggests that Caesar's is a slightly different business model than LVS, at least from how its being evaluated in pricing

correlation here. One important thing to note here is that only 6 of the R values are negative; that is to say, there is a clear trend of positive correlation between almost all of these stocks (which makes sense: as the economy in general improves, so does LVS). Interestingly, the two stocks who are most negatively correlated with LVS are DLTR (Dollar Tree) and DG (Dollar General). One explanation for this is that when people have less money, they tend to travel (to LVS less) and also spend more money at dollar stores.

To perform variable selection, we did stepwise variable selection, then compared the results of stepwise to the four other methods (CP, r squared, r squared adjusted, and press). S3 shows the stepwise variable selection output. The stepwise selected for about 28 out of the total variables, meaning it cut the variables in half. The other variable selection methods showed that somewhere in between 15-40 variables is probably best, but they are all quite similar and its hard to make a call as to which spot in the 15-40 is best. S1 (press method) picked 18. S2 (cp method) picked 28 (that was where the cp had the highest value, and the least amount of variables). S4 (r squared method) peaked at .9743, but after 23 it was .9700 or higher until the final 55 variables, so it really did not improve much from there and 23 was a good cutoff point. S5 (adjusted r squared method) oscillated between 33-40 variables the whole process without improving anywhere outside of .965-.975, and the picture shows a sample of that oscillation.

What this tells us is that although there is not agreement on precisely which number is ideal, there is an approximate guess between all the methods except adjusted r squared to limit it between 18-30 variables. This validates the stepwise selection method, because 28 falls within that range. I think its safe to go with the stepwise selection choice. While the specific variables chosen vary a little between each method, its pretty clear that after 20 variables or so it doesn't make a great difference which ones you pick in improving variability prediction statistics. With this in mind, we will go with the 28 variables that stepwise has selected.

Regression Analysis

Upon doing an initial multiple regression with these 28 variables, none of the individual residual plots suggest problems with interaction or needing a second order term. When looking for issues with multicollinearity, pictures S8-S12 reveal no problems as there are no patterns that emerge visually. Lets test stocks that are in a similar category by creating an interaction for MGM and Wynn as well as a 2nd order term for Wynn. S7 shows the output of these two terms added, and although the R² Adjusted and F-test values are comparable (and great) to the initial multiple regression, the t-tests for both are quite bad and suggest they should not be added.

S13+S14 show the final output of the multiple regression. For each dollar increase in the value of the price of LVS daily closing price, the following stocks shifted per dollar by:

- Amazon went up .039
- Aptiv went down .04
- AutoZone went down .002
- Bath and Body Works went down .07
- Burger King went down .004
- Booking Holdings went down .004
- BorgWarner went down .17
- CarMax went down .004
- Chipotle went down .004
- D.R. Horton went down .2
- Darden Restaurants went down .04

- Dollar General went down .03
- Domino's went down .012
- Expedia Group went up .02
- Ford Motor Company went down .22
- Genuine Parts Company went up .08
- Hasbro went up .06
- LKQ Corporation went up .25
- McDonald's went up .03
- Newell Brands went up .4
- Nike went up .04
- Norwegian Cruise Line went down .19
- Pool Corporation went down .04
- Pulte Group went up .42
- Starbucks went up .19
- Tapestry Inc. went up .12
- Whirlpool went down .046
- Wynn went up .5
- Yum Brands went up .04

Of the positive coefficients, Wynn, Tapestry Inc, Starbucks, Pulte Group, Newell Brands, LKQ Corporation experienced the largest dollar increases, with Wynn at a peak of .5 (that is, for each dollar increase in the stock price of LVS, there is a 50 cent increase in the stock price of Wynn, on average in this model). Of the stocks here you might not know, Tapestry is a luxury fashion holding company (you probably know their Coach brand); Pulte Group is a home building construction firm (you might have heard of Del Webb which they own); Newell Brands is a consumer goods company with many diverse brands like Rubbermaid, Expo, Graco, etc. LKQ is a provider of salvage, recycled OEM and OE aftermarket auto parts to repair and accessorize consumer automobiles.

Of the negative coefficients, Norwegian Cruise Lines, Ford, D.R. Horton, and BorgWarner had the largest dollar decreases, with the largest being to Ford: for each dollar increase in LVS stock, Ford goes down by 22 cents. D.R. Horton is a home construction company (with 3 times as many homes sold compared to Pulte Group, and over twice the revenue). Borg Warner makes vehicle components like engines and clutches.

Conclusion

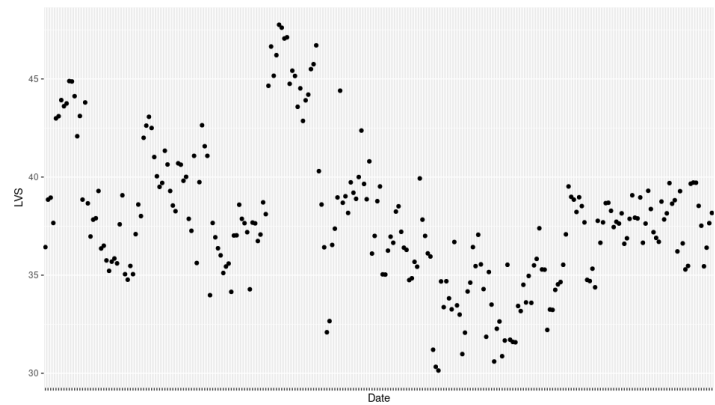
Based on our final model, the most important conclusion to draw is that a rise in LVS stock does not correspond to any particular industry of our overall dataset, but rather seems to match in terms of the type of consumer who would buy that product more. Of all the stocks who have high positive coefficients, they all seem to reflect consumers who are willing to pay a premium to get something that is nicer, whether that be a cup of coffee, a home, or auto accessories. We can also see the inverse of this reflected in the largest negative coefficients: Ford selling standard vehicles, D.R. Horton selling standard/cheaper homes, Borg Warner selling parts to standard vehicles, etc. However this conclusion is quite flimsy, as McDonald's also has a positive coefficient, albeit quite small compared to say Pulte Group, and its not clear that McDonald's would be considered a "luxury" fast food chain. In addition, there are numerous exceptions to this general notion: Norwegian Cruise Lines probably has a negative coefficient because they are direct competitors, for instance. I think that despite these limitations and exceptions, our final model does reveal that as LVS stock goes up in price, so do the other "luxury" consumer discretionary stocks.

One other thing to mention is the lack of clarity in this whole process. There were repeated issues where clear patterns never emerged, and even when patterns did emerge they were weak signals. This makes some sense: if there were clear patterns and the relationship between all these stocks was clear, investors would make easy decisions with neatly drawn lines between all these businesses. Of course, everyone knows investing sometimes feels like gambling, and the lines drawn between businesses are vague, or non-existent as businesses become larger. So maybe it's no surprise that it was hard to get strong signals one way or the other; if there were strong signals, investors would pick up on it and plan accordingly (and naturally businesses would perish based on those strong signals, leaving only the strong signal businesses to compete: thus returning to our impasse).

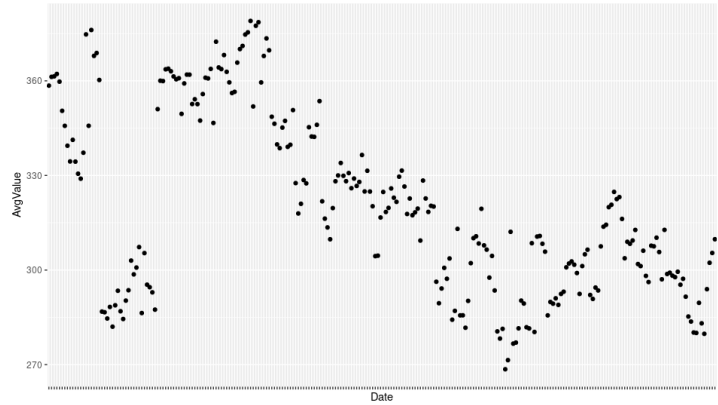
With this in mind, if I were to do this whole process again, I would use a stricter stepwise regression of maybe .05, instead of the default .15. Our variable selection process showed that the difference between 18-40 variables was minimal, and I suspect we could have removed at least 5 if not 10 of these variables and still had a very good model. Perhaps after removing these 5-10, we would have had a greater confirmation in the "luxury" vs. "standard" consumer discretionary spending stock categories.

Appendix

R1

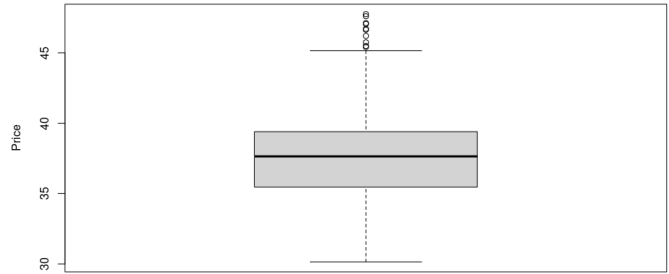


R2



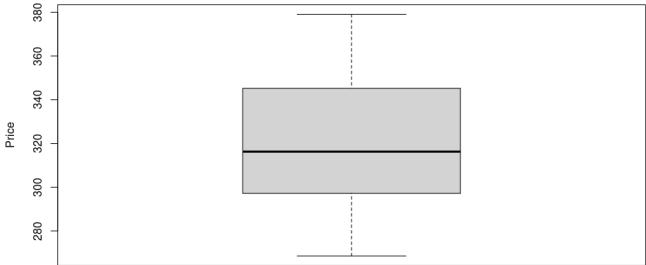
R3

LVS Daily Price Boxplot



R4

Average of All Other Stocks Boxplot



S1

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	5.918216	2.010722	2.94
APTV	1	-0.041994	0.010851	-3.87
BKNG	1	-0.003677	0.000685	-5.36
KMX	1	-0.041611	0.010202	-4.08
DHI	1	-0.202497	0.031199	-6.49
DG	1	-0.034164	0.005964	-5.73
EXPE	1	0.024532	0.007645	3.21
F	1	-0.188776	0.047155	-4.00
LKQ	1	0.171100	0.033475	5.11
MCD	1	0.029469	0.010097	2.92
NWL	1	0.386299	0.066915	5.77
POOL	1	-0.040989	0.004245	-9.65
PHM	1	0.538602	0.049866	10.80
RCL	1	-0.079453	0.011130	-7.14
SBUX	1	0.137752	0.020015	6.88
TPR	1	0.162210	0.037958	4.27
TSLA	1	0.007779	0.002717	2.86
WHR	1	-0.052706	0.011160	-4.72
WYNN	1	0.482600	0.015986	30.19

S4

22	0.9695	APTV BBWI BKNG DRI DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG DG F HD LEN LKQ NWL NCLH NVR POOL PHM RL RCL SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG DG F HD LEN LKQ MAR NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG DG F HD LEN LKQ NWL NCLH NVR POOL PHM ROST RCL SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BBY BKNG KMX DG F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	AMZN APTV BBWI BKNG KMX DG F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN YUM
22	0.9695	APTV BBWI BKNG DG F GRMN HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG KMX DHI DG F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG DG EXPE F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO VFC WHR WYNN
22	0.9695	APTV BBWI BKNG CZR KMX DG F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG KMX DG F HD LEN LKQ NWL NKE NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG KMX DG F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO VFC WHR WYNN
22	0.9695	APTV BBWI BKNG KMX DG DLTR F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG KMX CMG DG F HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG KMX DG F GPC HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
22	0.9695	APTV BBWI BKNG KMX DG F GM HD LEN LKQ NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9702	APTV BBWI BKNG BWA DG DPZ F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9702	APTV BBWI BKNG BWA DG F GPC HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9702	APTV BBWI BKNG BWA DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO ULTA WHR WYNN
23	0.9702	APTV BBWI BBY BKNG KMX DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9701	APTV BBWI BBY BKNG DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO VFC WHR WYNN
23	0.9701	APTV BBWI BKNG BWA DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO VFC WHR WYNN
23	0.9701	APTV AZO BBWI BKNG KMX DG F HD LEN LKQ MHK NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9701	BBWI BKNG BWA KMX DG DPZ F HD LEN LKQ MHK NWL NCLH NVR POOL PHM SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9701	APTV BBWI BKNG KMX DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO ULTA WHR WYNN
23	0.9701	APTV BBWI BKNG BWA CCL DG F HD LEN LKQ NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9701	BBWI BKNG BWA DG DPZ F HD LEN LKQ MHK NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO WHR WYNN
23	0.9701	APTV BBWI BBY BKNG DG F HD LEN LKQ MHK NWL NCLH NVR POOL PHM RCL SBUX TPR TGT TSLA TSCO ULTA WHR WYNN

S2

28	24.0721	0.9714	APTV AZO BBWI BWA CCL DG DLTR F HD LEN LKQ MAR NWL NCLH NVR ONLY POOL PHM RL RCL SBUX TPR TGT TSLA TSCO ULTA WHR WYNN
----	---------	--------	---

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	30	3166.58436	105.55281	249.52	<.0001
Error	221	93.48641	0.42302		
Corrected Total	251	3260.07077			

Root MSE	0.65040	R-Square	0.9713
Dependent Mean	37.83782	Adj R-Sq	0.9674
Coeff Var	1.71891		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-3.59686	3.72060	-0.97	0.3347
wynn2	1	-0.00034339	0.00058661	-0.59	0.5589
mgmwynn	1	-0.00060550	0.00063885	-0.95	0.3443
AMZN	1	0.03872	0.01209	3.20	0.0016
APTV	1	-0.04382	0.01637	-2.68	0.0080
AZO	1	-0.00291	0.00120	-2.42	0.0164
BBWI	1	-0.07415	0.02304	-3.22	0.0015
BKNG	1	-0.00413	0.00080571	-5.13	<.0001
BWA	1	-0.17429	0.08925	-1.95	0.0521
KMX	1	-0.02966	0.01262	-2.35	0.0196
CMG	1	-0.00353	0.00122	-2.90	0.0041
DHI	1	-0.19973	0.03726	-5.36	<.0001

Intercept	-0.69659	2.54646	0.03162	0.07	0.7847
AMZN	0.03903	0.01183	4.59798	10.88	0.0011
APTV	-0.04265	0.01633	2.88384	6.83	0.0096
AZO	-0.00265	0.00119	2.10707	4.99	0.0265
BBWI	-0.06862	0.02243	3.95328	9.36	0.0025
BKNG	-0.00409	0.00078908	11.35324	26.87	<.0001
BWA	-0.17367	0.08560	1.73918	4.12	0.0437
KMX	-0.02669	0.01240	1.95658	4.63	0.0325
CMG	-0.00375	0.00120	4.10744	9.72	0.0021
DHI	-0.19956	0.03724	12.13182	28.72	<.0001
DRI	-0.04043	0.02167	1.47121	3.48	0.0633
DG	-0.02970	0.00790	5.97631	14.15	0.0002
DPZ	-0.01247	0.00467	3.01531	7.14	0.0081
EXPE	0.02168	0.00903	2.43652	5.77	0.0171
F	-0.22367	0.07459	3.79878	8.99	0.0030
GPC	0.08401	0.02300	5.63646	13.34	0.0003
HAS	0.05624	0.02284	2.56186	6.06	0.0146
LKQ	0.25326	0.04732	12.10051	28.64	<.0001
MCD	0.02903	0.01330	2.01211	4.76	0.0301
NWL	0.40737	0.07051	14.10199	33.38	<.0001
NKE	0.03991	0.01888	1.88747	4.47	0.0357
NCLH	-0.18747	0.04731	6.63377	15.70	<.0001
POOL	-0.04010	0.00454	32.90576	77.89	<.0001
PHM	0.42268	0.07011	15.35612	36.35	<.0001
SBUX	0.19912	0.02563	25.49929	60.36	<.0001
TPR	0.12192	0.04612	2.95251	6.99	0.0088
WHR	-0.04613	0.01376	4.74681	11.24	0.0009
WYNN	0.49424	0.01549	430.17481	1018.21	<.0001
YUM	0.04396	0.02922	0.95586	2.26	0.1340

[illegible]

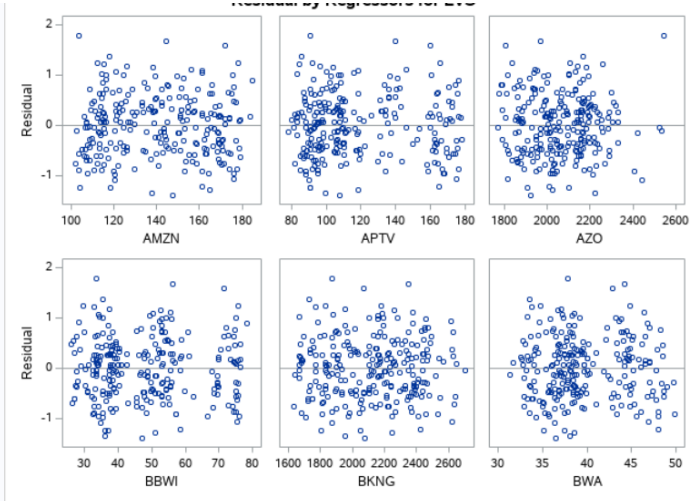
S6

Pearson Correlation Coefficients, N = 252
Prob > |r| under H0: Rho=0

	Date	AAP	AMZN	APTV	AZO	BBWI	BBY	BKNG	BWA	CZR	KMX	CCL	CMG	DHI	DRI	DG	DLTR	DPZ	EBAY	ETSY	EXPE	F	GRMN	GM	GPC	HAS	HLT	HD	LVS	LEN
LVS	-0.32075 0.0002	0.37601 <.0001	0.48373 <.0001	0.42213 <.0001	-0.24816 <.0001	0.35190 <.0001	0.40013 <.0001	0.51252 <.0001	0.49556 <.0001	0.46026 <.0001	0.24667 <.0001	0.49417 <.0001	0.36573 <.0001	0.38169 <.0001	0.57262 <.0001	-0.43225 <.0001	-0.42605 <.0001	0.28476 <.0001	0.38277 <.0001	0.34540 <.0001	0.50374 <.0001	0.50552 <.0001	0.33554 <.0001	0.44605 <.0001	-0.09505 0.1324	0.28367 <.0001	0.44770 <.0001	0.37216 <.0001	1.00000	0.39714 <.0001

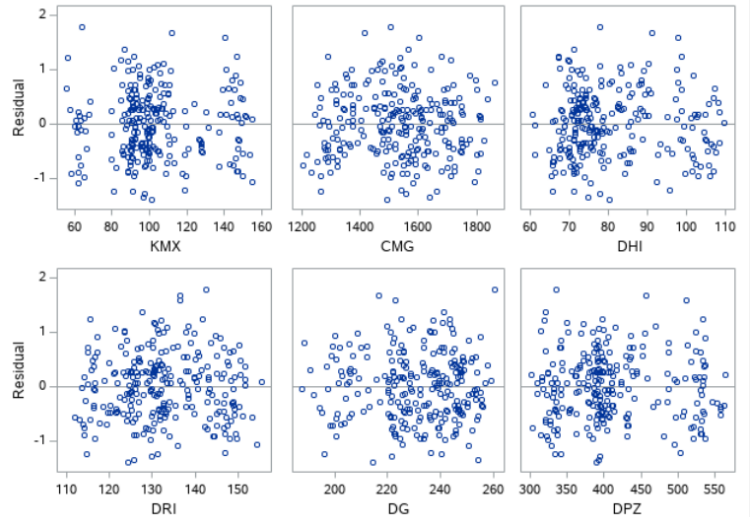
	LKQ	LOW	MAR	MCD	MGM	MHK	NWL	NKE	NCLH	NVR	ORLY	POOL	PHM	RL	ROST	RCL	SBUX	TPR	TGT	TSLA	TJX	TSCO	ULTA	VFC	WHR	WYNN	YUM
	0.23140 0.0002	0.45192 <.0001	0.30595 <.0001	0.19264 0.0021	0.58548 <.0001	0.23821 <.0001	0.34088 <.0001	0.38104 <.0001	0.49685 <.0001	0.43659 <.0001	0.11424 0.0702	0.29145 <.0001	0.42305 <.0001	0.56778 <.0001	0.33401 <.0001	0.49152 <.0001	0.45849 <.0001	0.51908 <.0001	0.32881 <.0001	0.39602 <.0001	0.46650 <.0001	0.27042 <.0001	-0.21189 0.0007	0.33473 <.0001	0.33148 <.0001	0.70402 <.0001	0.35943 <.0001

S8



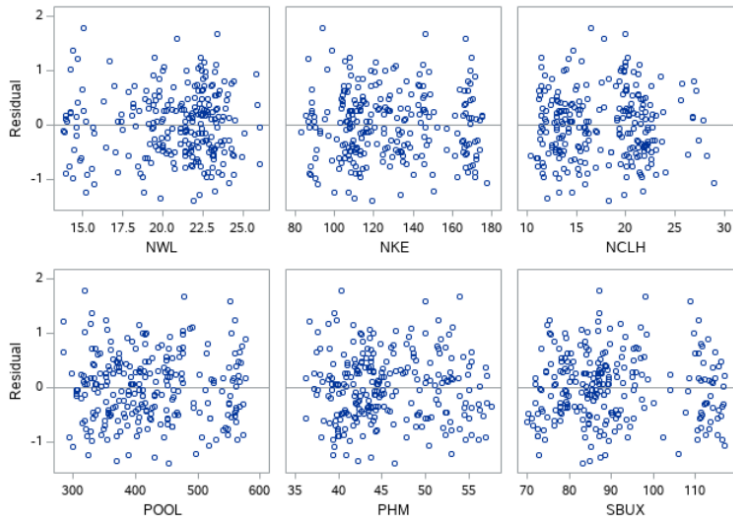
S9

Residual by Regressors for LVS



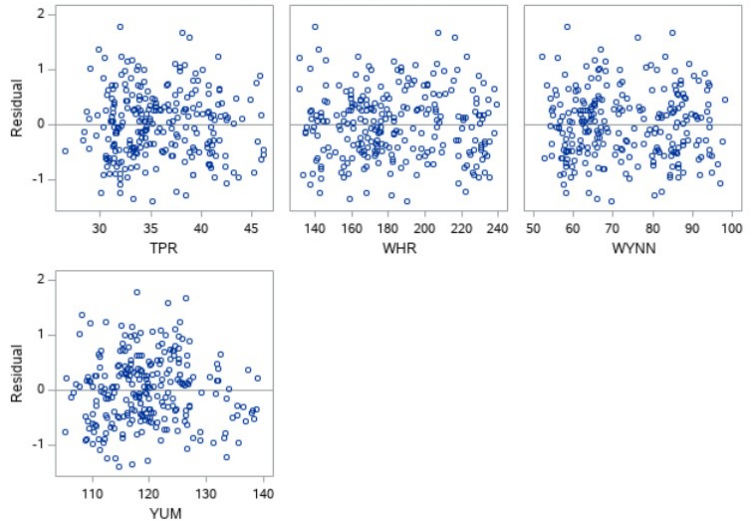
S11

Residual by Regressors for LVS

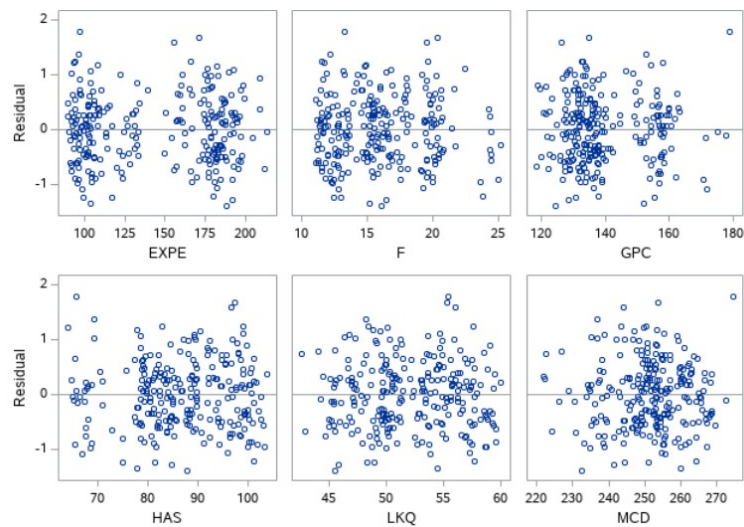


S12

Residual by Regressors for LVS



S10



Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	28	3165.85713	113.06633	267.62	<.0001
Error	223	94.21364	0.42248		
Corrected Total	251	3260.07077			

Root MSE	0.64999	R-Square	0.9711
Dependent Mean	37.83782	Adj R-Sq	0.9675
Coeff Var	1.71782		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-0.69659	2.54646	-0.27	0.7847
AMZN	1	0.03903	0.01183	3.30	0.0011
APTV	1	-0.04265	0.01633	-2.61	0.0096
AZO	1	-0.00265	0.00119	-2.23	0.0265
BBWI	1	-0.06862	0.02243	-3.06	0.0025
BKNG	1	-0.00409	0.00078908	-5.18	<.0001
BWA	1	-0.17367	0.08560	-2.03	0.0437
KMX	1	-0.02669	0.01240	-2.15	0.0325
CMG	1	-0.00375	0.00120	-3.12	0.0021

BKNG	1	-0.00409	0.00078908	-5.18	<.0001
BWA	1	-0.17367	0.08560	-2.03	0.0437
KMX	1	-0.02669	0.01240	-2.15	0.0325
CMG	1	-0.00375	0.00120	-3.12	0.0021
DHI	1	-0.19956	0.03724	-5.36	<.0001
DRI	1	-0.04043	0.02167	-1.87	0.0633
DG	1	-0.02970	0.00790	-3.76	0.0002
DPZ	1	-0.01247	0.00467	-2.67	0.0081
EXPE	1	0.02168	0.00903	2.40	0.0171
F	1	-0.22367	0.07459	-3.00	0.0030
GPC	1	0.08401	0.02300	3.65	0.0003
HAS	1	0.05624	0.02284	2.46	0.0146
LKQ	1	0.25326	0.04732	5.35	<.0001
MCD	1	0.02903	0.01330	2.18	0.0301
NWL	1	0.40737	0.07051	5.78	<.0001
NKE	1	0.03991	0.01888	2.11	0.0357
NCLH	1	-0.18747	0.04731	-3.96	<.0001
POOL	1	-0.04010	0.00454	-8.83	<.0001
PHM	1	0.42268	0.07011	6.03	<.0001
SBUX	1	0.19912	0.02563	7.77	<.0001
TPR	1	0.12192	0.04612	2.64	0.0088
WHR	1	-0.04613	0.01376	-3.35	0.0009
WYNN	1	0.49424	0.01549	31.91	<.0001
YUM	1	0.04396	0.02922	1.50	0.1340

S13/
S14

R code

```
# Packages
library(ggplot2)

#import data
spdat <- read.csv("/home/judge/Documents/Current/Folders/Class/Current/stp429/lab3/data/
prices.csv")

# subsetting data
spstockdat <- spdat[, -which(names(spdat) == "Date")]
lvsstockdat <- spstockdat[, which(names(spstockdat) == "LVS")]
spstockdat <- spstockdat[, -which(names(spstockdat) == "LVS")]
rmspstockdat <- data.frame(Date=spdat$Date,AvgValue=rowMeans(spstockdat))

#plotting
ggplot(spdat, aes(x = Date, y = LVS)) + geom_point() + theme(axis.text.x=element_blank())
ggplot(rmspstockdat, aes(x=Date,y=AvgValue)) + geom_point() +
theme(axis.text.x=element_blank())

boxplot(lvsstockdat, main='LVS Daily Price Boxplot',ylab='Price')
boxplot(rmspstockdat$AvgValue, main='Average of All Other Stocks Boxplot',ylab='Price')
```

SAS code

```
/* Generated Code (IMPORT) */
/* Source File: prices.csv */
/* Source Path: /home/u62117616/sasuser.v94 */
/* Code generated on: 11/25/22, 11:57 AM */
%web_drop_table(pr);
FILENAME REFFILE '/home/u62117616/sasuser.v94/prices.csv';
PROC IMPORT DATAFILE=REFFILE
      DBMS=CSV
      OUT=pr;
      GETNAMES=YES;
RUN;
proc corr data=pr;
      var _numeric_;
      with lvs;
      run;

*stepwise;
proc reg data=pr;
      model lvs=AAP AMZN APTV AZO BBWI BBY BKNG BWA CZR KMX CCL CMG DHI DRI DG
DLTR DPZ EBAY ETSY EXPE F GRMN GM GPC HAS HLT HD LEN LKQ LOW MAR MCD MGM MHK NWL
NKE NCLH NVR ORLY POOL PHM RL ROST RCL SBUX TPR TGT TSLA TJX TSCO ULTA VFC WHR
WYNN YUM/selection=stepwise;
      run;
*rsquare;
proc reg data=pr;
      model lvs=AAP AMZN APTV AZO BBWI BBY BKNG BWA CZR KMX CCL CMG DHI DRI DG
DLTR DPZ EBAY ETSY EXPE F GRMN GM GPC HAS HLT HD LEN LKQ LOW MAR MCD MGM MHK NWL
NKE NCLH NVR ORLY POOL PHM RL ROST RCL SBUX TPR TGT TSLA TJX TSCO ULTA VFC WHR
WYNN YUM/selection=rsquare;
      run;
*adjrsq;
proc reg data=pr;
```

```

        model lvs=AAP AMZN APTV AZO BBWI BBY BKNG BWA CZR KMX CCL CMG DHI DRI DG
DLTR DPZ EBAY ETSY EXPE F GRMN GM GPC HAS HLT HD LEN LKQ LOW MAR MCD MGM MHK NWL
NKE NCLH NVR ORLY POOL PHM RL ROST RCL SBUX TPR TGT TSLA TJX TSCO ULTA VFC WHR
WYNN YUM/selection=adjrsq;
        run;
*cp;
proc reg data=pr;
        model lvs=AAP AMZN APTV AZO BBWI BBY BKNG BWA CZR KMX CCL CMG DHI DRI DG
DLTR DPZ EBAY ETSY EXPE F GRMN GM GPC HAS HLT HD LEN LKQ LOW MAR MCD MGM MHK NWL
NKE NCLH NVR ORLY POOL PHM RL ROST RCL SBUX TPR TGT TSLA TJX TSCO ULTA VFC WHR
WYNN YUM/selection=cp;
        run;
*press;
proc glmselect data=pr;
        model lvs=AAP AMZN APTV AZO BBWI BBY BKNG BWA CZR KMX CCL CMG DHI DRI DG
DLTR DPZ EBAY ETSY EXPE F GRMN GM GPC HAS HLT HD LEN LKQ LOW MAR MCD MGM MHK NWL
NKE NCLH NVR ORLY POOL PHM RL ROST RCL SBUX TPR TGT TSLA TJX TSCO ULTA VFC WHR
WYNN YUM/selection=stepwise(choose=press);
        run;

*multiple regress;
proc reg data=pr;
        model lvs = AMZN APTV AZO BBWI BKNG BWA KMX CMG DHI DRI DG DPZ EXPE F
GPC HAS LKQ MCD NWL NKE NCLH POOL PHM SBUX TPR WHR WYNN YUM;
        run;

*create data step;
data pr2;
        set pr;
        wynn2 = wynn * wynn;
        mgmwynn = mgm * wynn;
*multiple regress to test interaction and quadratic;
proc reg data=pr2;
        model lvs = wynn2 mgmwynn AMZN APTV AZO BBWI BKNG BWA KMX CMG DHI DRI DG
DPZ EXPE F GPC HAS LKQ MCD NWL NKE NCLH POOL PHM SBUX TPR WHR WYNN YUM;
        run;
%web_open_table(pr);

```