Applied Regression Analysis – Final Project

Predicting Auto Sales in the United States

Akaash Bhalla, Thomas Klocanas, Andrew Taub, Carl Oskar Thoen, Alon Weber, Carl Ziade

Predicting Auto Sales: Background

Project Idea

Q: Why do we care about auto sales?

A: Car sales can indicate the overall state and direction of the economy. After home purchases, automobiles are the most expensive items a consumer owns. Auto sales is a strong proxy for consumer confidence.

Expectations

Q: What did we expect to be the predictors of auto sales?

A: Unemployment rate, consumer confidence index, and interest rates. These are industry standards used by analysts.

Experiment

Q: What other variables could be indicators of auto sales?

A: We considered variables in the auto-industry and other economic factors.

Variable Selection

Auto-Market

Q: What might affect consumer decisions?
A: Gas prices, interest rates on auto loans from banks, price inflation on new and used vehicles

Other auto industry factors could include inventory levels and total vehicle miles driven

7

Economic

Q: What economic factors could explain auto sales?
A: Disposable income, household savings and debt, inflation, population

Other indicators could be house price inflation to see how the level of home pricing could affect auto sales, as well as those employed in manufacturing, assuming that people employed in that sector might require a vehicle more compared to other industry positions

11

Time

Given seasonality, we included quarterly dummy variables as well as incremental units to capture time

March 1994 – November 2017 (monthly data)

5

Variable Overview

Data was reported
as CAR and TRUCK → Kept split to see if certain variables
explained CAR vs. TRUCK sales differently

| Prediction Variables	Denomination	Source		Economic Variables	Denomination	Source
New Car Sales	units	Ward's Auto				
New Truck Sales	units	Ward's Auto				
Total New Vehicle Sales	units	Ward's Auto				
Auto Market Variables	**Denomination**	**Source**		**Economic Variables**	**Denomination**	**Source**
Auto Inventories	units (SA)	Bureau of Economic Analysis		US Population	units	Bureau of Economic Analysis
New Vehicle Price Inflation	percentage	Bureau of Labor Statistics		Unemployment Rate	percentage	Bloomberg
Used Vehicle Price Inflation	percentage	Bureau of Labor Statistics		Disposible Income (Real)	USD / capita (SAAR)	Bureau of Economic Analysis
Gas Prices	USD / gallon	US Dept of Energy		Household Savings	USD / capita (SA)	Bureau of Economic Analysis
Consumer Confidence - Automobiles	percentage (SA)	Conference Board		Household Debt	USD / capita	Bloomberg
Miles Driven	units	US Dept of Transportation		House Price Inflation	Percentage	Federal Housing Finance Agency
New Car Interest Rate	percentage	Federal Reserve		Consumer Confidence	Index units	Conference Board
				GDP (Real)	USD / capita	Bureau of Economic Analysis
Other	**Denomination**	**Source**		Fed Funds Rate	percentage	Federal Reserve
Quarters (to capture seasonality)	dummy	Team analysis		Inflation	percentage	Bureau of Labor Statistics
Time	units	Team analysis		Non-Farm Manufacturing Payroll	units	Bureau of Labor Statistics

Summary Statistics

Car	Truck	Vehicle Type
Small	Pickup	Passenger
Midsize	Cross-over	Passenger
Large	Minivan	Passenger
Luxury	Luxury SUV	Passenger
	Small-Large SUV	Passenger
	Small-Large Van	Passenger/Commercial
	Light Truck	Commercial
	Heavy Truck	Commercial

	Car	Truck	Total
Min	315,446	339,241	656,693
Max	876,696	1,094,126	1,808,130
Mean	633,142	664,389	1,297,667
Std Dev	113,541	146,878	212,802
Correl	0.32		

	Percent of Sales	
	1994	**2017**
Car	60%	36%
Truck	40%	64%
Total	100%	100%

Car versus Truck Sales (units): 1994-2017
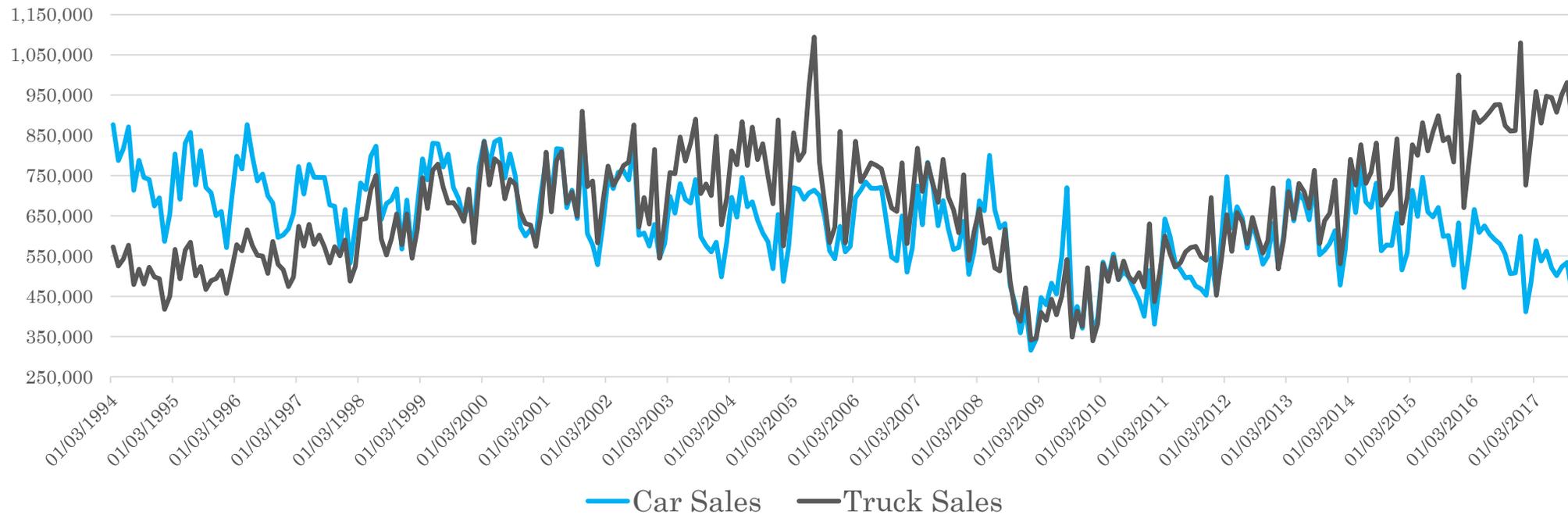
Correlation Matrix

	Car	Truck	Total	Time	Year	Q1	Q2	Q3	Inventories	CPI New Cars	CPI Used Cars	Gas Prices	Cons.Conf. Auto	Miles Driven	Bank Auto Interest	US Popl	U-Rate	Disp. Income	Savings	Household Debt	House Price	Cons.Conf	Real GDP	Fed Funds Rate	CPI Change
Truck	0.33																								
Total	0.76	0.86																							
Time	-0.54	0.34	-0.06																						
Year	-0.54	0.33	-0.06	1.00																					
Q1	-0.15	-0.17	-0.20	-0.01	0.02																				
Q2	0.32	0.12	0.26	-0.01	0.00	-0.33																			
Q3	0.09	0.06	0.09	0.01	0.00	-0.33	-0.34																		
Inventories	0.51	0.03	0.29	-0.59	-0.59	0.01	0.00	-0.01																	
CPI New Cars	-0.01	-0.47	-0.33	-0.14	-0.14	0.01	0.02	0.00	-0.07																
CPI Used Cars	0.09	-0.29	-0.15	-0.17	-0.17	0.01	0.02	-0.01	-0.03	0.58															
Gas Prices	-0.38	0.05	-0.17	0.80	0.80	-0.05	0.07	0.05	-0.67	0.14	0.04														
Cons. Conf. Auto	-0.07	0.40	0.24	0.54	0.54	0.00	-0.01	0.04	0.01	-0.10	-0.10	0.33													
Miles Driven	-0.13	0.56	0.31	0.72	0.72	-0.40	0.19	0.25	-0.52	-0.26	-0.23	0.60	0.28												
Bank Auto Interest Rate	0.44	-0.35	-0.01	-0.93	-0.93	0.01	0.00	-0.01	0.45	0.09	0.18	-0.72	-0.62	-0.63											
US Population	-0.55	0.33	-0.07	1.00	1.00	-0.01	-0.02	0.01	-0.62	-0.15	-0.17	0.81	0.52	0.73	-0.92										
U-Rate	-0.46	-0.48	-0.58	0.35	0.35	0.00	0.00	0.00	-0.56	0.43	0.26	0.52	0.01	0.13	-0.40	0.36									
Disposible Income	-0.53	0.40	0.00	0.97	0.97	-0.01	-0.01	0.00	-0.65	-0.23	-0.23	0.78	0.44	0.77	-0.87	0.98	0.27								
Saving / Capita	-0.25	-0.27	-0.32	0.36	0.36	0.02	0.06	-0.04	-0.15	0.18	-0.01	0.33	0.35	0.08	-0.46	0.35	0.58	0.28							
Household Debt	-0.56	0.18	-0.18	0.81	0.81	0.00	-0.01	0.00	-0.78	-0.14	-0.18	0.79	0.04	0.70	-0.64	0.83	0.35	0.88	0.12						
US House Price	0.42	0.65	0.67	-0.19	-0.19	-0.01	0.00	0.01	0.41	-0.37	-0.19	-0.39	0.15	0.01	0.11	-0.20	-0.70	-0.12	-0.50	-0.26					
Consumer Confidence	0.52	0.41	0.56	-0.44	-0.44	-0.02	0.03	0.02	0.60	-0.30	-0.09	-0.59	0.14	-0.24	0.45	-0.46	-0.82	-0.41	-0.48	-0.56	0.67				
Annualized Real GDP	-0.49	0.46	0.05	0.94	0.94	-0.01	-0.02	0.01	-0.63	-0.29	-0.25	0.76	0.41	0.78	-0.82	0.95	0.16	0.99	0.15	0.89	-0.03	-0.33			
Fed Funds Rate	0.54	-0.12	0.21	-0.81	-0.81	-0.01	0.00	0.01	0.58	0.07	0.19	-0.63	-0.38	-0.54	0.88	-0.81	-0.65	-0.76	-0.56	-0.62	0.32	0.70	-0.68		
US Urban CPI Change	0.27	-0.01	0.14	-0.34	-0.34	0.00	0.01	0.00	0.04	0.26	0.31	0.02	-0.23	-0.18	0.40	-0.33	-0.29	-0.27	-0.48	-0.12	0.17	0.18	-0.20	0.46	
Nonfarm Mfg Payroll	0.62	-0.11	0.25	-0.93	-0.93	-0.02	0.01	0.01	0.72	-0.03	0.08	-0.85	-0.33	-0.65	0.88	-0.94	-0.62	-0.91	-0.43	-0.86	0.38	0.70	-0.85	0.88	0.33

What did we learn from the correlation matrix?

CORREL

- Based on initial review of the correlation matrix, we saw that different variables having stronger relationships with CAR sales versus TRUCK sales.

- We decided to keep our analysis separate moving forward and try to build two models to see which set of indicators were the best for predicting CAR sales compared to TRUCK sales.

- Ultimately, we wanted to see if, and how, the sets of predictors for each model might be different.

NEXT

- As a next step, we generated scatter plots below to further understand the strength of each variable's relationship to CAR and TRUCK sales.

- We boxed the scatterplots in which the R-Square values were greater than 20%.

Scatter Plots: Car Sales and Auto Variables
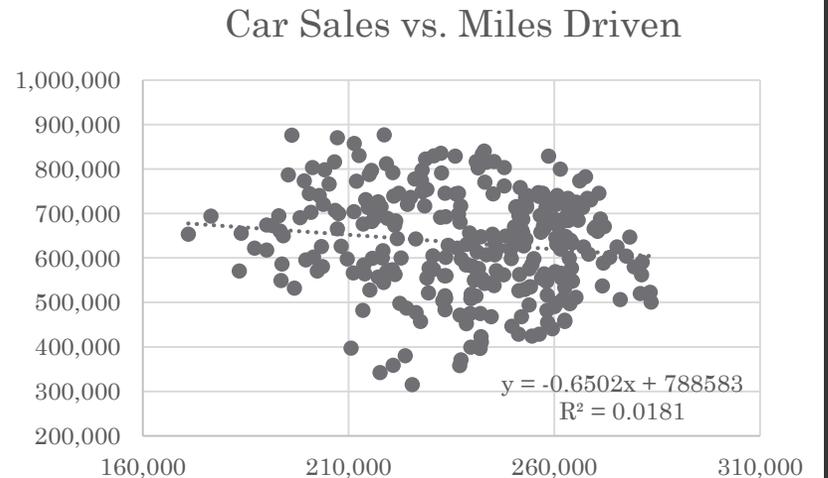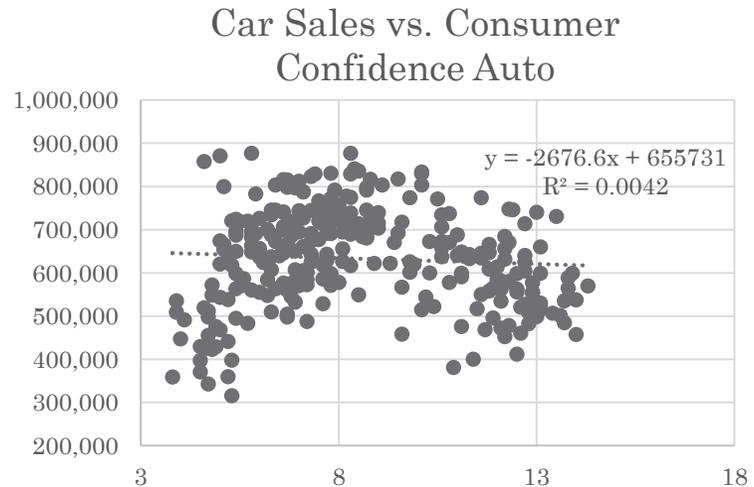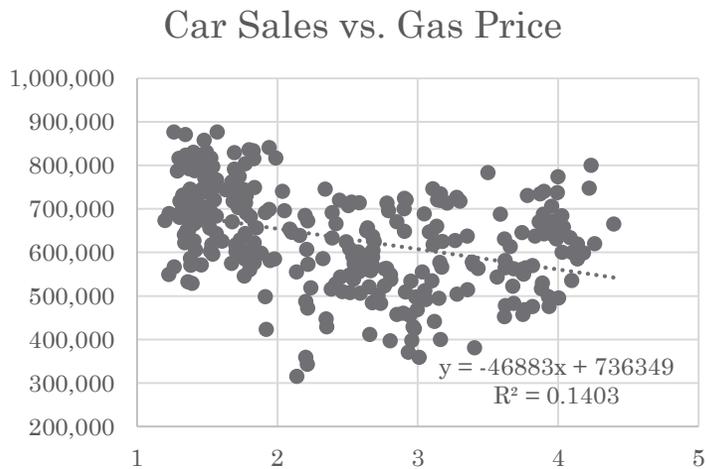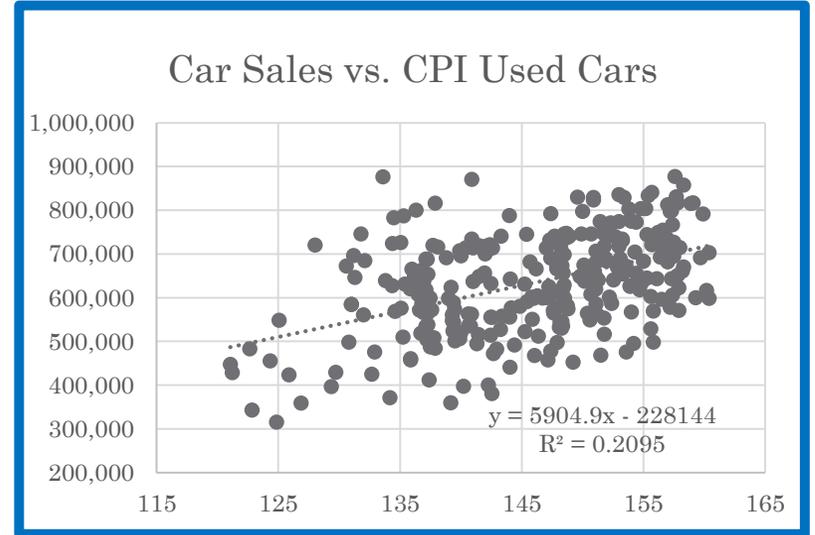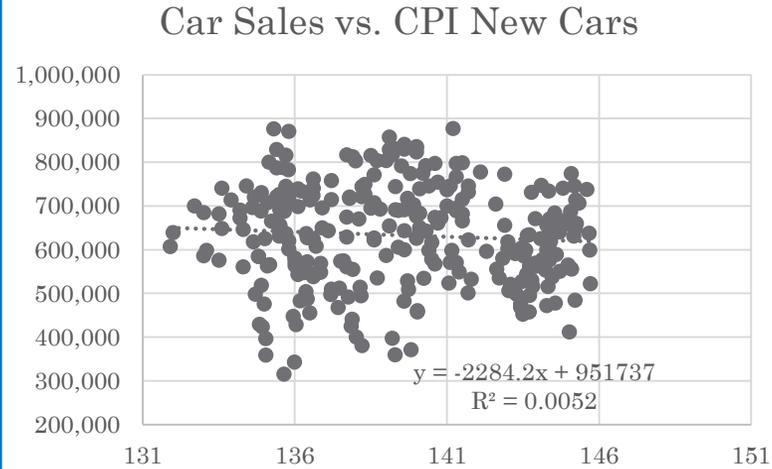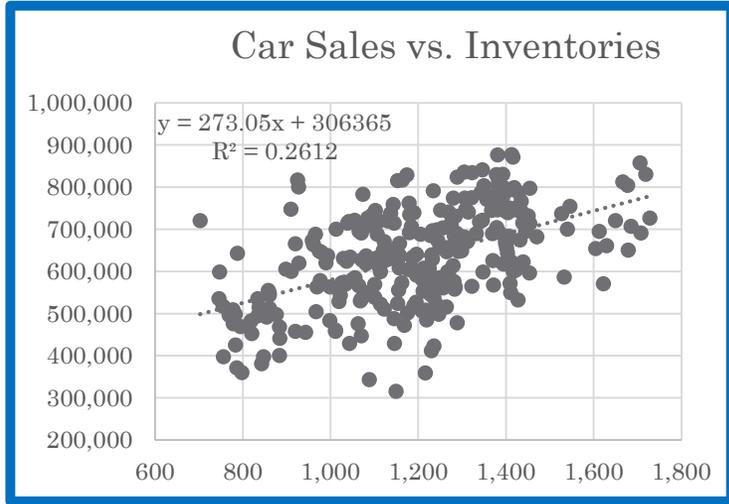
Scatter Plots: Car Sales and Auto Variables

Car Sales vs. Bank Auto Interest Rate

y = 28362x + 435681
R² = 0.1912
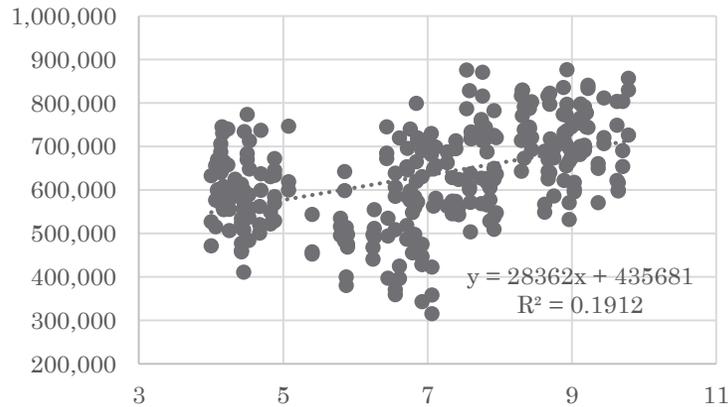
Car Sales vs. Population

y = -3383.1x + 2E+06
R² = 0.3061
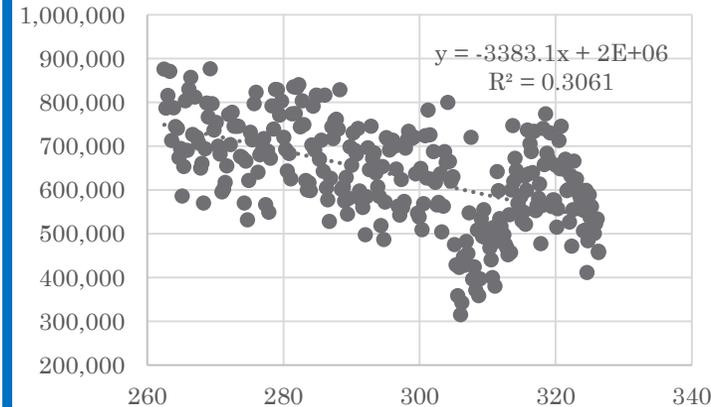
Car Sales vs. Unemployment

y = -32208x + 822674
R² = 0.2085

Car Sales vs. vs. Disposable Income / Capita Real

y = -16.076x + 1E+06
R² = 0.2763
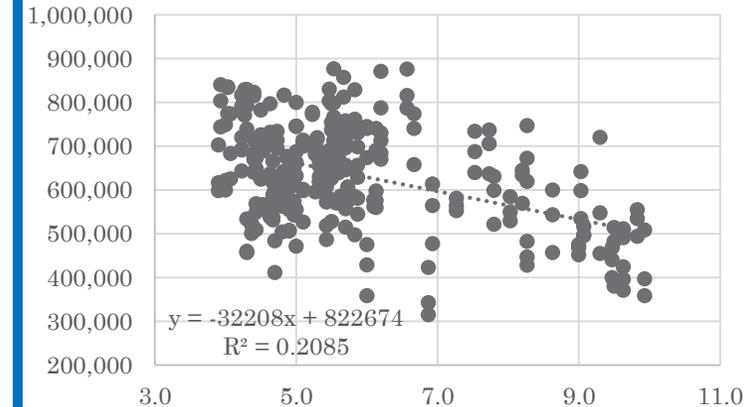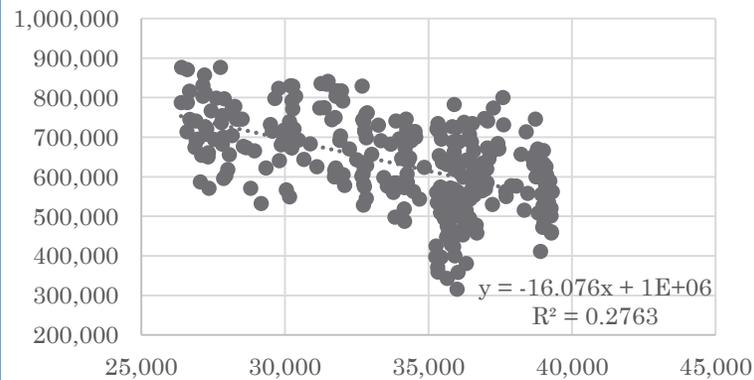
Car Sales vs. Saving / Capita

y = -58.786x + 733782
R² = 0.062
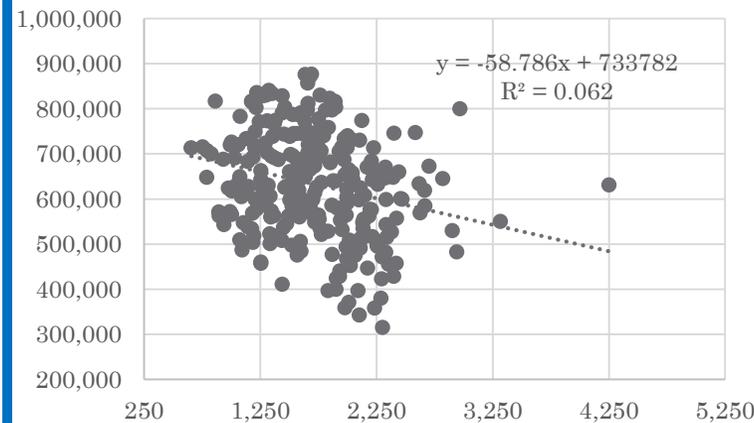
Cars Sales vs. Household Debt / Capita

y = -8.1069x + 928595
R² = 0.3121

Scatter Plots: Car Sales and Auto Variables

Scatter Plots: Truck Sales and Auto Variables

Truck Sales vs. Inventories
$y = 19.939x + 640527$
$R^2 = 0.0008$
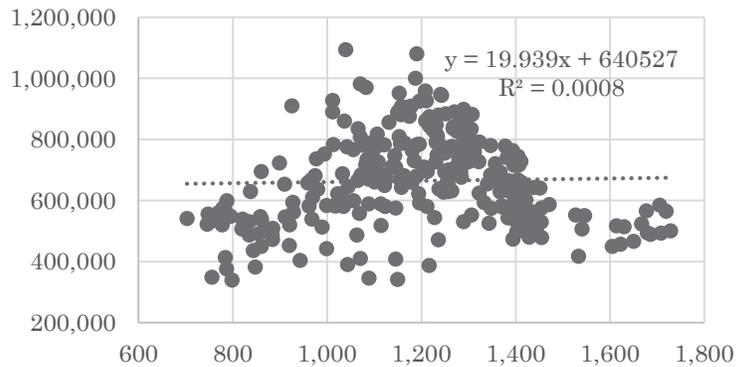
Truck Sales vs. CPI New Cars
$y = -47027x + 680144$
$R^2 = 0.2186$
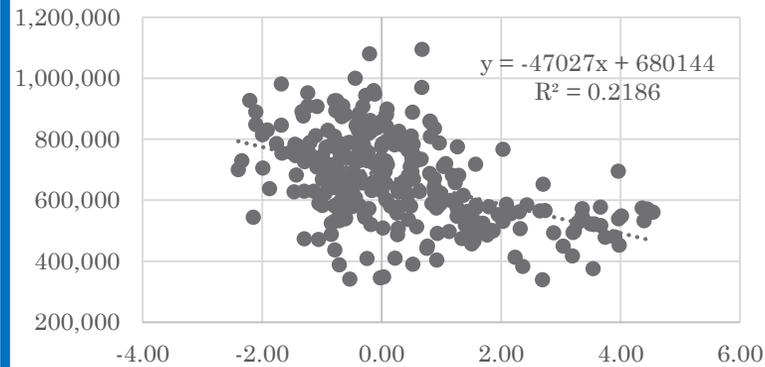
Truck Sales vs. CPI Used Cars
$y = -7725.8x + 666280$
$R^2 = 0.0829$

Truck Sales vs. Gas Price
$y = 7566x + 647733$
$R^2 = 0.0022$
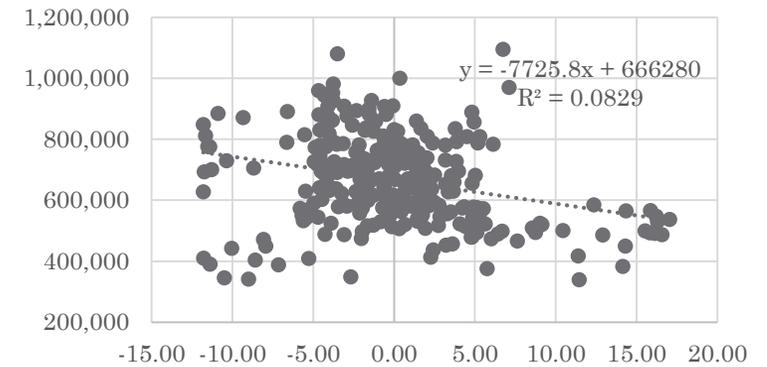
Truck Sales vs. Consumer Confidence Auto
$y = 21510x + 482864$
$R^2 = 0.1609$
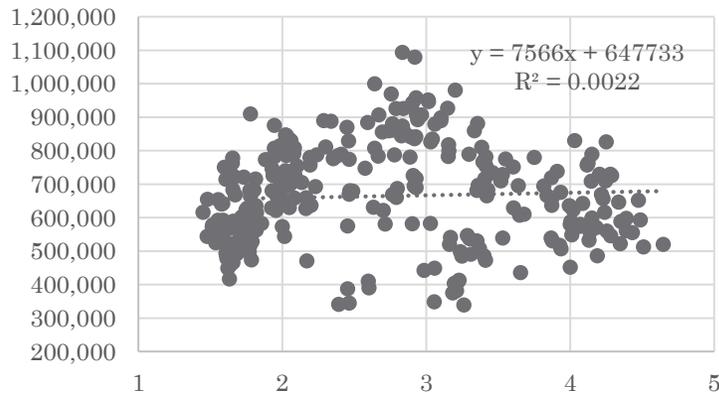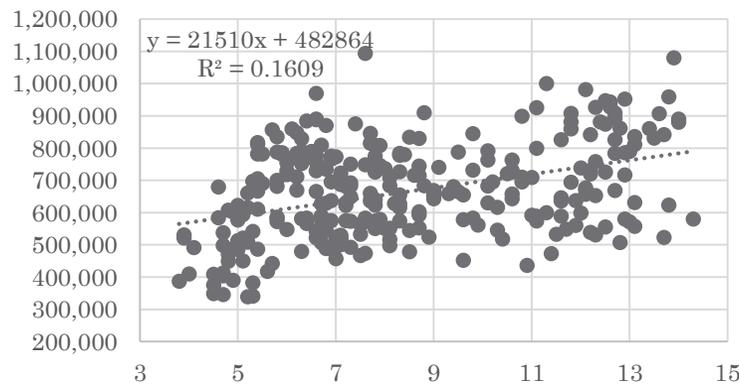
Truck Sales vs. Miles Driven
$y = 3.4966x - 171588$
$R^2 = 0.3125$
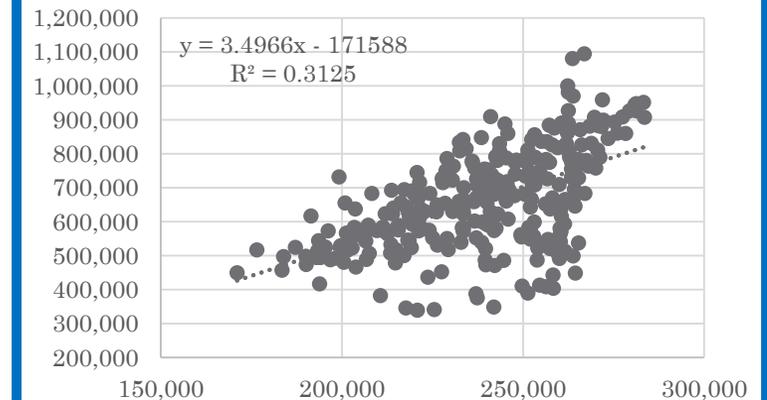
Scatter Plots: Truck Sales and Auto Variables

Truck Sales vs. Bank Auto Interest Rate — $y = -29449x + 869416$, $R^2 = 0.1232$
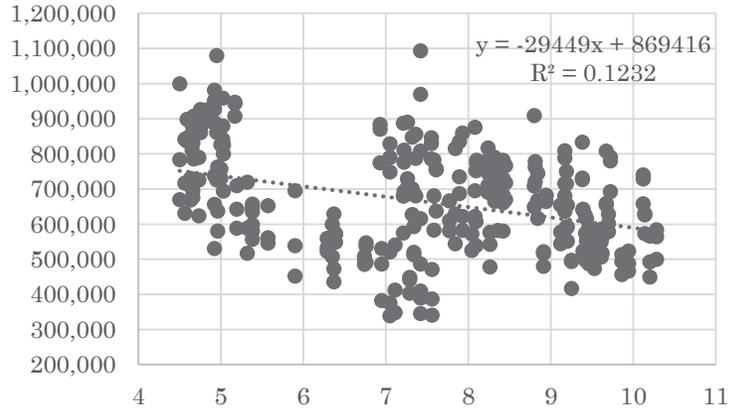
Truck Sales vs. Population — $y = 2602.6x - 107828$, $R^2 = 0.1082$
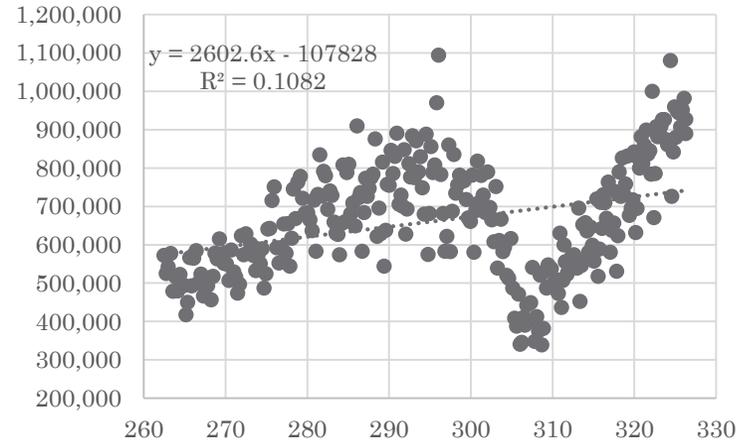
Truck Sales vs. Unemployment — $y = -44041x + 923552$, $R^2 = 0.233$
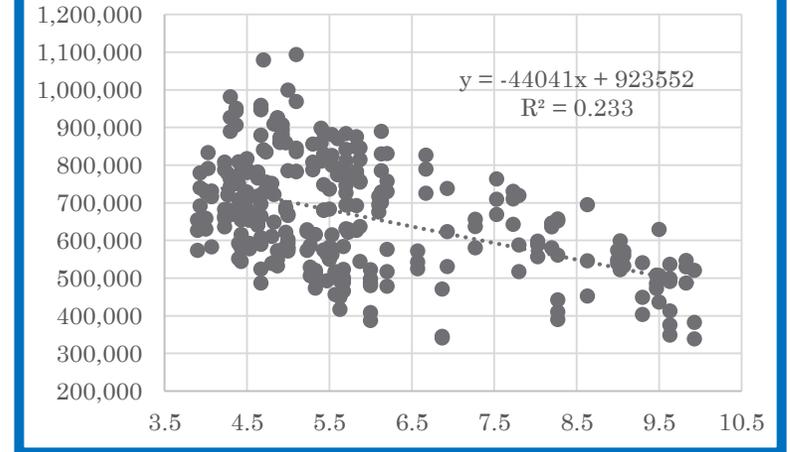
Truck Sales vs. Disposable Income / Capita Real — $y = 15.822x + 129022$, $R^2 = 0.1599$
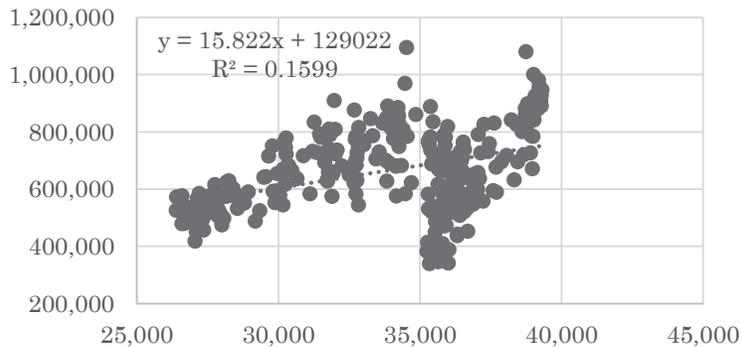
Truck Sales vs. Saving / Capita — $y = -81.981x + 804737$, $R^2 = 0.0721$
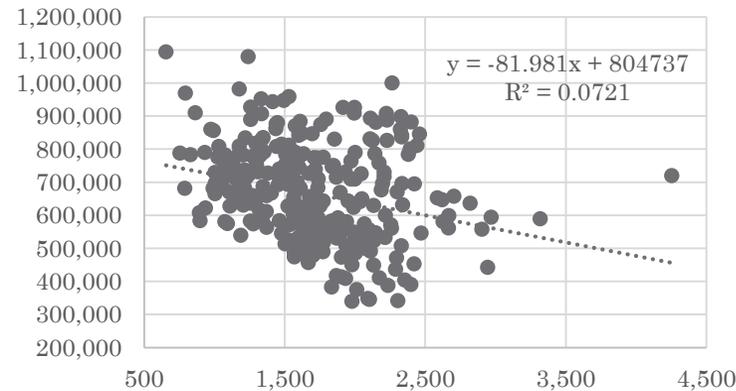
Truck Sales vs. Household Debt / Capita — $y = 3.2899x + 544490$, $R^2 = 0.0307$
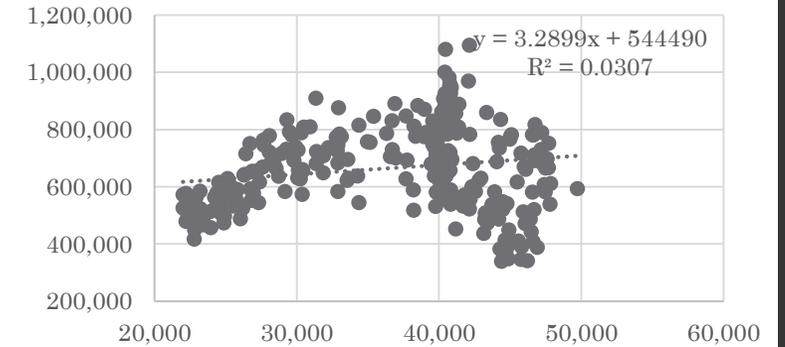
Scatter Plots: Truck Sales and Auto Variables

Truck Sales vs. US House Price Index
$y = 21250x + 589214$
$R^2 = 0.4256$

Truck Sales vs. Consumer Confidence
$y = 2273.1x + 448784$
$R^2 = 0.1671$
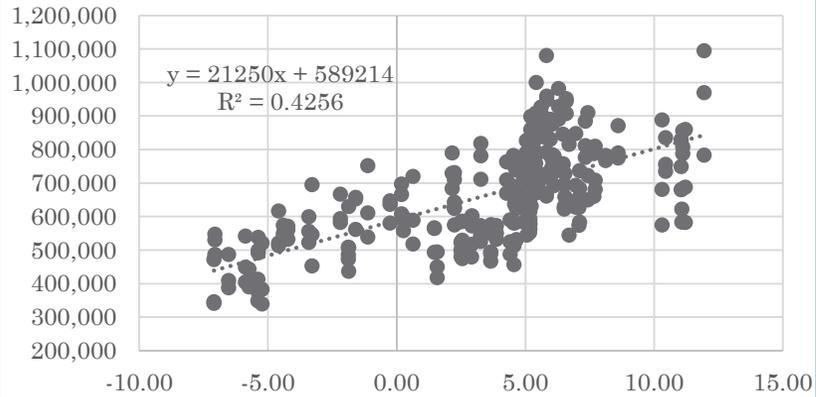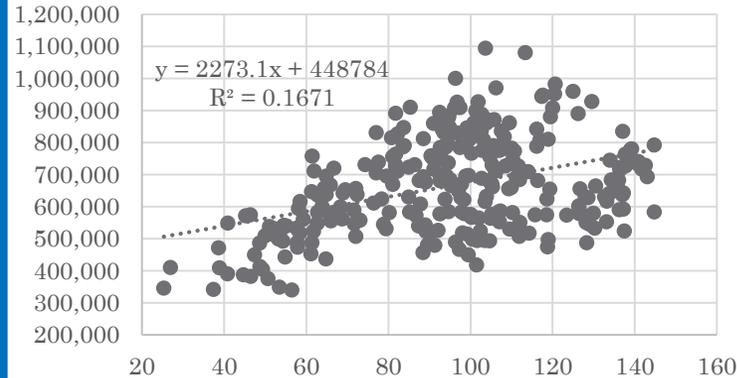
Truck Sales vs. Real GDP / Capita
$y = 16.085x - 76115$
$R^2 = 0.2086$
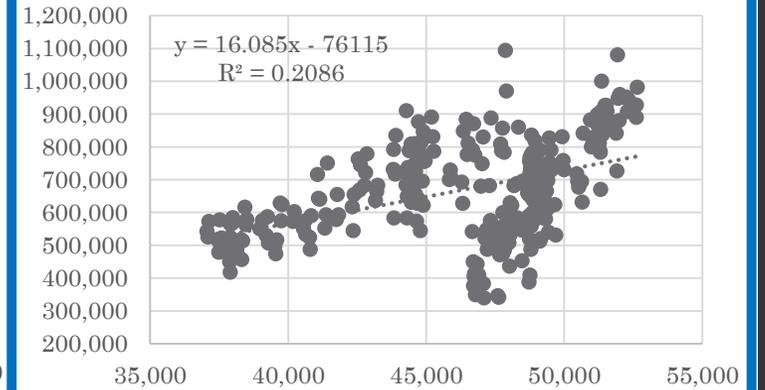
Truck Sales vs. Fed Funds Rate
$y = -7709.6x + 684573$
$R^2 = 0.0148$
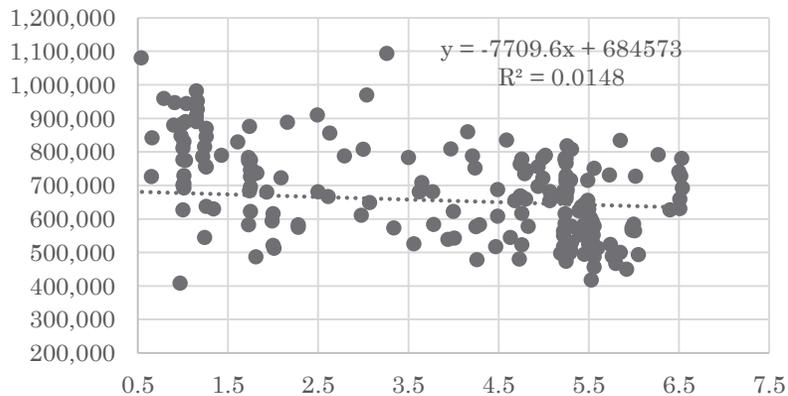
Truck Sales vs. U.S. Urban CPI Change
$y = -647.67x + 665,835.70$
$R^2 = 0.00$
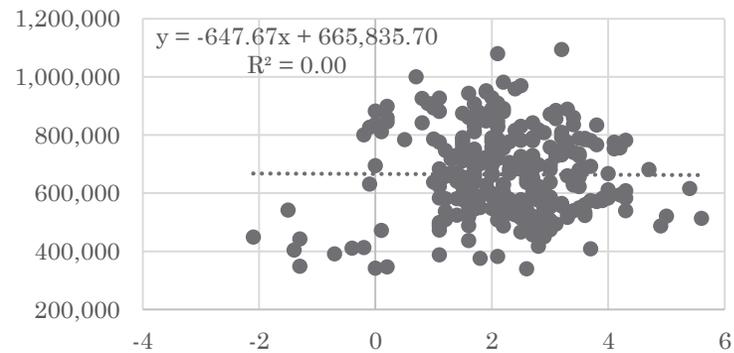
Truck Sales vs. Nonfarm Manufacturing Payrolls
$y = -7.3544x + 770220$
$R^2 = 0.0122$
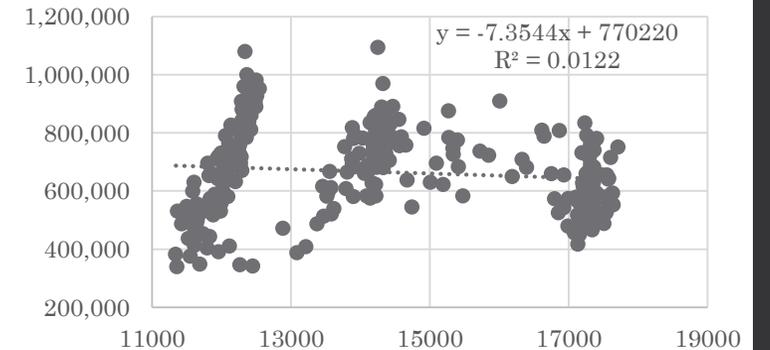
What did we learn from the scatter plots?

```
S O
C L E A R L Y                P P
A N O T H E R      C  C      O A
T E S T - - -      O  R      R R A
T O - - - - -      U  U      C S L
E V I N C E -      R  T      I N W
R E L A T I O N    T  C      N I A
                   S  H      P P Y
                            I    S
```

(Note: that says SCATTER read-down; not, "So clearly another test to evince relationships," read-across.)

- We validated that certain variables had different relationships between CAR sales and TRUCK sales

- We continued moving forward with building two separate models

But wait, we couldn't not make any histograms…
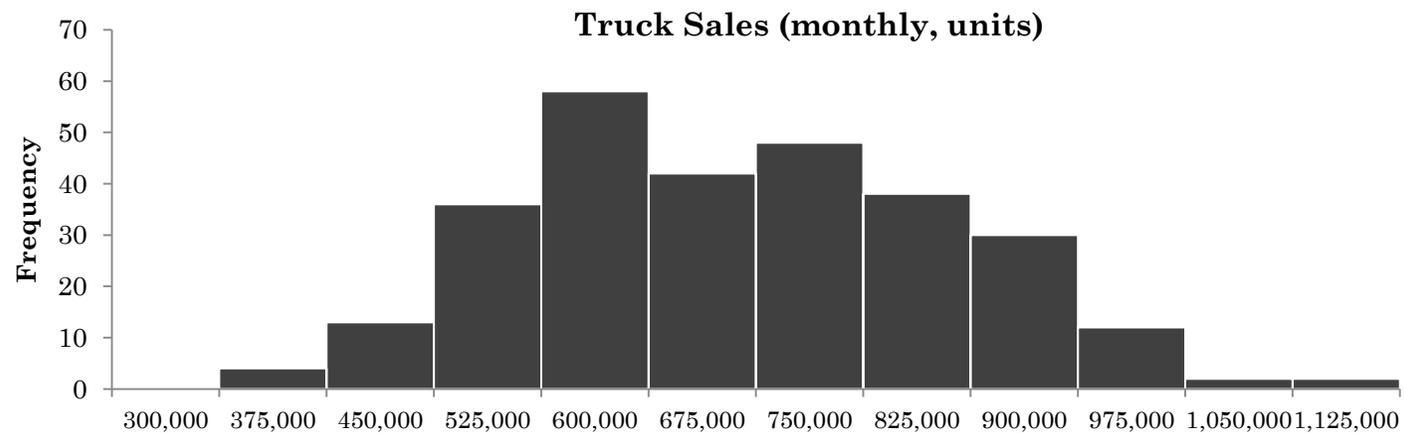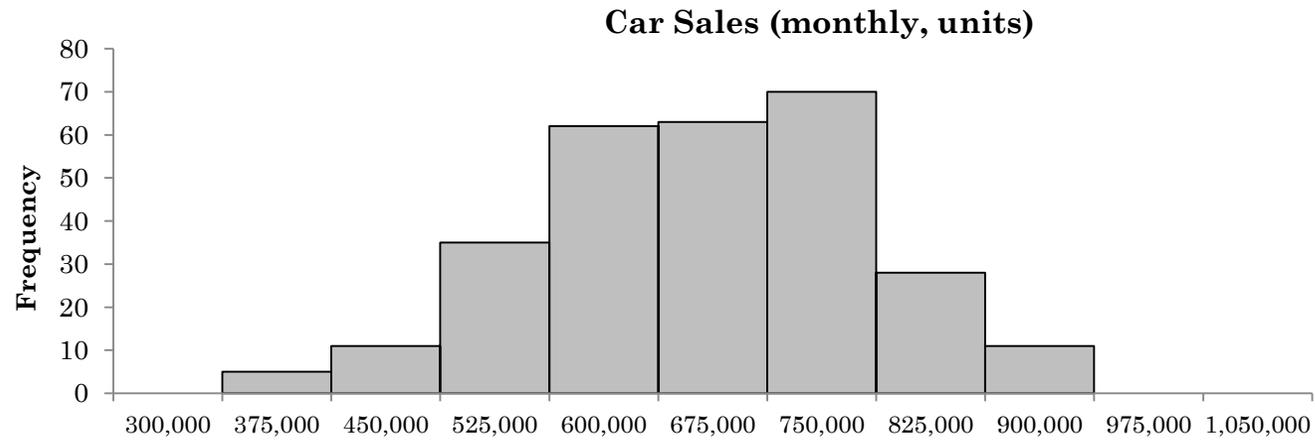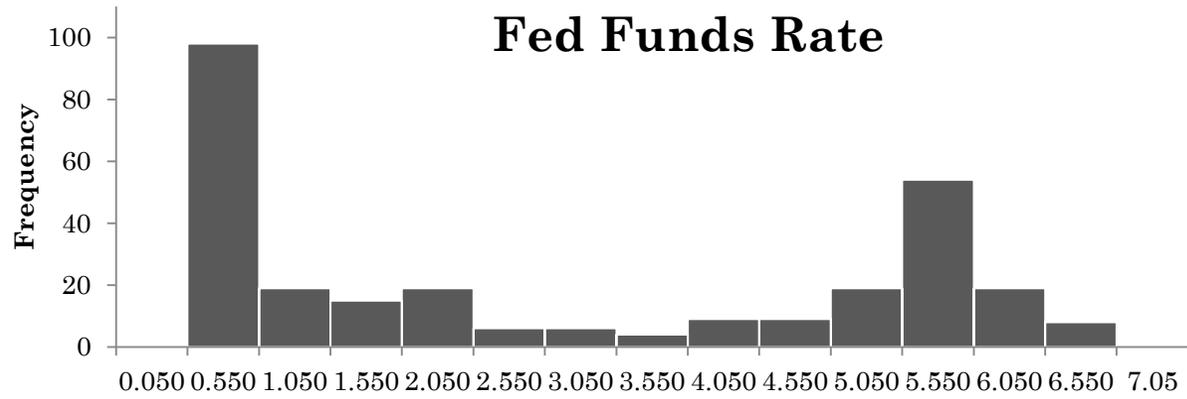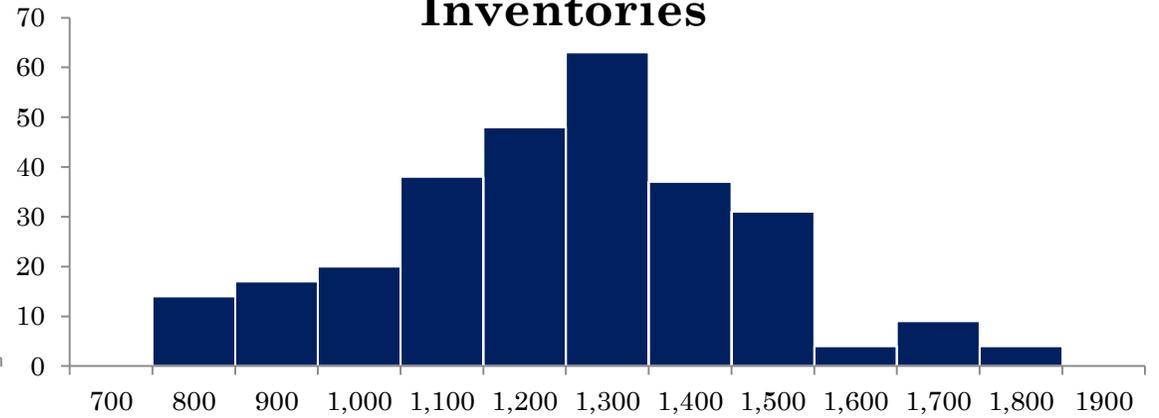
Car Sales (monthly, units)

Truck Sales (monthly, units)

Model 1: Car Sales – Best Subsets

Vars	= 18
R^2	= 84%
Adj. R^2	= 83%
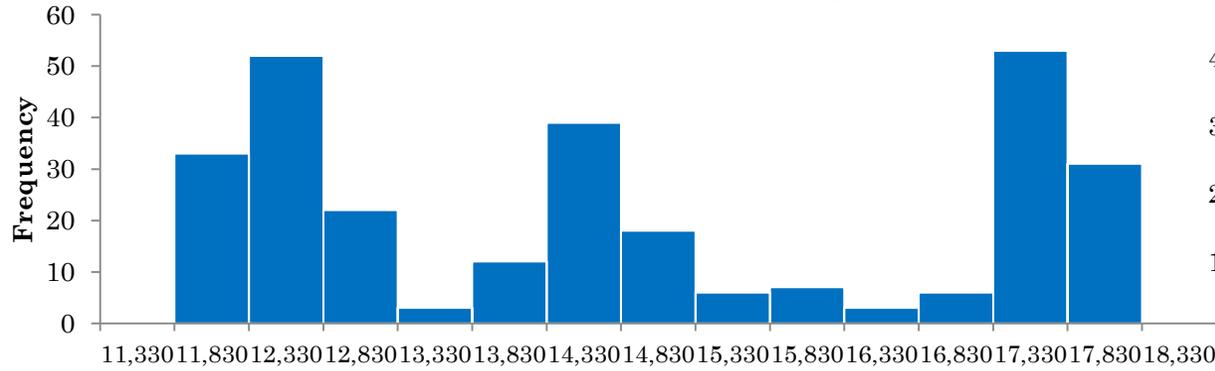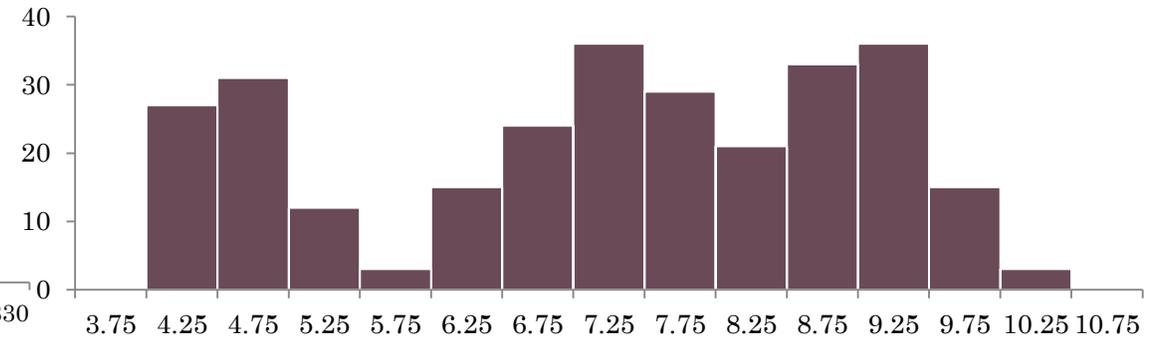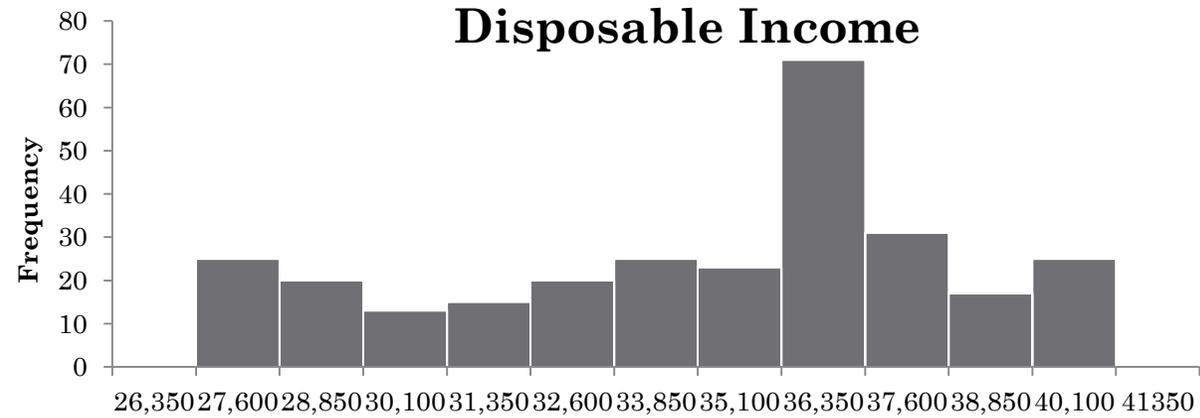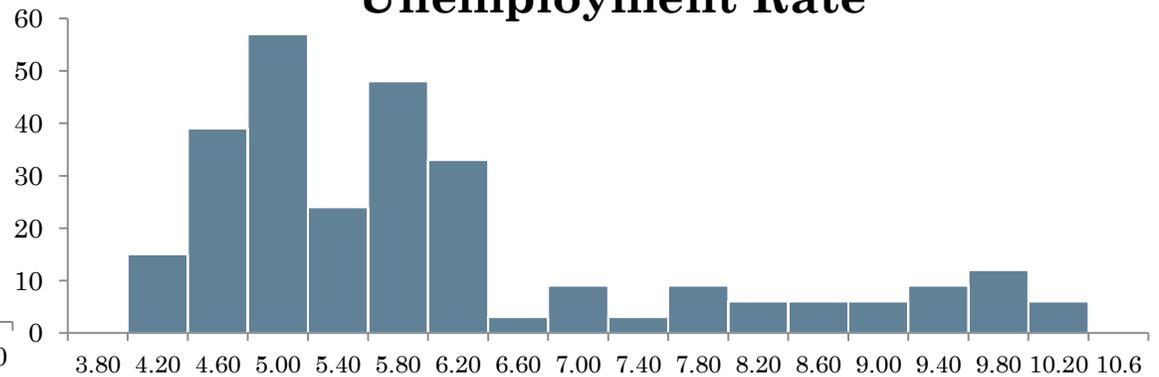
Model 1: Car Sales – Full Regression Model

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
49090.5	82.49%	81.37%	79.72%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1209530	2396802	0.50	0.614	
Time	26217	3961	6.62	0.000	12558.92
Year	-359204	44109	-8.14	0.000	10820.66
Q1	322391	33962	9.49	0.000	25.27
Q2	222727	23900	9.32	0.000	12.76
Q3	95366	14416	6.62	0.000	4.64
CPI Used Cars NSA YoY	2855	741	3.85	0.000	1.94
Gas Prices	134628	11385	11.83	0.000	12.61
Miles Driven	3.241	0.291	11.12	0.000	5.54
US Population (in mill.)	18	11689	0.00	0.999	5571.28
Unempl. Rate	-56658	13336	-4.25	0.000	54.50
Disposible Income Capita Real	53.9	16.4	3.28	0.001	439.09
Saving / Capita	-34.6	16.4	-2.10	0.036	7.40
Household Debt / Capita	-21.63	3.09	-7.00	0.000	69.11
Annualized RealGDP / Capita	-16.2	13.4	-1.21	0.226	367.13
Fed Funds Rate	-243	5413	-0.04	0.964	18.65
US Urban CPI Change NSA	-29778	4598	-6.48	0.000	3.43
Nonfarm Mfg Payrolls NSA	-61.6	25.3	-2.43	0.016	368.60

Residual Plots for Car
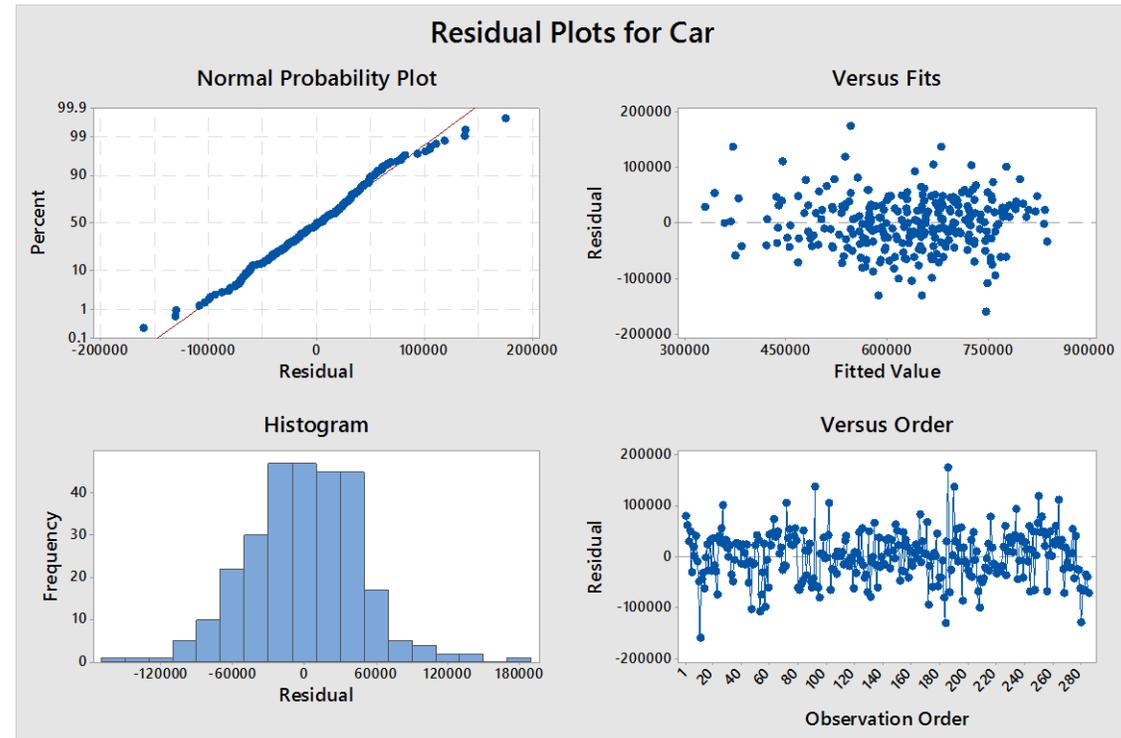
Model 1: Car Sales – Stepwise Method

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
54890.1	77.94%	76.71%	74.95%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	916081	346121	2.65	0.009	
Time	-3572	465	-7.68	0.000	138.49
Q1	54583	10471	5.21	0.000	1.92
Q2	44861	10287	4.36	0.000	1.89
Q3	6671	10205	0.65	0.514	1.86
CPI Used Cars NSA YoY	2903	827	3.51	0.001	1.94
Gas Prices	131265	12290	10.68	0.000	11.75
Miles Driven	3.237	0.324	10.00	0.000	5.47
Unempl. Rate	-53457	7489	-7.14	0.000	13.75
Disposible Income Capita Real	43.6	17.9	2.44	0.015	416.77
Saving / Capita	-22.9	18.1	-1.27	0.205	7.14
Household Debt / Capita	-20.67	2.73	-7.57	0.000	43.14
Annualized RealGDP / Capita	-8.2	13.8	-0.60	0.551	312.87
Fed Funds Rate	-1237	5397	-0.23	0.819	14.83
US Urban CPI Change NSA	-29099	4952	-5.88	0.000	3.18
Nonfarm Mfg Payrolls NSA	-54.4	18.7	-2.90	0.004	161.32

Residual Plots for Car

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
56513.0	76.36%	75.31%	73.84%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-280780	203365	-1.38	0.169	
Time	-2427	328	-7.40	0.000	64.97
Q1	62620	10576	5.92	0.000	1.85
Q2	50732	10416	4.87	0.000	1.83
Q3	7887	10468	0.75	0.452	1.85
CPI Used Cars NSA YoY	2308	752	3.07	0.002	1.51
Gas Prices	117427	10948	10.73	0.000	8.80
Miles Driven	3.255	0.331	9.83	0.000	5.40
Unempl. Rate	-25942	4032	-6.43	0.000	3.76
Disposible Income Capita Real	29.60	9.06	3.27	0.001	100.85
Saving / Capita	-13.8	12.0	-1.15	0.251	2.95
Household Debt / Capita	-15.77	1.76	-8.97	0.000	16.90
US Urban CPI Change NSA	-25586	4983	-5.14	0.000	3.04

Residual Plots for Car
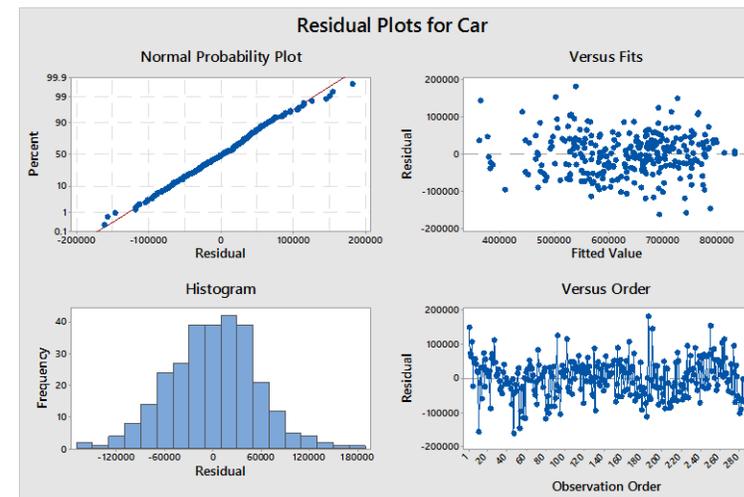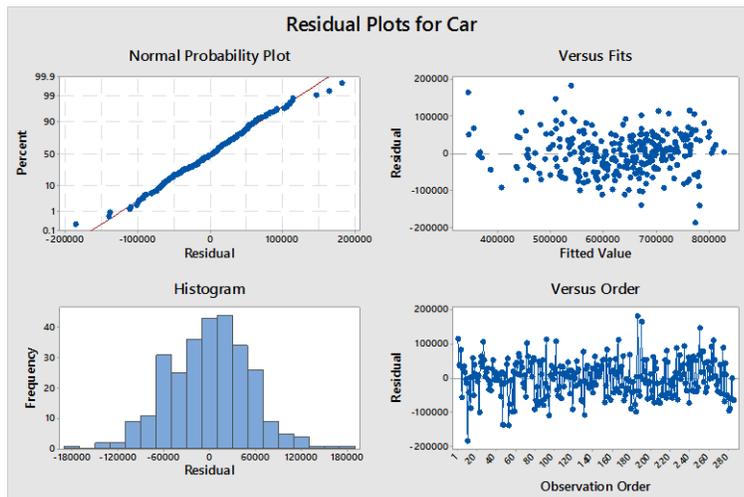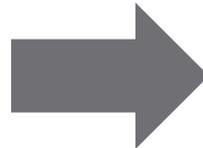
Model 1: Car Sales – Final Model (<u>without Q3</u>)

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
57354.8	75.38%	74.57%	73.43%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	347569	63148	5.50	0.000	
Time	-1433	112	-12.82	0.000	7.32
Q1	66181	10252	6.46	0.000	1.69
Q2	47015	8516	5.52	0.000	1.19
CPI Used Cars NSA YoY	1954	750	2.61	0.010	1.46
Gas Prices	106502	9848	10.81	0.000	6.91
Unempl. Rate	-31986	3297	-9.70	0.000	2.44
Household Debt / Capita	-11.272	0.900	-12.52	0.000	4.30
US Urban CPI Change NSA	-21775	4612	-4.72	0.000	2.53
Miles Driven	3.660	0.292	12.54	0.000	4.07

Regression Equation

Car   =   347569 - 1433 Time + 66181 Q1 + 47015 Q2 + 1954 CPI Used Cars NSA YoY
          + 106502 Gas Prices - 31986 Unempl. Rate - 11.272 Household Debt / Capita
          - 21775 US Urban CPI Change NSA + 3.660 Miles Driven
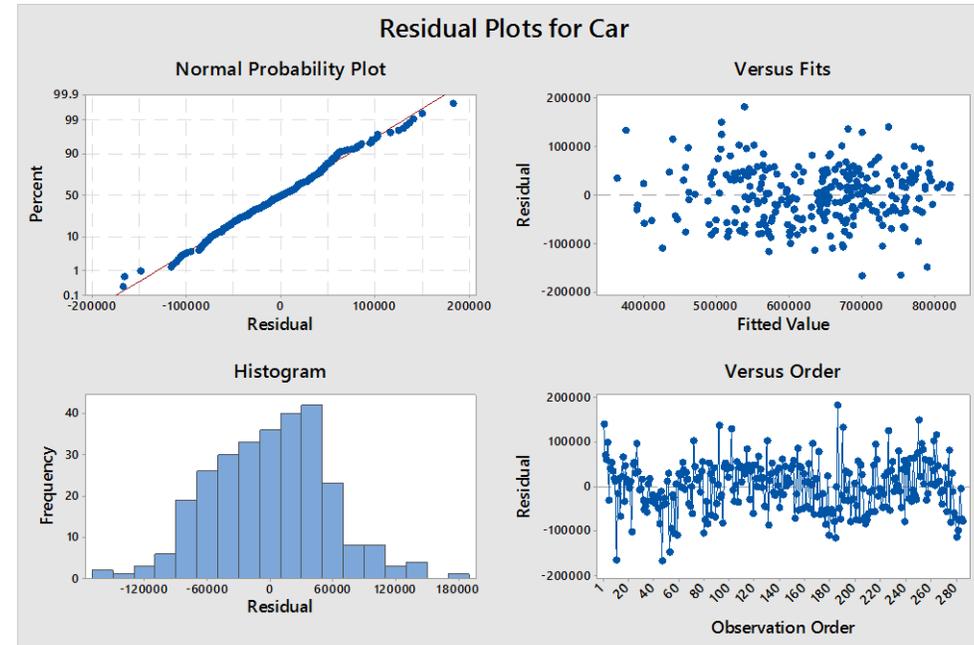
Model 1: Car Sales – Final Model (<u>with Q3</u>)

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
57424.1	75.41%	74.51%	73.29%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	356583	65107	5.48	0.000	
Time	-1414	116	-12.15	0.000	7.93
Q1	67777	10627	6.38	0.000	1.81
Q2	50625	10557	4.80	0.000	1.82
Q3	6161	10624	0.58	0.562	1.84
CPI Used Cars NSA YoY	1938	751	2.58	0.010	1.46
Gas Prices	105168	10125	10.39	0.000	7.29
Unempl. Rate	-31902	3304	-9.66	0.000	2.44
Household Debt / Capita	-11.157	0.923	-12.09	0.000	4.51
US Urban CPI Change NSA	-21429	4656	-4.60	0.000	2.57
Miles Driven	3.589	0.317	11.32	0.000	4.79

Residual Plots for Car

Regression Equation

Car = 356583 - 1414 Time + 67777 Q1 + 50625 Q2 + 6161 Q3 + 1938 CPI Used Cars NSA YoY
+ 105168 Gas Prices - 31902 Unempl. Rate - 11.157 Household Debt / Capita
- 21429 US Urban CPI Change NSA + 3.589 Miles Driven

Model 2: Truck Sales – Best Subsets

Vars = 19

R^2 = 88.6%

Adj. R^2 = 87.8%
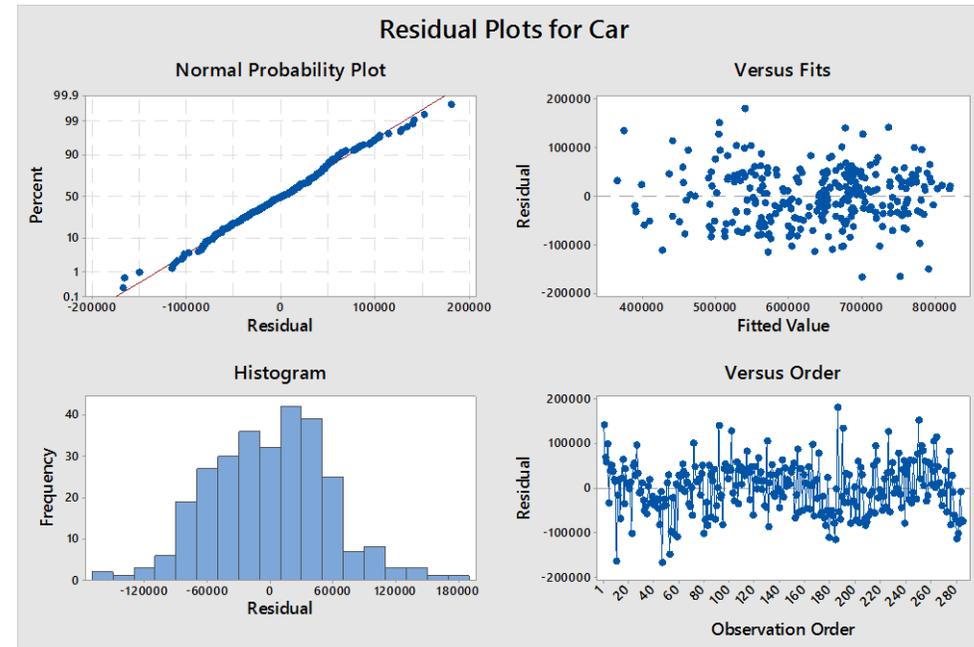
Model 2: Truck Sales – Full Regression Model

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
51381.9	88.62%	87.81%	86.74%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	627917	436949	1.44	0.152	
Time	32601	3824	8.53	0.000	10685.47
Q1	324297	35511	9.13	0.000	25.22
Q2	209719	24630	8.51	0.000	12.36
Q3	82972	14708	5.64	0.000	4.41
CPI Used Cars NSA YoY	1559	770	2.02	0.044	1.92
Unempl. Rate	-34790	9117	-3.82	0.000	23.25
Household Debt / Capita	-15.47	3.21	-4.82	0.000	68.07
US Urban CPI Change NSA	7297	4125	1.77	0.078	2.52
Miles Driven	3.262	0.305	10.71	0.000	5.52
Year	-429289	45991	-9.33	0.000	10737.78
CPI New Cars NSA YoY	-8614	3413	-2.52	0.012	2.68
Cons. Conf. Autos	4632	2695	1.72	0.087	5.88
Bank Auto Interest Rate	-35509	9452	-3.76	0.000	29.56
Disposible Income Capita Real	116.0	16.3	7.14	0.000	392.87
Saving / Capita	-114.4	17.0	-6.72	0.000	7.23
US House Price Inflation NSA	3808	1617	2.36	0.019	5.74
Consumer Confidence	1422	330	4.31	0.000	8.21
Annualized RealGDP / Capita	-48.5	13.8	-3.51	0.001	358.79
Nonfarm Mfg Payrolls NSA	-59.8	25.1	-2.38	0.018	330.81

Residual Plots for Truck
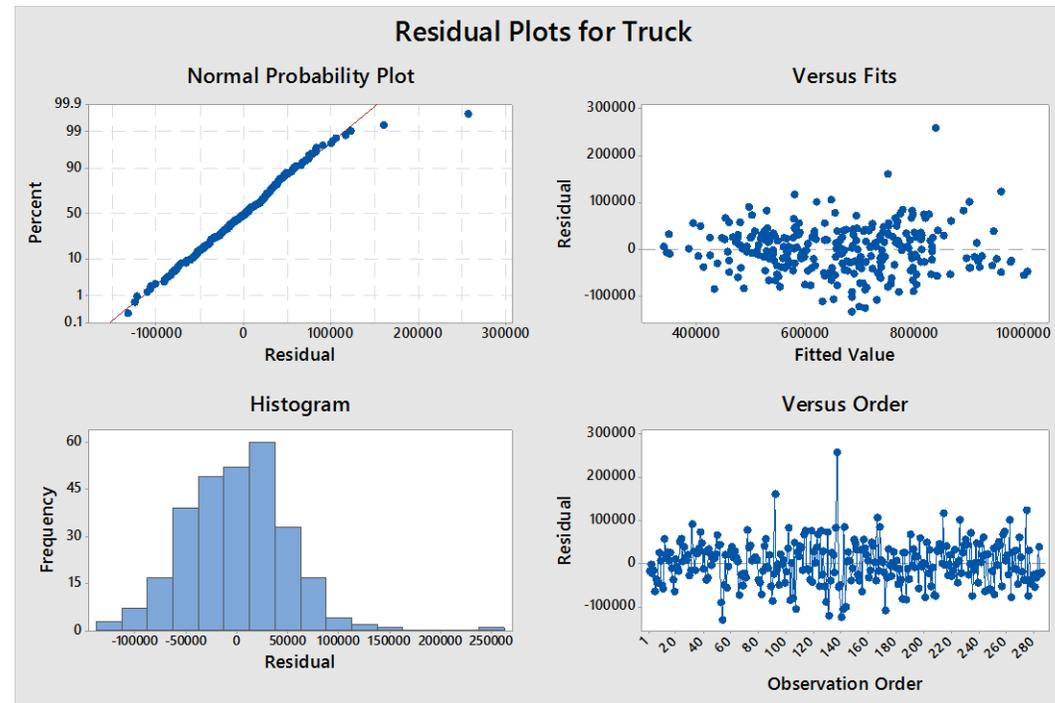
Model 2: Truck Sales – Stepwise Method

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
59118.0	84.88%	83.86%	82.53%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-26403	496226	-0.05	0.958	
Time	-2737	620	-4.41	0.000	212.43
Q1	6445	11590	0.56	0.579	2.03
Q2	-2827	10803	-0.26	0.794	1.80
Q3	-22574	10822	-2.09	0.038	1.80
CPI Used Cars NSA YoY	1319	886	1.49	0.138	1.92
Unempl. Rate	-23675	10399	-2.28	0.024	22.85
Household Debt / Capita	-14.55	3.69	-3.94	0.000	68.01
US Urban CPI Change NSA	8269	4744	1.74	0.082	2.52
Miles Driven	3.192	0.350	9.11	0.000	5.52
CPI New Cars NSA YoY	-6910	3922	-1.76	0.079	2.67
Cons. Conf. Autos	100	3050	0.03	0.974	5.69
Bank Auto Interest Rate	-42760	10839	-3.95	0.000	29.36
Disposible Income Capita Real	114.3	18.7	6.12	0.000	392.82
Saving / Capita	-107.6	19.6	-5.50	0.000	7.22
US House Price Inflation NSA	4376	1859	2.35	0.019	5.73
Consumer Confidence	1441	380	3.79	0.000	8.21
Annualized RealGDP / Capita	-45.7	15.9	-2.87	0.004	358.62
Nonfarm Mfg Payrolls NSA	-32.0	28.7	-1.12	0.265	326.18
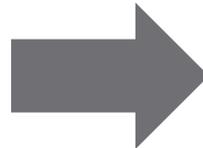
Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
64889.2	81.58%	80.55%	79.11%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	138631	127966	1.08	0.280	
Time	-179	291	-0.61	0.539	38.87
Q1	19187	12098	1.59	0.114	1.84
Q2	-2535	11708	-0.22	0.829	1.75
Q3	-28915	11820	-2.45	0.015	1.79
CPI Used Cars NSA YoY	647	965	0.67	0.503	1.89
Unempl. Rate	-16270	5608	-2.90	0.004	5.52
Household Debt / Capita	-3.49	1.78	-1.97	0.050	13.08
US Urban CPI Change NSA	6834	4817	1.42	0.157	2.16
Miles Driven	3.806	0.369	10.31	0.000	5.09
CPI New Cars NSA YoY	-8928	3859	-2.31	0.021	2.15
Cons. Conf. Autos	938	3274	0.29	0.775	5.44
Bank Auto Interest Rate	-33582	9388	-3.58	0.000	18.28
Saving / Capita	-23.6	13.0	-1.81	0.071	2.65
US House Price Inflation NSA	10123	1652	6.13	0.000	3.76
Consumer Confidence	904	364	2.48	0.014	6.26

Residual Plots for Truck
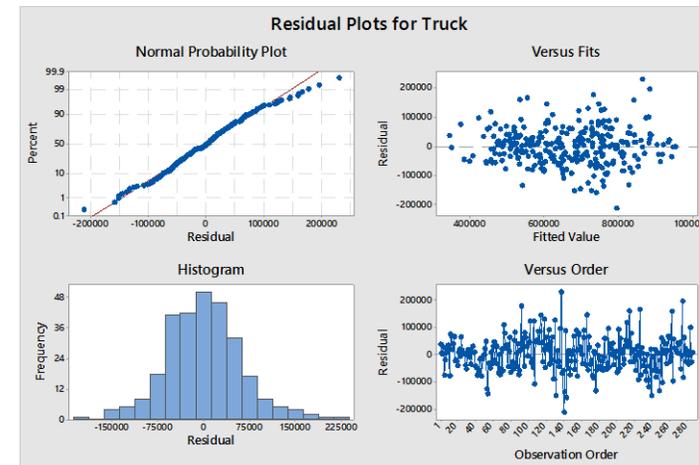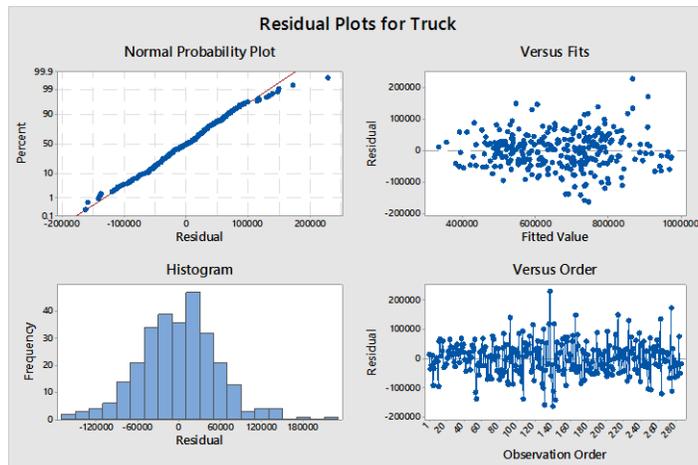
Model 2: Truck Sales – Stepwise Method

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
64664.5	81.57%	80.69%	79.42%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	149105	124090	1.20	0.231	
Time	-125	241	-0.52	0.604	26.70
Q1	19983	11548	1.73	0.085	1.68
Q3	-27232	9652	-2.82	0.005	1.20
CPI Used Cars NSA YoY	596	950	0.63	0.531	1.84
Unempl. Rate	-15966	5482	-2.91	0.004	5.31
Household Debt / Capita	-3.80	1.40	-2.71	0.007	8.22
US Urban CPI Change NSA	7303	4499	1.62	0.106	1.89
Miles Driven	3.773	0.345	10.92	0.000	4.48
CPI New Cars NSA YoY	-8993	3839	-2.34	0.020	2.14
Bank Auto Interest Rate	-33424	9320	-3.59	0.000	18.15
Saving / Capita	-23.1	12.6	-1.83	0.068	2.51
US House Price Inflation NSA	10202	1631	6.25	0.000	3.69
Consumer Confidence	931	340	2.74	0.006	5.48

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
64491.0	81.53%	80.79%	79.63%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	119300	115297	1.03	0.302	
Q1	19450	11446	1.70	0.090	1.66
Q3	-27153	9618	-2.82	0.005	1.20
Unempl. Rate	-14653	5161	-2.84	0.005	4.73
Household Debt / Capita	-4.32	1.05	-4.12	0.000	4.61
US Urban CPI Change NSA	8113	4339	1.87	0.063	1.77
Miles Driven	3.746	0.339	11.04	0.000	4.35
CPI New Cars NSA YoY	-8162	3552	-2.30	0.022	1.84
Bank Auto Interest Rate	-29032	4154	-6.99	0.000	3.62
Saving / Capita	-23.5	12.6	-1.87	0.062	2.50
US House Price Inflation NSA	10607	1429	7.42	0.000	2.84
Consumer Confidence	893	310	2.88	0.004	4.60

Residual Plots for Truck

Residual Plots for Truck
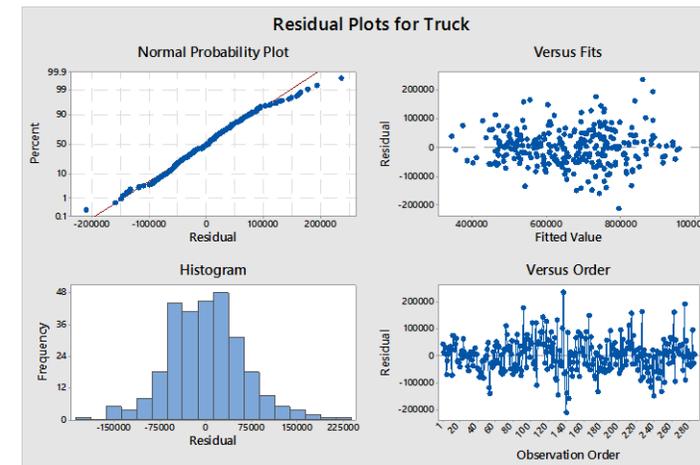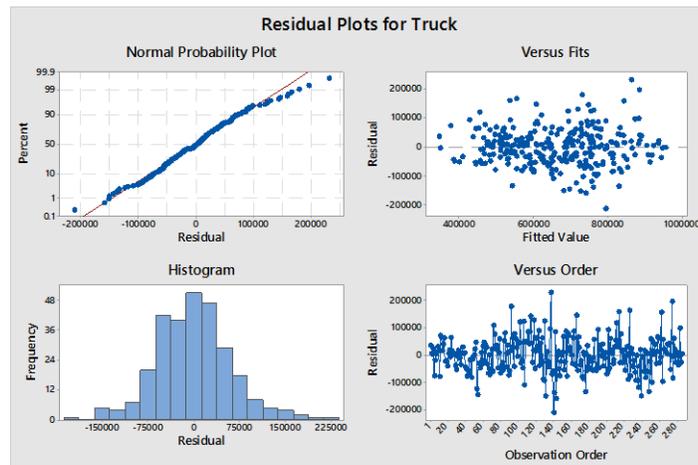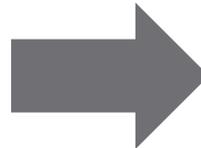
Model 2: Truck Sales – Final Model

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
64786.3	81.30%	80.61%	79.71%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	20501	103019	0.20	0.842	
Q1	19233	11497	1.67	0.095	1.66
Q3	-26240	9650	-2.72	0.007	1.19
Unempl. Rate	-15231	5176	-2.94	0.004	4.71
Household Debt / Capita	-3.662	0.993	-3.69	0.000	4.10
US Urban CPI Change NSA	10064	4232	2.38	0.018	1.67
Miles Driven	3.747	0.341	11.00	0.000	4.35
CPI New Cars NSA YoY	-8180	3568	-2.29	0.023	1.84
Bank Auto Interest Rate	-25515	3723	-6.85	0.000	2.88
US House Price Inflation NSA	11577	1338	8.65	0.000	2.47
Consumer Confidence	947	310	3.05	0.002	4.56

Residual Plots for Truck

Regression Equation

Truck   =   20501 + 19233 Q1 - 26240 Q3 - 15231 Unempl. Rate - 3.662 Household Debt / Capita
         + 10064 US Urban CPI Change NSA + 3.747 Miles Driven - 8180 CPI New Cars NSA YoY
         - 25515 Bank Auto Interest Rate + 11577 US House Price Inflation NSA
         + 947 Consumer Confidence

Model Comparison: Car Sales vs. Truck Sales

Model 1: Car Sales

R-sq	R-sq(adj)
75.38%	74.57%

Model 2: Truck Sales

R-sq	R-sq(adj)
81.30%	80.61%

- Truck Sales Model has greater explanatory power with higher R-square and Adjusted R-square

	Coefficient Comparison	Variables in Both Models	
1	Q1	66,181	19,233
2	U-Rate	-31,986	-15,231
3	Household Debt	-11.3	-3.7
4	CPI Change	-21,775	10,064
5	Miles Driven	3.7	3.7

Q1	Both CAR and TRUCK sales are significantly affected by seasonality. Comparing Q1 to Q4 (dummy), there is a strong increase in the number of cars sold (~66K). Truck sales (~19K) are impacted too, yet not as much as car sales.
U-Rate	For every increase in % of the unemployment rate, both CAR and TRUCK sales are negatively impacted. Car sales though are affected 2x more by an increase in unemployment than truck sales.
CPI Change	While both are affected by an increase in % of the inflation rate, CAR sales are negatively hit whereas TRUCK sales are positively impacted.
Debt & Miles Driven	Both CAR and TRUCK sales are negatively affected as household debt increases. For miles driven, the sales for both increases. The coefficient values may seem small, but the variable inputs have very high values.

Model Comparison: Car Sales vs. Truck Sales

Model 1: Car Sales		R-sq	R-sq(adj)
		75.38%	74.57%

Model 2: Truck Sales		R-sq	R-sq(adj)
		81.30%	80.61%

	Coefficient Comparison	Only in Model 1 - Car Sales	Only in Model 2 - Truck Sales
1	Time	-1,433	
2	Q2	47,015	
3	CPI Used Cars	1,954	
4	Gas Prices	106,502	
5	Q3		-26,240
6	CPI New Cars		-8,180
7	Bank Auto Interest Rate		-25,515
8	House Price Inflation		11,577
9	Consumer Confidence		947

Time — For CARs, there have been less sold over time, whereas for TRUCK sales, the coefficient wasn't statistically significant.

Q2 — Similar to Q1 for CARs, compared to Q4 (dummy), Q2 is more active in selling units.

Gas Prices — Surprisingly, against our intuition, CAR sales increase as gas prices increase by a large factor. The team hypothesized that fuel efficiency standards have greatly increased, as well as that this is not a causal relationship.

Bank Auto — As the interest rates on auto loans from banks increase, TRUCK sales are negatively impacted.

House Price — As the prices of homes increase, TRUCK sales are positively affected.

CPI New Cars — As the prices new cars increase, TRUCK sales are negatively affected.

Testing the Models

Step 1	Select variables from final models for CAR sales and TRUCK sales
Step 2	Using those variables, re-run both multivariate regressions, but only from **1994-2016**
Step 3	Predict monthly sales for CAR and TRUCK in **2017**

Model 1 — Car Sales

	Without Q3	With Q3
R-Square	75.38%	75.41%
Adj. R-Square	74.57%	74.51%
Variable Count	9	10
Variables	Time	Time
	Q1	Q1
	Q2	Q2
	CPI Used Cars	Q3
	Gas Prices	CPI Used Cars
	U-Rate	Gas Prices
	Household Debt	U-Rate
	CPI Change	Household Debt
	Miles Driven	CPI Change
		Miles Driven

Model 2 — Truck Sales

R-Square	81.30%
Adj. R-Square	80.61%
Variable Count	10
Variables	Q1
	Q3
	U-Rate
	Household Debt
	CPI Change
	Miles Driven
	CPI New Cars
	Bank Auto Interest Rate
	House Price Inflation
	Consumer Confidence

Model 1 Car Sales: Actual vs. Prediction

Car Sales (without Q3)

	Actual	Predicted
Mean	507,433	600,427
Deviation		18%

Date	Actual	Predicted
Jan-17	411,489	527,713
Feb-17	484,276	489,834
Mar-17	589,051	636,683
Apr-17	537,320	635,109
May-17	562,233	668,916
Jun-17	520,568	671,636
Jul-17	501,179	634,242
Aug-17	523,354	634,391
Sep-17	534,177	571,458
Oct-17	457,603	564,504
Nov-17	460,510	570,204
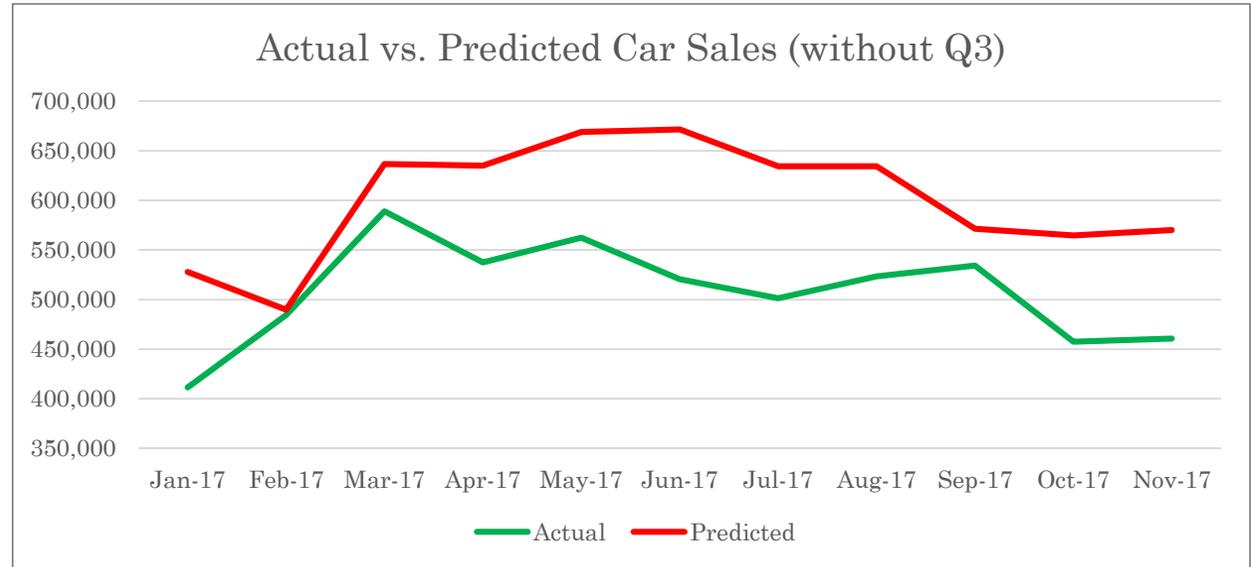
Car Sales (with Q3)

	Actual	Predicted
Mean	507,433	602,424
Deviation		19%

Date	Actual	Predicted
Jan-17	411,489	530,474
Feb-17	484,276	493,773
Mar-17	589,051	636,067
Apr-17	537,320	637,344
May-17	562,233	670,006
Jun-17	520,568	672,801
Jul-17	501,179	639,117
Aug-17	523,354	639,340
Sep-17	534,177	578,546
Oct-17	457,603	561,813
Nov-17	460,510	567,383

Actual vs. Predicted Car Sales (without Q3)

Actual vs. Predicted Car Sales (with Q3)
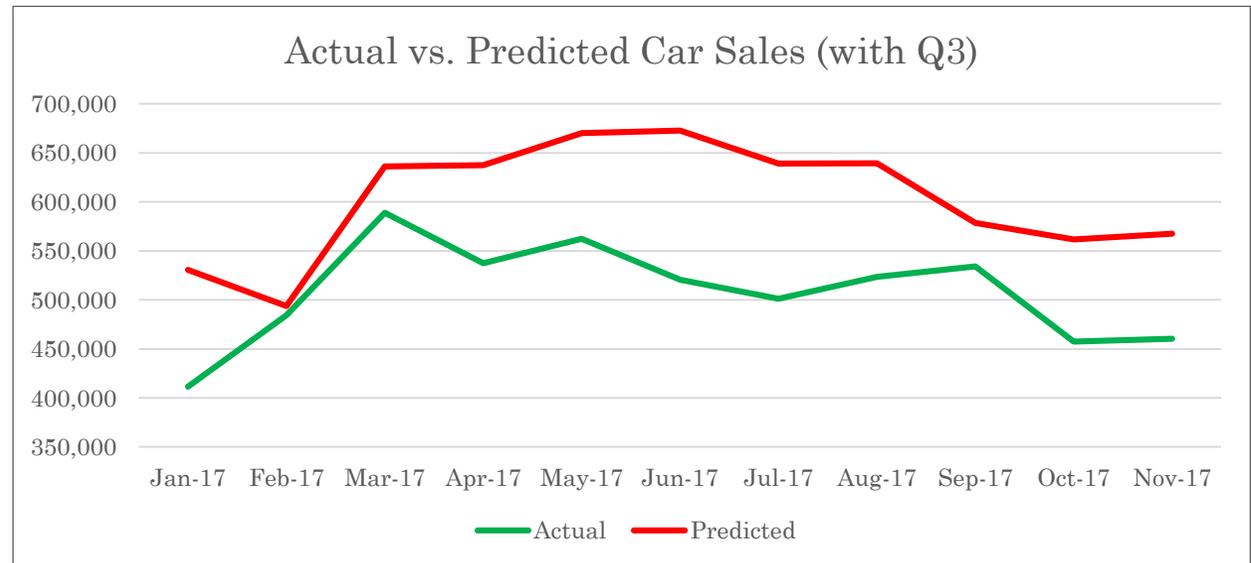
Model 2 Truck Sales: Actual vs. Prediction

Truck Sales		

Date	Actual	Predicted
Jan-17	726,293	805,520
Feb-17	841,623	779,147
Mar-17	958,993	932,186
Apr-17	880,042	903,137
May-17	947,400	932,022
Jun-17	943,709	942,702
Jul-17	907,305	934,519
Aug-17	951,676	943,351
Sep-17	981,582	868,822
Oct-17	889,899	902,564
Nov-17	927,449	906,313

	Actual	Predicted
Mean	905,088	895,480
Deviation	-1%	

Actual vs. Predicted Truck Sales
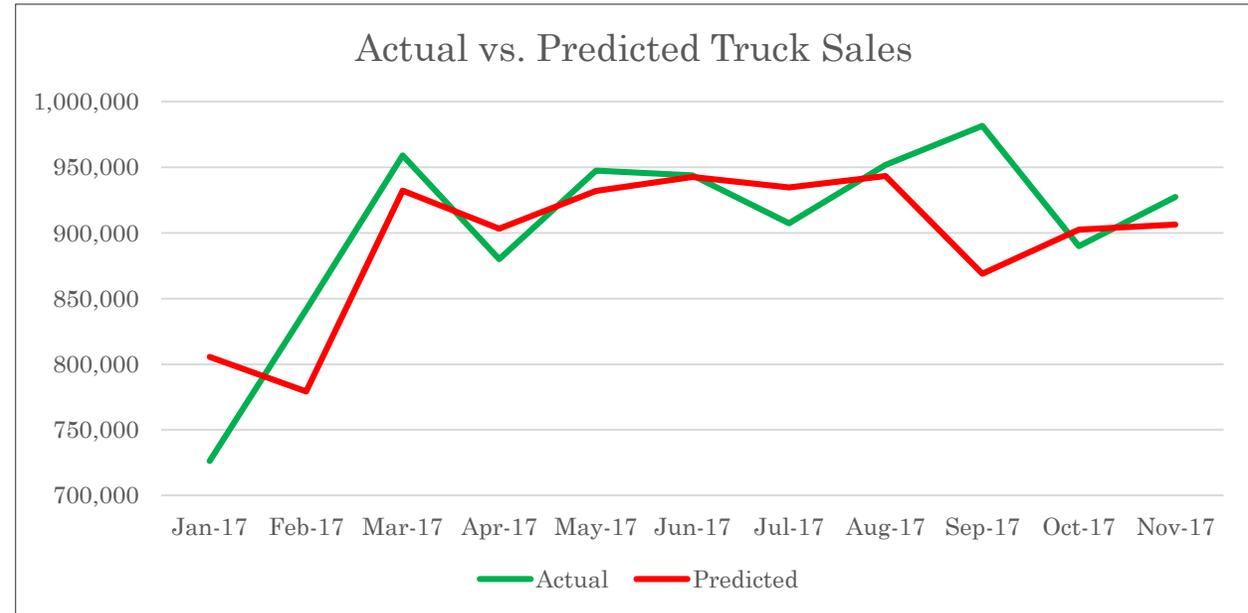
Truck Sales Model: A Stronger Economic Indicator

Takeaways

Q: What did we learn?

- Given the **low correlation** between **CARS and TRUCKS**, we conducted two **separate analyses** to increase our models' predictive power.

- While many variables overlapped between models, **only the TRUCKS** model **includes** macro-economic factors such as **CPI**, **new vehicle price inflation**, **bank auto interest rate, US house inflation**, **consumer confidence**, US **house price inflation**

- Our **CAR** sales **model overpredicted** and **did not capture demographic and behavioral changes** which may impact purchasing decisions.

- Beyond our models' performances, we observe that **TRUCKS are closer to representing state of economy**, both as passenger and commercial vehicles.

- Buying a truck might also be a **more deliberate purchase** than a consumer's choice to purchase a car.

Model Improvement

Models

Q: What additional variables might increase the models' predictive power?

A1: We could **incorporate geographic mobility** and look for monthly data reporting the number of people living in urban versus rural areas. This would support the notion that as more people move to cities, they rely more on public transportation and are less likely to need a car.

A2: We could further **isolate commercial trucks** in our data set. By only focusing on commercial vehicles versus a blend of both commercial and passenger trucks, we could test if the model improves.

A3: We could **segment customers by age** and explore whether there are trends in the type of people who are purchasing vehicles.

Appendix: Auto Sales: Predictor of Recessions

Q & A