

BUAN 6337.5U1

Predictive Analytics Using SAS

Summer 2019

Research Question

How do vehicle attributes affect its price?

Mentored By

Dr. Sourav Chatterjee

Prepared By

Pavan Sai Krishna Gorantla

Himaja Barla

Nirlep Shah

Ashley Creek

Siril Sai Alpuri

Contents

1. Abstract	3
2. Data Description	3
2.1. Outliers and Missing Values	4
2.2. Exploratory Data Analysis	6
2.3. Correlations among Predictors	10
3. Regression Models	13
3.1. Pricing Model with Fuel, Transmission and Title	13
3.2. Pricing Model with Manufacturer	14
3.3. Pricing Model Condition, Cylinders and Drive	15
3.4. Pricing Model with Type and Color	16
4. Conclusion	16
5. Appendix	17
5.1. SAS Code	17
5.2. SAS Output	21

1. Abstract

Craigslist is a popular platform used for buying and selling new and used goods by either private individuals or businesses. A dataset has been scraped from this website containing information about vehicles that are posted for sale on craigslist. The data has information on the sale price and other attributes of the vehicle. In this study, we focus on how the price depends on the attributes of vehicle and which attribute has significant effects on its sale price. Since most of the variables are categorical in nature, we run generalized linear regression models and interpret the effect shown on price. We aim to find best combination of attributes that can fetch highest price.

2. Data Description

We have cross-sectional dataset for 1,723,065 individual vehicles posted over several years span. Following below are variable descriptions and the green highlighted ones are vehicle attributes. We will be using only the highlighted variables and will be ignoring the rest of all other variables.

Variable	Definition
<i>url</i>	Website URL containing the vehicle details (unique identifier)
<i>city</i>	City in which the vehicle is available
<i>price</i>	Sale price of the vehicle
<i>year</i>	Year in which the vehicle was manufactured
<i>manufacturer</i>	Manufacturer name
<i>make</i>	Detailed description of make and model
<i>condition</i>	Describes vehicle condition (new, excellent, good, fair, like new, salvage)
<i>cylinders</i>	Number of cylinders in the vehicle (3, 4, 5, 6, 8, 10, 12, other)
<i>fuel</i>	Fuel type of the vehicle (diesel, electric, gas, hybrid, other)
<i>odometer</i>	Current odometer reading of the vehicle
<i>title_status</i>	Title about past history (clean, lien, missing, parts only, rebuilt, salvage)
<i>transmission</i>	Transmission Type of the vehicle (automatic, manual, other)
<i>vin</i>	Vehicle Identification Number
<i>drive</i>	Drive type (front wheel drive, rear wheel drive, four wheel drive)
<i>size</i>	Vehicle size (compact, full-size, mid-size, sub-compact)
<i>type</i>	Vehicle type (suv, bus, convertible, coupe, hatchback, mini-van, offroad, pickup, sedan, truck, van, wagon, other)
<i>paint_color</i>	Vehicle color (black, blue, brown, green, grey, orange, purple, red, silver, white, yellow, custom)
<i>image_url</i>	URL containing the vehicle's image
<i>lat</i>	Latitude of the location where the vehicle is available
<i>long</i>	Longitude of location where the vehicle is available
<i>county_fips</i>	Federal Information Processing Standard codes for the county
<i>county_name</i>	Name of the county
<i>state_fips</i>	Federal Information Processing Standard codes for the state
<i>state_code</i>	Two letter state code
<i>state_name</i>	Name of the state
<i>weather</i>	Weather code of the state

2.1. Outliers and Missing Values

Quantiles (Definition 5)		Quantiles (Definition 5)		Quantiles (Definition 5)	
Level	Quantile	Level	Quantile	Level	Quantile
100% Max	9999999	100% Max	2019	100% Max	10000000
99%	51999	99%	2018	99%	300000
95%	32950	95%	2017	95%	222000
90%	25000	90%	2015	90%	194000
75% Q3	14999	75% Q3	2012	75% Q3	152000
50% Median	7000	50% Median	2007	50% Median	107000
25% Q1	3295	25% Q1	2002	25% Q1	58000
10%	1500	10%	1994	10%	23456
5%	850	5%	1980	5%	8000
1%	1	1%	1955	1%	41
0% Min	1	0% Min	302	0% Min	0

Extreme Observations		Extreme Observations		Extreme Observations		Extreme Observations	
Lowest		Highest		Lowest		Highest	
Value	Obs	Value	Obs	Value	Obs	Value	Obs
302	7039	2019	1.72E6	0	1.72E6	10000000	1.43E6
718	614121	2019	1.72E6	0	1.72E6	10000000	1.47E6
1553	86238	2019	1.72E6	0	1.72E6	10000000	1.58E6
1740	1.3E6	2019	1.72E6	0	1.72E6	10000000	1.6E6
1796	354711	2019	1.72E6	0	1.72E6	10000000	1.72E6

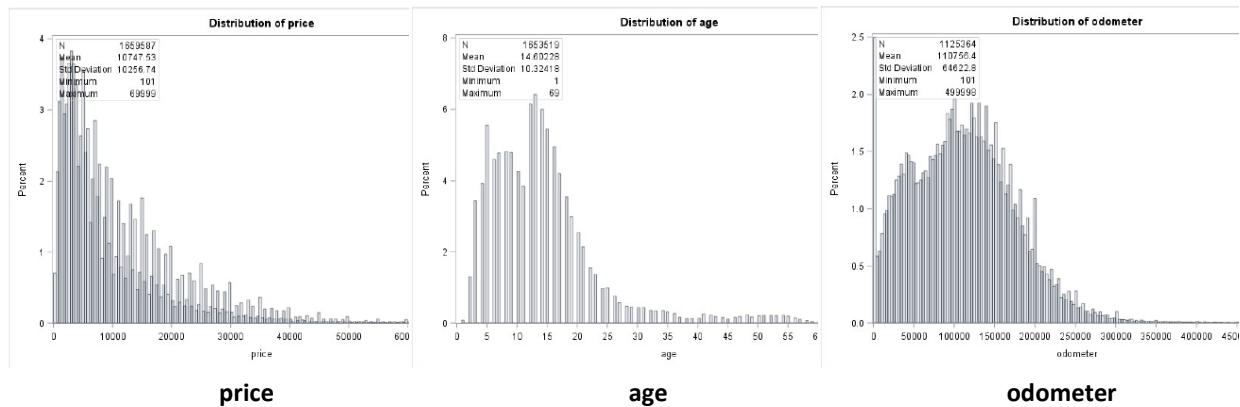
Missing Values				Missing Values			
Missing Value	Count	Percent Of	All Obs	Missing Value	Count	Percent Of	All Obs
.	6315	0.37	100.00	.	564054	32.74	100.00

price

year

odometer

To remove illogical values, we filtered rows with *price* between 100 and 70,000 and *odometer* between 500 and 500,000 and *year* more than 1950. The *odometer* has 32.74% missing values. And we converted *year* into *age* by subtracting it from 2020 and new distribution is as follows:



Since, all graphs are right-skewed, its better to normalize the data by taking the natural logarithm.

Among non-numerical variables, *make* is very long descriptive text and we will be ignoring that. Below are other variables frequency distributions using which we can find missing values share.

fuel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	9915	0.61	9915	0.61
diesel	114241	6.98	124156	7.59
electric	2181	0.13	126337	7.72
gas	1456237	89.00	1582574	96.72
hybrid	10553	0.64	1593127	97.37
other	43100	2.63	1636227	100.00

condition	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	666626	40.74	666626	40.74
excellent	422244	25.81	1088870	66.55
fair	69834	4.27	1158704	70.82
good	360895	22.06	1519599	92.87
like new	105124	6.42	1624723	99.30
new	6271	0.38	1630994	99.68
salvage	5233	0.32	1636227	100.00

title_status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	2515	0.15	2515	0.15
clean	1523614	93.12	1526129	93.27
lien	20845	1.27	1546974	94.55
missing	8870	0.54	1555844	95.09
parts onl	3558	0.22	1559402	95.30
rebuilt	46617	2.85	1606019	98.15
salvage	30208	1.85	1636227	100.00

cylinders	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	654185	39.98	654185	39.98
10 cylinders	4383	0.27	658568	40.25
12 cylinders	651	0.04	659219	40.29
3 cylinders	1700	0.10	660919	40.39
4 cylinders	281654	17.21	942573	57.61
5 cylinders	10025	0.61	952598	58.22
6 cylinders	346366	21.17	1298964	79.39
8 cylinders	303249	18.53	1602213	97.92
other	34014	2.08	1636227	100.00

transmission	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	8736	0.53	8736	0.53
automatic	1411711	86.28	1420447	86.81
manual	185549	11.34	1605996	98.15
other	30231	1.85	1636227	100.00

drive	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	620627	37.93	620627	37.93
4wd	429128	26.23	1049755	64.16
fwd	358030	21.88	1407785	86.04
rwd	228442	13.96	1636227	100.00

paint_color	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	652376	39.87	652376	39.87
black	191832	11.72	844208	51.59
blue	111105	6.79	955313	58.39
brown	25803	1.58	981116	59.96
custom	23860	1.46	1004976	61.42
green	40712	2.49	1045688	63.91
grey	96722	5.91	1142410	69.82
orange	6658	0.41	1149068	70.23
purple	3627	0.22	1152695	70.45
red	110222	6.74	1262917	77.18
silver	142611	8.72	1405528	85.90
white	220399	13.47	1625927	99.37
yellow	10300	0.63	1636227	100.00

type	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	658700	40.26	658700	40.26
SUV	237461	14.51	896161	54.77
bus	1952	0.12	898113	54.89
converti	29267	1.79	927380	56.68
coupe	67229	4.11	994609	60.79
hatchbac	37066	2.27	1031675	63.05
mini-van	24108	1.47	1055783	64.53
offroad	4712	0.29	1060495	64.81
other	21141	1.29	1081636	66.11
pickup	118086	7.22	1199722	73.32
sedan	260127	15.90	1459849	89.22
truck	127399	7.79	1587248	97.01
van	24427	1.49	1611675	98.50
wagon	24552	1.50	1636227	100.00

size	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	1066052	65.15	1066052	65.15
compact	89148	5.45	1155200	70.60
full-size	305165	18.65	1460365	89.25
mid-size	164189	10.03	1624554	99.29
sub-compact	11673	0.71	1636227	100.00

manufacturer	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	120818	7.38	120818	7.38
acura	17416	1.06	138234	8.45
alfa	73	0.00	138307	8.45
alfa-romeo	130	0.01	138437	8.46
aston	33	0.00	138470	8.46

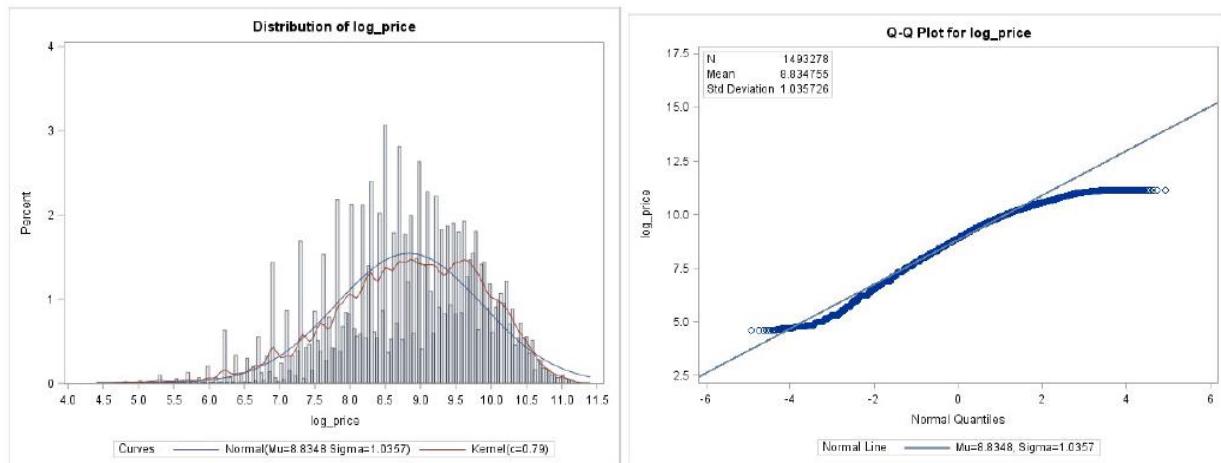
The *manufacturer* has 7% missing values but *size* has 65% missing values. As two-thirds of this variable is empty, we are ignoring the *size* variable. Also, *odometer*, *condition*, *cylinders*, *drive*, *type*, *paint_color* have 30-40% missing data. Hence, they are separated from other variables into a new dataset with smaller number of rows, upon deleting the missing rows mentioned as below:

Dataset	Variables	Rows	Rows %
With 4 category variables (4cat)	<i>price</i> , <i>age</i> , <i>log_price</i> , <i>log_age</i> , <i>fuel</i> , <i>title_status</i> , <i>transmission</i> , <i>manufacturer</i>	1493278	86.66 %
With 9 category variables (9cat)	<i>price</i> , <i>age</i> , <i>odometer</i> , <i>log_price</i> , <i>log_age</i> , <i>log_odometer</i> , <i>fuel</i> , <i>title_status</i> , <i>transmission</i> , <i>manufacturer</i> , <i>condition</i> , <i>cylinders</i> , <i>drive</i> , <i>type</i> , <i>paint_color</i>	441328	25.61 %

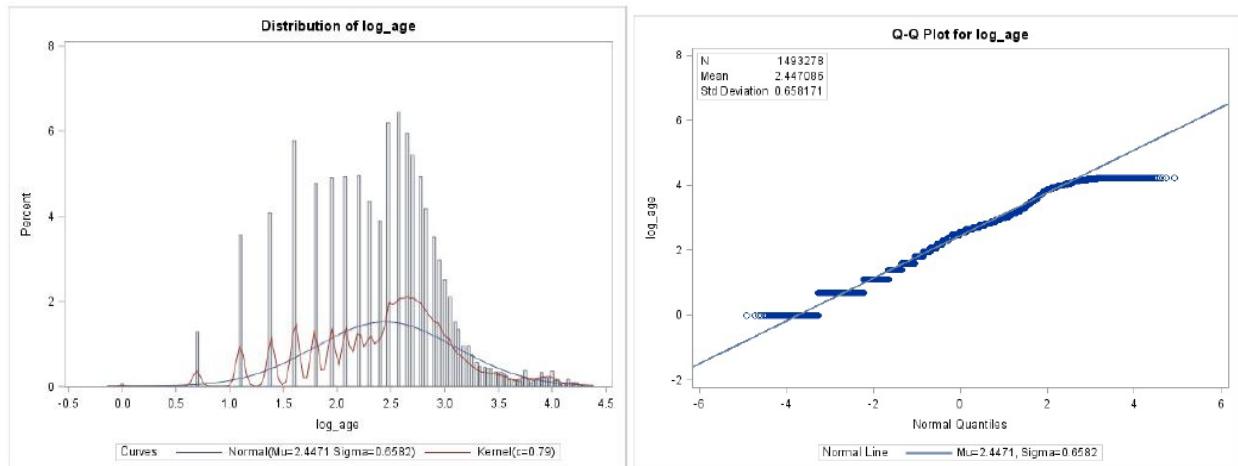
We will be using first table while using *fuel*, *title_status*, *transmission*, *manufacturer* as predictors and the second table while using *condition*, *cylinders*, *drive*, *type*, *paint_color* as the predictors.

2.2. Exploratory Data Analysis

After removing the missing value rows, *log_price* and *log_age* (in 4cat) are normally distributed.

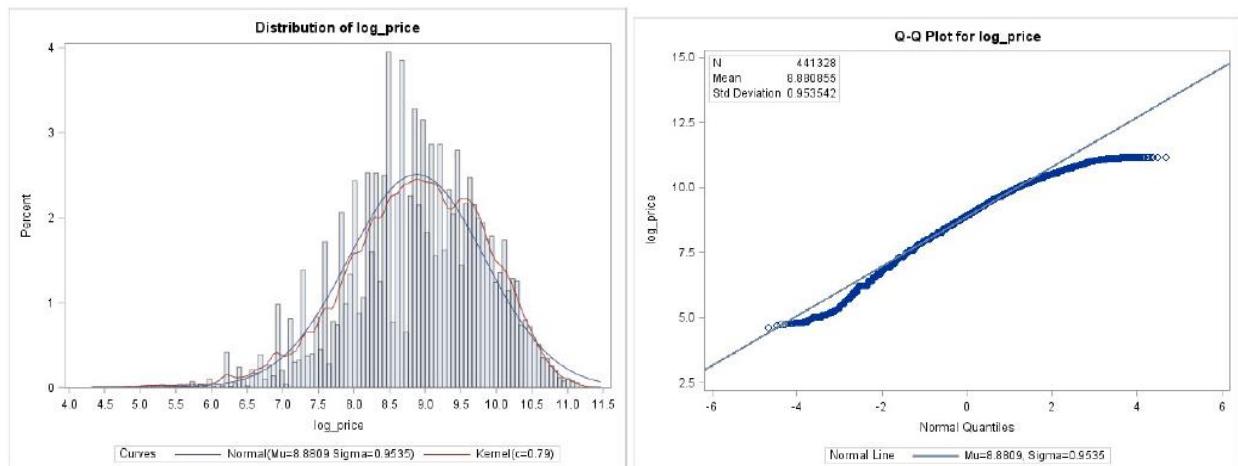


4cat Dataset - Distribution of *log_price*

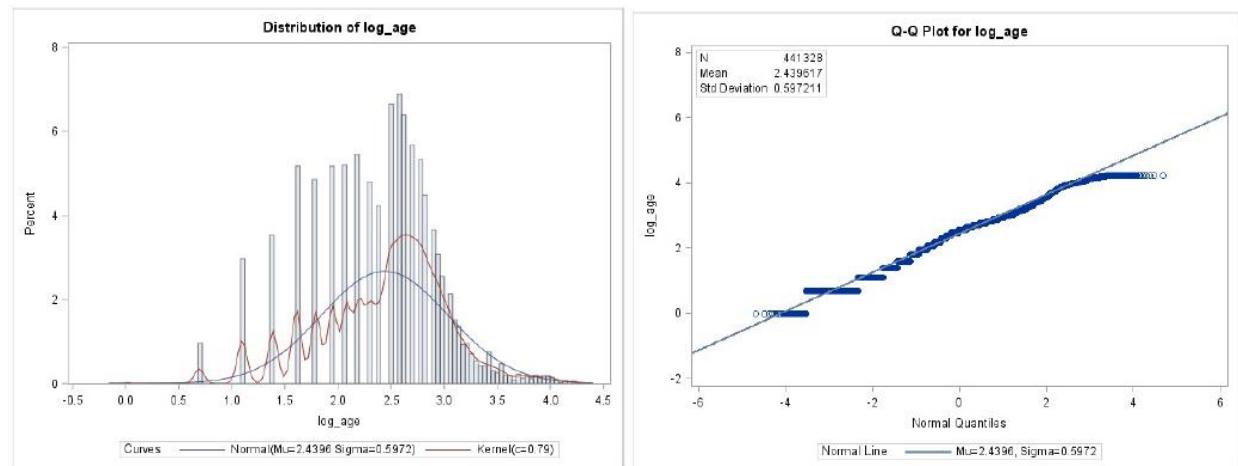


4cat Dataset - Distribution of log_age

After removing the missing value rows, *log_price* and *log_age* (in 9cat) are normally distributed.

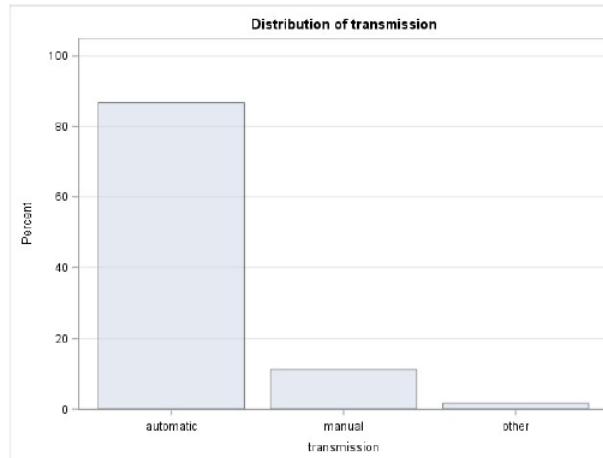


9cat Dataset - Distribution of log_price

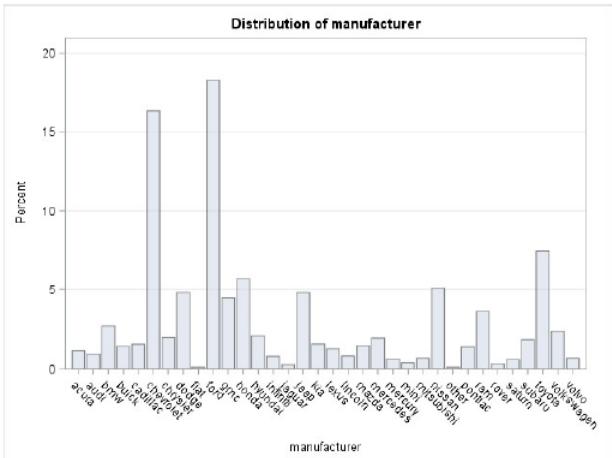


9cat Dataset - Distribution of log_age

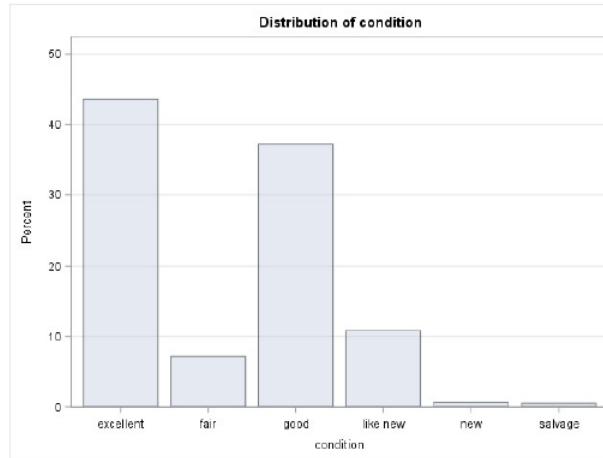
Below are the vehicles frequency distributions plots of all major categorical variables in the data



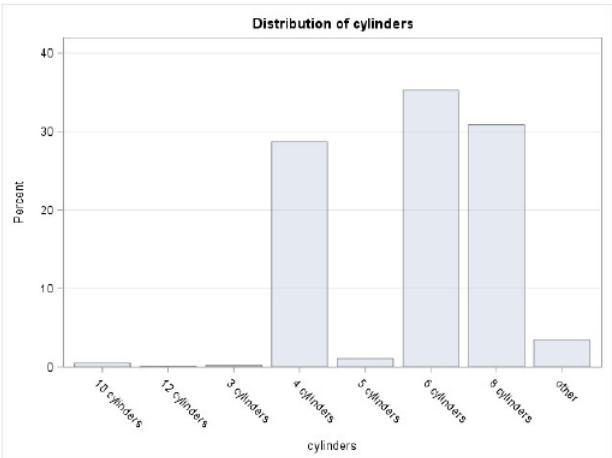
Most vehicles have automatic transmission



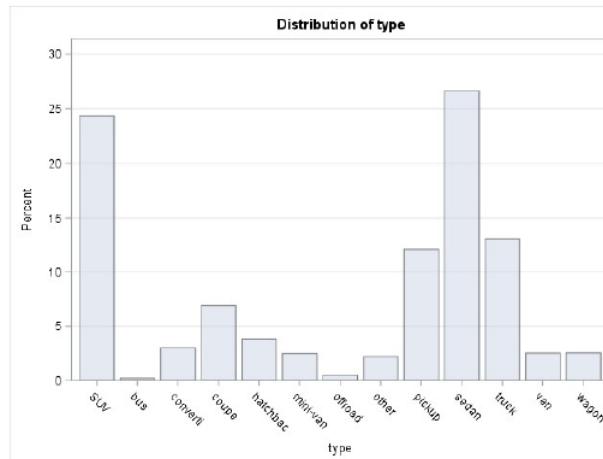
Top vehicles brands are ford and chevrolet



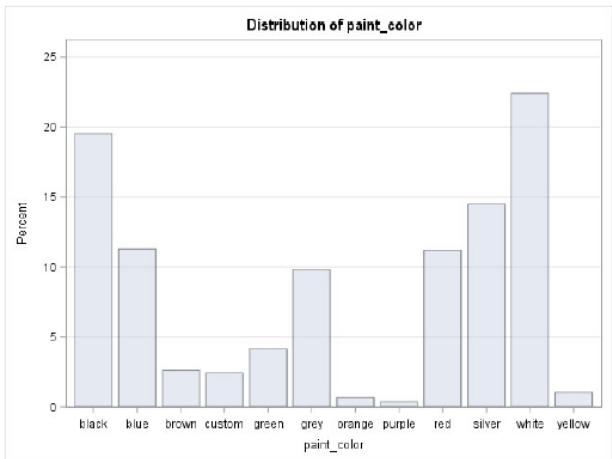
Most vehicles are in excellent or good condition



Most vehicles have either 4, 6 or 8 cylinders

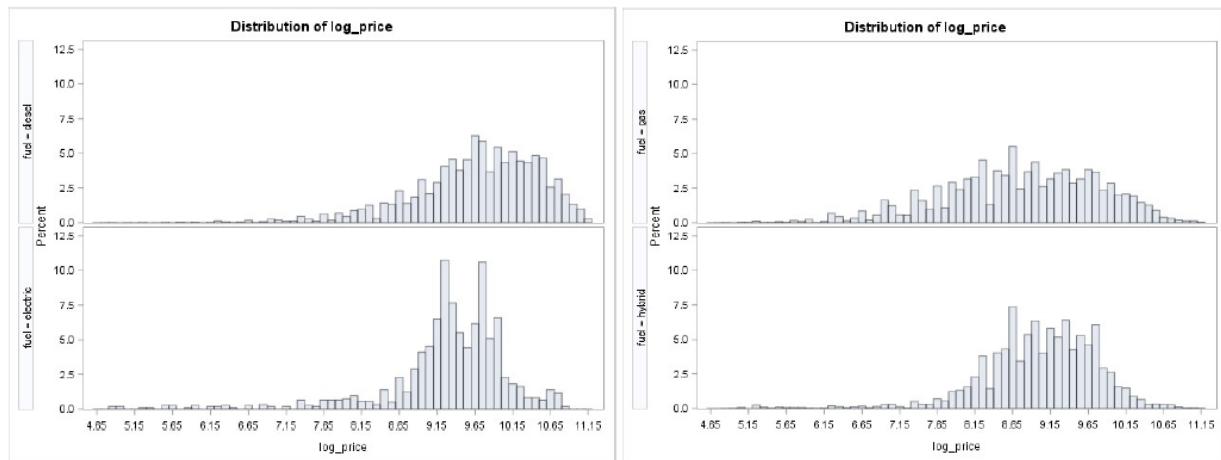


Top vehicles types are sedan, suv, truck, pickup

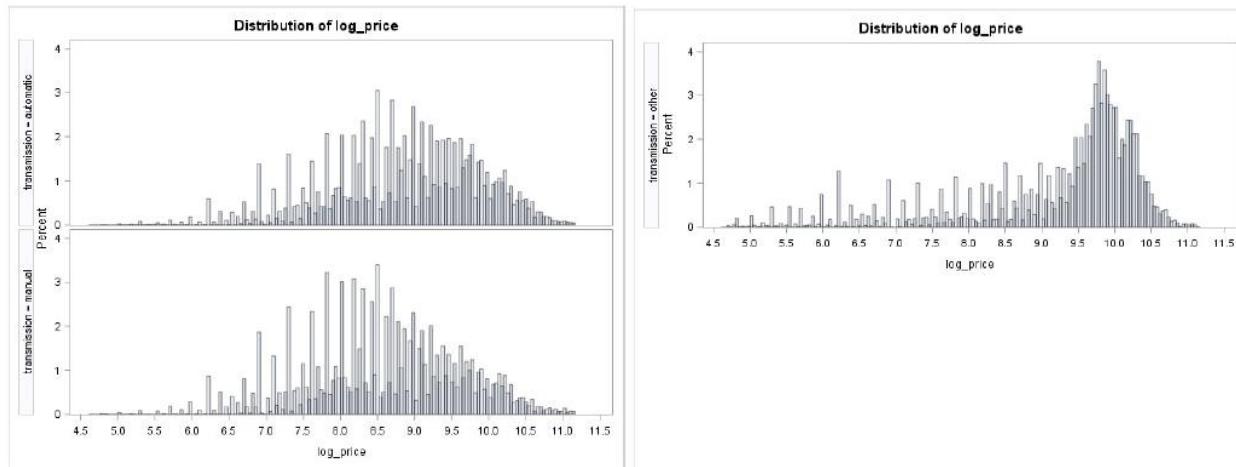


Most used vehicle colors are black, white, silver

Below are distribution plots of *log_price* split by some category variables like *fuel*, *transmission*

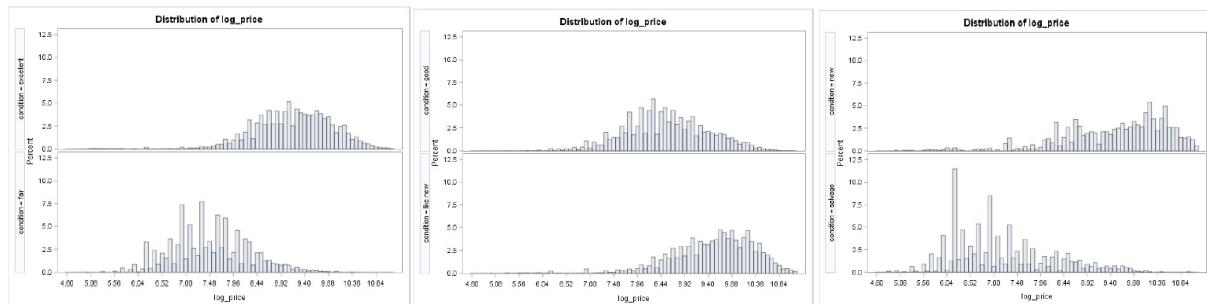


4cat Dataset - Variance of *log_price* is more when *fuel* is gas



4cat Dataset - *log_price* is more left skewed when transmission type is neither automatic nor manual

Below are the histogram distribution plots of *log_price* split by the category variable *condition*



left-top-excellent, left-down-fair, middle-top-good, middle-down-like new, right-top-new, right-down-salvage

As we expect, mean of *price* is higher when *condition* is excellent, like new or new. For salvage, the *price* is scattered all over with high variance as people are skeptical about salvaged vehicles.

2.3. Correlations among predictors

As we have continuous and categorical variables, we need to use Pearson linear correlation for continuous, chi-square test and Cramer for categorical, anova test for continuous and categorical. Pearson correlation coefficients between *price* and *age* in *4cat*, *price*, *age* and *odometer* in *9cat*

2 Variables:		log_price	log_age	3 Variables:			log_price	log_age	log_odometer
Pearson Correlation Coefficients, N = 1493278					Pearson Correlation Coefficients, N = 441328				
Prob > r under H0: Rho=0					Prob > r under H0: Rho=0				
		log_price	log_age			log_price	log_age	log_odometer	
log_price		1.00000	-0.57035 <.0001			1.00000	-0.60279 <.0001	-0.51291 <.0001	
log_age		-0.57035 <.0001	1.00000			-0.60279 <.0001	1.00000	0.55334 <.0001	
				log_odometer		-0.51291 <.0001	0.55334 <.0001	1.00000	

Mantel-Haenszel chi-square test and cramer's v score between categorical variables in *4cat* table

Statistics for Table of fuel by title_status

Statistic	DF	Value	Prob
Chi-Square	20	5781.8500	<.0001
Likelihood Ratio Chi-Square	20	6632.9706	<.0001
Mantel-Haenszel Chi-Square	1	26.4642	<.0001
Phi Coefficient		0.0622	
Contingency Coefficient		0.0621	
Cramer's V		0.0311	

Statistics for Table of fuel by transmission

Statistic	DF	Value	Prob
Chi-Square	8	13178.6336	<.0001
Likelihood Ratio Chi-Square	8	8073.7620	<.0001
Mantel-Haenszel Chi-Square	1	355.8534	<.0001
Phi Coefficient		0.0939	
Contingency Coefficient		0.0935	
Cramer's V		0.0664	

Statistics for Table of fuel by manufacturer

Statistic	DF	Value	Prob
Chi-Square	132	200966	<.0001
Likelihood Ratio Chi-Square	132	165398	<.0001
Mantel-Haenszel Chi-Square	1	12.59120	0.0004
Phi Coefficient		0.36685	
Contingency Coefficient		0.34441	
Cramer's V		0.18343	

Statistics for Table of title_status by transmission

Statistic	DF	Value	Prob
Chi-Square	10	11082.7679	<.0001
Likelihood Ratio Chi-Square	10	6420.7090	<.0001
Mantel-Haenszel Chi-Square	1	25.8547	<.0001
Phi Coefficient		0.0861	
Contingency Coefficient		0.0858	
Cramer's V		0.0609	

Statistics for Table of title_status by manufacturer

Statistic	DF	Value	Prob
Chi-Square	165	12775.3827	<.0001
Likelihood Ratio Chi-Square	165	11944.4375	<.0001
Mantel-Haenszel Chi-Square	1	155.8603	<.0001
Phi Coefficient		0.0925	
Contingency Coefficient		0.0921	
Cramer's V		0.0414	

Statistics for Table of transmission by manufacturer

Statistic	DF	Value	Prob
Chi-Square	66	76708.8443	<.0001
Likelihood Ratio Chi-Square	66	68250.3476	<.0001
Mantel-Haenszel Chi-Square	1	4326.6619	<.0001
Phi Coefficient		0.2266	
Contingency Coefficient		0.2210	
Cramer's V		0.1603	

Mantel-Haenszel chi-square test and Cramer's V score between categorical variables in *9cat* table

Statistics for Table of condition by cylinders

Statistic	DF	Value	Prob
Chi-Square	35	16092.8445	<.0001
Likelihood Ratio Chi-Square	35	18192.1653	<.0001
Mantel-Haenszel Chi-Square	1	1298.1592	<.0001
Phi Coefficient		0.1910	
Contingency Coefficient		0.1876	
Cramer's V		0.0854	

Statistics for Table of condition by drive

Statistic	DF	Value	Prob
Chi-Square	10	1283.1305	<.0001
Likelihood Ratio Chi-Square	10	1273.3594	<.0001
Mantel-Haenszel Chi-Square	1	238.2943	<.0001
Phi Coefficient		0.0539	
Contingency Coefficient		0.0538	
Cramer's V		0.0381	

Statistics for Table of condition by type

Statistic	DF	Value	Prob
Chi-Square	60	5529.9988	<.0001
Likelihood Ratio Chi-Square	60	5534.1236	<.0001
Mantel-Haenszel Chi-Square	1	238.2314	<.0001
Phi Coefficient		0.1119	
Contingency Coefficient		0.1112	
Cramer's V		0.0501	

Statistics for Table of condition by paint_color

Statistic	DF	Value	Prob
Chi-Square	55	7070.8815	<.0001
Likelihood Ratio Chi-Square	55	6807.0878	<.0001
Mantel-Haenszel Chi-Square	1	0.3827	0.5362
Phi Coefficient		0.1266	
Contingency Coefficient		0.1256	
Cramer's V		0.0566	

Statistics for Table of cylinders by drive

Statistic	DF	Value	Prob
Chi-Square	14	135371	<.0001
Likelihood Ratio Chi-Square	14	154404	<.0001
Mantel-Haenszel Chi-Square	1	1189	<.0001
Phi Coefficient		0.55384	
Contingency Coefficient		0.48449	
Cramer's V		0.39162	

Statistics for Table of cylinders by type

Statistic	DF	Value	Prob
Chi-Square	84	191688	<.0001
Likelihood Ratio Chi-Square	84	189312	<.0001
Mantel-Haenszel Chi-Square	1	85.97904	<.0001
Phi Coefficient		0.65905	
Contingency Coefficient		0.55029	
Cramer's V		0.24910	

Statistics for Table of cylinders by paint_color

Statistics for Table of drive by type

Statistic	DF	Value	Prob
Chi-Square	77	13034.0112	<.0001
Likelihood Ratio Chi-Square	77	13515.4863	<.0001
Mantel-Haenszel Chi-Square	1	163.9275	<.0001
Phi Coefficient		0.1719	
Contingency Coefficient		0.1694	
Cramer's V		0.0650	

Statistic	DF	Value	Prob
Chi-Square	24	266314	<.0001
Likelihood Ratio Chi-Square	24	290685	<.0001
Mantel-Haenszel Chi-Square	1	9507	<.0001
Phi Coefficient		0.77681	
Contingency Coefficient		0.61347	
Cramer's V		0.54929	

Statistics for Table of drive by paint_color

Statistics for Table of type by paint_color

Statistic	DF	Value	Prob
Chi-Square	22	10150.7946	<.0001
Likelihood Ratio Chi-Square	22	9928.4219	<.0001
Mantel-Haenszel Chi-Square	1	871.7726	<.0001
Phi Coefficient		0.1517	
Contingency Coefficient		0.1499	
Cramer's V		0.1072	

Statistic	DF	Value	Prob
Chi-Square	132	36600.8363	<.0001
Likelihood Ratio Chi-Square	132	33884.4508	<.0001
Mantel-Haenszel Chi-Square	1	1231.0941	<.0001
Phi Coefficient		0.2880	
Contingency Coefficient		0.2767	
Cramer's V		0.0868	

Anova between *log_age*, *log_odometer* and other categorical variables in *4cat* and *9cat* datasets

Dependent Variable: log_age						Dependent Variable: log_age					
Source	DF	Type I SS	Mean Square	F Value	Pr > F	Source	DF	Type I SS	Mean Square	F Value	Pr > F
fuel	4	5239.610384	1309.902596	3048.54	<.0001	title_status	5	12562.24688	2512.44938	5914.73	<.0001
fuel	4	5239.610384	1309.902596	3048.54	<.0001	title_status	5	12562.24688	2512.44938	5914.73	<.0001
Dependent Variable: log_age						Dependent Variable: log_age					
Source	DF	Type I SS	Mean Square	F Value	Pr > F	Source	DF	Type I SS	Mean Square	F Value	Pr > F
transmission	2	34783.62185	17391.81092	42429.7	<.0001	manufacturer	33	32668.86096	989.96548	2406.79	<.0001
transmission	2	34783.62185	17391.81092	42429.7	<.0001	manufacturer	33	32668.86096	989.96548	2406.79	<.0001
Dependent Variable: log_age						Dependent Variable: log_odometer					
Source	DF	Type I SS	Mean Square	F Value	Pr > F	Source	DF	Type I SS	Mean Square	F Value	Pr > F
condition	5	29575.30171	5915.06034	20421.4	<.0001	condition	5	41734.70179	8346.94036	17889.3	<.0001
condition	5	29575.30171	5915.06034	20421.4	<.0001	condition	5	41734.70179	8346.94036	17889.3	<.0001
Dependent Variable: log_age						Dependent Variable: log_odometer					
Source	DF	Type I SS	Mean Square	F Value	Pr > F	Source	DF	Type I SS	Mean Square	F Value	Pr > F
cylinders	7	9055.787797	1293.683971	3848.57	<.0001	cylinders	7	4004.745567	572.106510	1036.27	<.0001
cylinders	7	9055.787797	1293.683971	3848.57	<.0001	cylinders	7	4004.745567	572.106510	1036.27	<.0001
Dependent Variable: log_age						Dependent Variable: log_odometer					
Source	DF	Type I SS	Mean Square	F Value	Pr > F	Source	DF	Type I SS	Mean Square	F Value	Pr > F
drive	2	8556.166704	4278.083352	12684.3	<.0001	drive	2	1076.209379	538.104689	963.11	<.0001
drive	2	8556.166704	4278.083352	12684.3	<.0001	drive	2	1076.209379	538.104689	963.11	<.0001
Dependent Variable: log_age						Dependent Variable: log_odometer					
Source	DF	Type I SS	Mean Square	F Value	Pr > F	Source	DF	Type I SS	Mean Square	F Value	Pr > F
type	12	6407.875171	533.989598	1560.69	<.0001	type	12	5151.924153	429.327013	781.32	<.0001
type	12	6407.875171	533.989598	1560.69	<.0001	type	12	5151.924153	429.327013	781.32	<.0001
Dependent Variable: log_age						Dependent Variable: log_odometer					
Source	DF	Type I SS	Mean Square	F Value	Pr > F	Source	DF	Type I SS	Mean Square	F Value	Pr > F
paint_color	11	7961.516914	723.774265	2137.36	<.0001	paint_color	11	2433.773670	221.252152	398.19	<.0001
paint_color	11	7961.516914	723.774265	2137.36	<.0001	paint_color	11	2433.773670	221.252152	398.19	<.0001

3. Regression Models

We are building the following 4 regression models listed below using the GLM procedure in SAS.

Model 1 - *log_price* as the dependent and *fuel*, *transmission*, *title_status* as predictors on *4cat*

Model 2 - *log_price* as the dependent and *manufacturer* as predictors on *4cat*

Model 3 - *log_price* as the dependent and *condition*, *cylinders*, *drive* as predictors on *9cat*

Model 4 - *log_price* as the dependent and *type*, *paint_color* as predictors on *9cat*

3.1. Pricing Model with Fuel, Transmission, Title

Dependent Variable: log_price					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	615985.178	51332.098	77749.1	<.0001
Error	1.49E6	985894.560	0.660		
Corrected Total	1.49E6	1601879.738			
R-Square	Coeff Var	Root MSE	log_price Mean		
0.384539	9.197130	0.812544	8.834755		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
log_age	1	521096.7980	521096.7980	789269	<.0001
fuel	4	74424.0306	18606.0076	28181.2	<.0001
title_status	5	17138.0637	3427.6127	5191.56	<.0001
transmission	2	3326.2856	1663.1428	2519.05	<.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
log_age	1	481193.7959	481193.7959	728830	<.0001
fuel	4	71929.7263	17982.4316	27236.7	<.0001
title_status	5	17401.7723	3480.3545	5271.45	<.0001
transmission	2	3326.2856	1663.1428	2519.05	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	9.612392773	B	0.01664976	577.33 <.0001
log_age	-0.898533306	0.00105250	-853.72 <.0001	
fuel diesel	0.641082792	B	0.00498403	128.63 <.0001
fuel electric	-0.316091499	B	0.02366759	-13.36 <.0001
fuel gas	-0.244395652	B	0.00429852	-56.86 <.0001
fuel hybrid	-0.252637438	B	0.00912867	-27.68 <.0001
fuel other	0.000000000	B	.	.
title_status clean	1.652348549	B	0.01556788	106.14 <.0001
title_status lien	1.934548839	B	0.01663475	116.30 <.0001
title_status missing	0.977651556	B	0.01834601	53.29 <.0001
title_status rebuilt	1.481138926	B	0.01604784	92.30 <.0001
title_status salvage	1.255873236	B	0.01630047	77.05 <.0001
title_status parts onl	0.000000000	B	.	.
transmission automatic	-0.052636951	B	0.00528566	-9.96 <.0001
transmission manual	0.103643411	B	0.00567423	18.27 <.0001
transmission other	0.000000000	B	.	.

All estimates are found to be significant at 1% and R^2 is 38%. From the above, we can interpret:

When the age of the vehicle increases by 10% then the sale price of the vehicle decreases by 9%.
If the vehicle fuel type is diesel, then price is more by 64% when compared to “other” fuel type.
If the vehicle fuel type is electric, then price is less by 31% when compared to “other” fuel type.
If the vehicle fuel type is gas, then price is less by 24% when compared to the “other” fuel type.
If the vehicle fuel type is hybrid, then price is less by 25% when compared with “other” fuel type.
If the vehicle title status is clean, then price is more by 165% when compared to “parts only” title.
If the vehicle title status is lien, then price is less by 193% when compared to a “parts only” title.
If the vehicle title status is missing, then price is less by 98% when compared to “parts only” title.
If the vehicle title status is rebuilt, then price is less by 148% when compared to “parts only” title.
If the vehicle title status is salvage then price is less by 125% when compared to “parts only” title.
If vehicle transmission is automatic, price is less by 5% when compared to “other” transmission.
If vehicle transmission is manual, price is more by 10% when compared to “other” transmission.

3.2. Pricing Model with Manufacturer

Dependent Variable: log_price					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	34	619030.784	18206.788	27661.6	<.0001
Error	1.49E6	982848.954	0.658		
Corrected Total	1.49E6	1601879.738			
R-Square	Coeff Var	Root MSE	log_price Mean		
0.386440	9.182980	0.811294	8.834755		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
log_age	1	521096.7980	521096.7980	791703	<.0001
manufacturer	33	97933.9858	2967.6965	4508.82	<.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
log_age	1	510626.1385	510626.1385	775795	<.0001
manufacturer	33	97933.9858	2967.6965	4508.82	<.0001
Parameter	Estimate		Standard Error	t Value	Pr > t
Intercept	11.67998703		B	0.01934005	603.93 <.0001
log_age	-0.91179136		B	0.00103519	-880.79 <.0001
manufacturer acura	-0.86742663		B	0.02006342	-43.23 <.0001
manufacturer audi	-0.55984240		B	0.02026168	-27.63 <.0001
manufacturer bmw	-0.46365622		B	0.01950542	-23.77 <.0001
manufacturer buick	-0.89020379		B	0.01986936	-44.80 <.0001
manufacturer cadillac	-0.47306003		B	0.01979686	-23.90 <.0001
manufacturer chevrolet	-0.47596619		B	0.01915690	-24.85 <.0001
manufacturer chrysler	-1.02437091		B	0.01965864	-52.11 <.0001
manufacturer dodge	-0.78151840		B	0.01933049	-40.43 <.0001
manufacturer fiat	-0.94782378		B	0.02682164	-35.34 <.0001
manufacturer ford	-0.49559572		B	0.01915195	-25.88 <.0001
manufacturer gmc	-0.28737663		B	0.01934987	-14.85 <.0001
manufacturer honda	-0.94177631		B	0.01929223	-48.82 <.0001
manufacturer hyundai	-1.04862674		B	0.01964236	-53.39 <.0001
manufacturer infiniti	-0.57187957		B	0.02048624	-27.92 <.0001
manufacturer jaguar	-0.49632065		B	0.02303129	-21.55 <.0001
manufacturer jeep	-0.43443853		B	0.01933194	-22.47 <.0001
manufacturer kia	-1.05157563		B	0.01982138	-53.05 <.0001
manufacturer lexus	-0.39497382		B	0.01996102	-19.79 <.0001
manufacturer lincoln	-0.68051203		B	0.02044438	-33.29 <.0001
manufacturer mazda	-0.95731289		B	0.01984601	-48.24 <.0001
manufacturer mercedes	-0.32901687		B	0.01966647	-16.73 <.0001
manufacturer mercury	-1.10232235		B	0.02078932	-53.02 <.0001
manufacturer mini	-0.70779579		B	0.02171974	-32.59 <.0001
manufacturer mitsubishi	-1.07686045		B	0.02069795	-52.03 <.0001
manufacturer nissan	-0.94655933		B	0.01932818	-48.97 <.0001
manufacturer pontiac	-0.92643823		B	0.01987464	-46.61 <.0001
manufacturer ram	-0.03984956		B	0.01941631	-2.05 0.0401
manufacturer rover	-0.13645735		B	0.02223451	-6.14 <.0001
manufacturer saturn	-1.41639493		B	0.02092782	-67.68 <.0001
manufacturer subaru	-0.72266733		B	0.01970657	-36.67 <.0001
manufacturer toyota	-0.61944080		B	0.01924773	-32.18 <.0001
manufacturer volkswagen	-0.81911298		B	0.01956001	-41.88 <.0001
manufacturer volvo	-0.91483996		B	0.02067453	-44.25 <.0001
manufacturer other	0.00000000		B	.	.

The conclusions about relation between price and age still hold valid in this regression model too. All estimates are found to be significant at 5% and R^2 is 38%. From the above, we can interpret:

Except Ram, all estimates are significant at 1%. Cheapest one is Saturn and costliest one is Ram. Other cheap vehicles include manufacturers like Mercury, Mitsubishi, Kia, Hyundai, and Chrysler. Other costly vehicles include manufacturers like Rover, GMC, Mercedes, Lexus, Jeep and BMW.

3.3. Pricing Model with Condition, Cylinders and Drive

Dependent Variable: log_price						
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F	
Model	16	253396.0195	15837.2512	47263.2	<.0001	
Error	441311	147877.4001	0.3351			
Corrected Total	441327	401273.4197				
R-Square	Coeff Var	Root MSE	log_price Mean			
0.631480	6.518141	0.578867	8.880855			
Source	DF	Type I SS	Mean Square	F Value	Pr > F	
log_age	1	145804.0439	145804.0439	435123	<.0001	
log_odometer	1	18605.6277	18605.6277	55524.8	<.0001	
condition	5	32753.2641	6550.6528	19549.1	<.0001	
cylinders	7	40269.5719	5752.7960	17168.1	<.0001	
drive	2	15963.5120	7981.7560	23820.0	<.0001	
Source	DF	Type III SS	Mean Square	F Value	Pr > F	
log_age	1	43881.38498	43881.38498	130955	<.0001	
log_odometer	1	11849.27691	11849.27691	35361.8	<.0001	
condition	5	29544.41793	5908.88359	17633.9	<.0001	
cylinders	7	11866.83760	1695.26251	5059.18	<.0001	
drive	2	15963.51200	7981.75600	23820.0	<.0001	

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	12.82102336	B	0.02155361	594.84 <.0001
log_age	-0.70050183		0.00193574	-361.88 <.0001
log_odometer	-0.27698836		0.00147297	-188.05 <.0001
condition excellent	1.46748612	B	0.01405023	104.45 <.0001
condition fair	0.39144406	B	0.01441897	27.15 <.0001
condition good	1.16247241	B	0.01404423	82.77 <.0001
condition like new	1.44691340	B	0.01427400	101.37 <.0001
condition new	1.12326102	B	0.01975394	56.86 <.0001
condition salvage	0.00000000	B	.	.
cylinders 10 cylinders	0.31015242	B	0.01329205	23.33 <.0001
cylinders 12 cylinders	0.38009359	B	0.03495810	10.87 <.0001
cylinders 3 cylinders	-0.61063134	B	0.02568106	-23.78 <.0001
cylinders 4 cylinders	-0.32594400	B	0.00540527	-60.30 <.0001
cylinders 5 cylinders	-0.26402011	B	0.01001218	-26.37 <.0001
cylinders 6 cylinders	-0.18016605	B	0.00533035	-33.80 <.0001
cylinders 8 cylinders	0.15128080	B	0.00544797	27.77 <.0001
cylinders other	0.00000000	B	.	.
drive fwd	-0.49930729	B	0.00229314	-217.74 <.0001
drive rwd	-0.13494144	B	0.00238495	-56.58 <.0001
drive 4wd	0.00000000	B	.	.

All estimates are found to be significant at 1% and R^2 is 63%. From the above, we can interpret:

When the age of the vehicle increases by 10% then the sale price of the vehicle decreases by 7%.
When odometer reading increases by 10% then the sale price of the vehicle decreases by 2.77%.
If the vehicle condition is excellent, price is more by 146% when compared to a salvaged vehicle.
If the vehicle condition is fair, then price is more by 39% when compare with a salvaged vehicle.
If the vehicle condition is good, then price is more by 116% when compared to salvaged vehicle.
If the vehicle condition is like new, price is more by 144% when compared with a salvaged vehicle.
If the vehicle condition is new, then price is more by 112% when compared with salvaged vehicle.
If the vehicle has 3 cylinders, then price is less by 61% when compared to “other” cylinders count.
If the vehicle has 4 cylinders, then price is less by 32% when compared to “other” cylinders count.
If the vehicle has 5 cylinders, then price is less by 26% when compared to “other” cylinders count.
If the vehicle has 6 cylinders, then price is less by 18% when compared to “other” cylinders count.
If vehicle has 8 cylinders, then price is more by 15% when compared to “other” cylinders count.
If vehicle has 10 cylinders, then price is more by 31% when compared to “other” cylinders count.
If vehicle has 12 cylinders, then price is more by 38% when compared to “other” cylinders count.
If vehicle has front wheel drive, price is less by 50% when compared to four wheel drive vehicle.
If vehicle has rear wheel drive, price is less by 13% when compared to a four wheel drive vehicle.
As number of cylinders goes up, vehicle gets costly. Front wheel drive vehicles are more cheaper.

3.4. Price Model with Type and Color

Dependent Variable: log_price					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	210097.4023	8403.8961	19399.2	<.0001
Error	441302	191176.0173	0.4332		
Corrected Total	441327	401273.4197			
R-Square	Coeff Var	Root MSE	log_price Mean		
0.523577	7.411293	0.658186	8.880855		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
log_age	1	145804.0439	145804.0439	336567	<.0001
log_odometer	1	18605.6277	18605.6277	42948.4	<.0001
type	12	43969.8856	3664.1571	8458.17	<.0001
paint_color	11	1717.8451	156.1677	360.49	<.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
log_age	1	56460.50868	56460.50868	130331	<.0001
log_odometer	1	17356.60834	17356.60834	40065.2	<.0001
type	12	42323.63850	3526.96987	8141.50	<.0001
paint_color	11	1717.84510	156.16774	360.49	<.0001
Parameter	Estimate		Standard Error	t Value	Pr > t
Intercept	14.68913103		B	0.02042256	719.26 <.0001
log_age	-0.77123760		B	0.00213631	-361.01 <.0001
log_odometer	-0.33074449		B	0.00165238	-200.16 <.0001
type SUV	-0.01912623	B	0.01131892	-1.69	0.0911
type bus	0.37551053	B	0.03102759	12.10	<.0001
type converti	0.26779246	B	0.01251803	21.39	<.0001
type coupe	-0.03582146	B	0.01171610	-3.06	0.0022
type hatchbac	-0.43384369	B	0.01218340	-35.61	<.0001
type mini-van	-0.37052350	B	0.01257901	-29.46	<.0001
type offroad	0.37621092	B	0.01742038	21.60	<.0001
type pickup	0.33788954	B	0.01153014	29.30	<.0001
type sedan	-0.42728121	B	0.01129058	-37.84	<.0001
type truck	0.40331506	B	0.01142219	35.31	<.0001
type van	-0.10845854	B	0.01273539	-8.52	<.0001
type wagon	-0.29423580	B	0.01300717	-22.62	<.0001
type other	0.00000000	B		-	-
paint_color black	0.00435113	B	0.00681937	0.64	0.5234
paint_color blue	-0.11978805	B	0.00703158	-17.04	<.0001
paint_color brown	-0.13654849	B	0.00856618	-15.94	<.0001
paint_color green	-0.24409450	B	0.00799297	-30.54	<.0001
paint_color grey	-0.04522703	B	0.00706922	-6.40	<.0001
paint_color orange	0.00812714	B	0.01430908	0.57	0.5701
paint_color purple	-0.23734662	B	0.01707889	-13.90	<.0001
paint_color red	-0.12058141	B	0.00704129	-17.12	<.0001
paint_color silver	-0.08364177	B	0.00689901	-12.12	<.0001
paint_color white	-0.02182426	B	0.00677128	-3.22	0.0013
paint_color yellow	0.08264198	B	0.01258758	6.57	<.0001
paint_color custom	0.00000000	B		-	-

All the estimates except SUV type, Black and Orange colors, are significant at 1% and R^2 is 52%.

When the age of the vehicle increases by 10% then the sale price of the vehicle decreases by 8%. When odometer reading increases by 10% then the sale price of the vehicle decreases by 3.31%. If vehicle type is Bus, Convertible, Offroad, Pickup, Truck then sale price is more than “other” type. If vehicle is SUV, Coupe, Hatchback, Mini-Van, Sedan, Van, Wagon, price is less than “other” type. All the above vehicle colors except for Yellow color have prices less than Custom colored vehicles. Purple and Green color vehicles have least sale price when compared to Custom colored vehicles.

4. Conclusion

According to our regression results, a vehicle with below attributes fetches best possible price:

A Pickup, Truck or Offroad type vehicle that is in New or Excellent condition with Diesel fuel type and Clean title status, having Manual transmission, more than 6 cylinders and Four wheel drive, manufactured by Ram, Rover, GMC, Mercedes, Lexus, Jeep or BMW will get the highest sale price.

```
/* clearing log,output,results*/
dm 'clear log'; dm 'clear output'; dm 'odsresults; clear';
/* assigning library */
libname data "E:\pcg180000\final"; title;

/* DATA IMPORT */
proc import out = data.vehicles_raw
    datafile = "E:\pcg180000\final\craigslistVehiclesFull.csv"
    dbms = tab replace;
    delimiter = ",";
    getnames = yes;
    datarow = 2;
run;

/* copying data to work library */
proc datasets; copy in = data out = work; select vehicles_raw; run;

/* DATA CLEANING */
proc contents data = vehicles_raw varnum; run;

/* changing formats for numeric variables */
data vehicles_raw;
    set vehicles_raw;
    price_new = input(price,7.);
    year_new = input(year,6.);
    odometer_new = input(odometer,8.);
    drop price;
    drop year;
    drop odometer;
    rename price_new = price;
    rename year_new = year;
    rename odometer_new = odometer;
run;

/* numeric variables distribution */
proc univariate data = vehicles_raw; var price year odometer; histogram; inset n mean std min max; run

/* removing illogical values */
proc sql;
    create table vehicles_tmp as
    select price,year,fuel,title_status,transmission,manufacturer,
           odometer,condition,cylinders,drive,size,type,paint_color
    from vehicles_raw
    where (year > 1950 or year = .)
    and ((price > 100 and price < 70000) or (price = .))
    and ((odometer > 500 and odometer < 500000) or (odometer = .));
quit;

/* converting year into age */
data vehicles_tmp;
    set vehicles_tmp;
    age = 2020-year;
    drop year;
run;

/* numeric variable logarithms */
```

```

data vehicles_tmp;
  set vehicles_tmp;
  log_price = log(price);
  log_age = log(age);
  log_odometer = log(odometer);
run;

/* numeric variables distribution */
proc univariate data = vehicles_tmp; var price age odometer; histogram; inset n mean std min max; run;
proc univariate data = vehicles_tmp; var log_price log_age log_odometer; histogram; inset n mean std m;

/* categorical variables frequency */
proc freq data = vehicles_tmp;
  tables fuel title_status transmission manufacturer condition cylinders drive size type paint_color;
run;

/* correcting errors and categorizing less frequent manufacturers as other */
proc sql; update vehicles_tmp set manufacturer = 'chevrolet' where manufacturer = 'chev'; quit;
proc sql; update vehicles_tmp set manufacturer = 'chevrolet' where manufacturer = 'chevy'; quit;
proc sql; update vehicles_tmp set manufacturer = 'infiniti' where manufacturer = 'infinity'; quit;
proc sql; update vehicles_tmp set manufacturer = 'mercedes' where manufacturer = 'mercedes-be'; quit;
proc sql; update vehicles_tmp set manufacturer = 'mercedes' where manufacturer = 'mercedesben'; quit;
proc sql; update vehicles_tmp set manufacturer = 'volkswagen' where manufacturer = 'vw'; quit;
proc sql; update vehicles_tmp set manufacturer = 'other'
  where manufacturer in ('alfa','alfa-romeo','aston','aston-marti','datsun','ferrari','harley-davidson',
                         'hennessey','landrover','land rover','morgan','noble','porsche','renault','subaru','toyota','volvo')
quit;

/* dataset with price,age and 4 categorical variables */
proc sql;
  create table vehicles_4cat as
  select price,age,log_price,log_age,fuel,title_status,transmission,manufacturer from vehicles_tmp;
quit;

/* dataset with price,age,odometer and 9 categorical variables */
proc sql;
  create table vehicles_9cat as
  select price,age,odometer,log_price,log_age,log_odometer,
         fuel,title_status,transmission,manufacturer,
         condition,cylinders,drive,type,paint_color
  from vehicles_tmp;
quit;

/* removing all rows with missing values */
data vehicles_4cat; set vehicles_4cat; if cmiss(of _all_) then delete; run;
data vehicles_9cat; set vehicles_9cat; if cmiss(of _all_) then delete; run;

/* backup data to own library */
proc datasets; copy in = work out = data; select vehicles_4cat vehicles_9cat; run;
proc export data = vehicles_4cat outfile = "E:\pcg180000\final\vehicles_4cat.csv" dbms = csv replace;
proc export data = vehicles_9cat outfile = "E:\pcg180000\final\vehicles_9cat.csv" dbms = csv replace;

/* EXPLORATORY DATA ANALYSIS */
proc univariate data = vehicles_4cat;
  var log_price log_age;
  histogram / normal kernel;

```

```

qqplot / normal(mu=est sigma=est);
inset n mean std; run;
proc univariate data = vehicles_9cat;
var log_price log_age log_odometer;
histogram / normal kernel;
qqplot / normal(mu=est sigma=est);
inset n mean std; run;

/* price distribution by categories */
proc univariate data = vehicles_4cat; class fuel; var log_price; histogram; run;
proc univariate data = vehicles_4cat; class title_status; var log_price; histogram; run;
proc univariate data = vehicles_4cat; class transmission; var log_price; histogram; run;

proc univariate data = vehicles_9cat; class condition; var log_price; histogram; run;
proc univariate data = vehicles_9cat; class cylinders; var log_price; histogram; run;
proc univariate data = vehicles_9cat; class drive; var log_price; histogram; run;

proc freq data = vehicles_tmp;
tables fuel title_status transmission manufacturer condition cylinders drive type paint_color
      / plots(only)=freqplot(scale=percent);
run;

/* INDEPENDENT VARIABLE CORRELATIONS */
/* pearson correlations */
proc corr data = vehicles_4cat nosimple; var log_price log_age; run;
proc corr data = vehicles_9cat nosimple; var log_price log_age log_odometer; run;

/* chi square correlations */
proc freq data = vehicles_4cat; tables fuel*(title_status transmission manufacturer) / chisq; run;
proc freq data = vehicles_4cat; tables title_status*(transmission manufacturer) / chisq; run;
proc freq data = vehicles_4cat; tables transmission*manufacturer / chisq; run;

proc freq data = vehicles_9cat; tables condition*(cylinders drive type paint_color) / chisq; run;
proc freq data = vehicles_9cat; tables cylinders*(drive type paint_color) / chisq; run;
proc freq data = vehicles_9cat; tables drive*(type paint_color) / chisq; run;
proc freq data = vehicles_9cat; tables type*paint_color / chisq; run;

/* anova test between age,odometer and others */
proc glm data=vehicles_4cat; class fuel; model log_age=fuel; lsmeans fuel/adjust=tukey; run;
proc glm data=vehicles_4cat; class title_status; model log_age=title_status; lsmeans title_status/adjust=tukey;
proc glm data=vehicles_4cat; class transmission; model log_age=transmission; lsmeans transmission/adjust=tukey;
proc glm data=vehicles_4cat; class manufacturer; model log_age=manufacturer; lsmeans manufacturer/adjust=tukey;

proc glm data=vehicles_9cat; class condition; model log_age=condition; lsmeans condition/adjust=tukey;
proc glm data=vehicles_9cat; class cylinders; model log_age=cylinders; lsmeans cylinders/adjust=tukey;
proc glm data=vehicles_9cat; class drive; model log_age=drive; lsmeans drive/adjust=tukey; run;
proc glm data=vehicles_9cat; class type; model log_age=type; lsmeans type/adjust=tukey; run;
proc glm data=vehicles_9cat; class paint_color; model log_age=paint_color; lsmeans paint_color/adjust=tukey;

proc glm data=vehicles_9cat; class condition; model log_odometer=condition; lsmeans condition/adjust=tukey;
proc glm data=vehicles_9cat; class cylinders; model log_odometer=cylinders; lsmeans cylinders/adjust=tukey;
proc glm data=vehicles_9cat; class drive; model log_odometer=drive; lsmeans drive/adjust=tukey; run;
proc glm data=vehicles_9cat; class type; model log_odometer=type; lsmeans type/adjust=tukey; run;
proc glm data=vehicles_9cat; class paint_color; model log_odometer=paint_color; lsmeans paint_color/adjust=tukey;

/* GENERALIZED LINEAR REGRESSION */

```

```
ods graphics on;
proc glm data = vehicles_4cat plots(only)=(contourfit);
  class fuel(ref='other') title_status(ref='parts onl') transmission(ref='other');
  model log_price = log_age fuel title_status transmission / solution;
run;
proc glm data = vehicles_4cat plots(only)=(contourfit);
  class manufacturer(ref='other');
  model log_price = log_age manufacturer / solution;
run;
proc glm data = vehicles_9cat plots(only)=(contourfit);
  class condition(ref='salvage') cylinders(ref='other') drive(ref='4wd');
  model log_price = log_age log_odometer condition cylinders drive / solution;
run;
proc glm data = vehicles_9cat plots(only)=(contourfit);
  class type(ref='other') paint_color(ref='custom');
  model log_price = log_age log_odometer type paint_color / solution;
run;
ods graphics off;
quit;
```

Directory	
Libref	
Engine	
Physical Name	
Filename	
Owner Name	
File Size	
File Size (bytes)	

#	Name	Member Type	File Size	Last Modified
1	SASGOPT	CATALOG	5KB	07/29/2019 21:28:13
2	SASMAC3	CATALOG	13KB	07/30/2019 04:55:27

The CONTENTS Procedure

Data Set Name	WORK.VEHICLES_RAW	Observations	1723065
Member Type	DATA	Variables	26
Engine	V9	Indexes	0
Created	07/30/2019 04:55:27	Observation Length	386
Last Modified	07/30/2019 04:55:27	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_64		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information	
Data Set Page Size	65536
Number of Data Set Pages	10196
First Data Page	1
Max Obs per Page	169
Obs in First Data Page	160
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	E:\SAS Temporary Files\PCG180000_TD5956_SMVSASCLASSC_\vehicles_raw.sas7bdat
Release Created	9.0401M4
Host Created	X64_SR12R2
Owner Name	CAMPUS\PCG180000
File Size	637MB
File Size (bytes)	668270592

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	url	Char	85	\$85.	\$85.
2	city	Char	10	\$10.	\$10.
3	price	Char	7	\$7.	\$7.
4	year	Char	6	\$6.	\$6.
5	manufacturer	Char	11	\$11.	\$11.
6	make	Char	23	\$23.	\$23.
7	condition	Char	11	\$11.	\$11.
8	cylinders	Char	14	\$14.	\$14.
9	fuel	Char	10	\$10.	\$10.
10	odometer	Char	8	\$8.	\$8.
11	title_status	Char	9	\$9.	\$9.
12	transmission	Char	11	\$11.	\$11.
13	vin	Char	19	\$19.	\$19.
14	drive	Char	5	\$5.	\$5.
15	size	Char	13	\$13.	\$13.
16	type	Char	8	\$8.	\$8.
17	paint_color	Char	7	\$7.	\$7.
18	image_url	Char	61	\$61.	\$61.
19	lat	Char	11	\$11.	\$11.
20	long	Char	13	\$13.	\$13.
21	county_fips	Char	7	\$7.	\$7.
22	county_name	Char	11	\$11.	\$11.
23	state_fips	Char	4	\$4.	\$4.
24	state_code	Char	4	\$4.	\$4.
25	state_name	Char	14	\$14.	\$14.
26	weather	Char	4	\$4.	\$4.

The UNIVARIATE Procedure
Variable: price

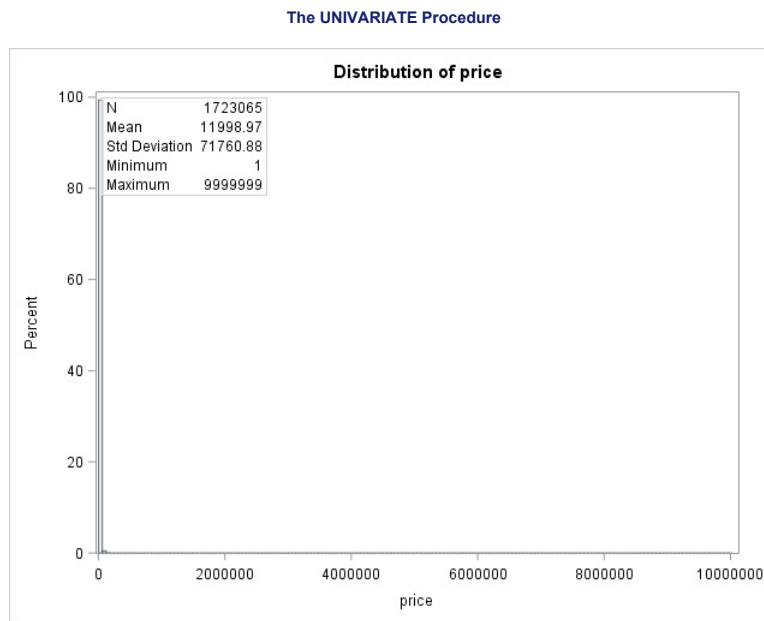
Moments			
N	1723065	Sum Weights	1723065
Mean	11998.9699	Sum Observations	2.0675E10
Std Deviation	71760.883	Variance	5149624323
Skewness	105.622505	Kurtosis	13188.6776
Uncorrected SS	9.12121E15	Corrected SS	8.87313E15
Coeff Variation	598.058695	Std Error Mean	54.668473

Basic Statistical Measures			
Location		Variability	
Mean	11998.97	Std Deviation	71761
Median	7000.00	Variance	5149624323
Mode	2500.00	Range	9999998
		Interquartile Range	11704

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	219.4861	Pr > t <.0001
Sign	M	861532.5	Pr >= M <.0001
Signed Rank	S	7.422E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	9999999
99%	51999
95%	32950
90%	25000
75% Q3	14999
50% Median	7000
25% Q1	3295
10%	1500
5%	850
1%	1
0% Min	1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	1.72E6	9999999	1.38E6
1	1.72E6	9999999	1.53E6
1	1.72E6	9999999	1.55E6
1	1.72E6	9999999	1.63E6
1	1.72E6	9999999	1.64E6



The UNIVARIATE Procedure
Variable: year

Moments			
N	1716750	Sum Weights	1716750
Mean	2004.84084	Sum Observations	3441810513
Std Deviation	12.0877163	Variance	146.112885
Skewness	-4.9795921	Kurtosis	315.319101
Uncorrected SS	6.90053E12	Corrected SS	250839149
Coeff Variation	0.60292648	Std Error Mean	0.00922552

Basic Statistical Measures			
Location		Variability	
Mean	2004.841	Std Deviation	12.08772
Median	2007.000	Variance	146.11288
Mode	2007.000	Range	1717
		Interquartile Range	10.00000

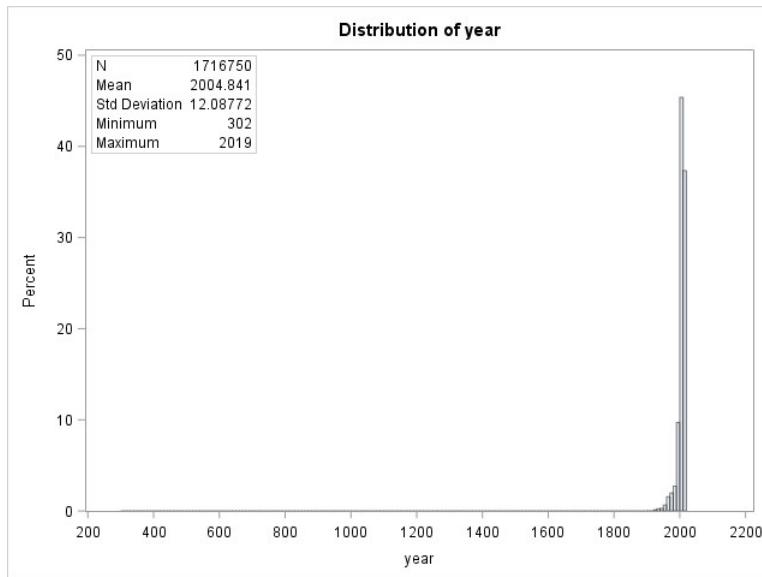
Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	217314.7	Pr > t <.0001
Sign	M	858375	Pr => M <.0001
Signed Rank	S	7.368E11	Pr => S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	2019
99%	2018
95%	2017
90%	2015
75% Q3	2012
50% Median	2007
25% Q1	2002
10%	1994
5%	1980
1%	1955
0% Min	302

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
302	7039	2019	1.72E6
718	614121	2019	1.72E6
1553	86238	2019	1.72E6
1740	1.3E6	2019	1.72E6
1796	354711	2019	1.72E6

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	6315	0.37	100.00

The UNIVARIATE Procedure



The UNIVARIATE Procedure
Variable: odometer

Moments			
N	1159011	Sum Weights	1159011
Mean	113638.653	Sum Observations	1.31708E11
Std Deviation	136860.29	Variance	1.87307E10
Skewness	35.0466154	Kurtosis	2157.41688
Uncorrected SS	3.66763E16	Corrected SS	2.17091E16
Coeff Variation	120.434629	Std Error Mean	127.125793

Basic Statistical Measures			
Location		Variability	
Mean	113638.7	Std Deviation	136860
Median	107000.0	Variance	1.87307E10
Mode	150000.0	Range	10000000
		Interquartile Range	94000

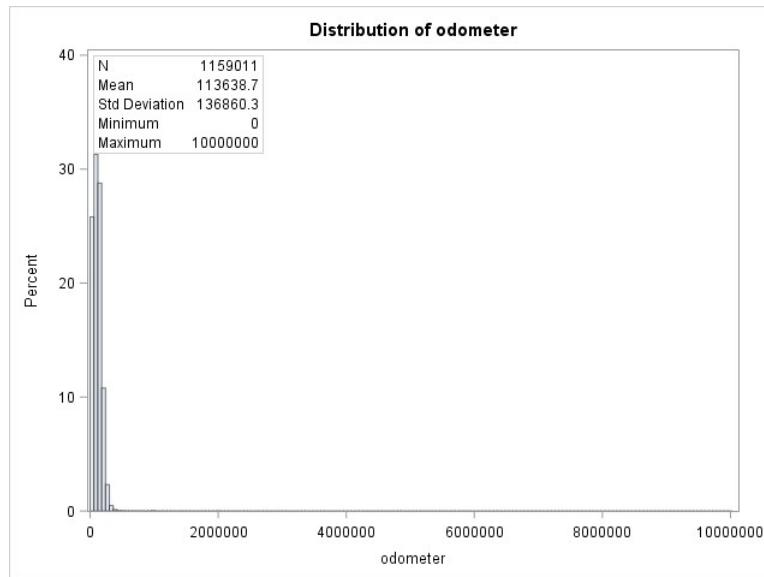
Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	893.9071	Pr > t <.0001
Sign	M	576775.5	Pr >= M <.0001
Signed Rank	S	3.327E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	10000000
99%	300000
95%	222000
90%	194000
75% Q3	152000
50% Median	107000
25% Q1	58000
10%	23456
5%	8000
1%	41
0% Min	0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	1.72E6	10000000	1.43E6
0	1.72E6	10000000	1.47E6
0	1.72E6	10000000	1.58E6
0	1.72E6	10000000	1.6E6
0	1.72E6	10000000	1.72E6

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	564054	32.74	100.00

The UNIVARIATE Procedure



The UNIVARIATE Procedure
Variable: price

Moments			
N	1636227	Sum Weights	1636227
Mean	10802.471	Sum Observations	1.76753E10
Std Deviation	10268.4302	Variance	105440658
Skewness	1.75919055	Kurtosis	3.79078232
Uncorrected SS	3.63462E14	Corrected SS	1.72525E14
Coeff Variation	95.0563081	Std Error Mean	8.02753627

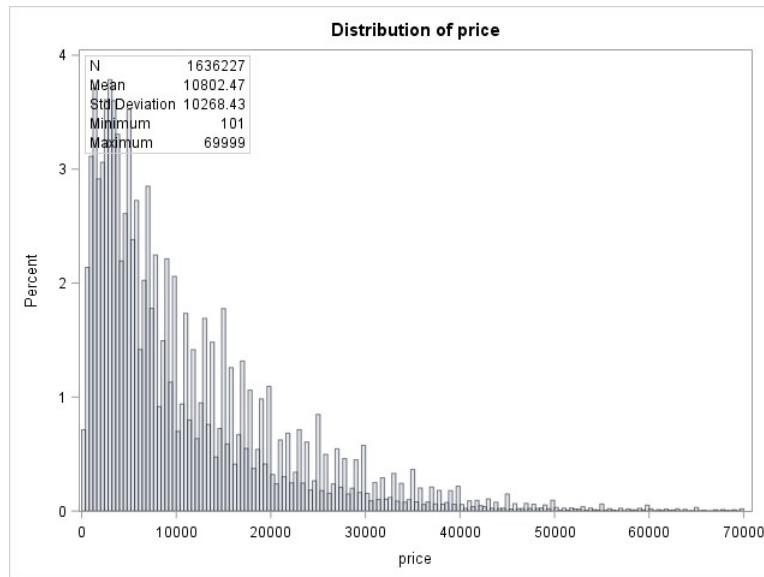
Basic Statistical Measures			
Location		Variability	
Mean	10802.47	Std Deviation	10268
Median	7300.00	Variance	105440658
Mode	2500.00	Range	69898
		Interquartile Range	11499

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	1345.677	Pr > t <.0001
Sign	M	818113.5	Pr >= M <.0001
Signed Rank	S	6.693E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	69999
99%	47000
95%	31987
90%	24995
75% Q3	14999
50% Median	7300
25% Q1	3500
10%	1600
5%	1000
1%	495
0% Min	101

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
101	1.45E6	69999	1.34E6
101	1.45E6	69999	1.41E6
101	1.45E6	69999	1.41E6
101	1.19E6	69999	1.43E6
101	668053	69999	1.43E6

The UNIVARIATE Procedure



The UNIVARIATE Procedure
Variable: age

Moments			
N	1630255	Sum Weights	1630255
Mean	14.5860691	Sum Observations	23779012
Std Deviation	10.3596211	Variance	107.321749
Skewness	2.10509727	Kurtosis	5.86185929
Uncorrected SS	521804022	Corrected SS	174961711
Coeff Variation	71.0240783	Std Error Mean	0.00811365

Basic Statistical Measures			
Location		Variability	
Mean	14.58607	Std Deviation	10.35962
Median	13.00000	Variance	107.32175
Mode	13.00000	Range	68.00000
		Interquartile Range	10.00000

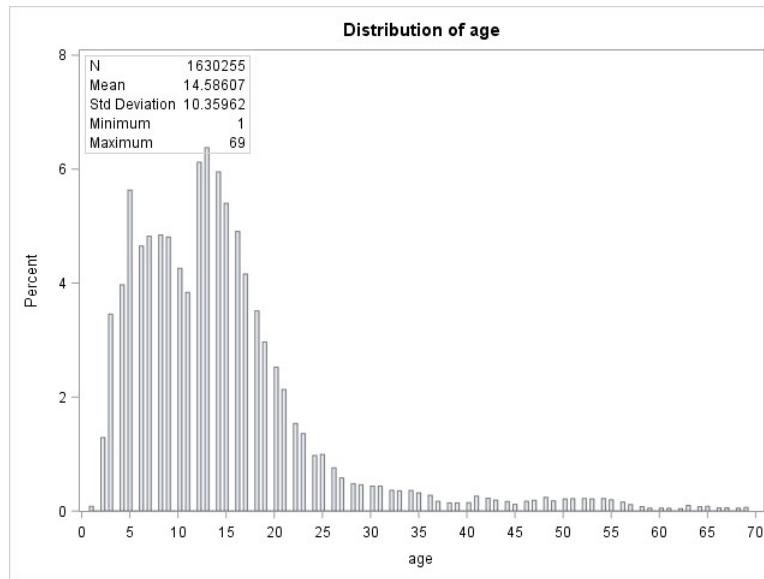
Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	1797.72	Pr > t 	<.0001
Sign	M	815127.5	Pr >= M 	<.0001
Signed Rank	S	6.644E11	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	69
99%	56
95%	35
90%	25
75% Q3	18
50% Median	13
25% Q1	8
10%	5
5%	4
1%	2
0% Min	1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	1.63E6	69	1.63E6
1	1.63E6	69	1.63E6
1	1.63E6	69	1.63E6
1	1.63E6	69	1.64E6
1	1.63E6	69	1.64E6

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	5972	0.36	100.00

The UNIVARIATE Procedure



The UNIVARIATE Procedure
Variable: odometer

Moments			
N	1102004	Sum Weights	1102004
Mean	113100.588	Sum Observations	1.24637E11
Std Deviation	63244.8342	Variance	3999909054
Skewness	0.56953899	Kurtosis	0.6161862
Uncorrected SS	1.85045E16	Corrected SS	4.40791E15
Coeff Variation	55.9191029	Std Error Mean	60.2467291

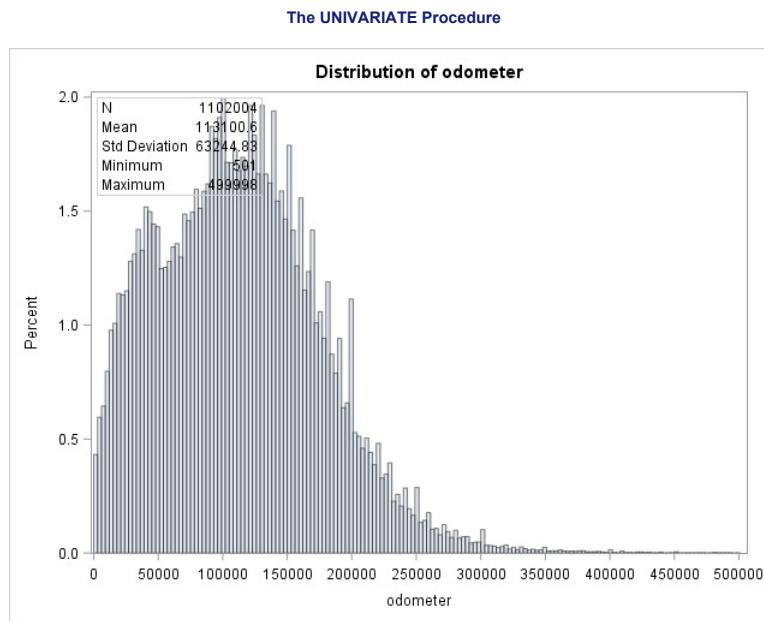
Basic Statistical Measures			
Location		Variability	
Mean	113100.6	Std Deviation	63245
Median	110000.0	Variance	3999909054
Mode	150000.0	Range	499497
		Interquartile Range	89166

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	1877.29	Pr > t 	<.0001
Sign	M	551002	Pr >= M 	<.0001
Signed Rank	S	3.036E11	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	499998
99%	283000
95%	220014
90%	194344
75% Q3	153823
50% Median	110000
25% Q1	64657
10%	31863
5%	19400
1%	5836
0% Min	501

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
501	18572	499453	277293
502	1.4E6	499453	448142
502	996230	499595	744148
502	156690	499936	540285
504	1.06E6	499998	1.27E6

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	534223	32.65	100.00



The UNIVARIATE Procedure
Variable: log_price

Moments			
N	1636227	Sum Weights	1636227
Mean	8.82132672	Sum Observations	14433693
Std Deviation	1.05212507	Variance	1.10696717
Skewness	-0.4650401	Kurtosis	-0.0240436
Uncorrected SS	129135570	Corrected SS	1811248.46
Coeff Variation	11.9270616	Std Error Mean	0.00082252

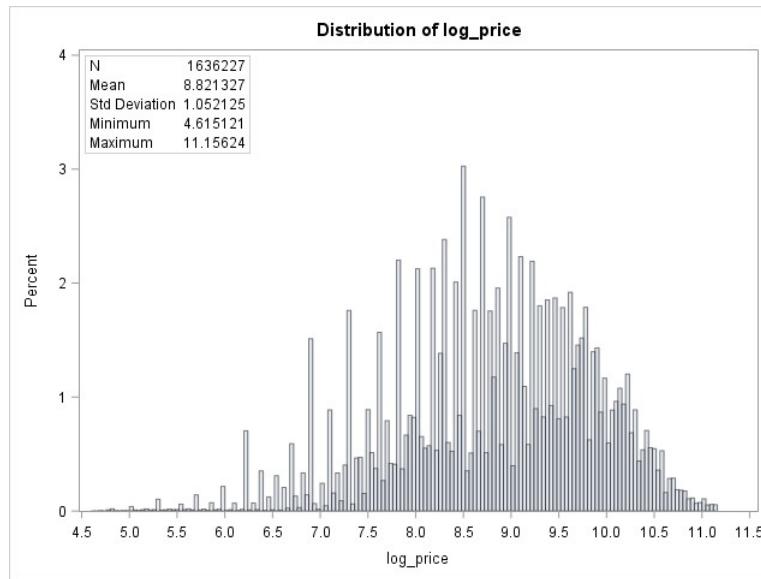
Basic Statistical Measures			
Location		Variability	
Mean	8.821327	Std Deviation	1.05213
Median	8.895630	Variance	1.10697
Mode	7.824046	Range	6.54112
		Interquartile Range	1.45522

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	10724.78	Pr > t <.0001
Sign	M	818113.5	Pr >= M <.0001
Signed Rank	S	6.693E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.75790
95%	10.37308
90%	10.12643
75% Q3	9.61574
50% Median	8.89563
25% Q1	8.16052
10%	7.37776
5%	6.90776
1%	6.20456
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	1.45E6	11.1562	1.34E6
4.61512	1.45E6	11.1562	1.41E6
4.61512	1.45E6	11.1562	1.41E6
4.61512	1.19E6	11.1562	1.43E6
4.61512	668053	11.1562	1.43E6

The UNIVARIATE Procedure



The UNIVARIATE Procedure
Variable: log_age

Moments			
N	1630255	Sum Weights	1630255
Mean	2.4651533	Sum Observations	4018828.49
Std Deviation	0.6691238	Variance	0.44772666
Skewness	-0.2238893	Kurtosis	0.20730163
Uncorrected SS	10636936.5	Corrected SS	729908.173
Coeff Variation	27.1432936	Std Error Mean	0.00052406

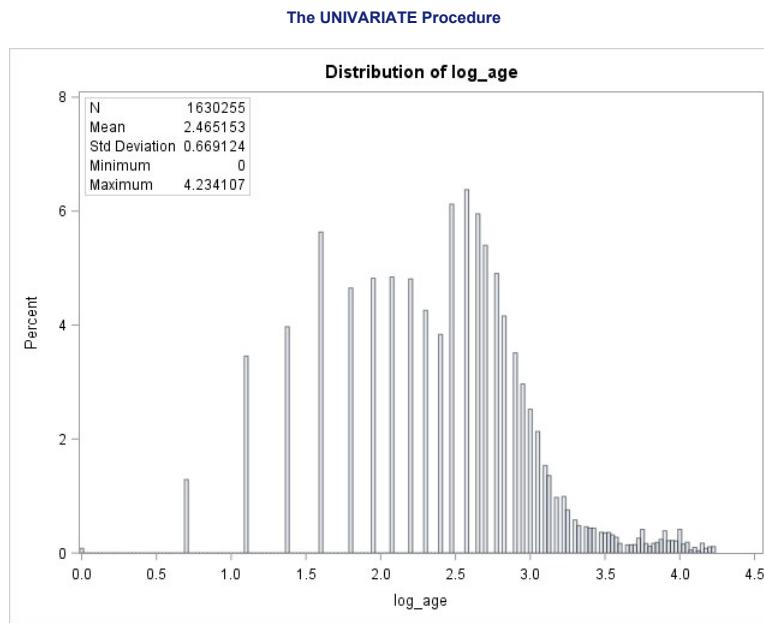
Basic Statistical Measures			
Location		Variability	
Mean	2.465153	Std Deviation	0.66912
Median	2.564949	Variance	0.44773
Mode	2.564949	Range	4.23411
		Interquartile Range	0.81093

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	4703.977	Pr > t <.0001
Sign	M	814460.5	Pr >= M <.0001
Signed Rank	S	6.633E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	4.234107
99%	4.025352
95%	3.555348
90%	3.218876
75% Q3	2.890372
50% Median	2.564949
25% Q1	2.079442
10%	1.609438
5%	1.386294
1%	0.693147
0% Min	0.000000

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	1.63E6	4.23411	1.63E6
0	1.63E6	4.23411	1.63E6
0	1.63E6	4.23411	1.63E6
0	1.63E6	4.23411	1.64E6
0	1.63E6	4.23411	1.64E6

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	5972	0.36	100.00



The UNIVARIATE Procedure
Variable: log_odometer

Moments			
N	1102004	Sum Weights	1102004
Mean	11.4110594	Sum Observations	12575033.1
Std Deviation	0.79336625	Variance	0.62943001
Skewness	-1.6257042	Kurtosis	4.10205556
Uncorrected SS	144188084	Corrected SS	693633.761
Coeff Variation	6.9526082	Std Error Mean	0.00075576

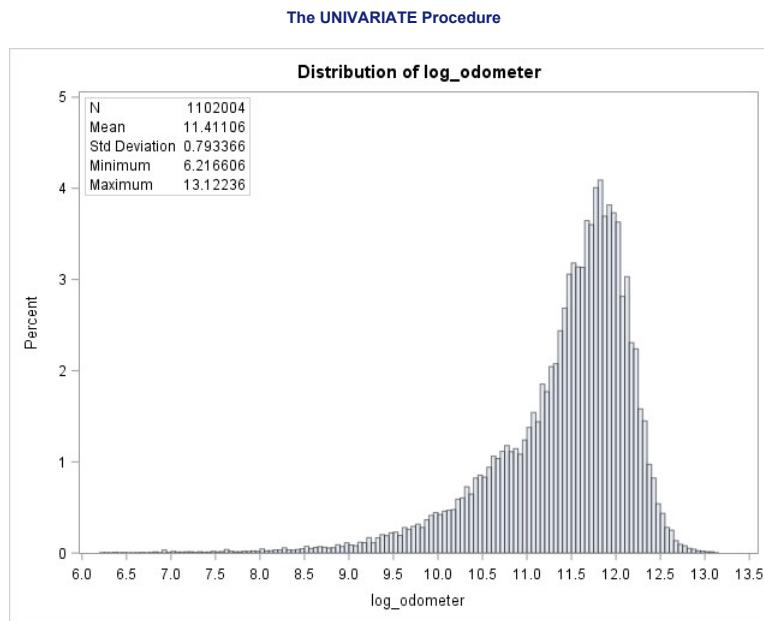
Basic Statistical Measures			
Location		Variability	
Mean	11.41106	Std Deviation	0.79337
Median	11.60824	Variance	0.62943
Mode	11.91839	Range	6.90575
		Interquartile Range	0.86671

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	15098.85	Pr > t <.0001
Sign	M	551002	Pr >= M <.0001
Signed Rank	S	3.036E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	13.12236
99%	12.55320
95%	12.30145
90%	12.17739
75% Q3	11.94356
50% Median	11.60824
25% Q1	11.07685
10%	10.36920
5%	9.87303
1%	8.67180
0% Min	6.21661

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
6.21661	18572	13.1213	277293
6.21860	1.4E6	13.1213	448142
6.21860	996230	13.1216	744148
6.21860	158690	13.1222	540285
6.22258	1.06E6	13.1224	1.27E6

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	534223	32.65	100.00



The FREQ Procedure

fuel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	9915	0.61	9915	0.61
diesel	114241	6.98	124156	7.59
electric	2181	0.13	126337	7.72
gas	1456237	89.00	1582574	96.72
hybrid	10553	0.64	1593127	97.37
other	43100	2.63	1636227	100.00

title_status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	2515	0.15	2515	0.15
clean	1523614	93.12	1526129	93.27
lien	20845	1.27	1546974	94.55
missing	8870	0.54	1555844	95.09
parts onl	3558	0.22	1559402	95.30
rebuilt	46617	2.85	1606019	98.15
salvage	30208	1.85	1636227	100.00

transmission	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	8736	0.53	8736	0.53
automatic	1411711	86.28	1420447	86.81
manual	185549	11.34	1605996	98.15
other	30231	1.85	1636227	100.00

manufacturer	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	120818	7.38	120818	7.38
acura	17416	1.06	138234	8.45
alfa	73	0.00	138307	8.45
alfa-romeo	130	0.01	138437	8.46
aston	33	0.00	138470	8.46
aston-marti	39	0.00	138509	8.47
audi	14524	0.89	153033	9.35
bmw	41732	2.55	194765	11.90
buick	21789	1.33	216554	13.23
cadillac	24167	1.48	240721	14.71
chev	788	0.05	241509	14.76
chevrolet	208882	12.77	450391	27.53
chevy	38067	2.33	488458	29.85
chrysler	30330	1.85	518788	31.71
datsun	577	0.04	519365	31.74
dodge	72392	4.42	591757	36.17
ferrari	57	0.00	591814	36.17
fiat	1882	0.12	593696	36.28
ford	277397	16.95	871093	53.24
gmc	67145	4.10	938238	57.34
harley	227	0.01	938465	57.36
harley-davi	515	0.03	938980	57.39
hennessey	1	0.00	938981	57.39
honda	86023	5.26	1025004	62.64
hyundai	32286	1.97	1057290	64.62
infiniti	11594	0.71	1068884	65.33
infinity	499	0.03	1069383	65.36
jaguar	3997	0.24	1073380	65.60
jeep	72333	4.42	1145713	70.02
kia	24266	1.48	1169979	71.50
land rover	62	0.00	1170041	71.51

landrover	39	0.00	1170080	71.51
lexus	19601	1.20	1189681	72.71
lincoln	12365	0.76	1202046	73.46
mazda	22674	1.39	1224720	74.85
mercedes	6810	0.42	1231530	75.27
mercedes-be	22974	1.40	1254504	76.67
mercedesben	4	0.00	1254508	76.67
mercury	9713	0.59	1264221	77.26
mini	6265	0.38	1270486	77.65
mitsubishi	10470	0.64	1280956	78.29
morgan	8	0.00	1280964	78.29
nissan	76920	4.70	1357884	82.99
noble	2	0.00	1357886	82.99
pontiac	21384	1.31	1379270	84.30
porche	109	0.01	1379379	84.30
ram	54540	3.33	1433919	87.64
rover	5154	0.31	1439073	87.95
saturn	8973	0.55	1448046	88.50
subaru	28601	1.75	1476647	90.25
toyota	112361	6.87	1589008	97.11
volkswagen	32440	1.98	1621448	99.10
volvo	10507	0.64	1631955	99.74
vw	4272	0.26	1636227	100.00

condition	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	666626	40.74	666626	40.74
excellent	422244	25.81	1088870	66.55
fair	69834	4.27	1158704	70.82
good	360895	22.06	1519599	92.87
like new	105124	6.42	1624723	99.30
new	6271	0.38	1630994	99.68
salvage	5233	0.32	1636227	100.00

cylinders	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	654185	39.98	654185	39.98
10 cylinders	4383	0.27	658568	40.25
12 cylinders	651	0.04	659219	40.29
3 cylinders	1700	0.10	660919	40.39
4 cylinders	281654	17.21	942573	57.61
5 cylinders	10025	0.61	952598	58.22
6 cylinders	346366	21.17	1298964	79.39
8 cylinders	303249	18.53	1602213	97.92
other	34014	2.08	1636227	100.00

drive	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	620627	37.93	620627	37.93
4wd	429128	26.23	1049755	64.16
fwd	358030	21.88	1407785	86.04
rwd	228442	13.96	1636227	100.00

size	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	1066052	65.15	1066052	65.15
compact	89148	5.45	1155200	70.60
full-size	305165	18.65	1460365	89.25
mid-size	164189	10.03	1624554	99.29
sub-compact	11673	0.71	1636227	100.00

			Cumulative	Cumulative
--	--	--	------------	------------

type	Frequency	Percent	Frequency	Percent
	658700	40.26	658700	40.26
SUV	237461	14.51	896161	54.77
bus	1952	0.12	898113	54.89
converti	29267	1.79	927380	56.68
coupe	67229	4.11	994609	60.79
hatchbac	37066	2.27	1031675	63.05
mini-van	24108	1.47	1055783	64.53
offroad	4712	0.29	1060495	64.81
other	21141	1.29	1081636	66.11
pickup	118086	7.22	1199722	73.32
sedan	260127	15.90	1459849	89.22
truck	127399	7.79	1587248	97.01
van	24427	1.49	1611675	98.50
wagon	24552	1.50	1636227	100.00

paint_color	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	652376	39.87	652376	39.87
black	191832	11.72	844208	51.59
blue	111105	6.79	955313	58.39
brown	25803	1.58	981116	59.96
custom	23860	1.46	1004976	61.42
green	40712	2.49	1045688	63.91
grey	96722	5.91	1142410	69.82
orange	6658	0.41	1149068	70.23
purple	3627	0.22	1152695	70.45
red	110222	6.74	1262917	77.18
silver	142611	8.72	1405528	85.90
white	220399	13.47	1625927	99.37
yellow	10300	0.63	1636227	100.00

Directory				
Libref	WORK			
Engine	V9			
Physical Name	E:\SAS Temporary Files\PCG180000_TD5956_SMVSASCLASSC_			
Filename	E:\SAS Temporary Files\PCG180000_TD5956_SMVSASCLASSC_			
Owner Name	CAMPUS\PCG180000			
File Size	48KB			
File Size (bytes)	49152			
1	SASGOPT	CATALOG	5KB	07/29/2019 21:28:13
2	SASMAC1	CATALOG	13KB	07/30/2019 04:56:25
3	SASMAC3	CATALOG	13KB	07/30/2019 04:55:27
4	VEHICLES_4CAT	DATA	114MB	07/30/2019 04:57:29
5	VEHICLES_9CAT	DATA	57MB	07/30/2019 04:57:30
6	VEHICLES_RAW	DATA	645MB	07/30/2019 04:56:03
7	VEHICLES_TMP	DATA	238MB	07/30/2019 04:57:20

The UNIVARIATE Procedure
Variable: log_price

Moments			
N	1493278	Sum Weights	1493278
Mean	8.83475458	Sum Observations	13192744.7
Std Deviation	1.03572573	Variance	1.07272779
Skewness	-0.4646812	Kurtosis	-0.0124669
Uncorrected SS	118156541	Corrected SS	1601879.74
Coeff Variation	11.7233107	Std Error Mean	0.00084757

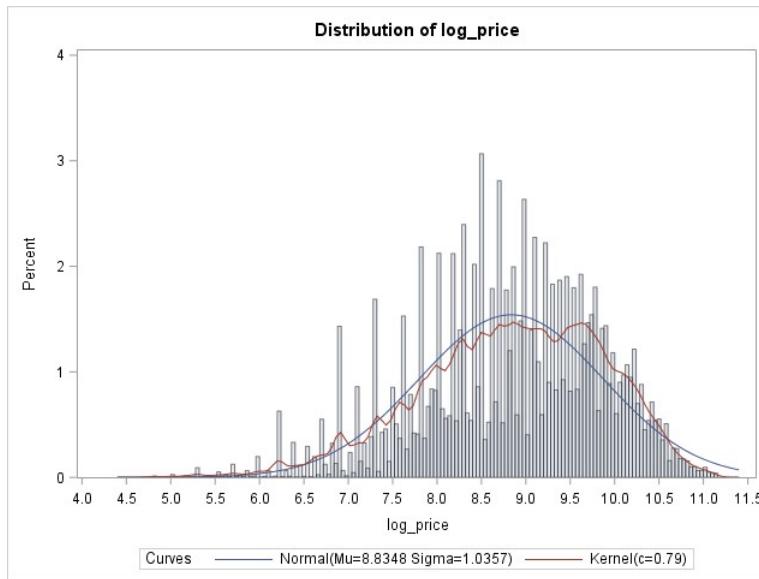
Basic Statistical Measures			
Location		Variability	
Mean	8.834755	Std Deviation	1.03573
Median	8.921991	Variance	1.07273
Mode	7.824046	Range	6.54112
		Interquartile Range	1.45522

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	10423.66	Pr > t <.0001
Sign	M	746639	Pr >= M <.0001
Signed Rank	S	5.575E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.73637
95%	10.35946
90%	10.12623
75% Q3	9.61574
50% Median	8.92199
25% Q1	8.16052
10%	7.46737
5%	6.90776
1%	6.21461
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	1.32E6	11.1562	696850
4.61512	1.32E6	11.1562	1.28E6
4.61512	1.32E6	11.1562	1.29E6
4.61512	1.08E6	11.1562	1.3E6
4.61512	608170	11.1562	1.3E6

The UNIVARIATE Procedure

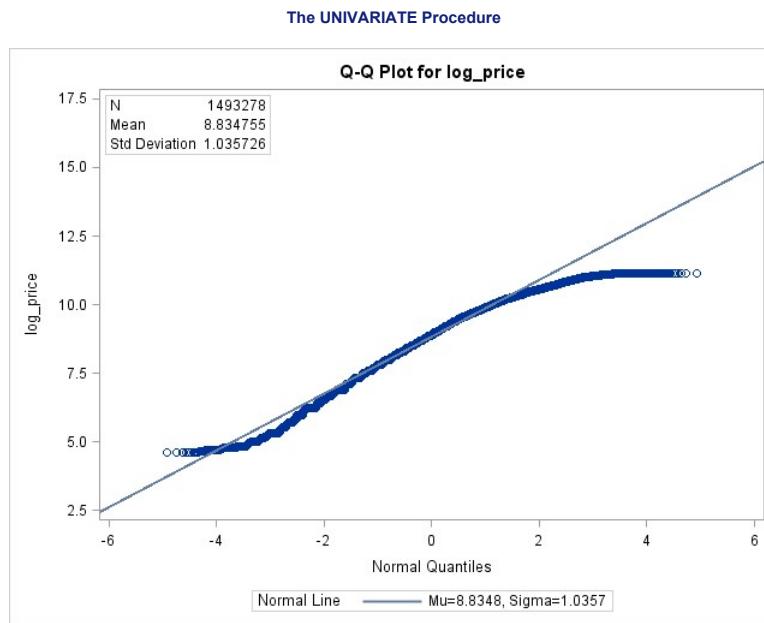


The UNIVARIATE Procedure
Fitted Normal Distribution for log_price

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	8.834755
Std Dev	Sigma	1.035726

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.04364	Pr > D	<0.010
Cramer-von Mises	W-Sq	702.46446	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	4899.49095	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	6.21461	6.42530
5.0	6.90776	7.13114
10.0	7.46737	7.50742
25.0	8.16052	8.13617
50.0	8.92199	8.83475
75.0	9.61574	9.53334
90.0	10.12623	10.16209
95.0	10.35946	10.53837
99.0	10.73637	11.24421



The UNIVARIATE Procedure
Variable: log_age

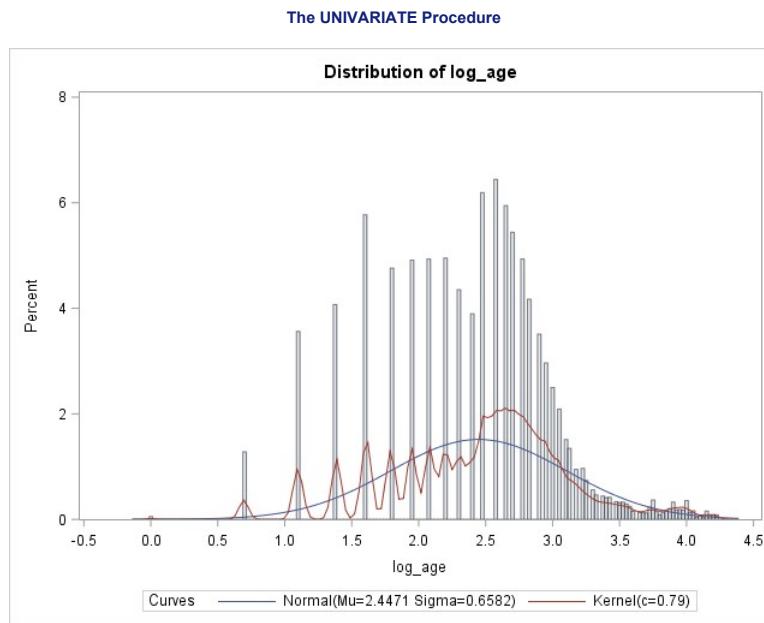
Moments			
N	1493278	Sum Weights	1493278
Mean	2.44708586	Sum Observations	3654179.48
Std Deviation	0.65817142	Variance	0.43318961
Skewness	-0.239101	Kurtosis	0.18490611
Uncorrected SS	9588963.05	Corrected SS	646872.089
Coeff Variation	26.8961309	Std Error Mean	0.0005386

Basic Statistical Measures			
Location		Variability	
Mean	2.447086	Std Deviation	0.65817
Median	2.564949	Variance	0.43319
Mode	2.564949	Range	4.23411
		Interquartile Range	0.75377

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	4543.395	Pr > t <.0001
Sign	M	746243	Pr >= M <.0001
Signed Rank	S	5.569E11	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	4.234107
99%	4.007333
95%	3.496508
90%	3.178054
75% Q3	2.833213
50% Median	2.564949
25% Q1	2.079442
10%	1.609438
5%	1.386294
1%	0.693147
0% Min	0.000000

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	1.49E6	4.23411	1.49E6
0	1.49E6	4.23411	1.49E6
0	1.49E6	4.23411	1.49E6
0	1.48E6	4.23411	1.49E6
0	1.48E6	4.23411	1.49E6

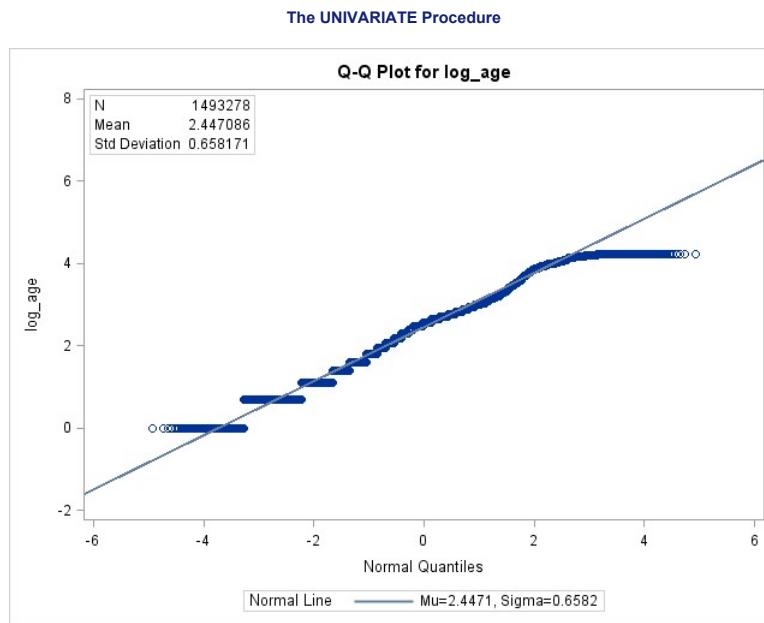


The UNIVARIATE Procedure
Fitted Normal Distribution for log_age

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	2.447086
Std Dev	Sigma	0.658171

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.0975	Pr > D	<0.010
Cramer-von Mises	W-Sq	1901.8862	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	10138.0011	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	0.69315	0.91595
5.0	1.38629	1.36449
10.0	1.60944	1.60361
25.0	2.07944	2.00316
50.0	2.56495	2.44709
75.0	2.83321	2.89102
90.0	3.17805	3.29057
95.0	3.49651	3.52968
99.0	4.00733	3.97822



The UNIVARIATE Procedure
Variable: log_price

Moments			
N	441328	Sum Weights	441328
Mean	8.88085539	Sum Observations	3919370.15
Std Deviation	0.95354226	Variance	0.90924285
Skewness	-0.4272486	Kurtosis	0.1376484
Uncorrected SS	35208632.9	Corrected SS	401273.42
Coeff Variation	10.7370543	Std Error Mean	0.00143535

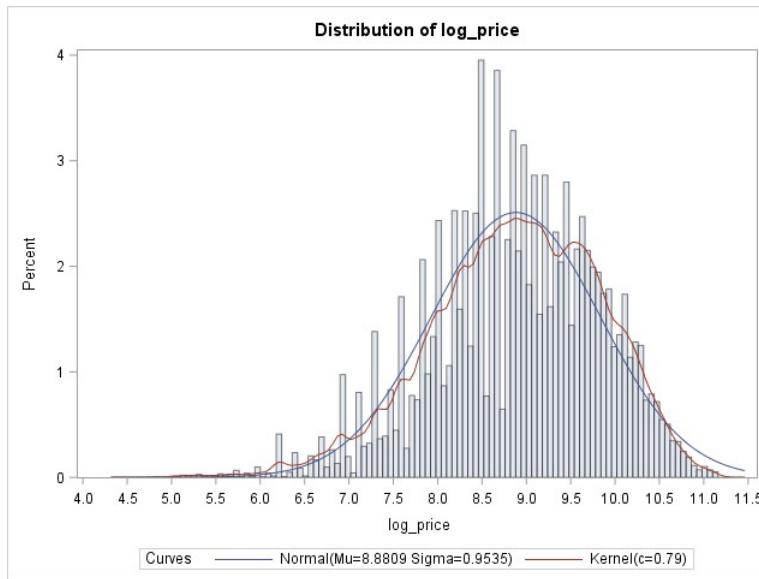
Basic Statistical Measures			
Location		Variability	
Mean	8.880855	Std Deviation	0.95354
Median	8.922658	Variance	0.90924
Mode	7.824046	Range	6.53126
		Interquartile Range	1.31406

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	6187.22	Pr > t <.0001
Sign	M	220664	Pr >= M <.0001
Signed Rank	S	4.869E10	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.71440
95%	10.30892
90%	10.08539
75% Q3	9.59553
50% Median	8.92266
25% Q1	8.28147
10%	7.60090
5%	7.24423
1%	6.39526
0% Min	4.62497

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.62497	365541	11.1562	345237
4.70953	159856	11.1562	347087
4.74493	91598	11.1562	55891
4.74493	85116	11.1562	56794
4.75359	359837	11.1562	199801

The UNIVARIATE Procedure

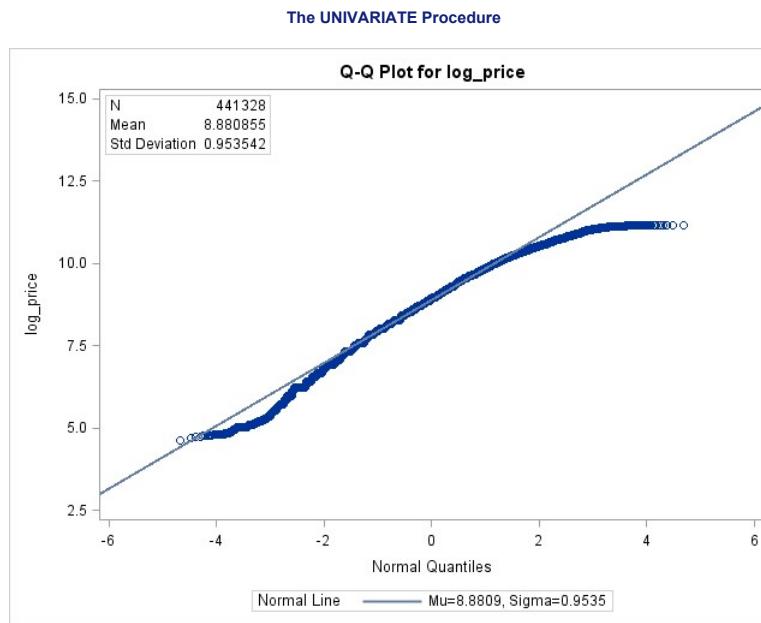


The UNIVARIATE Procedure
Fitted Normal Distribution for log_price

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	8.880855
Std Dev	Sigma	0.953542

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.032166	Pr > D	<0.010
Cramer-von Mises	W-Sq	118.614154	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	882.986089	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	6.39526	6.66258
5.0	7.24423	7.31242
10.0	7.60090	7.65884
25.0	8.28147	8.23770
50.0	8.92266	8.88086
75.0	9.59553	9.52401
90.0	10.08539	10.10287
95.0	10.30892	10.44929
99.0	10.71440	11.09913



The UNIVARIATE Procedure
Variable: log_age

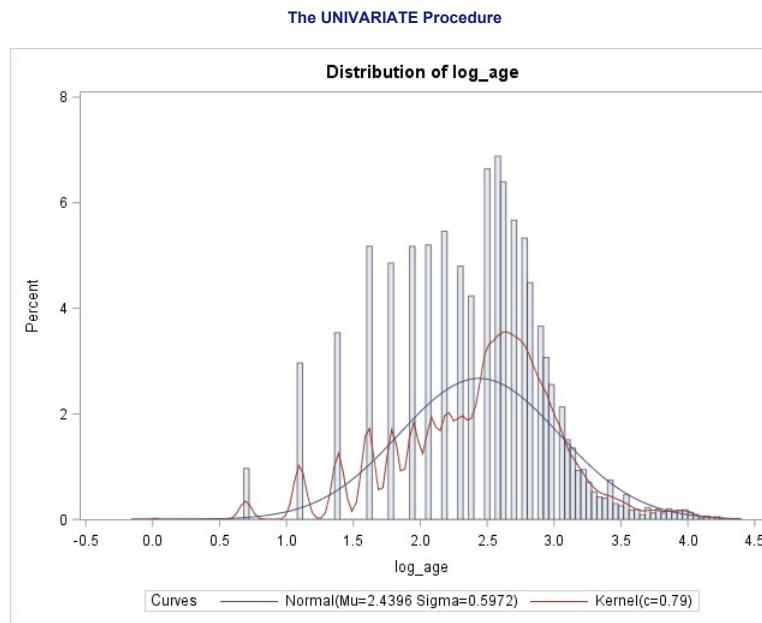
Moments			
N	441328	Sum Weights	441328
Mean	2.43961694	Sum Observations	1076671.27
Std Deviation	0.59721099	Variance	0.35666097
Skewness	-0.3909869	Kurtosis	0.31551385
Uncorrected SS	2784069.58	Corrected SS	157404.116
Coeff Variation	24.4797035	Std Error Mean	0.00089897

Basic Statistical Measures			
Location		Variability	
Mean	2.439617	Std Deviation	0.59721
Median	2.564949	Variance	0.35666
Mode	2.564949	Range	4.23411
		Interquartile Range	0.75377

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	2713.78	Pr > t <.0001
Sign	M	220619.5	Pr >= M <.0001
Signed Rank	S	4.867E10	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	4.23411
99%	3.87120
95%	3.29584
90%	3.09104
75% Q3	2.83321
50% Median	2.56495
25% Q1	2.07944
10%	1.60944
5%	1.38629
1%	1.09861
0% Min	0.00000

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	436421	4.23411	413677
0	435803	4.23411	414276
0	435175	4.23411	421223
0	434416	4.23411	427865
0	432270	4.23411	433379

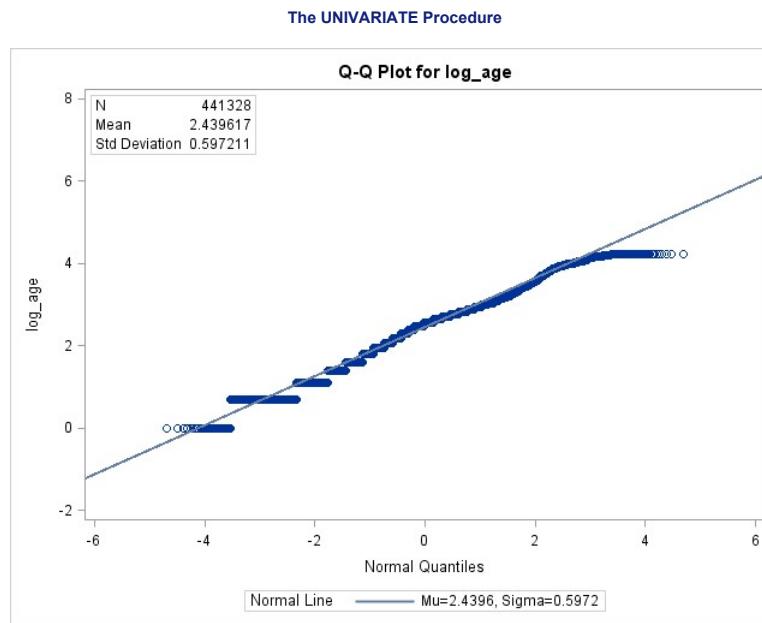


The UNIVARIATE Procedure
Fitted Normal Distribution for log_age

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	2.439617
Std Dev	Sigma	0.597211

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.10632	Pr > D	<0.010
Cramer-von Mises	W-Sq	689.47853	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	3717.83189	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	1.09861	1.05030
5.0	1.38629	1.45729
10.0	1.60944	1.67426
25.0	2.07944	2.03680
50.0	2.56495	2.43962
75.0	2.83321	2.84243
90.0	3.09104	3.20497
95.0	3.29584	3.42194
99.0	3.87120	3.82894



The UNIVARIATE Procedure
Variable: log_odometer

Moments			
N	441328	Sum Weights	441328
Mean	11.5099709	Sum Observations	5079672.44
Std Deviation	0.74909932	Variance	0.56114979
Skewness	-1.831209	Kurtosis	5.30995693
Uncorrected SS	58714532.4	Corrected SS	247650.552
Coeff Variation	6.50826422	Std Error Mean	0.00112761

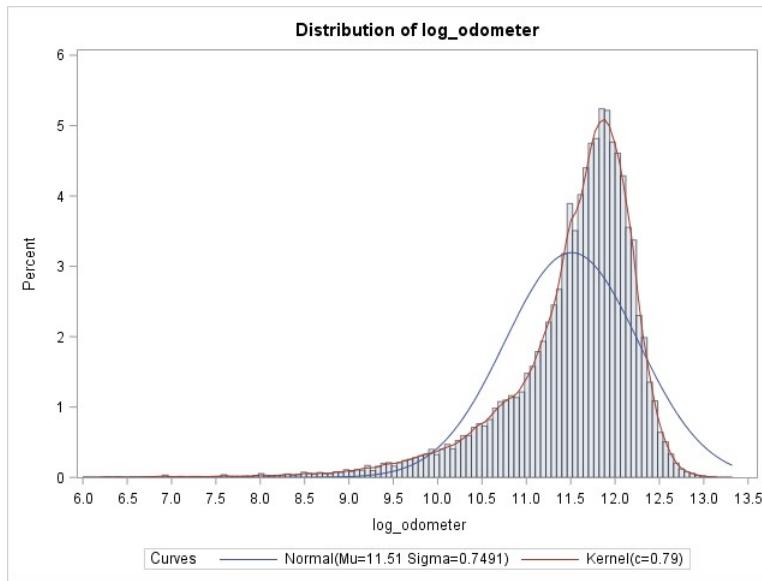
Basic Statistical Measures			
Location		Variability	
Mean	11.50997	Std Deviation	0.74910
Median	11.69525	Variance	0.56115
Mode	11.91839	Range	6.90575
		Interquartile Range	0.75473

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	10207.41	Pr > t <.0001
Sign	M	220664	Pr >= M <.0001
Signed Rank	S	4.869E10	Pr >= S <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	13.12236
99%	12.57764
95%	12.34218
90%	12.21006
75% Q3	11.99322
50% Median	11.69525
25% Q1	11.23849
10%	10.55347
5%	10.04325
1%	8.85066
0% Min	6.21661

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
6.21661	5185	13.1204	111963
6.22456	248727	13.1213	72706
6.22456	248708	13.1213	119052
6.22456	248475	13.1222	143388
6.22456	242617	13.1224	333314

The UNIVARIATE Procedure

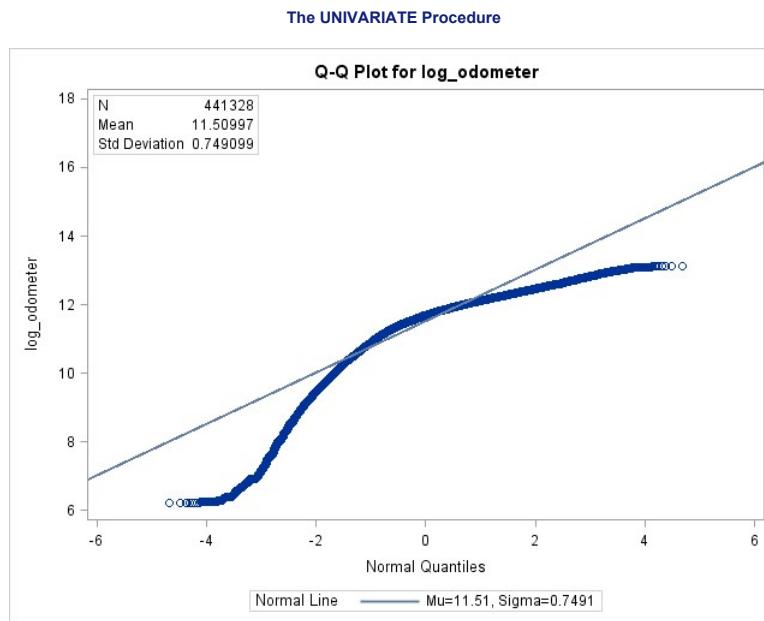


The UNIVARIATE Procedure
Fitted Normal Distribution for log_odometer

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	11.50997
Std Dev	Sigma	0.749099

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.1265	Pr > D	<0.010
Cramer-von Mises	W-Sq	2556.5039	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	14569.5933	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	8.85066	9.76731
5.0	10.04325	10.27781
10.0	10.55347	10.54996
25.0	11.23849	11.00471
50.0	11.69525	11.50997
75.0	11.99322	12.01523
90.0	12.21006	12.46998
95.0	12.34218	12.74213
99.0	12.57764	13.25264



The UNIVARIATE Procedure
Variable: log_price
fuel = diesel

Moments			
N	96724	Sum Weights	96724
Mean	9.66526717	Sum Observations	934863.302
Std Deviation	0.83676533	Variance	0.70017621
Skewness	-1.0139988	Kurtosis	1.74142485
Uncorrected SS	9103426.73	Corrected SS	67723.1439
Coeff Variation	8.65744643	Std Error Mean	0.00269052

Basic Statistical Measures			
Location		Variability	
Mean	9.665267	Std Deviation	0.83677
Median	9.769956	Variance	0.70018
Mode	9.615805	Range	6.44671
		Interquartile Range	1.06454

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	3592.339	Pr > t	<.0001
Sign	M	48362	Pr >= M	<.0001
Signed Rank	S	2.3389E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	11.00210
95%	10.78932
90%	10.62128
75% Q3	10.27488
50% Median	9.76996
25% Q1	9.21034
10%	8.55641
5%	8.15909
1%	7.09008
0% Min	4.70953

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.70953	979434	11.1562	191188
4.70953	226032	11.1562	1.28E6
4.74493	818210	11.1562	1.29E6
4.74493	295048	11.1562	1.3E6
4.78749	1.33E6	11.1562	1.3E6

The UNIVARIATE Procedure
Variable: log_price
fuel = electric

Moments			
N	1218	Sum Weights	1218
Mean	9.30468868	Sum Observations	11333.1108
Std Deviation	0.82769509	Variance	0.68507916
Skewness	-1.9857136	Kurtosis	6.46831128
Uncorrected SS	106284.809	Corrected SS	833.741335
Coeff Variation	8.89546244	Std Error Mean	0.02371629

Basic Statistical Measures			
Location		Variability	
Mean	9.304689	Std Deviation	0.82770
Median	9.392245	Variance	0.68508
Mode	9.797849	Range	6.04501
		Interquartile Range	0.70563

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	392.3332	Pr > t	<.0001
Sign	M	609	Pr >= M	<.0001
Signed Rank	S	371185.5	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	10.87332
99%	10.75769
95%	10.30226
90%	10.02127
75% Q3	9.79256
50% Median	9.39225
25% Q1	9.08693
10%	8.47637
5%	7.71869
1%	5.85793
0% Min	4.82831

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.82831	824440	10.7643	196296
4.82831	714032	10.7975	177008
4.96284	25058	10.7975	196159
4.96981	928571	10.8198	1.49E6
5.29832	481641	10.8733	43475

The UNIVARIATE Procedure
 Variable: log_price
 fuel = gas

Moments			
N	1348019	Sum Weights	1348019
Mean	8.76078486	Sum Observations	11809704.4
Std Deviation	1.02144539	Variance	1.04335068
Skewness	-0.4204588	Kurtosis	-0.0701855
Uncorrected SS	104868735	Corrected SS	1406455.49
Coeff Variation	11.6592908	Std Error Mean	0.00087977

Basic Statistical Measures			
Location		Variability	
Mean	8.760785	Std Deviation	1.02145
Median	8.838552	Variance	1.04335
Mode	7.824046	Range	6.54112
		Interquartile Range	1.44629

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	9958.086	Pr > t	<.0001
Sign	M	674009.5	Pr >= M	<.0001
Signed Rank	S	4.543E11	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.64542
95%	10.27505
90%	10.04303
75% Q3	9.54646
50% Median	8.83855
25% Q1	8.10016
10%	7.37776
5%	6.90776
1%	6.21461
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	1.08E6	11.1562	1.23E6
4.61512	608170	11.1562	152431
4.61512	238902	11.1562	187687
4.62497	1.26E6	11.1562	695988
4.62497	1.08E6	11.1562	696850

The UNIVARIATE Procedure
 Variable: log_price
 fuel = hybrid

Moments			
N	10092	Sum Weights	10092
Mean	9.01597926	Sum Observations	90989.2627
Std Deviation	0.75941945	Variance	0.57671791
Skewness	-1.1100428	Kurtosis	3.48433981
Uncorrected SS	826176.965	Corrected SS	5819.66041
Coeff Variation	8.42303906	Std Error Mean	0.0075595

Basic Statistical Measures			
Location		Variability	
Mean	9.015979	Std Deviation	0.75942
Median	9.088173	Variance	0.57672
Mode	8.612503	Range	6.34264
		Interquartile Range	0.93295

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	1192.669	Pr > t	<.0001
Sign	M	5046	Pr >= M	<.0001
Signed Rank	S	25464639	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15482
99%	10.51325
95%	10.06900
90%	9.85193
75% Q3	9.54528
50% Median	9.08817
25% Q1	8.61232
10%	8.16052
5%	7.82405
1%	6.39693
0% Min	4.81218

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.81218	154401	11.0744	529182
4.94876	1.47E6	11.0818	527452
4.97673	1.11E6	11.0821	1.45E6
5.01064	1.41E6	11.1548	813351
5.01064	1.11E6	11.1548	814007

The UNIVARIATE Procedure
Variable: log_price
fuel = other

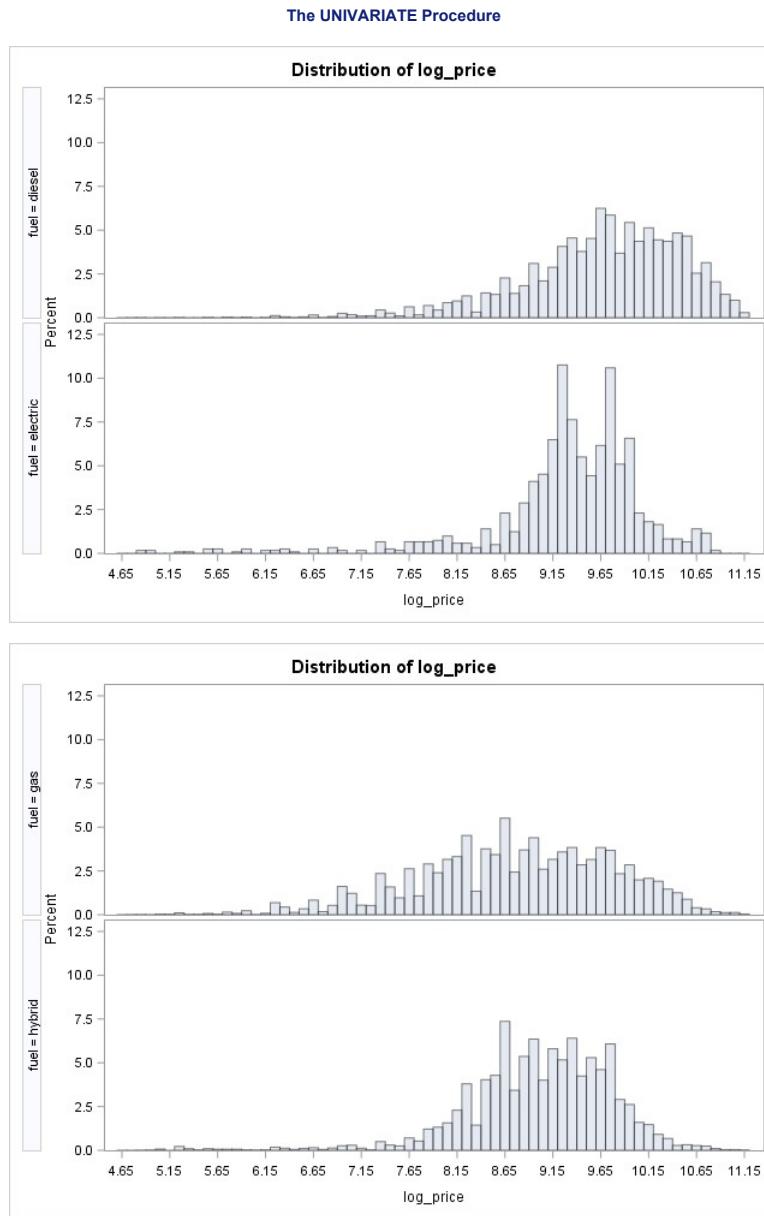
Moments			
N	37225	Sum Weights	37225
Mean	9.29092085	Sum Observations	345854.529
Std Deviation	1.01844698	Variance	1.03723424
Skewness	-1.9219998	Kurtosis	4.84215054
Uncorrected SS	3251917.06	Corrected SS	38610.0075
Coeff Variation	10.9617442	Std Error Mean	0.00527863

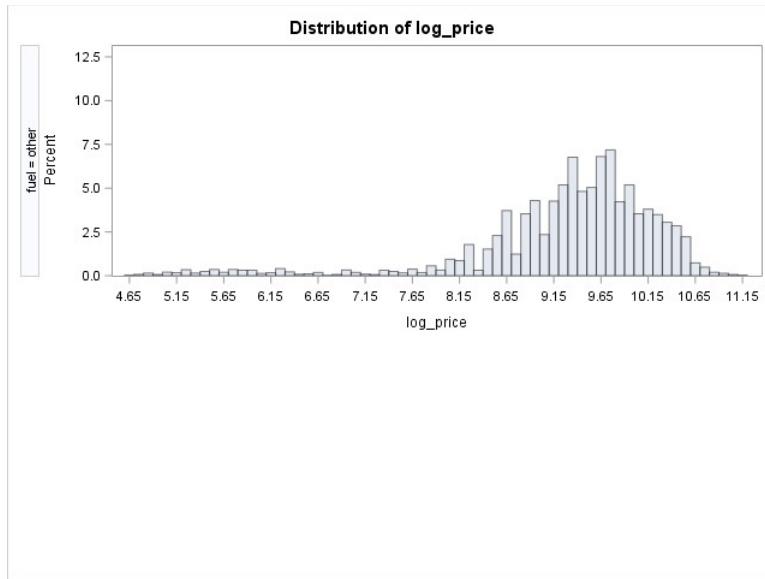
Basic Statistical Measures			
Location		Variability	
Mean	9.290921	Std Deviation	1.01845
Median	9.472320	Variance	1.03723
Mode	9.209840	Range	6.53953
		Interquartile Range	0.92862

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	1760.101	Pr > t	<.0001
Sign	M	18612.5	Pr >= M	<.0001
Signed Rank	S	3.4643E8	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15465
99%	10.68965
95%	10.45737
90%	10.29215
75% Q3	9.90324
50% Median	9.47232
25% Q1	8.97462
10%	8.29280
5%	7.30653
1%	5.29832
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	1.32E6	11.1404	741314
4.61512	1.32E6	11.1418	1.04E6
4.61512	1.32E6	11.1418	1.05E6
4.61512	323245	11.1491	213387
4.61512	228941	11.1546	612927





The UNIVARIATE Procedure
Variable: log_price
title_status = clean

Moments			
N	1392304	Sum Weights	1392304
Mean	8.84437826	Sum Observations	12314063.2
Std Deviation	1.03217477	Variance	1.06538476
Skewness	-0.4487876	Kurtosis	-0.0443961
Uncorrected SS	110393572	Corrected SS	1483338.4
Coeff Variation	11.6704051	Std Error Mean	0.00087475

Basic Statistical Measures			
Location		Variability	
Mean	8.844378	Std Deviation	1.03217
Median	8.922658	Variance	1.06538
Mode	7.824046	Range	6.54112
		Interquartile Range	1.45529

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	10110.7	Pr > t	<.0001
Sign	M	696152	Pr >= M	<.0001
Signed Rank	S	4.846E11	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.74721
95%	10.37036
90%	10.12651
75% Q3	9.61581
50% Median	8.92266
25% Q1	8.16052
10%	7.49554
5%	6.95655
1%	6.21461
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	1.32E6	11.1562	696850
4.61512	1.32E6	11.1562	1.28E6
4.61512	1.32E6	11.1562	1.29E6
4.61512	323245	11.1562	1.3E6
4.61512	228941	11.1562	1.3E6

The UNIVARIATE Procedure
 Variable: log_price
 title_status = lien

Moments			
N	19408	Sum Weights	19408
Mean	9.49416864	Sum Observations	184262.825
Std Deviation	0.88240323	Variance	0.77863546
Skewness	-1.3829685	Kurtosis	2.95825766
Uncorrected SS	1764533.31	Corrected SS	15110.9783
Coeff Variation	9.29416004	Std Error Mean	0.00633398

Basic Statistical Measures			
Location		Variability	
Mean	9.494169	Std Deviation	0.88240
Median	9.615805	Variance	0.77864
Mode	9.615805	Range	6.32651
		Interquartile Range	0.98083

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	1498.926	Pr > t	<.0001
Sign	M	9704	Pr >= M	<.0001
Signed Rank	S	94172468	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15482
99%	10.87805
95%	10.59663
90%	10.44871
75% Q3	10.08581
50% Median	9.61581
25% Q1	9.10498
10%	8.46590
5%	7.82405
1%	6.21461
0% Min	4.82831

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.82831	824440	11.1419	1.27E6
5.01064	1.36E6	11.1419	1.31E6
5.01064	262540	11.1491	424504
5.01064	120577	11.1491	1.07E6
5.01064	79056	11.1548	228314

The UNIVARIATE Procedure
Variable: log_price
title_status = missing

Moments			
N	6916	Sum Weights	6916
Mean	7.2879251	Sum Observations	50403.2907
Std Deviation	0.99697092	Variance	0.99395102
Skewness	0.48516562	Kurtosis	0.11989059
Uncorrected SS	374208.584	Corrected SS	6873.17133
Coeff Variation	13.6797634	Std Error Mean	0.01198823

Basic Statistical Measures			
Location		Variability	
Mean	7.287925	Std Deviation	0.99697
Median	7.237059	Variance	0.99395
Mode	6.214608	Range	6.37261
		Interquartile Range	1.38629

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	607.9235	Pr > t	<.0001
Sign	M	3458	Pr >= M	<.0001
Signed Rank	S	11959493	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.08214
99%	9.90349
95%	9.09381
90%	8.61250
75% Q3	7.93737
50% Median	7.23706
25% Q1	6.55108
10%	6.21461
5%	5.85793
1%	5.29832
0% Min	4.70953

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.70953	646318	10.9151	430340
4.81218	1.47E6	10.9151	1.13E6
4.81218	1.35E6	11.0268	430216
4.81218	486376	11.0268	434138
4.81218	475997	11.0821	813359

The UNIVARIATE Procedure
Variable: log_price
title_Status = parts onl

Moments			
N	2745	Sum Weights	2745
Mean	6.69352543	Sum Observations	18373.7273
Std Deviation	0.95068269	Variance	0.90379758
Skewness	0.79806632	Kurtosis	1.13897317
Uncorrected SS	125465.031	Corrected SS	2480.02056
Coeff Variation	14.2030191	Std Error Mean	0.01814531

Basic Statistical Measures			
Location		Variability	
Mean	6.693525	Std Deviation	0.95068
Median	6.551080	Variance	0.90380
Mode	6.214608	Range	6.31001
		Interquartile Range	1.02165

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	368.8846	Pr > t 	<.0001
Sign	M	1372.5	Pr >= M 	<.0001
Signed Rank	S	1884443	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	10.92513
99%	9.54681
95%	8.51719
90%	7.97247
75% Q3	7.13090
50% Median	6.55108
25% Q1	6.10925
10%	5.61677
5%	5.29832
1%	4.82831
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	238902	10.2400	1.45E6
4.70048	1.42E6	10.2751	90212
4.70048	1E6	10.4773	83317
4.70953	1.4E6	10.8198	115562
4.70953	1.06E6	10.9251	867695

The UNIVARIATE Procedure
 Variable: log_price
 title_status = rebuilt

Moments			
N	43940	Sum Weights	43940
Mean	8.86304664	Sum Observations	389442.269
Std Deviation	0.79200381	Variance	0.62727003
Skewness	-0.5234545	Kurtosis	0.73723485
Uncorrected SS	3479206.61	Corrected SS	27561.6179
Coeff Variation	8.93602211	Std Error Mean	0.00377831

Basic Statistical Measures			
Location		Variability	
Mean	8.863047	Std Deviation	0.79200
Median	8.921991	Variance	0.62727
Mode	8.612503	Range	6.53970
		Interquartile Range	1.02157

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	2345.772	Pr > t 	<.0001
Sign	M	21970	Pr >= M 	<.0001
Signed Rank	S	4.8269E8	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15482
99%	10.43412
95%	10.08539
90%	9.82553
75% Q3	9.43340
50% Median	8.92199
25% Q1	8.41183
10%	7.82405
5%	7.49554
1%	6.68461
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	1.08E6	11.1419	246113
4.62497	1.08E6	11.1419	760075
4.70953	1.11E6	11.1419	1.34E6
4.70953	1.06E6	11.1548	269915
4.81218	1.34E6	11.1548	1.26E6

The UNIVARIATE Procedure
 Variable: log_price
 title_status = salvage

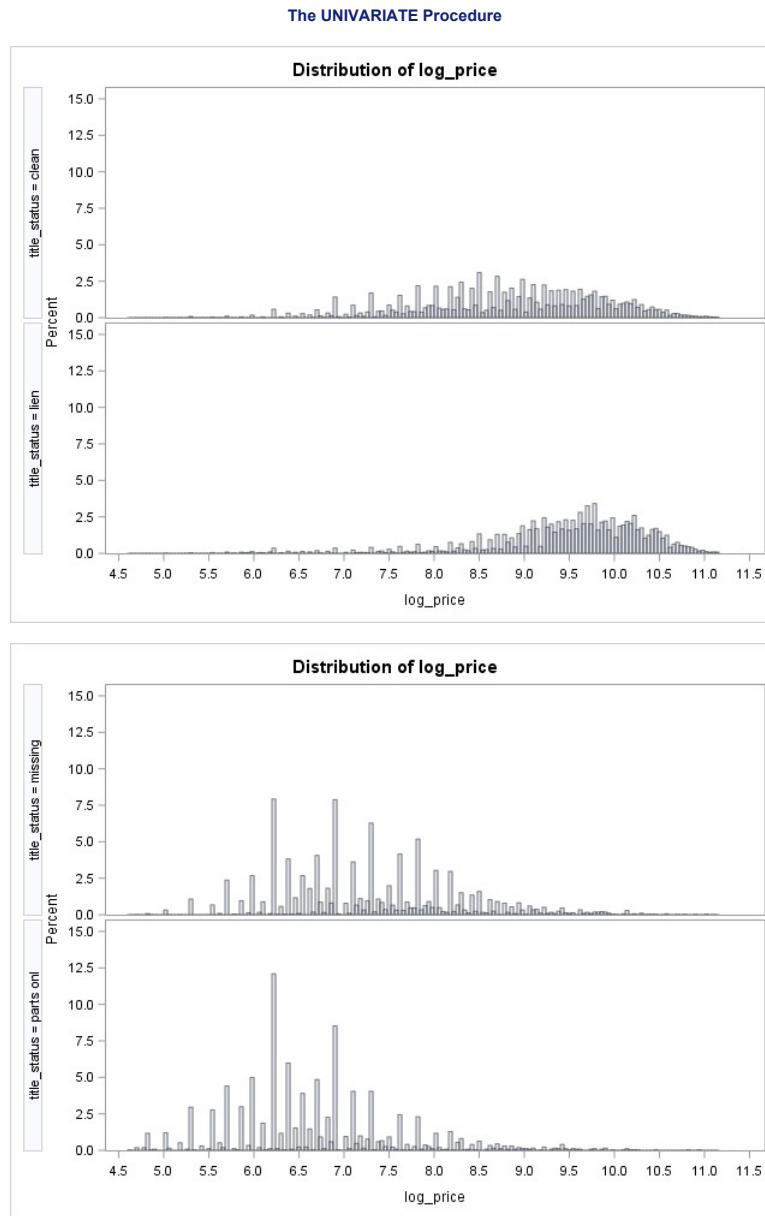
Moments			
N	27965	Sum Weights	27965
Mean	8.4462474	Sum Observations	236199.309
Std Deviation	0.93712549	Variance	0.87820418
Skewness	-0.3982401	Kurtosis	0.02919644
Uncorrected SS	2019555.9	Corrected SS	24558.1016
Coeff Variation	11.0951698	Std Error Mean	0.0056039

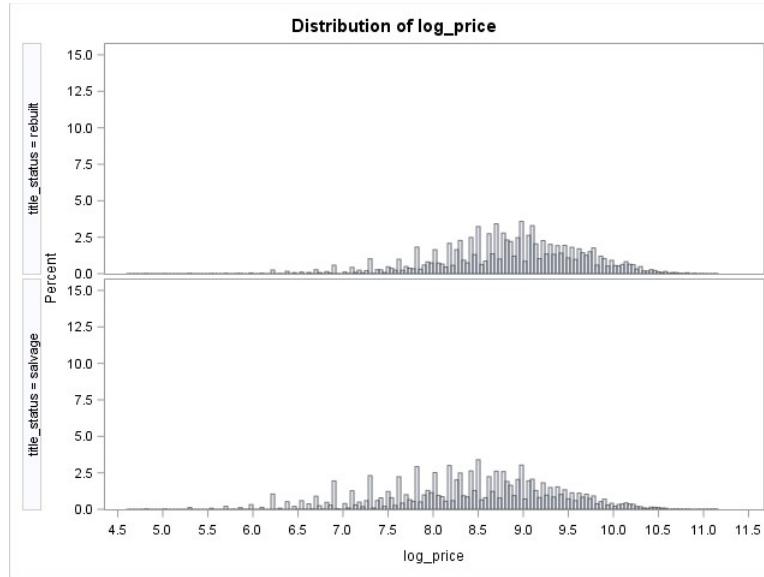
Basic Statistical Measures			
Location		Variability	
Mean	8.446247	Std Deviation	0.93713
Median	8.517193	Variance	0.87820
Mode	7.824046	Range	6.53396
		Interquartile Range	1.27536

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	1507.209	Pr > t	<.0001
Sign	M	13982.5	Pr >= M	<.0001
Signed Rank	S	1.9552E8	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.14908
99%	10.30879
95%	9.85214
90%	9.60912
75% Q3	9.09941
50% Median	8.51719
25% Q1	7.82405
10%	7.17012
5%	6.74524
1%	6.10925
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	608170	11.1491	13095
4.78749	1.12E6	11.1491	86107
4.78749	559335	11.1491	97457
4.81218	1.1E6	11.1491	897222
4.81218	1.06E6	11.1491	1.49E6





The UNIVARIATE Procedure
 Variable: log_price
 transmission = automatic

Moments			
N	1310215	Sum Weights	1310215
Mean	8.8573327	Sum Observations	11605010.2
Std Deviation	1.02666533	Variance	1.0540417
Skewness	-0.475595	Kurtosis	0.01519697
Uncorrected SS	104170456	Corrected SS	1381020.19
Coeff Variation	11.5911343	Std Error Mean	0.00089693

Basic Statistical Measures			
Location		Variability	
Mean	8.857333	Std Deviation	1.02667
Median	8.922658	Variance	1.05404
Mode	7.824046	Range	6.54112
		Interquartile Range	1.43269

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	9875.187	Pr > t	<.0001
Sign	M	655107.5	Pr >= M	<.0001
Signed Rank	S	4.292E11	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.74074
95%	10.37318
90%	10.12655
75% Q3	9.61581
50% Median	8.92266
25% Q1	8.18312
10%	7.49554
5%	6.99393
1%	6.21461
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	1.32E6	11.1562	696850
4.61512	1.32E6	11.1562	1.28E6
4.61512	1.32E6	11.1562	1.29E6
4.61512	1.08E6	11.1562	1.3E6
4.61512	608170	11.1562	1.3E6

The UNIVARIATE Procedure
 Variable: log_price
 transmission = manual

Moments			
N	158561	Sum Weights	158561
Mean	8.59155465	Sum Observations	1362285.5
Std Deviation	1.03492399	Variance	1.07106767
Skewness	-0.2167112	Kurtosis	-0.2006338
Uncorrected SS	11873978.8	Corrected SS	169828.49
Coeff Variation	12.0458291	Std Error Mean	0.00259902

Basic Statistical Measures			
Location		Variability	
Mean	8.591555	Std Deviation	1.03492
Median	8.612503	Variance	1.07107
Mode	7.824046	Range	6.54112
		Interquartile Range	1.44910

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	3305.685	Pr > t	<.0001
Sign	M	79280.5	Pr >= M	<.0001
Signed Rank	S	6.2854E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.71442
95%	10.22915
90%	9.94750
75% Q3	9.35010
50% Median	8.61250
25% Q1	7.90101
10%	7.27932
5%	6.85646
1%	6.10925
0% Min	4.61512

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.61512	238902	11.1562	869522
4.70953	646318	11.1562	1.23E6
4.74493	315672	11.1562	152431
4.78749	1.2E6	11.1562	187687
4.78749	1.07E6	11.1562	695988

The UNIVARIATE Procedure
 Variable: log_price
 transmission = other

Moments			
N	24502	Sum Weights	24502
Mean	9.20124845	Sum Observations	225448.99
Std Deviation	1.24034679	Variance	1.53846015
Skewness	-1.4666941	Kurtosis	1.55232183
Uncorrected SS	2112105.98	Corrected SS	37693.8121
Coeff Variation	13.480201	Std Error Mean	0.00792396

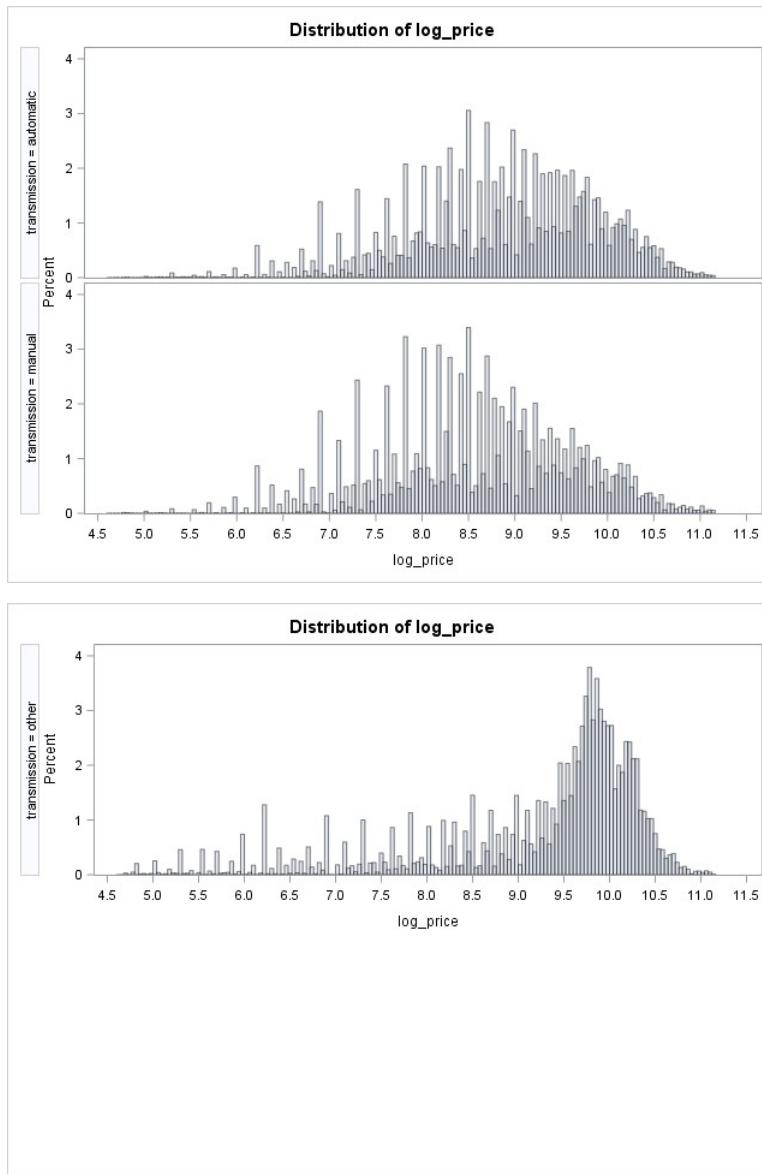
Basic Statistical Measures			
Location		Variability	
Mean	9.201248	Std Deviation	1.24035
Median	9.680031	Variance	1.53846
Mode	6.214608	Range	6.44138
		Interquartile Range	1.30833

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	1161.193	Pr > t	<.0001
Sign	M	12251	Pr >= M	<.0001
Signed Rank	S	1.5009E8	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.14186
99%	10.71431
95%	10.42754
90%	10.28534
75% Q3	10.00785
50% Median	9.68003
25% Q1	8.69951
10%	7.31322
5%	6.21461
1%	5.29832
0% Min	4.70048

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.70048	1.42E6	11.1124	265854
4.70048	1.34E6	11.1124	275766
4.70048	1.04E6	11.1124	1.12E6
4.70048	1E6	11.1415	587755
4.70048	666937	11.1419	63196

The UNIVARIATE Procedure



The UNIVARIATE Procedure
 Variable: log_price
 condition = excellent

Moments			
N	201812	Sum Weights	201812
Mean	9.17064483	Sum Observations	1850746.17
Std Deviation	0.78403596	Variance	0.61471238
Skewness	-0.6439586	Kurtosis	1.75319265
Uncorrected SS	17096591.5	Corrected SS	124055.721
Coeff Variation	8.54940926	Std Error Mean	0.00174527

Basic Statistical Measures			
Location		Variability	
Mean	9.170645	Std Deviation	0.78404
Median	9.208338	Variance	0.61471
Mode	8.411833	Range	6.41125
		Interquartile Range	1.05224

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	5254.572	Pr > t	<.0001
Sign	M	100906	Pr >= M	<.0001
Signed Rank	S	1.018E10	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15618
99%	10.71442
95%	10.35774
90%	10.13630
75% Q3	9.73495
50% Median	9.20834
25% Q1	8.68271
10%	8.21609
5%	7.93737
1%	7.09008
0% Min	4.74493

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.74493	91598	11.1562	104849
4.77912	433279	11.1562	144595
4.77912	18383	11.1562	298038
4.79579	430337	11.1562	345237
4.81218	430791	11.1562	347087

The UNIVARIATE Procedure
 Variable: log_price
 condition = fair

Moments			
N	26164	Sum Weights	26164
Mean	7.50067685	Sum Observations	196247.709
Std Deviation	0.78367213	Variance	0.614142
Skewness	0.29235374	Kurtosis	0.00975693
Uncorrected SS	1488058.44	Corrected SS	16067.7973
Coeff Variation	10.4480188	Std Error Mean	0.00484487

Basic Statistical Measures			
Location		Variability	
Mean	7.500677	Std Deviation	0.78367
Median	7.494708	Variance	0.61414
Mode	7.313220	Range	6.26996
		Interquartile Range	1.09861

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	1548.168	Pr > t	<.0001
Sign	M	13082	Pr >= M	<.0001
Signed Rank	S	1.7115E8	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.08214
99%	9.47232
95%	8.85367
90%	8.51719
75% Q3	8.00637
50% Median	7.49471
25% Q1	6.90776
10%	6.55108
5%	6.21461
1%	5.85793
0% Min	4.81218

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.81218	370741	10.6573	356284
4.81218	277716	10.7144	60990
5.01064	433798	11.0021	256093
5.01064	332714	11.0021	349188
5.01064	271456	11.0821	34165

The UNIVARIATE Procedure
 Variable: log_price
 condition = good

Moments			
N	158320	Sum Weights	158320
Mean	8.54807139	Sum Observations	1353330.66
Std Deviation	0.8398736	Variance	0.70538766
Skewness	-0.0147621	Kurtosis	-0.0696211
Uncorrected SS	11680043.4	Corrected SS	111676.27
Coeff Variation	9.82529932	Std Error Mean	0.00211079

Basic Statistical Measures			
Location		Variability	
Mean	8.548071	Std Deviation	0.83987
Median	8.516993	Variance	0.70539
Mode	7.824046	Range	6.53126
		Interquartile Range	1.10687

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	4049.693	Pr > t	<.0001
Sign	M	79160	Pr >= M	<.0001
Signed Rank	S	6.2663E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.40411
95%	9.97581
90%	9.68003
75% Q3	9.10487
50% Median	8.51699
25% Q1	7.99800
10%	7.49554
5%	7.20786
1%	6.62007
0% Min	4.62497

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.62497	365541	11.1411	377999
4.70953	159856	11.1419	71326
4.74493	85116	11.1419	218406
4.75359	359837	11.1505	78990
4.77912	82681	11.1562	55891

The UNIVARIATE Procedure
 Variable: log_price
 condition = like new

Moments			
N	51552	Sum Weights	51552
Mean	9.50449285	Sum Observations	489975.615
Std Deviation	0.83004809	Variance	0.68897983
Skewness	-0.9377766	Kurtosis	1.89271164
Uncorrected SS	4692487.33	Corrected SS	35517.5991
Coeff Variation	8.73321808	Std Error Mean	0.00365578

Basic Statistical Measures			
Location		Variability	
Mean	9.504493	Std Deviation	0.83005
Median	9.615405	Variance	0.68898
Mode	9.047821	Range	6.34405
		Interquartile Range	1.13555

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	2599.851	Pr > t	<.0001
Sign	M	25776	Pr >= M	<.0001
Signed Rank	S	6.6442E8	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.96820
95%	10.64531
90%	10.46296
75% Q3	10.12262
50% Median	9.61541
25% Q1	8.98707
10%	8.47637
5%	8.10016
1%	6.90776
0% Min	4.81218

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.81218	375525	11.1562	236142
4.81218	69614	11.1562	239235
4.96981	269973	11.1562	239416
4.99721	91595	11.1562	252610
5.01064	406090	11.1562	56794

The UNIVARIATE Procedure
 Variable: log_price
 condition = new

Moments			
N	1761	Sum Weights	1761
Mean	9.50391397	Sum Observations	16736.3925
Std Deviation	1.08017289	Variance	1.16677348
Skewness	-0.9775142	Kurtosis	1.00290607
Uncorrected SS	161114.756	Corrected SS	2053.52133
Coeff Variation	11.3655584	Std Error Mean	0.0257403

Basic Statistical Measures			
Location		Variability	
Mean	9.503914	Std Deviation	1.08017
Median	9.729134	Variance	1.16677
Mode	8.698681	Range	6.14560
		Interquartile Range	1.53016

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	369.2231	Pr > t	<.0001
Sign	M	880.5	Pr >= M	<.0001
Signed Rank	S	775720.5	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	11.08214
95%	10.86952
90%	10.70324
75% Q3	10.30895
50% Median	9.72913
25% Q1	8.77879
10%	8.15909
5%	7.31322
1%	6.07074
0% Min	5.01064

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
5.01064	373760	11.1491	432054
5.19296	193599	11.1491	434416
5.26269	9304	11.1491	435803
5.29832	73215	11.1548	65808
5.52146	368891	11.1562	199801

The UNIVARIATE Procedure
 Variable: log_price
 condition = salvage

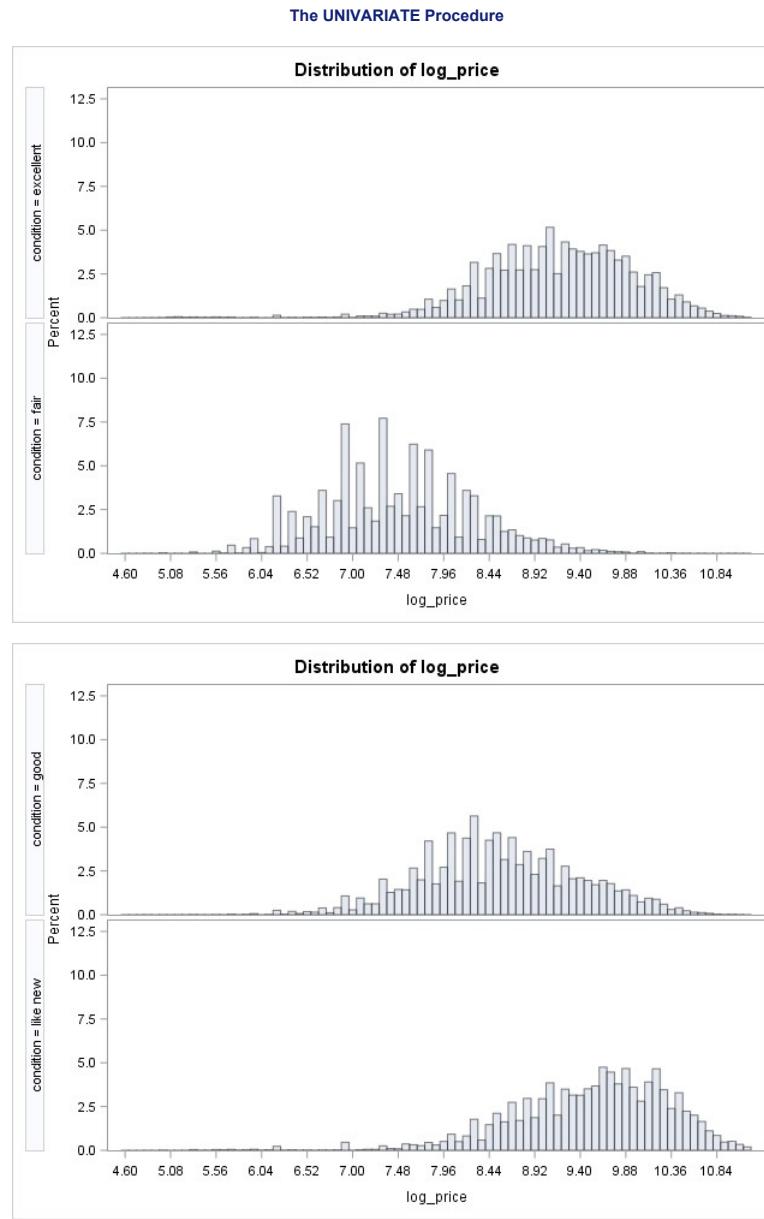
Moments			
N	1719	Sum Weights	1719
Mean	7.17486631	Sum Observations	12333.5952
Std Deviation	1.03646369	Variance	1.07425698
Skewness	0.68845436	Kurtosis	-0.1046446
Uncorrected SS	90337.4701	Corrected SS	1845.57349
Coeff Variation	14.4457561	Std Error Mean	0.02499863

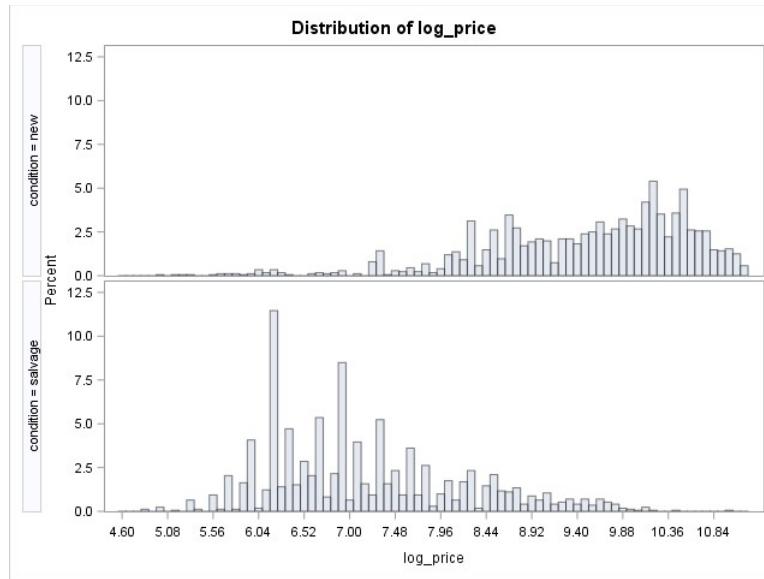
Basic Statistical Measures			
Location		Variability	
Mean	7.174866	Std Deviation	1.03646
Median	6.907755	Variance	1.07426
Mode	6.214608	Range	6.15601
		Interquartile Range	1.42712

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	287.0104	Pr > t	<.0001
Sign	M	859.5	Pr >= M	<.0001
Signed Rank	S	739170	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	10.96820
99%	9.76996
95%	9.20533
90%	8.69951
75% Q3	7.82405
50% Median	6.90776
25% Q1	6.39693
10%	5.99146
5%	5.85793
1%	5.29832
0% Min	4.81218

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.81218	196141	10.1064	231651
4.81218	192870	10.1266	307112
5.01064	214203	10.2400	102782
5.01064	209442	10.4043	146134
5.01064	98665	10.9682	167661





The UNIVARIATE Procedure
 Variable: log_price
 cylinders = 10 cylinders

Moments			
N	2248	Sum Weights	2248
Mean	9.20258287	Sum Observations	20687.4063
Std Deviation	0.81576213	Variance	0.66546785
Skewness	-0.2574251	Kurtosis	0.44173576
Uncorrected SS	191872.877	Corrected SS	1495.30625
Coeff Variation	8.86449096	Std Error Mean	0.01720542

Basic Statistical Measures			
Location		Variability	
Mean	9.202583	Std Deviation	0.81576
Median	9.185023	Variance	0.66547
Mode	8.612503	Range	5.72676
		Interquartile Range	1.11737

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	534.8652	Pr > t	<.0001
Sign	M	1124	Pr >= M	<.0001
Signed Rank	S	1263938	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.09739
99%	10.80973
95%	10.54402
90%	10.30892
75% Q3	9.76996
50% Median	9.18502
25% Q1	8.65259
10%	8.23616
5%	7.97247
1%	7.17012
0% Min	5.37064

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
5.37064	77347	11.0509	367063
5.52146	61328	11.0666	394578
5.85793	113855	11.0700	242258
5.91080	341960	11.0744	394823
6.21461	297529	11.0974	287054

The UNIVARIATE Procedure
 Variable: log_price
 cylinders = 12 cylinders

Moments			
N	281	Sum Weights	281
Mean	9.08169517	Sum Observations	2551.95634
Std Deviation	0.88502133	Variance	0.78326275
Skewness	-0.0853058	Kurtosis	-0.2111983
Uncorrected SS	23395.4032	Corrected SS	219.313571
Coeff Variation	9.74511159	Std Error Mean	0.05279595

Basic Statistical Measures			
Location		Variability	
Mean	9.081695	Std Deviation	0.88502
Median	9.047821	Variance	0.78326
Mode	8.699515	Range	4.74493
		Interquartile Range	1.31464

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t 172.015	Pr > t	<.0001
Sign	M 140.5	Pr >= M	<.0001
Signed Rank	S 19810.5	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.14186
99%	11.10646
95%	10.46310
90%	10.25766
75% Q3	9.76996
50% Median	9.04782
25% Q1	8.45532
10%	7.93737
5%	7.60090
1%	7.00307
0% Min	6.39693

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
6.39693	381490	11.0509	7113
6.85646	388788	11.0821	171431
7.00307	354804	11.1065	355796
7.09008	357381	11.1419	156439
7.24423	360859	11.1419	368399

The UNIVARIATE Procedure
 Variable: log_price
 cylinders = 3 cylinders

Moments			
N	530	Sum Weights	530
Mean	8.45711165	Sum Observations	4482.26917
Std Deviation	1.01177256	Variance	1.02368371
Skewness	-0.7160143	Kurtosis	0.7851244
Uncorrected SS	38448.5795	Corrected SS	541.528684
Coeff Variation	11.963571	Std Error Mean	0.04394859

Basic Statistical Measures			
Location		Variability	
Mean	8.457112	Std Deviation	1.01177
Median	8.612503	Variance	1.02368
Mode	7.090077	Range	5.71775
		Interquartile Range	1.13118

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	192.4319	Pr > t	<.0001
Sign	M	265	Pr >= M	<.0001
Signed Rank	S	70357.5	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	10.75471
99%	10.35774
95%	10.03889
90%	9.61574
75% Q3	9.10365
50% Median	8.61250
25% Q1	7.97247
10%	7.09008
5%	6.55108
1%	5.16479
0% Min	5.03695

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
5.03695	430437	10.4602	40478
5.11799	437809	10.5051	310274
5.11799	433175	10.5187	76874
5.11799	18388	10.5941	83198
5.12396	430368	10.7547	10965

The UNIVARIATE Procedure
 Variable: log_price
 cylinders = 4 cylinders

Moments			
N	128419	Sum Weights	128419
Mean	8.68497185	Sum Observations	1115315.4
Std Deviation	0.88585885	Variance	0.78474591
Skewness	-0.6665132	Kurtosis	0.62787152
Uncorrected SS	9787258.35	Corrected SS	100775.5
Coeff Variation	10.1999047	Std Error Mean	0.00247201

Basic Statistical Measures			
Location		Variability	
Mean	8.684972	Std Deviation	0.88586
Median	8.779557	Variance	0.78475
Mode	7.824046	Range	6.53126
		Interquartile Range	1.16315

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	3513.327	Pr > t	<.0001
Sign	M	64209.5	Pr >= M	<.0001
Signed Rank	S	4.1229E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.30862
95%	9.92818
90%	9.73979
75% Q3	9.32367
50% Median	8.77956
25% Q1	8.16052
10%	7.52294
5%	7.09008
1%	6.21461
0% Min	4.62497

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.62497	365541	11.1346	371222
4.74493	91598	11.1419	433106
4.77912	18383	11.1548	235208
4.78749	61480	11.1548	235436
4.79579	430337	11.1562	55891

The UNIVARIATE Procedure
Variable: log_price
cylinders = 5 cylinders

Moments			
N	4543	Sum Weights	4543
Mean	8.4828423	Sum Observations	38537.5526
Std Deviation	0.7505778	Variance	0.56336703
Skewness	-0.5018728	Kurtosis	0.88872165
Uncorrected SS	329466.794	Corrected SS	2558.81307
Coeff Variation	8.84818759	Std Error Mean	0.01113587

Basic Statistical Measures			
Location		Variability	
Mean	8.482842	Std Deviation	0.75058
Median	8.517193	Variance	0.56337
Mode	7.824046	Range	5.64837
		Interquartile Range	0.91504

Note: The mode displayed is the smallest of 2 modes with a count of 110.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	761.758	Pr > t 	<.0001
Sign	M	2271.5	Pr >= M 	<.0001
Signed Rank	S	5160848	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.00023
99%	10.18490
95%	9.58190
90%	9.30556
75% Q3	8.98595
50% Median	8.51719
25% Q1	8.07091
10%	7.49554
5%	7.09008
1%	6.30992
0% Min	5.35186

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
5.35186	360706	10.7144	91396
5.35186	354210	10.7144	139321
5.35186	354068	10.7364	370046
5.50939	155417	10.9151	9114
5.52146	275154	11.0002	433671

The UNIVARIATE Procedure
 Variable: log_price
 cylinders = 6 cylinders

Moments			
N	158218	Sum Weights	158218
Mean	8.79071325	Sum Observations	1390849.07
Std Deviation	0.98135741	Variance	0.96306237
Skewness	-0.2959257	Kurtosis	-0.1372369
Uncorrected SS	12378928.2	Corrected SS	152372.84
Coeff Variation	11.1635698	Std Error Mean	0.00246717

Basic Statistical Measures			
Location		Variability	
Mean	8.790713	Std Deviation	0.98136
Median	8.809564	Variance	0.96306
Mode	7.824046	Range	6.37711
		Interquartile Range	1.38544

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	3563.074	Pr > t	<.0001
Sign	M	79109	Pr >= M	<.0001
Signed Rank	S	6.2583E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.65726
95%	10.30561
90%	10.08560
75% Q3	9.54596
50% Median	8.80956
25% Q1	8.16052
10%	7.54961
5%	7.09008
1%	6.21461
0% Min	4.77912

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.77912	433279	11.1491	326587
4.77912	82681	11.1548	171189
4.80402	169606	11.1562	104504
4.80402	155614	11.1562	104849
4.81218	430791	11.1562	56794

The UNIVARIATE Procedure
 Variable: log_price
 cylinders = 8 cylinders

Moments			
N	133876	Sum Weights	133876
Mean	9.13416851	Sum Observations	1222845.94
Std Deviation	0.93430502	Variance	0.87292587
Skewness	-0.4907605	Kurtosis	0.06762066
Uncorrected SS	11286543.9	Corrected SS	116862.951
Coeff Variation	10.2286816	Std Error Mean	0.00255351

Basic Statistical Measures			
Location		Variability	
Mean	9.134169	Std Deviation	0.93431
Median	9.209840	Variance	0.87293
Mode	8.160518	Range	6.44671
		Interquartile Range	1.30853

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	3577.105	Pr > t	<.0001
Sign	M	66938	Pr >= M	<.0001
Signed Rank	S	4.4807E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.83958
95%	10.50507
90%	10.27505
75% Q3	9.82553
50% Median	9.20984
25% Q1	8.51699
10%	7.88796
5%	7.49554
1%	6.68461
0% Min	4.70953

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.70953	159856	11.1562	252610
4.74493	85116	11.1562	298038
4.75359	359837	11.1562	345237
4.78749	202298	11.1562	347087
4.81218	217430	11.1562	199801

The UNIVARIATE Procedure
 Variable: log_price
 cylinders = other

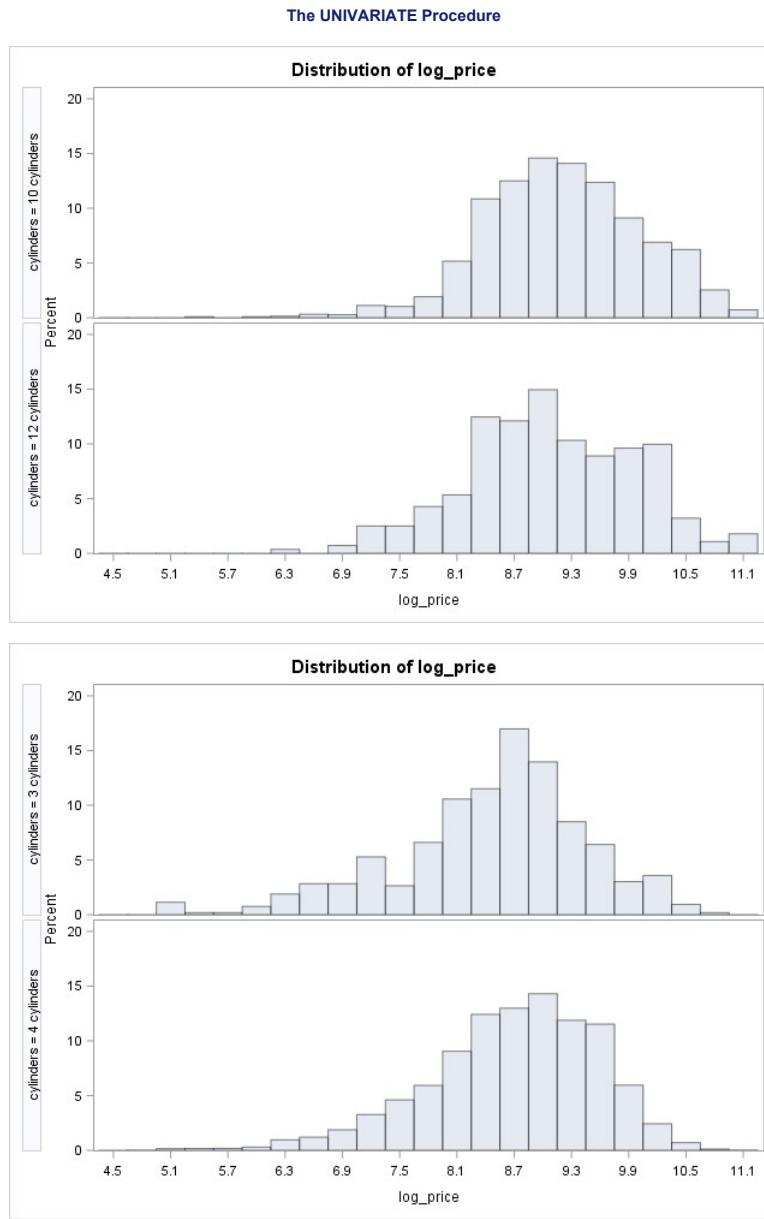
Moments			
N	13213	Sum Weights	13213
Mean	9.39230693	Sum Observations	124100.551
Std Deviation	0.73453463	Variance	0.53954112
Skewness	-0.8068939	Kurtosis	2.00651351
Uncorrected SS	1172718.89	Corrected SS	7128.41724
Coeff Variation	7.82059862	Std Error Mean	0.00639016

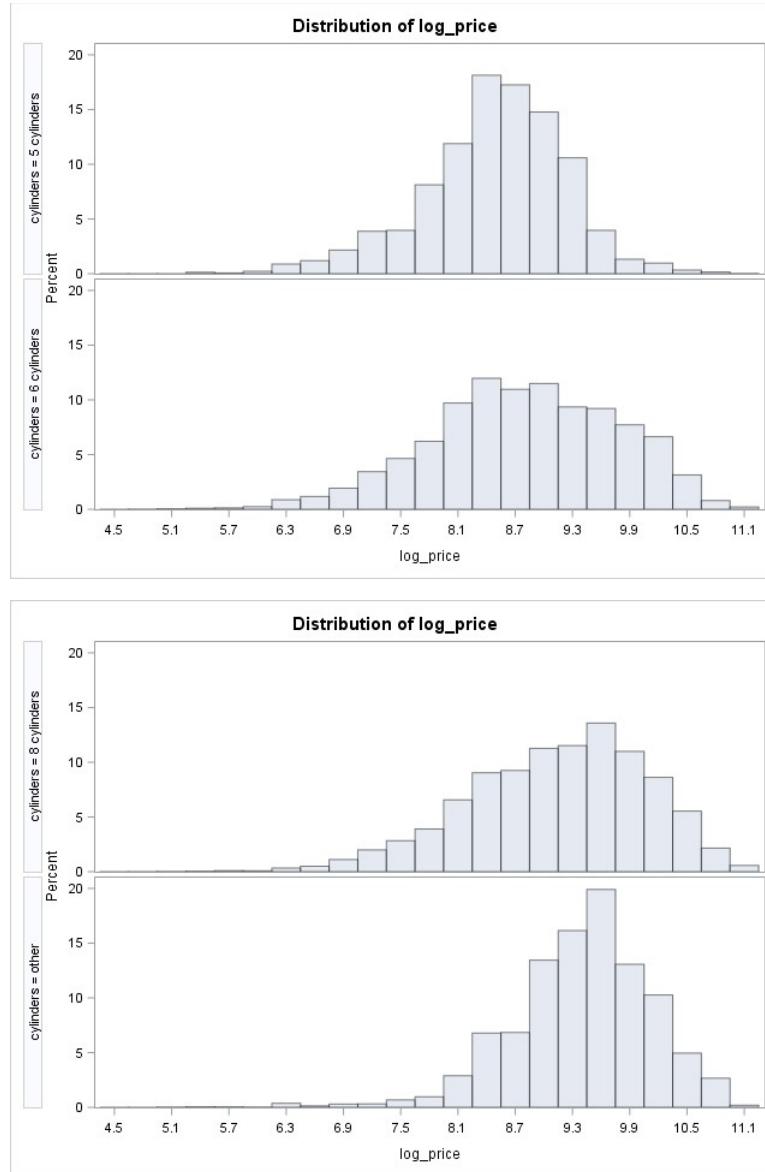
Basic Statistical Measures			
Location		Variability	
Mean	9.392307	Std Deviation	0.73453
Median	9.470780	Variance	0.53954
Mode	9.209840	Range	5.93778
		Interquartile Range	0.91315

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	1469.808	Pr > t	<.0001
Sign	M	6606.5	Pr >= M	<.0001
Signed Rank	S	43649146	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.14179
99%	10.81968
95%	10.49114
90%	10.23992
75% Q3	9.89848
50% Median	9.47078
25% Q1	8.98532
10%	8.51519
5%	8.18730
1%	7.16627
0% Min	5.20401

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
5.20401	30154	11.0898	125786
5.21494	30153	11.1418	64175
5.24175	358605	11.1418	70604
5.28827	356272	11.1418	317013
5.42935	356476	11.1418	318519





The UNIVARIATE Procedure
Variable: log_price
drive = 4wd

Moments			
N	180454	Sum Weights	180454
Mean	9.19098056	Sum Observations	1658549.21
Std Deviation	0.9308303	Variance	0.86644505
Skewness	-0.6206143	Kurtosis	0.30095242
Uncorrected SS	15400046.1	Corrected SS	156352.61
Coeff Variation	10.1276496	Std Error Mean	0.00219123

Basic Statistical Measures			
Location		Variability	
Mean	9.190981	Std Deviation	0.93083
Median	9.301095	Variance	0.86645
Mode	8.160518	Range	6.44671
		Interquartile Range	1.28939

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	4194.446	Pr > t	<.0001
Sign	M	90227	Pr >= M	<.0001
Signed Rank	S	8.141E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.81778
95%	10.49127
90%	10.30728
75% Q3	9.90098
50% Median	9.30109
25% Q1	8.61159
10%	7.97247
5%	7.54961
1%	6.68461
0% Min	4.70953

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.70953	159856	11.1562	236142
4.74493	85116	11.1562	239235
4.77912	433279	11.1562	239416
4.78749	61480	11.1562	298038
4.81218	375544	11.1562	56794

The UNIVARIATE Procedure
Variable: log_price
drive = fwd

Moments			
N	161004	Sum Weights	161004
Mean	8.5401619	Sum Observations	1375000.23
Std Deviation	0.87163919	Variance	0.75975488
Skewness	-0.5569099	Kurtosis	0.42516481
Uncorrected SS	11865047.4	Corrected SS	122322.815
Coeff Variation	10.2063544	Std Error Mean	0.00217229

Basic Statistical Measures			
Location		Variability	
Mean	8.540162	Std Deviation	0.87164
Median	8.612503	Variance	0.75975
Mode	7.824046	Range	6.53126
		Interquartile Range	1.15335

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	3931.404	Pr > t	<.0001
Sign	M	80502	Pr >= M	<.0001
Signed Rank	S	6.4806E9	Pr >= S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.18490
95%	9.79807
90%	9.61414
75% Q3	9.15905
50% Median	8.61250
25% Q1	8.00570
10%	7.37776
5%	6.90776
1%	6.21461
0% Min	4.62497

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.62497	365541	11.1491	279727
4.74493	91598	11.1491	432054
4.77912	82681	11.1491	434416
4.77912	18383	11.1491	435803
4.79579	430337	11.1562	55891

The UNIVARIATE Procedure
 Variable: log_price
 drive = rwd

Moments			
N	99870	Sum Weights	99870
Mean	8.86973782	Sum Observations	885820.716
Std Deviation	0.93088919	Variance	0.86655468
Skewness	-0.376743	Kurtosis	0.06369345
Uncorrected SS	7943539.45	Corrected SS	86541.9494
Coeff Variation	10.495115	Std Error Mean	0.00294565

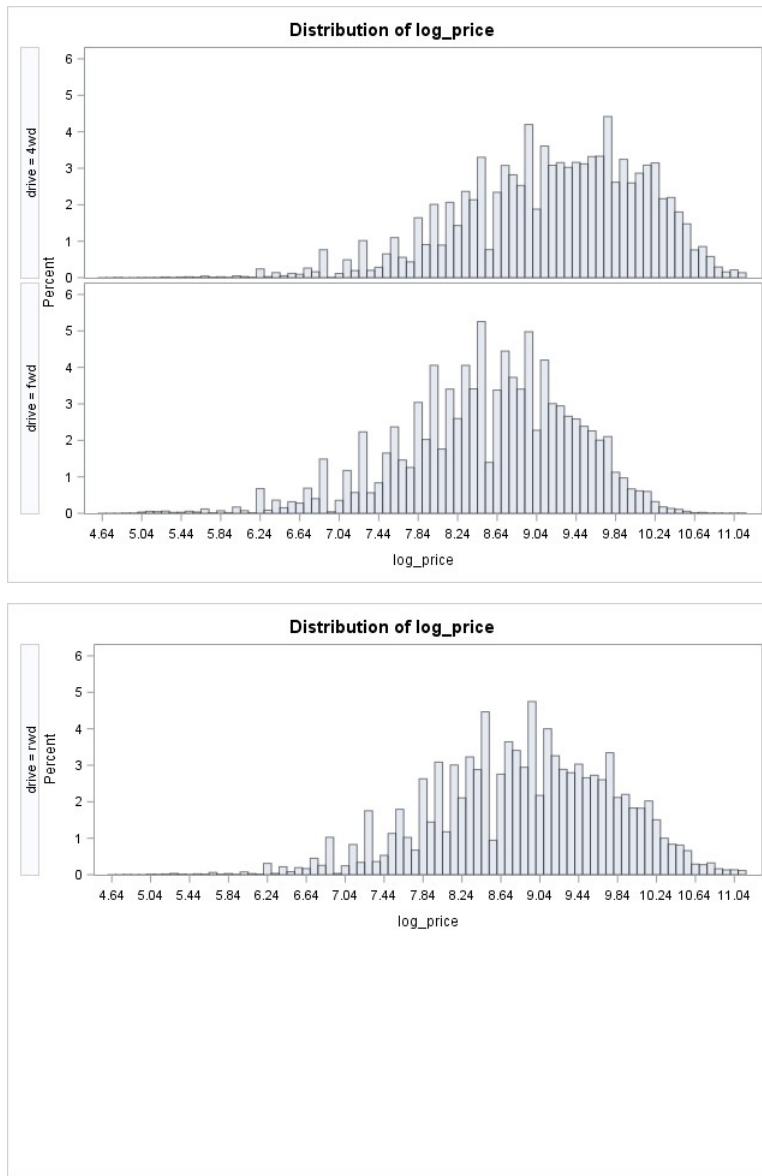
Basic Statistical Measures			
Location		Variability	
Mean	8.869738	Std Deviation	0.93089
Median	8.922658	Variance	0.86655
Mode	8.160518	Range	6.40265
		Interquartile Range	1.27808

Tests for Location: Mu0=0				
Test		Statistic		p Value
Student's t	t	3011.136	Pr > t 	<.0001
Sign	M	49935	Pr >= M 	<.0001
Signed Rank	S	2.4935E9	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	11.15624
99%	10.71442
95%	10.27160
90%	10.04297
75% Q3	9.54681
50% Median	8.92266
25% Q1	8.26873
10%	7.60090
5%	7.31322
1%	6.47697
0% Min	4.75359

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
4.75359	359837	11.1562	232297
4.78749	202298	11.1562	252610
4.80402	169606	11.1562	345237
4.80402	155614	11.1562	347087
4.81218	430791	11.1562	199801

The UNIVARIATE Procedure

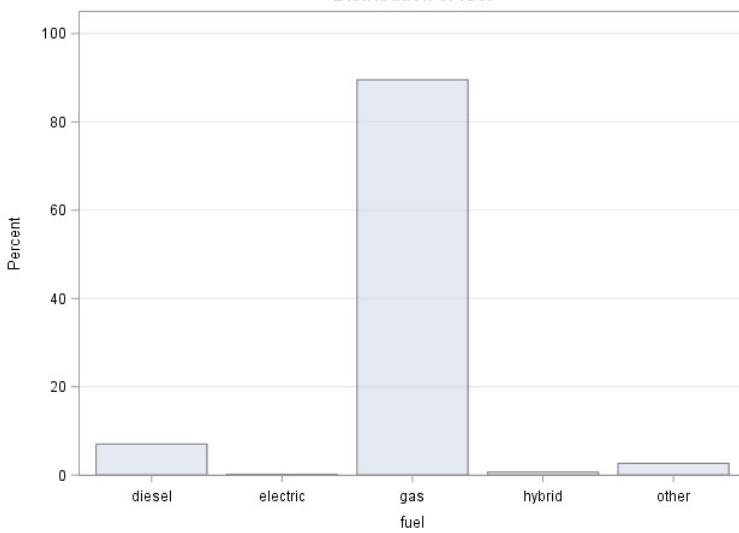


The FREQ Procedure

fuel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
diesel	114241	7.02	114241	7.02
electric	2181	0.13	116422	7.16
gas	1456237	89.54	1572659	96.70
hybrid	10553	0.65	1583212	97.35
other	43100	2.65	1626312	100.00

Frequency Missing = 9915

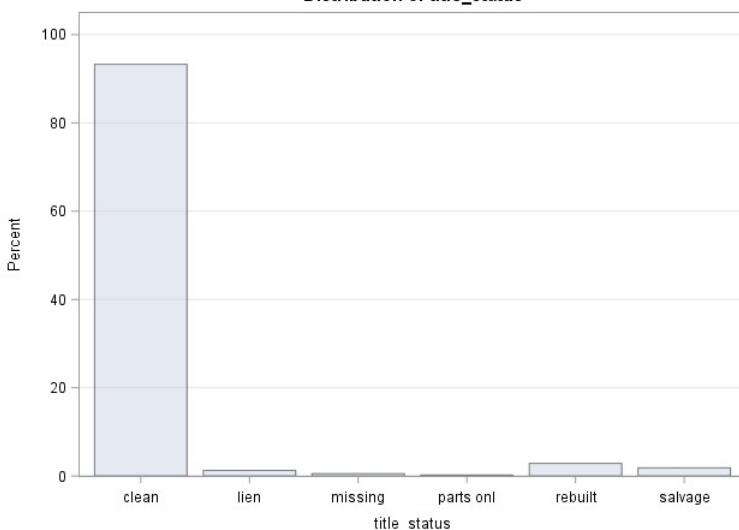
Distribution of fuel



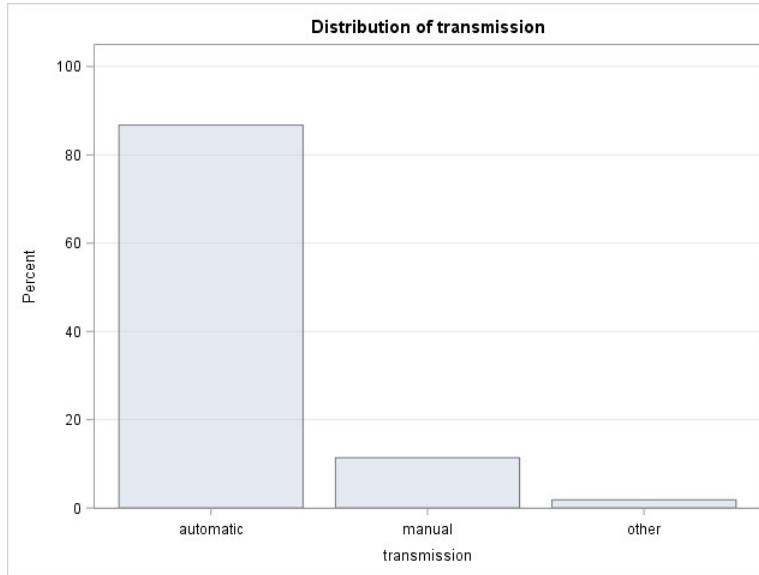
title_status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
clean	1523614	93.26	1523614	93.26
lien	20845	1.28	1544459	94.54
missing	8870	0.54	1553329	95.08
parts onl	3558	0.22	1556887	95.30
rebuilt	46617	2.85	1603504	98.15
salvage	30208	1.85	1633712	100.00

Frequency Missing = 2515

Distribution of title_status

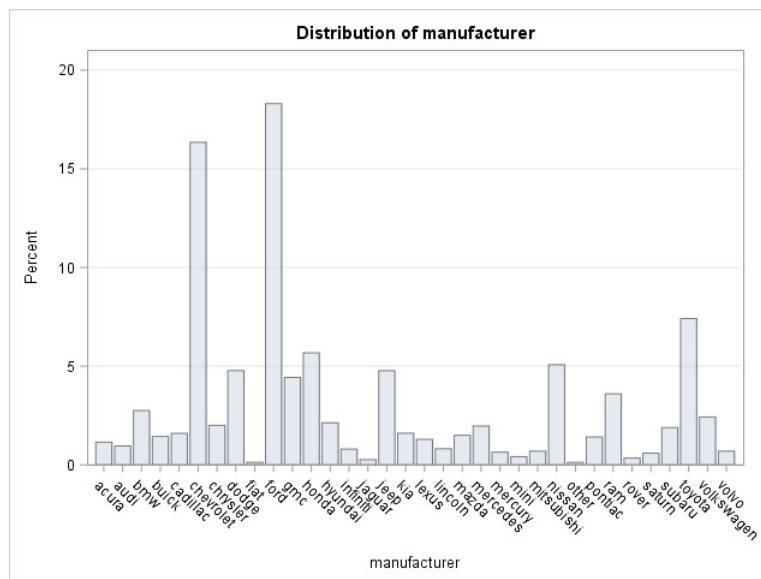


transmission	Frequency	Percent	Cumulative Frequency	Cumulative Percent
automatic	1411711	86.74	1411711	86.74
manual	185549	11.40	1597260	98.14
other	30231	1.86	1627491	100.00
Frequency Missing = 8736				

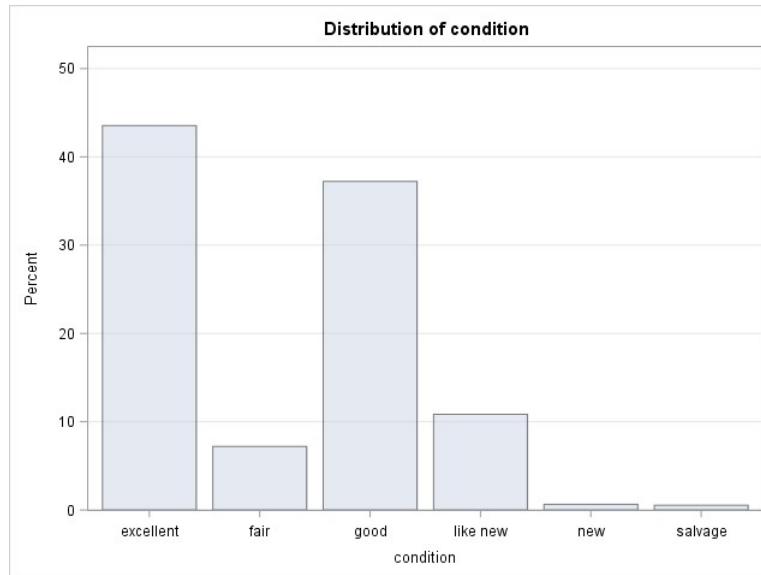


manufacturer	Frequency	Percent	Cumulative Frequency	Cumulative Percent
acura	17416	1.15	17416	1.15
audi	14524	0.96	31940	2.11
bmw	41732	2.75	73672	4.86
buick	21789	1.44	95461	6.30
cadillac	24167	1.59	119628	7.89
chevrolet	247737	16.35	367365	24.24
chrysler	30330	2.00	397695	26.24
dodge	72392	4.78	470087	31.02
fiat	1882	0.12	471969	31.14
ford	277397	18.31	749366	49.45
gmc	67145	4.43	816511	53.88
honda	86023	5.68	902534	59.56
hyundai	32286	2.13	934820	61.69
infiniti	12093	0.80	946913	62.49
jaguar	3997	0.26	950910	62.75
jeep	72333	4.77	1023243	67.52
kia	24266	1.60	1047509	69.12
lexus	19601	1.29	1067110	70.42
lincoln	12365	0.82	1079475	71.23
mazda	22674	1.50	1102149	72.73
mercedes	29788	1.97	1131937	74.70
mercury	9713	0.64	1141650	75.34
mini	6265	0.41	1147915	75.75
mitsubishi	10470	0.69	1158385	76.44
nissan	76920	5.08	1235305	81.52
other	1872	0.12	1237177	81.64
pontiac	21384	1.41	1258561	83.05
ram	54540	3.60	1313101	86.65
rover	5154	0.34	1318255	86.99
saturn	8973	0.59	1327228	87.58
subaru	28601	1.89	1355829	89.47

toyota	112361	7.41	1468190	96.88
volkswagen	36712	2.42	1504902	99.31
volvo	10507	0.69	1515409	100.00
Frequency Missing = 120818				

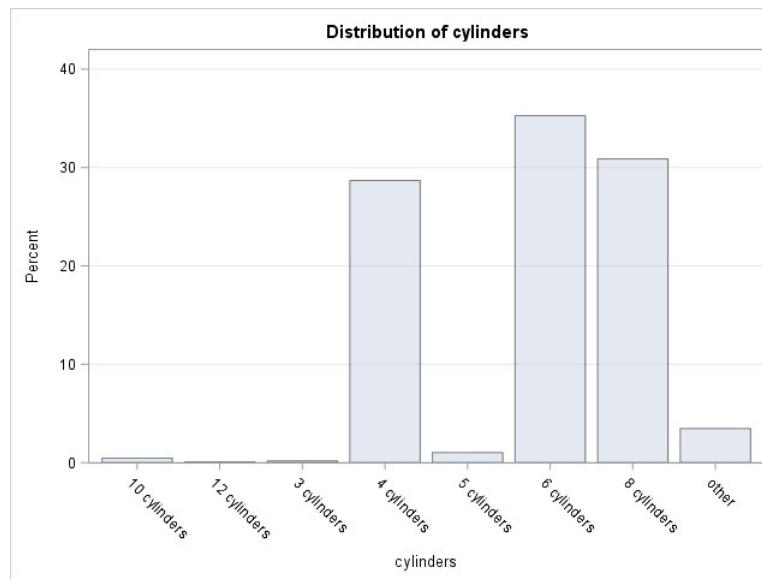


condition	Frequency	Percent	Cumulative Frequency	Cumulative Percent
excellent	422244	43.55	422244	43.55
fair	69834	7.20	492078	50.75
good	360895	37.22	852973	87.97
like new	105124	10.84	958097	98.81
new	6271	0.65	964368	99.46
salvage	5233	0.54	969601	100.00
Frequency Missing = 666626				

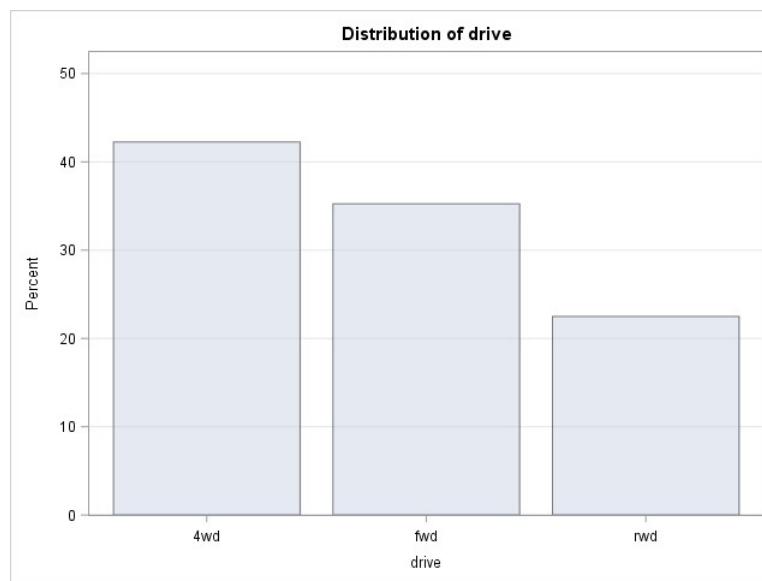


cylinders	Frequency	Percent	Cumulative Frequency	Cumulative Percent
10 cylinders	4383	0.45	4383	0.45
12 cylinders	651	0.07	5034	0.51
3 cylinders	1700	0.17	6734	0.69
4 cylinders	281654	28.68	288388	29.37

5 cylinders	10025	1.02	298413	30.39
6 cylinders	346366	35.27	644779	65.66
8 cylinders	303249	30.88	948028	96.54
other	34014	3.46	982042	100.00
Frequency Missing = 654185				

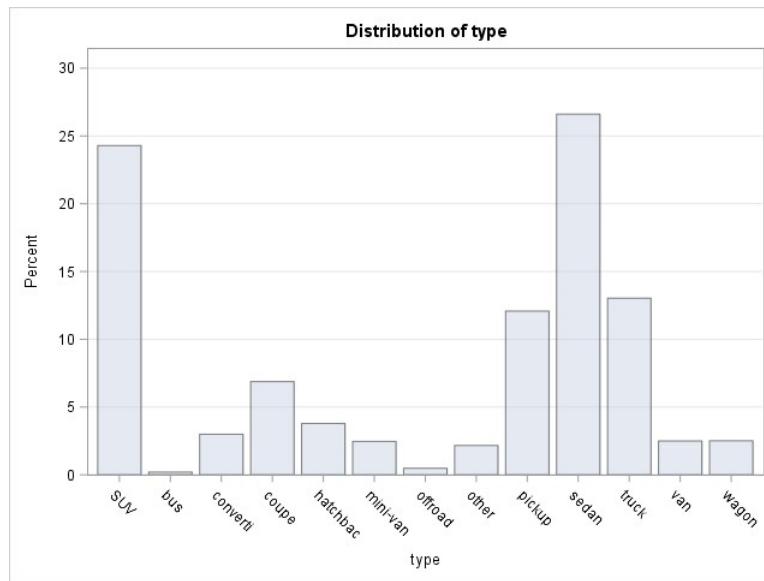


drive	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4wd	429128	42.25	429128	42.25
fwd	358030	35.25	787158	77.51
rwd	228442	22.49	1015600	100.00
Frequency Missing = 620627				

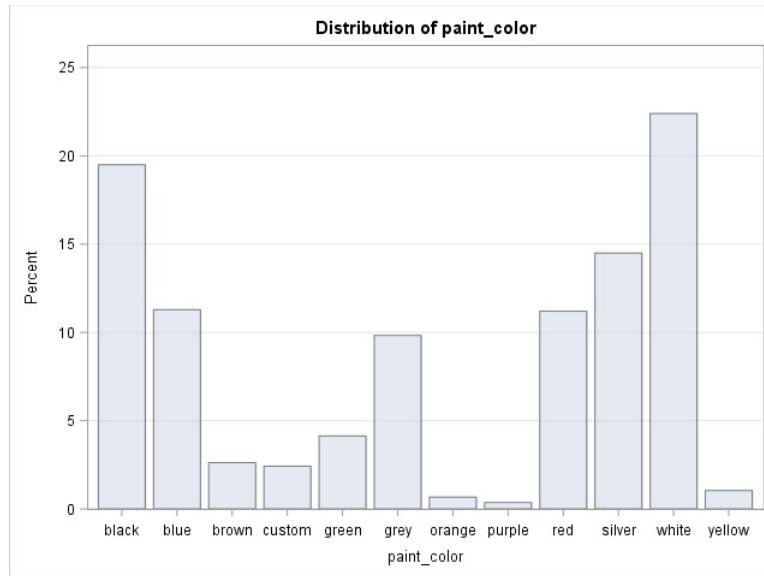


type	Frequency	Percent	Cumulative Frequency	Cumulative Percent
SUV	237461	24.29	237461	24.29
bus	1952	0.20	239413	24.49
converti	29267	2.99	268680	27.49
coupe	67229	6.88	335909	34.36
hatchbac	37066	3.79	372975	38.15
mini-van	24108	2.47	397083	40.62

offroad	4712	0.48	401795	41.10
other	21141	2.16	422936	43.27
pickup	118086	12.08	541022	55.35
sedan	260127	26.61	801149	81.96
truck	127399	13.03	928548	94.99
van	24427	2.50	952975	97.49
wagon	24552	2.51	977527	100.00
Frequency Missing = 658700				



paint_color	Frequency	Percent	Cumulative Frequency	Cumulative Percent
black	191832	19.50	191832	19.50
blue	111105	11.29	302937	30.79
brown	25803	2.62	328740	33.41
custom	23860	2.43	352600	35.84
green	40712	4.14	393312	39.98
grey	96722	9.83	490034	49.81
orange	6658	0.68	496692	50.48
purple	3627	0.37	500319	50.85
red	110222	11.20	610541	62.06
silver	142611	14.50	753152	76.55
white	220399	22.40	973551	98.95
yellow	10300	1.05	983851	100.00
Frequency Missing = 652376				



The CORR Procedure

2 Variables: log_price log_age

Pearson Correlation Coefficients, N = 1493278		
	Prob > r under H0: Rho=0	
	log_price	log_age
log_price	1.00000	-0.57035 <.0001
log_age	-0.57035 <.0001	1.00000

The CORR Procedure

3 Variables: log_price log_age log_odometer

Pearson Correlation Coefficients, N = 441328 Prob > r under H0: Rho=0			
	log_price	log_age	log_odometer
log_price	1.00000 <.0001	-0.60279 <.0001	-0.51291 <.0001
log_age	-0.60279 <.0001	1.00000	0.55334 <.0001
log_odometer	-0.51291 <.0001	0.55334 <.0001	1.00000

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of fuel by title_status						
	fuel	title_status					
		clean	lien	missing	parts onl	rebuilt	salvage
diesel	91183	2594		165	111	1573	1098
	6.11	0.17		0.01	0.01	0.11	0.07
	94.27	2.68		0.17	0.11	1.63	1.14
	6.55	13.37		2.39	4.04	3.58	3.93
electric	1144	41		7	2	10	14
	0.08	0.00		0.00	0.00	0.00	0.00
	93.92	3.37		0.57	0.16	0.82	1.15
	0.08	0.21		0.10	0.07	0.02	0.05
gas	1254312	16551		6525	2328	41818	26485
	84.00	1.11		0.44	0.16	2.80	1.77
	93.05	1.23		0.48	0.17	3.10	1.96
	90.09	85.28		94.35	84.81	95.17	94.71
hybrid	9134	183		8	11	435	321
	0.61	0.01		0.00	0.00	0.03	0.02
	90.51	1.81		0.08	0.11	4.31	3.18
	0.66	0.94		0.12	0.40	0.99	1.15
other	36531	39		211	293	104	47
	2.45	0.00		0.01	0.02	0.01	0.00
	98.14	0.10		0.57	0.79	0.28	0.13
	2.62	0.20		3.05	10.67	0.24	0.17
Total		1392304	19408	6916	2745	43940	27965
		93.24	1.30	0.46	0.18	2.94	1.87
							100.00

Statistics for Table of fuel by title_status

Statistic	DF	Value	Prob
Chi-Square	20	5781.8500	<.0001
Likelihood Ratio Chi-Square	20	6632.9706	<.0001
Mantel-Haenszel Chi-Square	1	26.4642	<.0001
Phi Coefficient		0.0622	
Contingency Coefficient		0.0621	
Cramer's V		0.0311	

Sample Size = 1493278

Frequency Percent Row Pct Col Pct	Table of fuel by transmission				
	fuel	transmission			
		automatic	manual	other	Total
diesel	84167	11675	882	96724	
	5.64	0.78	0.06	6.48	
	87.02	12.07	0.91		
	6.42	7.36	3.60		
electric	977	37	204	1218	
	0.07	0.00	0.01	0.08	
	80.21	3.04	16.75		
	0.07	0.02	0.83		
gas	1183570	144243	20206	1348019	
	79.26	9.66	1.35	90.27	
	87.80	10.70	1.50		
	90.33	90.97	82.47		
hybrid	9681	194	217	10092	
	0.65	0.01	0.01	0.68	
	95.93	1.92	2.15		
	0.74	0.12	0.89		
other	31820	2412	2993	37225	
	2.13	0.16	0.20	2.49	
	85.48	6.48	8.04		
	2.43	1.52	12.22		
Total		1310215	158561	24502	1493278
		87.74	10.62	1.64	100.00

Statistics for Table of fuel by transmission

Statistic	DF	Value	Prob

Chi-Square	8	13178.6336	<.0001
Likelihood Ratio Chi-Square	8	8073.7620	<.0001
Mantel-Haenszel Chi-Square	1	355.8534	<.0001
Phi Coefficient		0.0939	
Contingency Coefficient		0.0935	
Cramer's V		0.0664	

Sample Size = 1493278

Frequency Percent Row Pct Col Pct	fuel	Table of fuel by manufac																			
		acura	audi	bmw	buick	cadillac	chevrolet	chrysler	dodge	fiat	ford	gmc	honda	hyundai	infiniti	jaguar	jeep	kia	lexus	lin	
	diesel	17 0.00	329 0.02	1104 0.07	28 0.00	27 0.00	13757 0.92	42 0.00	5457 0.37	0 0.00	38112 2.55	7263 0.49	99 0.01	29 0.00	18 0.00	4 0.00	471 0.03	18 0.00	25 0.00	1	
		0.02 0.34	0.02 1.14	0.03 0.03	0.13 0.03	0.11 0.11	14.22 5.62	0.04 0.14	5.64 7.62	0.00 0.00	39.40 13.94	7.51 10.99	0.10 0.12	0.03 0.09	0.02 0.15	0.00 0.10	0.49 0.66	0.02 0.08	0.03 0.13	0.00 0.13	1
	electric	1 0.00	0 0.00	183 0.01	1 0.00	1 0.00	269 0.02	3 0.00	3 0.01	96 0.01	105 0.01	5 0.00	5 0.00	1 0.00	1 0.00	0 0.00	4 0.00	22 0.00	1 0.00	1 0.00	1
		0.08 0.01	0.08 0.00	15.02 0.44	0.08 0.00	0.08 0.00	22.09 0.11	0.25 0.01	0.25 0.00	7.88 5.17	8.62 0.04	0.41 0.01	0.41 0.01	0.08 0.01	0.08 0.01	0.00 0.01	0.33 0.01	1.81 0.09	0.08 0.01	0.00 0.01	1
	gas	16785 1.12	13656 0.91	38788 2.60	21047 1.41	23142 1.55	223782 14.99	29199 1.96	64308 4.31	1724 0.12	226375 15.16	56697 3.80	81779 5.48	30578 2.05	11558 0.77	3901 0.26	68540 4.59	23085 1.55	18144 1.22	11 1	
		1.25 1.01	1.25 1.01	2.88 2.88	1.56 1.56	1.72 1.72	16.60 5.31	2.17 0.19	4.77 0.17	0.13 0.00	16.79 12.48	4.21 1.39	6.07 10.25	2.27 2.41	0.86 0.43	0.29 0.00	5.08 0.17	1.71 1.77	1.35 6.15	1 1	
		97.33 95.30	97.33 95.30	94.19 94.19	97.63 97.63	97.10 97.10	91.50 97.32	97.32 89.84	92.79 89.84	82.77 85.77	85.77 96.46	96.46 96.19	96.78 96.78	98.56 98.56	96.28 96.28	96.79 96.79	93.79 93.79	91 91			
	hybrid	17 0.00	15 0.00	86 0.01	55 0.00	35 0.00	536 0.04	19 0.00	17 0.00	0 0.00	1259 0.08	140 0.01	1034 0.07	243 0.02	43 0.00	0 0.00	17 0.00	179 0.01	621 0.04	1 1	
		0.07 0.15	0.07 0.15	0.85 0.85	0.54 0.54	0.35 0.35	5.31 0.22	0.19 0.06	0.17 0.00	0.00 0.00	12.48 2.51	1.39 2.05	10.25 2.79	2.41 3.02	0.43 0.21	0.00 1.22	0.17 0.76	1.77 0.36	6.15 0.00	3.21 0.75	1 1
	other	426 0.03	329 0.02	1020 0.07	428 0.03	628 0.04	6239 0.42	741 0.05	1799 0.12	38 0.00	7634 0.51	1999 0.13	1866 0.12	937 0.06	323 0.02	53 0.00	2154 0.14	546 0.04	554 0.04	1 1	
		1.14 2.47	0.88 2.30	2.74 2.48	1.15 1.99	1.69 2.64	16.76 2.55	1.99 2.47	4.83 2.51	0.10 2.05	20.51 2.79	5.37 3.02	5.01 2.20	2.52 2.95	0.87 2.70	0.14 1.34	5.79 3.03	1.47 2.29	1.49 2.86	1 1	
	Total	17246 1.15	14329 0.96	41181 2.76	21559 1.44	23833 1.60	244583 16.38	30004 2.01	71584 4.79	1858 0.12	273485 18.31	66104 4.43	84783 5.68	31788 2.13	11943 0.80	3958 0.27	71186 4.77	23850 1.60	19345 1.30	12 1	

Statistics for Table of fuel by manufacturer

Statistic	DF	Value	Prob
Chi-Square	132	200966	<.0001
Likelihood Ratio Chi-Square	132	165398	<.0001
Mantel-Haenszel Chi-Square	1	12.59120	0.0004
Phi Coefficient		0.36685	
Contingency Coefficient		0.34441	
Cramer's V		0.18343	

Sample Size = 1493278

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of title_status by transmission				
	title_status	transmission			
		automatic	manual	other	Total
		1223915 81.96 87.91 93.41	145447 9.74 10.45 91.73	22942 1.54 1.65 93.63	1392304 93.24
clean	16613 1.11 85.60 1.27	2610 0.17 13.45 1.65	185 0.01 0.95 0.76		19408 1.30
lien	4215 0.28 60.95 0.32	2203 0.15 31.85 1.39	498 0.03 7.20 2.03		6916 0.46
missing	1704 0.11 62.08 0.13	528 0.04 19.23 0.33	513 0.03 18.69 2.09		2745 0.18
parts onl	39546 2.65 90.00 3.02	4214 0.28 9.59 2.66	180 0.01 0.41 0.73		43940 2.94
rebuilt	24222 1.62 86.62 1.85	3559 0.24 12.73 2.24	184 0.01 0.66 0.75		27965 1.87
salvage	1310215 87.74	158561 10.62	24502 1.64		1493278 100.00
Total					

Statistics for Table of title_status by transmission

Statistic	DF	Value	Prob
Chi-Square	10	11082.7679	<.0001
Likelihood Ratio Chi-Square	10	6420.7090	<.0001
Mantel-Haenszel Chi-Square	1	25.8547	<.0001
Phi Coefficient		0.0861	
Contingency Coefficient		0.0858	
Cramer's V		0.0609	

Sample Size = 1493278

Frequency Percent Row Pct Col Pct	title_status	Table of title_status by man																	
		acura	audi	bmw	buick	cadillac	chevrolet	chrysler	dodge	fiat	ford	gmc	honda	hyundai	infiniti	jaguar	jeep	kia	lexus
clean	15780 1.06 1.13 91.50	13230 0.89 0.95 92.33	38204 2.56 2.74 92.77	20297 1.36 1.46 94.15	22480 1.51 1.61 94.32	228333 15.29 16.40 93.36	28126 1.88 2.02 93.74	66557 4.46 4.78 92.98	1644 0.11 0.12 88.48	257286 17.23 18.48 94.08	62683 4.20 4.50 94.82	76956 5.15 5.53 90.77	29770 1.99 2.14 93.65	10574 0.71 0.76 88.54	3667 0.25 0.26 92.65	66738 4.47 4.79 93.75	22255 1.49 1.60 93.31	18105 1.21 1.30 93.59	
lien	125 0.01 0.64 0.72	198 0.01 1.02 1.38	467 0.03 2.41 1.13	138 0.01 0.71 0.64	219 0.01 1.13 0.92	3429 0.23 17.67 1.40	299 0.02 1.54 1.00	1052 0.07 5.42 1.47	34 0.00 0.18 1.83	3935 0.26 20.28 1.44	1147 0.08 5.91 1.74	693 0.05 3.57 0.82	325 0.05 1.67 1.02	129 0.01 0.66 1.08	38 0.09 0.20 0.96	1364 0.02 7.03 1.92	263 0.01 1.36 1.10	189 0.97 0.98 0.98	
missing	33 0.00 0.48 0.19	27 0.00 0.01 0.39	83 0.01 0.01 1.20	122 0.01 0.01 1.76	115 0.01 0.12 1.66	1788 0.12 25.85 0.39	117 0.01 1.69 0.39	440 0.03 6.36 0.61	28 0.00 0.40 1.51	1548 0.10 22.38 0.57	236 0.02 3.41 0.36	240 0.02 3.47 0.28	34 0.00 0.49 0.11	12 0.00 0.17 0.10	21 0.00 0.30 0.53	347 0.02 0.30 0.49	33 0.00 0.48 0.14	30 0.00 0.43 0.16	
parts onl	22 0.00 0.80 0.13	12 0.00 0.44 0.08	52 0.00 1.89 0.13	38 0.00 1.38 0.13	31 0.04 1.13 0.25	623 0.04 22.70 0.16	48 0.01 1.75 0.24	170 0.00 6.19 0.32	6 0.04 0.22 0.21	576 0.01 20.98 0.13	88 0.01 3.21 0.17	140 0.01 5.10 0.17	31 0.00 1.13 0.10	10 0.00 0.36 0.08	9 0.00 0.33 0.23	148 0.00 5.39 0.21	12 0.00 0.44 0.05	18 0.00 0.66 0.09	
rebuilt	734 0.05 1.67 4.26	554 0.04 1.26 3.87	1413 0.09 3.22 3.43	616 0.04 1.40 2.86	635 0.04 1.45 2.66	6519 0.44 14.84 2.67	905 0.06 2.06 3.02	2201 0.15 5.01 3.07	102 0.01 0.23 5.49	6135 0.01 13.96 2.24	1230 0.08 2.80 1.86	3720 0.08 8.47 4.39	1071 0.07 2.44 3.37	800 0.05 1.82 6.70	129 0.05 1.82 3.26	1668 0.11 2.44 2.34	845 0.06 1.92 3.54	583 0.04 1.33 3.01	
salvage	552 0.04 1.97 3.20	308 0.02 1.10 2.15	962 0.06 3.44 2.34	348 0.02 1.24 1.61	353 0.02 1.26 1.48	3891 0.26 13.91 1.59	509 0.03 1.82 1.70	1164 0.08 4.16 1.63	44 0.00 0.16 2.37	4005 0.27 14.32 1.46	720 0.05 2.57 1.09	3034 0.20 10.85 3.58	557 0.05 1.99 1.75	418 0.04 1.49 3.50	94 0.03 0.34 2.37	921 0.06 0.34 1.29	442 0.03 1.58 1.85	420 0.03 1.50 2.17	

Total	17246 1.15	14329 0.96	41181 2.76	21559 1.44	23833 1.60	244583 16.38	30004 2.01	71584 4.79	1858 0.12	273485 18.31	66104 4.43	84783 5.68	31788 2.13	11943 0.80	3958 0.27	71186 4.77	23850 1.60	19345 1.30
-------	---------------	---------------	---------------	---------------	---------------	-----------------	---------------	---------------	--------------	-----------------	---------------	---------------	---------------	---------------	--------------	---------------	---------------	---------------

Statistics for Table of title_status by manufacturer

Statistic	DF	Value	Prob
Chi-Square	165	12775.3827	<.0001
Likelihood Ratio Chi-Square	165	11944.4375	<.0001
Mantel-Haenszel Chi-Square	1	155.8603	<.0001
Phi Coefficient		0.0925	
Contingency Coefficient		0.0921	
Cramer's V		0.0414	

Sample Size = 1493278

The FREQ Procedure

Frequency Percent Row Pct Col Pct	transmission	Table of transmission																	
		acura	audi	bmw	buick	cadillac	chevrolet	chrysler	dodge	fiat	ford	gmc	honda	hyundai	infiniti	jaguar	jeep	kia	lexu
	automatic	15398 1.03 1.18 89.28	11998 0.80 0.92 83.73	35270 2.36 2.69 85.65	21041 1.41 1.61 97.60	23209 1.55 1.77 97.38	220308 14.75 16.81 90.07	28985 1.94 2.21 96.60	64101 4.29 4.89 89.55	1076 0.07 0.08 57.91	239023 16.01 18.24 87.40	62456 4.18 4.77 94.48	72568 4.86 5.54 85.59	28487 1.91 2.17 89.62	11020 0.74 0.84 92.27	3808 0.26 0.29 96.21	55419 3.71 4.23 77.85	21304 1.43 1.63 89.32	1895 1.2 1.4 97.9
	manual	1626 0.11 1.03 9.43	2126 0.14 1.34 14.84	5021 0.34 3.17 12.19	223 0.01 0.14 1.03	408 0.03 0.26 1.71	19540 1.31 12.32 7.99	719 0.05 0.45 2.40	5955 0.40 3.76 8.32	716 0.05 0.45 38.54	30449 2.04 19.20 11.13	2583 0.17 1.63 3.91	11588 0.78 7.31 13.67	2421 0.16 1.53 7.62	704 0.05 0.44 5.89	116 0.01 0.07 2.93	14475 0.97 9.13 20.33	1883 0.13 1.19 7.90	24 0.0 0.1 1.2
	other	222 0.01 0.91 1.29	205 0.01 0.84 1.43	890 0.06 3.63 2.16	295 0.02 1.20 1.37	216 0.01 0.88 0.91	4735 0.32 19.32 1.94	300 0.02 1.22 1.00	1528 0.10 6.24 2.13	66 0.00 0.27 3.55	4013 0.27 16.38 1.47	1065 0.07 4.35 1.61	627 0.04 2.56 0.74	880 0.06 3.59 2.77	219 0.01 0.89 1.83	34 0.00 0.14 0.86	1292 0.09 5.27 1.81	663 0.04 2.71 2.78	14 0.0 0.5 0.7
	Total	17246 1.15	14329 0.96	41181 2.76	21559 1.44	23833 1.60	244583 16.38	30004 2.01	71584 4.79	1858 0.12	273485 18.31	66104 4.43	84783 5.68	31788 2.13	11943 0.80	3958 0.27	71186 4.77	23850 1.60	1934 1.3

Statistics for Table of transmission by manufacturer

Statistic	DF	Value	Prob
Chi-Square	66	76708.8443	<.0001
Likelihood Ratio Chi-Square	66	68250.3476	<.0001
Mantel-Haenszel Chi-Square	1	4326.6619	<.0001
Phi Coefficient		0.2266	
Contingency Coefficient		0.2210	
Cramer's V		0.1603	

Sample Size = 1493278

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of condition by cylinders									
	condition	cylinders								
		10 cylinders	12 cylinders	3 cylinders	4 cylinders	5 cylinders	6 cylinders	8 cylinders	other	Total
excellent	792	117	194	58844	2109	70075	57188	12493	201812	45.73
	0.18	0.03	0.04	13.33	0.48	15.88	12.96	2.83		
	0.39	0.06	0.10	29.16	1.05	34.72	28.34	6.19		
	35.23	41.64	36.60	45.82	46.42	44.29	42.72	94.55		
fair	182	19	44	6720	301	9518	9288	92	26164	5.93
	0.04	0.00	0.01	1.52	0.07	2.16	2.10	0.02		
	0.70	0.07	0.17	25.68	1.15	36.38	35.50	0.35		
	8.10	6.76	8.30	5.23	6.63	6.02	6.94	0.70		
good	1053	115	166	42557	1801	59260	52934	434	158320	35.87
	0.24	0.03	0.04	9.64	0.41	13.43	11.99	0.10		
	0.67	0.07	0.10	26.88	1.14	37.43	33.43	0.27		
	46.84	40.93	31.32	33.14	39.64	37.45	39.54	3.28		
like new	211	29	113	19006	304	18111	13614	164	51552	11.68
	0.05	0.01	0.03	4.31	0.07	4.10	3.08	0.04		
	0.41	0.06	0.22	36.87	0.59	35.13	26.41	0.32		
	9.39	10.32	21.32	14.80	6.69	11.45	10.17	1.24		
new	3	1	6	652	6	601	483	9	1761	0.40
	0.00	0.00	0.00	0.15	0.00	0.14	0.11	0.00		
	0.17	0.06	0.34	37.02	0.34	34.13	27.43	0.51		
	0.13	0.36	1.13	0.51	0.13	0.38	0.36	0.07		
salvage	7	0	7	640	22	653	369	21	1719	0.39
	0.00	0.00	0.00	0.15	0.00	0.15	0.08	0.00		
	0.41	0.00	0.41	37.23	1.28	37.99	21.47	1.22		
	0.31	0.00	1.32	0.50	0.48	0.41	0.28	0.16		
Total	2248	281	530	128419	4543	158218	133876	13213	441328	
	0.51	0.06	0.12	29.10	1.03	35.85	30.33	2.99	100.00	

Statistics for Table of condition by cylinders

Statistic	DF	Value	Prob
Chi-Square	35	16092.8445	<.0001
Likelihood Ratio Chi-Square	35	18192.1653	<.0001
Mantel-Haenszel Chi-Square	1	1298.1592	<.0001
Phi Coefficient		0.1910	
Contingency Coefficient		0.1876	
Cramer's V		0.0854	

Sample Size = 441328

Frequency Percent Row Pct Col Pct	Table of condition by drive				
	condition	drive			
		4wd	fwd	rwd	Total
excellent	85143	74328	42341	201812	45.73
	19.29	16.84	9.59		
	42.19	36.83	20.98		
	47.18	46.17	42.40		
fair	10068	9067	7029	26164	5.93
	2.28	2.05	1.59		
	38.48	34.65	26.87		
	5.58	5.63	7.04		
good	62656	56402	39262	158320	35.87
	14.20	12.78	8.90		
	39.58	35.63	24.80		
	34.72	35.03	39.31		
like new	21367	19773	10412	51552	11.68
	4.84	4.48	2.36		
	41.45	38.36	20.20		
	11.84	12.28	10.43		
new	660	675	426	1761	0.40
	0.15	0.15	0.10		
	37.48	38.33	24.19		
	0.37	0.42	0.43		
salvage	560	759	400	1719	0.39
	0.13	0.17	0.09		
	32.58	44.15	23.27		
	0.31	0.47	0.40		

Total	180454 40.89	161004 36.48	99870 22.63	441328 100.00
--------------	-----------------	-----------------	----------------	------------------

Statistics for Table of condition by drive

Statistic	DF	Value	Prob
Chi-Square	10	1283.1305	<.0001
Likelihood Ratio Chi-Square	10	1273.3594	<.0001
Mantel-Haenszel Chi-Square	1	238.2943	<.0001
Phi Coefficient		0.0539	
Contingency Coefficient		0.0538	
Cramer's V		0.0381	

Sample Size = 441328

Frequency Percent Row Pct Col Pct	Table of condition by type														
	condition	type													
		SUV	bus	converti	coupe	hatchbac	mini-van	offroad	other	pickup	sedan	truck	van	Total	
excellent	excellent	52931 11.99 26.23 49.33	169 0.04 0.08 32.69	6615 1.50 3.28 50.64	13432 3.04 6.66 43.26	8095 1.83 4.01 45.78	5721 1.30 2.83 44.73	882 0.20 0.44 36.67	1191 0.27 0.59 33.84	20981 4.75 10.40 44.92	57017 12.92 28.25 47.17	26547 6.02 13.15 41.17	4084 0.93 2.02 35.87	4147 0.94 2.05 43.35	201812 45.73
		4689 1.06 17.92 4.37	46 0.01 0.18 8.90	581 0.13 2.22 4.45	2124 0.48 8.12 6.84	949 0.22 3.63 5.37	899 0.20 3.44 7.03	219 0.05 0.84 9.11	301 0.07 1.15 8.55	3701 0.84 14.15 7.92	6441 1.46 24.62 5.33	4759 1.08 18.19 7.38	795 0.18 3.04 6.98	660 0.15 2.52 6.90	26164 5.93
		35342 8.01 22.32 32.94	250 0.06 0.16 48.36	4057 0.92 2.56 31.06	11337 2.57 7.16 36.52	6021 1.36 3.80 34.05	4821 1.09 3.05 37.69	993 0.23 0.63 41.29	1746 0.40 1.10 49.62	17480 3.96 11.04 37.42	40624 9.20 25.66 33.61	26559 6.02 16.78 41.19	5234 1.19 3.31 45.97	3856 0.87 2.44 40.31	158320 35.87
		13684 3.10 26.54 12.75	51 0.01 0.10 9.86	1680 0.38 3.26 12.86	3806 0.86 7.38 12.26	2443 0.55 4.74 13.81	1286 0.29 2.49 10.05	275 0.06 0.53	246 0.06 0.48	4227 0.96 8.20	15666 3.55 30.39	6189 1.40 12.01	1182 0.27 2.29	817 0.19 1.58	51552
like new	like new	364 0.08 20.67 0.34	0 0.00 0.00 0.00	81 0.02 4.60 0.62	160 0.04 9.09 0.52	80 0.02 4.54 0.45	30 0.01 1.70 0.23	14 0.00 0.80 0.58	12 0.00 0.68 0.34	145 0.03 8.23	572 0.13 32.48	236 0.05 13.40	46 0.01 2.61	21 0.00 1.19	1761 0.40
		284 0.06 16.52 0.26	1 0.00 0.06 0.19	49 0.01 2.85 0.38	188 0.04 10.94 0.61	96 0.02 5.58 0.54	33 0.01 1.92 0.26	22 1.28 0.91	23 1.34 0.65	176 10.24 0.38	549 31.94 10.99	189 21.94 2.56	44 0.04 0.29	65 0.01 3.78	1719 0.39
		Total	107294 24.31	517 0.12	13063 2.96	31047 7.03	17684 4.01	12790 2.90	2405 0.54	3519 0.80	46710 10.58	120869 27.39	64479 14.61	11385 2.58	9566 2.17

Statistics for Table of condition by type

Statistic	DF	Value	Prob
Chi-Square	60	5529.9988	<.0001
Likelihood Ratio Chi-Square	60	5534.1236	<.0001
Mantel-Haenszel Chi-Square	1	238.2314	<.0001
Phi Coefficient		0.1119	
Contingency Coefficient		0.1112	
Cramer's V		0.0501	

Sample Size = 441328

Frequency Percent Row Pct Col Pct	Table of condition by paint_color													
	condition	paint_color												
		black	blue	brown	custom	green	grey	orange	purple	red	silver	white	yellow	Total
excellent	excellent	39312 8.91 19.48 48.43	22246 5.04 11.02 44.13	5421 1.23 2.69 40.79	4008 0.91 1.99 37.63	6442 1.46 3.19 34.24	24118 5.46 11.95 48.81	1198 0.27 0.59 45.14	647 0.15 0.32 37.46	22173 5.02 10.99 44.14	32094 7.27 15.90 48.56	42466 9.62 21.04 45.58	1687 0.38 0.84 45.45	201812 45.73
		4032 0.91 15.41	3586 0.81 13.71	1067 0.24 4.08	571 0.13 2.18	2379 0.54 9.09	2100 0.48 8.03	132 0.03 0.50	203 0.05 0.78	3689 14.10 11.86	3103 0.70 19.17	5015 1.14 19.17	287 0.07 1.10	26164 5.93

	4.97	7.11	8.03	5.36	12.65	4.25	4.97	11.75	7.34	4.69	5.38	7.73	
good	26849 6.08 16.96 33.08	19353 4.39 12.22 38.39	5434 1.23 3.43 40.89	4835 1.10 3.05 45.40	8761 1.99 5.53 46.57	15476 3.51 9.78 31.32	973 0.22 0.61 36.66	700 0.16 0.44 40.53	18392 4.17 11.62 36.61	23139 5.24 14.62 35.01	33054 7.49 20.88 35.48	1354 0.31 0.86 36.48	158320 35.87
like new	10289 2.33 19.96 12.68	4803 1.09 9.32 9.53	1306 0.30 2.53 9.83	1089 0.25 2.11 10.23	1067 1.66 2.07 5.67	7346 14.25 14.87	327 0.07 12.32	161 0.04 9.32	5559 1.26 10.78	7280 1.65 14.12	11981 2.71 23.24	344 0.08 0.67	51552 11.68
new	424 0.10 24.08 0.52	175 0.04 9.94 0.35	18 0.00 5.51 0.14	97 0.02 1.93 0.91	34 0.01 11.81 0.18	208 0.05 0.51 0.42	9 0.00 0.12	2 0.00 0.37	186 0.04 0.39	257 0.06 14.59	340 0.08 19.31	11 0.00 0.62	1761 0.40
salvage	267 0.06 15.53 0.33	252 0.06 14.66 0.50	44 0.01 2.56 0.33	50 0.01 2.91 0.47	129 0.03 7.50 0.69	163 0.04 9.48 0.33	15 0.00 0.87 0.57	14 0.00 0.81 0.81	232 0.05 13.50 0.46	220 0.05 12.80 0.33	304 0.07 17.68 0.33	29 0.01 1.69 0.78	1719 0.39
Total	81173 18.39	50415 11.42	13290 3.01	10650 2.41	18812 4.26	49411 11.20	2654 0.60	1727 0.39	50231 11.38	66093 14.98	93160 21.11	3712 0.84	441328 100.00

Statistics for Table of condition by paint_color

Statistic	DF	Value	Prob
Chi-Square	55	7070.8815	<.0001
Likelihood Ratio Chi-Square	55	6807.0878	<.0001
Mantel-Haenszel Chi-Square	1	0.3827	0.5362
Phi Coefficient		0.1266	
Contingency Coefficient		0.1256	
Cramer's V		0.0566	

Sample Size = 441328

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of cylinders by drive				
	cylinders	drive			
		4wd	fwd	rwd	Total
		1241 0.28 55.20 0.69	61 0.01 2.71 0.04	946 0.21 42.08 0.95	2248 0.51
10 cylinders	10 cylinders	38 0.01 13.52 0.02	15 0.00 5.34 0.01	228 0.05 81.14 0.23	281 0.06
		58 0.01 10.94 0.03	395 0.09 74.53 0.25	77 0.02 14.53 0.08	530 0.12
		25872 5.86 20.15 14.34	92210 20.89 71.80 57.27	10337 2.34 8.05 10.35	128419 29.10
12 cylinders	12 cylinders	1343 0.30 29.56 0.74	2636 0.60 58.02 1.64	564 0.13 12.41 0.56	4543 1.03
		68090 15.43 43.04 37.73	55119 12.49 34.84 34.23	35009 7.93 22.13 35.05	158218 35.85
		78147 17.71 58.37 43.31	5268 1.19 3.93 3.27	50461 11.43 37.69 50.53	133876 30.33
3 cylinders	3 cylinders	5665 1.28 42.87 3.14	5300 1.20 40.11 3.29	2248 0.51 17.01 2.25	13213 2.99
		Total	180454 40.89	161004 36.48	99870 22.63
					441328 100.00

Statistics for Table of cylinders by drive

Statistic	DF	Value	Prob
Chi-Square	14	135371	<.0001
Likelihood Ratio Chi-Square	14	154404	<.0001
Mantel-Haenszel Chi-Square	1	1189	<.0001
Phi Coefficient		0.55384	
Contingency Coefficient		0.48449	
Cramer's V		0.39162	

Sample Size = 441328

Frequency Percent Row Pct Col Pct	Table of cylinders by type														
	cylinders	type													
		SUV	bus	converti	coupe	hatchbac	mini-van	offroad	other	pickup	sedan	truck	van	wagon	Total
10 cylinders	10 cylinders	350 0.08 15.57 0.33	104 0.02 4.63 20.12	80 0.02 3.56 0.61	48 0.01 2.14 0.15	6 0.00 0.27 0.03	0 0.00 0.00 0.00	7 0.01 1.96 1.25	44 0.09 17.17 0.83	386 0.03 6.05 0.11	136 0.22 43.33 1.51	974 0.02 4.45 0.88	100 0.00 0.58 0.14	13 0.00 0.58 0.14	2248 0.51
		10 0.00 3.56 0.01	0 0.00 19.93 0.00	56 0.01 28.11 0.43	79 0.02 0.36 0.25	1 0.00 0.01 0.01	0 0.00 0.00 0.00	0 0.00 1.07 0.09	3 0.00 2.14 0.01	6 0.00 37.72 0.09	106 0.02 6.41 0.03	18 0.00 0.71 0.02	2 0.00 0.00 0.00	0 0.00 0.00 0.00	281 0.06
		37 0.01 6.98 0.03	1 0.00 0.19 0.19	17 0.00 3.21 0.13	52 0.01 9.81 0.17	215 0.05 40.57 1.22	15 0.00 2.83 0.12	3 0.00 2.83 0.43	15 0.00 2.45 0.03	13 0.00 24.53 0.11	130 0.00 3.40 0.03	18 0.00 1.89 0.09	10 0.00 1.89 0.04	4 0.00 0.75 0.04	530 0.12
12 cylinders	12 cylinders	24343 5.52 18.96 22.69	51 0.01 0.04 9.86	3688 0.84 2.87 28.23	9728 2.20 7.58 31.33	14657 3.32 11.41 82.88	822 0.19 0.64 6.43	341 0.08 0.27 14.18	826 0.19 0.64 23.47	2505 0.57 1.95 5.36	62058 14.06 48.32 51.34	2636 0.60 2.05 4.09	872 0.20 0.68 7.66	5892 1.34 4.59 61.59	128419 29.10

5 cylinders	297 0.07 6.54 0.28	2 0.00 0.04 1.22	160 0.04 3.52 0.67	207 0.05 4.56 0.77	313 0.07 6.89 1.77	35 0.01 0.77 0.27	8 0.00 0.18 0.33	12 0.00 0.26 0.34	325 0.07 7.15 0.70	2047 0.46 45.06 1.69	524 0.12 11.53 0.81	82 0.02 1.80 0.72	531 0.12 11.69 5.55	4543 1.03
6 cylinders	53292 12.08 33.68 49.67	48 0.01 0.03 9.28	4620 1.05 2.92 35.37	10432 2.36 6.59 33.60	1255 0.28 0.79 7.10	11463 2.60 7.25 89.62	1566 0.35 0.99 65.11	1223 0.28 0.77 34.75	11067 2.51 6.99 23.69	40924 9.27 25.87 33.86	14813 3.36 9.36 22.97	5226 1.18 3.30 45.90	2289 0.52 1.45 23.93	158218 35.85
8 cylinders	26254 5.95 19.61 24.47	305 0.07 0.23 58.99	4206 0.95 3.14 32.20	9822 2.23 7.34 31.64	569 0.13 0.43 3.22	54 0.01 0.04 0.42	470 0.11 0.35 19.54	1203 0.27 0.90 34.19	28504 6.46 21.29 61.02	11797 2.67 8.81 9.76	44991 10.19 33.61 69.78	5078 1.15 3.79 44.60	623 0.14 0.47 6.51	133876 30.33
other	2711 0.61 20.52 2.53	6 0.00 0.05 1.16	236 0.05 1.79 1.81	679 0.15 5.14 2.19	668 0.15 5.06 3.78	401 0.09 3.03 3.14	10 0.00 0.08 0.42	193 0.04 1.46 5.48	3904 29.55 27.78 8.36	3671 27.78 3.82 3.04	505 0.88 0.11 0.78	15 0.83 0.11 0.13	214 1.15 1.62 2.24	13213 2.99
Total	107294 24.31	517 0.12	13063 2.96	31047 7.03	17684 4.01	12790 2.90	2405 0.54	3519 0.80	46710 10.58	120869 27.39	64479 14.61	11385 2.58	9566 2.17	441328 100.00

Statistics for Table of cylinders by type

Statistic	DF	Value	Prob
Chi-Square	84	191688	<.0001
Likelihood Ratio Chi-Square	84	189312	<.0001
Mantel-Haenszel Chi-Square	1	85.97904	<.0001
Phi Coefficient		0.65905	
Contingency Coefficient		0.55029	
Cramer's V		0.24910	

Sample Size = 441328

Frequency Percent Row Pct Col Pct	Table of cylinders by paint_color												
	cylinders	paint_color											
		black	blue	brown	custom	green	grey	orange	purple	red	silver	white	yellow
10 cylinders	362 0.08 16.10 0.45	225 0.05 10.01 0.45	82 0.02 3.65 0.62	46 0.01 2.05 0.43	148 0.03 6.58 0.79	101 0.02 4.49 0.20	13 0.00 0.58 0.49	5 0.00 0.22 0.29	224 0.05 9.96 0.45	131 0.03 5.83 0.20	890 0.20 39.59 0.96	21 0.00 0.93 0.57	2248 0.51
	108 0.02 38.43 0.13	38 0.01 13.52 0.08	2 0.00 0.71 0.02	10 0.00 3.56 0.09	12 0.01 4.27 0.06	24 0.00 8.54 0.05	1 0.00 0.36 0.04	0 0.00 0.00 0.00	22 0.00 7.83 0.04	32 0.01 11.39 0.05	30 0.01 10.68 0.03	2 0.00 0.71 0.05	281 0.06
	69 0.02 13.02 0.09	64 0.01 12.08 0.13	6 0.00 1.13 0.05	5 0.00 0.94 0.05	22 0.02 4.15 0.12	74 0.02 13.96 0.15	8 0.00 1.51 0.30	9 0.00 1.70 0.52	73 0.02 13.77 0.15	84 0.02 15.85 0.13	108 0.02 20.38 0.12	8 0.00 1.51 0.22	530 0.12
	22181 5.03 17.27 27.33	16721 3.79 13.02 33.17	3265 0.74 2.54 24.57	2395 0.54 1.86 22.49	5076 1.15 3.95 26.98	17601 3.99 13.71 35.62	1070 0.24 0.83 40.32	564 0.13 0.44 32.66	14739 3.34 11.48 29.34	22888 5.19 17.82 34.63	20891 4.73 16.27 22.42	1028 0.23 0.80 27.69	128419 29.10
5 cylinders	825 0.19 18.16 1.02	568 0.13 12.50 1.13	90 0.02 1.98 0.68	64 0.01 1.41 0.60	187 0.04 4.12 0.99	557 0.13 12.26 1.13	51 0.01 1.12 1.92	15 0.00 0.33 0.87	525 0.12 11.56 1.05	856 0.19 18.84 1.30	741 0.17 16.31 0.80	64 0.01 1.41 1.72	4543 1.03
	29220 6.62 18.47 36.00	17550 3.98 11.09 34.81	5295 1.20 3.35 39.84	4193 0.95 2.65 39.37	7172 1.63 4.53 38.12	18554 4.20 11.73 37.55	746 0.17 0.47 28.11	659 0.15 0.42 38.16	16943 3.84 10.71 33.73	26569 6.02 16.79 40.20	30261 6.86 19.13 32.48	1056 0.24 0.67 28.45	158218 35.85
	25370 5.75 18.95 31.25	13685 3.10 10.22 27.14	4507 1.02 3.37 33.91	3889 0.88 2.90 36.52	5892 1.34 4.40 31.32	11006 2.49 8.22 22.27	712 0.16 0.53 26.83	440 0.10 0.26 25.48	16351 3.70 12.21 32.55	13518 3.06 10.10 20.45	37022 8.39 27.65 39.74	1484 0.34 1.11 39.98	133876 30.33
	3038 0.69 22.99 3.74	1564 0.35 11.84 3.10	43 0.01 0.33 0.32	48 0.01 0.36 0.45	303 0.07 2.29 1.61	1494 0.34 11.31 3.02	53 0.01 0.40 2.00	35 0.01 0.26 2.03	1354 0.31 10.25 2.70	2015 0.46 15.25 3.05	3217 0.73 15.25 3.45	49 0.01 0.37 1.32	13213 2.99
Total	81173 18.39	50415 11.42	13290 3.01	10650 2.41	18812 4.26	49411 11.20	2654 0.60	1727 0.39	50231 11.38	66093 14.98	93160 21.11	3712 0.84	441328 100.00

Statistics for Table of cylinders by paint_color

Statistic	DF	Value	Prob
Chi-Square	77	13034.0112	<.0001
Likelihood Ratio Chi-Square	77	13515.4863	<.0001
Mantel-Haenszel Chi-Square	1	163.9275	<.0001
Phi Coefficient		0.1719	
Contingency Coefficient		0.1694	
Cramer's V		0.0650	

Sample Size = 441328

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of drive by type														
	drive	type													
		SUV	bus	converti	coupe	hatchbac	mini-van	offroad	other	pickup	sedan	truck	van	wagon	Total
4wd	77374 17.53 42.88 72.11	14 0.00 0.01 2.71	623 0.14 0.35 4.77	1241 0.28 0.69 4.00	1938 0.44 1.07 10.96	564 0.13 0.31 4.41	2265 0.51 1.26 94.18	994 0.23 0.55 28.25	32387 7.34 17.95 69.34	12643 2.86 7.01 10.46	44852 10.16 24.86 69.56	658 0.15 0.36 5.78	4901 1.11 2.72 51.23	180454 40.89	
fwd	20226 4.58 12.56 18.85	41 0.01 0.03 7.93	3260 0.74 2.02 24.96	12172 2.76 7.56 39.21	14500 3.29 9.01 82.00	11703 2.65 7.27 91.50	59 0.01 0.04 2.45	1125 0.25 0.70 31.97	1071 0.24 0.67 2.29	87367 19.80 54.26 72.28	1589 0.36 0.99 2.46	4094 0.93 2.54 35.96	3797 0.86 2.36 39.69	161004 36.48	
rwd	9694 2.20 9.71 9.03	462 0.10 0.46 89.36	9180 2.08 9.19 70.27	17634 4.00 17.66 56.80	1246 0.28 1.25 7.05	523 0.12 0.52 4.09	81 0.02 0.08 3.37	1400 0.32 1.40 39.78	13252 3.00 13.27 28.37	20859 4.73 20.89 17.26	18038 4.09 18.06 27.97	6633 1.50 6.64 58.26	868 0.20 0.87 9.07	99870 22.63	
Total	107294 24.31	517 0.12	13063 2.96	31047 7.03	17684 4.01	12790 2.90	2405 0.54	3519 0.80	46710 10.58	120869 27.39	64479 14.61	11385 2.58	9566 2.17	441328 100.00	

Statistics for Table of drive by type

Statistic	DF	Value	Prob
Chi-Square	24	266314	<.0001
Likelihood Ratio Chi-Square	24	290685	<.0001
Mantel-Haenszel Chi-Square	1	9507	<.0001
Phi Coefficient		0.77681	
Contingency Coefficient		0.61347	
Cramer's V		0.54929	

Sample Size = 441328

Frequency Percent Row Pct Col Pct	Table of drive by paint_color														
	drive	paint_color													
		black	blue	brown	custom	green	grey	orange	purple	red	silver	white	yellow	Total	
4wd	36879 8.36 20.44 45.43	19693 4.46 10.91 39.06	5669 1.28 3.14 42.66	3963 0.90 2.20 37.21	8709 1.97 4.83 46.29	19605 4.44 10.86 39.68	918 0.21 0.51 34.59	589 0.13 0.33 34.11	20262 4.59 11.23 40.34	24336 5.51 13.49 36.82	38787 8.79 21.49 41.63	1044 0.24 0.58	180454 40.89		
fwd	26874 6.09 16.69 33.11	20139 4.56 12.51 39.95	4943 1.12 3.07 37.19	3866 0.88 2.40 36.30	5835 1.32 3.62 31.02	21637 4.90 13.44 43.79	938 0.21 0.58 35.34	778 0.18 0.48 45.05	17470 3.96 10.85 34.78	30311 6.87 18.83 45.86	27223 6.17 16.91 29.22	990 0.22	161004 36.48		
rwd	17420 3.95 17.44 21.46	10583 2.40 10.60 20.99	2678 0.61 2.68 20.15	2821 0.64 2.82 26.49	4268 0.97 4.27 22.69	8169 1.85 8.18 16.53	798 0.18 0.80 30.07	360 0.08 20.85	12499 2.83 24.88 21.32	11446 2.59 12.52 17.32	27150 6.15 11.46 29.14	1678 0.38	99870 22.63		
Total	81173 18.39	50415 11.42	13290 3.01	10650 2.41	18812 4.26	49411 11.20	2654 0.60	1727 0.39	50231 11.38	66093 14.98	93160 21.11	3712 0.84	441328 100.00		

Statistics for Table of drive by paint_color

Statistic	DF	Value	Prob
Chi-Square	22	10150.7946	<.0001
Likelihood Ratio Chi-Square	22	9928.4219	<.0001
Mantel-Haenszel Chi-Square	1	871.7726	<.0001
Phi Coefficient		0.1517	
Contingency Coefficient		0.1499	
Cramer's V		0.1072	

Sample Size = 441328

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of type by paint_color													
	type	paint_color												
		black	blue	brown	custom	green	grey	orange	purple	red	silver	white	yellow	Total
SUV	22838 5.17 21.29 28.13	11463 2.60 10.68 22.74	3962 0.90 3.69 29.81	2620 0.59 2.44 24.60	5054 1.15 4.71 26.87	12471 2.83 11.62 25.24	624 0.14 0.58 23.51	410 0.09 0.38 23.74	9934 2.25 9.26 19.78	17350 3.93 16.17 26.25	20001 4.53 18.64 21.47	567 0.13 0.53 15.27	107294 24.31	
bus	23 0.01 4.45 0.03	24 0.00 4.64 0.05	9 0.00 1.74 0.07	13 0.00 2.51 0.12	12 0.00 2.32 0.06	9 0.00 1.74 0.02	12 0.00 2.32 0.45	1 0.00 0.19 0.06	12 0.00 1.35 0.02	7 0.00 1.35 0.01	354 0.08 68.47 0.38	41 0.01 7.93 1.10	517 0.12	
converti	2559 0.58 19.59 3.15	1555 0.35 11.90 3.08	158 0.04 1.21 1.19	336 0.08 2.57 3.15	585 0.13 4.48 3.11	813 0.18 6.22 1.65	123 0.03 0.94 4.63	72 0.02 0.55 4.17	2700 0.61 20.67 5.38	1949 0.44 14.92 2.95	1783 0.40 13.65 1.91	430 0.10 3.29 11.58	13063 2.96	
coupe	6755 1.53 21.76 8.32	3785 0.86 12.19 7.51	416 0.09 1.34 3.13	648 0.15 2.09 6.08	1191 0.27 3.84 6.33	2909 0.66 9.37 5.89	523 0.12 1.68 19.71	185 0.04 10.71	5404 1.22 17.41 10.76	4110 0.93 13.24 6.22	4271 0.97 13.76 4.58	850 0.19 2.74 22.90	31047 7.03	
hatchbac	2988 0.68 16.90 3.68	2637 0.60 14.91 5.23	204 0.05 1.15 1.53	205 0.05 1.16 1.92	741 0.17 4.19 3.94	2260 0.51 12.78 4.57	295 0.07 1.67 11.12	125 0.03 0.71	2552 1.22 14.43	2857 0.65 16.16	2580 0.58 14.59	240 0.05 1.36	17684 4.01	
mini-van	1081 0.24 17.79 1.33	2275 0.52 4.21 4.51	539 0.12 2.25 4.06	288 0.07 4.32 2.70	553 0.13 13.15 2.94	1682 0.38 0.08 3.40	10 0.00 0.38	57 0.01 3.30	1197 0.27 2.38	2591 0.59 3.92	2483 0.56 2.67	34 0.01 0.92	12790 2.90	
offroad	482 0.11 20.04 0.59	248 0.06 10.31 0.49	82 0.02 3.41 0.62	86 0.02 3.58 0.81	262 0.06 10.89 1.39	163 0.04 6.78 0.33	50 0.01 2.08	8 0.00 0.33	394 0.09 0.46	192 0.04 0.78	355 0.08 0.29	83 0.02 0.34	2405 0.54	
other	649 0.15 18.44 0.80	330 0.07 9.38 0.65	97 0.02 2.76 0.73	189 0.04 5.37 1.77	186 0.05 5.29 0.99	211 0.38 6.00 0.43	34 0.01 0.97	19 0.54	408 11.59 11.59	385 10.94 10.94	919 0.21 0.58	92 0.02 0.99	3519 0.80	
pickup	7599 1.72 16.27 9.36	4991 1.13 10.69 9.90	1441 0.33 3.08 10.84	844 0.19 1.81 7.92	2349 0.53 5.03 12.49	4288 0.97 9.18 8.68	231 0.05 0.49	139 0.30	6792 14.54 13.52	4481 9.59 8.70	13282 28.44 13.52	273 0.58	46710 10.58	
sedan	23954 5.43 19.82 29.51	13953 3.16 11.54 27.68	3787 0.86 3.13 28.50	3207 0.73 2.65 30.11	3918 0.89 3.24 20.83	17256 3.91 14.28 34.92	345 0.08 0.29	499 0.11 0.41	10602 2.40 8.77	23077 5.23 19.09	19782 4.48 16.37	489 0.11 0.40	120869 27.39	
truck	10354 2.35 16.06 12.76	6593 1.49 10.23 13.08	1995 0.45 3.09 15.01	1545 0.35 2.40 14.51	2793 0.63 4.33 14.85	5546 1.26 8.60 11.22	326 0.07 0.51	156 0.04 0.24	8751 1.98 13.57	6239 1.41 9.68	19696 4.46 30.55	485 0.11 0.75	64479 14.61	
van	586 0.13 5.15 0.72	1070 0.24 9.40 2.12	291 0.07 2.56 2.19	415 0.09 3.65 3.90	439 0.10 3.86 2.33	651 0.15 5.72 1.32	29 0.01 0.25	24 0.01 0.21	612 0.14 1.22	1026 0.23 1.55	6169 1.40 6.62	73 0.02 0.64	11385 2.58	
wagon	1305 0.30 13.64 1.61	1491 0.34 15.59 2.96	309 0.07 3.23 2.33	254 0.06 2.66 3.88	729 0.17 7.62 3.23	1152 0.26 12.04 2.33	52 0.01 0.54	32 0.01 0.33	873 0.20 1.74	1829 0.41 2.77	1485 0.34 1.59	55 0.01 0.57	9566 2.17	
Total	81173 18.39	50415 11.42	13290 3.01	10650 2.41	18812 4.26	49411 11.20	2654 0.60	1727 0.39	50231 11.38	66093 14.98	93160 21.11	3712 0.84	441328 100.00	

Statistics for Table of type by paint_color

Statistic	DF	Value	Prob
Chi-Square	132	36600.8363	<.0001
Likelihood Ratio Chi-Square	132	33884.4508	<.0001
Mantel-Haenszel Chi-Square	1	1231.0941	<.0001
Phi Coefficient		0.2880	
Contingency Coefficient		0.2767	
Cramer's V		0.0868	

Sample Size = 441328

The GLM Procedure

Class Level Information		
Class	Levels	Values
fuel	5	diesel electric gas hybrid other

Number of Observations Read	1493278
Number of Observations Used	1493278

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	5239.6104	1309.9026	3048.54	<.0001
Error	1.49E6	641632.4781	0.4297		
Corrected Total	1.49E6	646872.0885			

R-Square	Coeff Var	Root MSE	log_age Mean
0.008100	26.78702	0.655501	2.447086

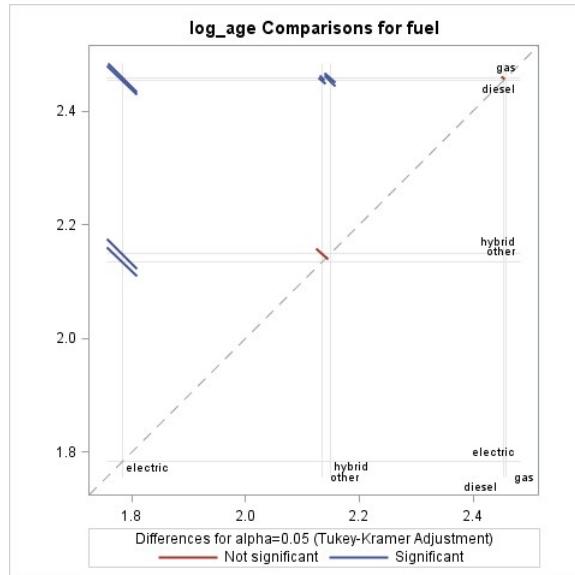
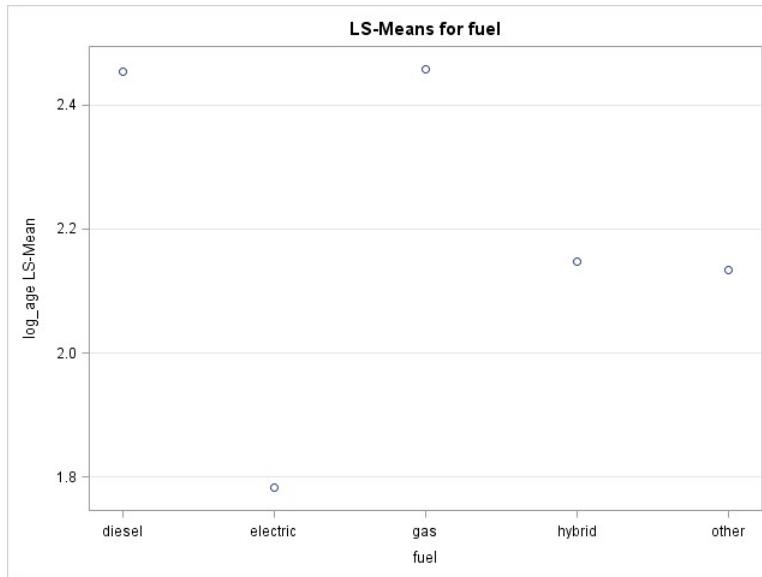
Source	DF	Type I SS	Mean Square	F Value	Pr > F
fuel	4	5239.610384	1309.902596	3048.54	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
fuel	4	5239.610384	1309.902596	3048.54	<.0001

The GLM Procedure
 Least Squares Means
 Adjustment for Multiple Comparisons: Tukey-Kramer

fuel	log_age LSMEAN	LSMEAN Number
diesel	2.45344230	1
electric	1.78272009	2
gas	2.45809471	3
hybrid	2.14840613	4
other	2.13462157	5

Least Squares Means for effect fuel					
Pr > t for H0: LSMean(i)=LSMean(j)					
Dependent Variable: log_age					
i\j	1	2	3	4	5
1		<.0001	0.2064	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001
3	0.2064	<.0001		<.0001	<.0001
4	<.0001	<.0001	<.0001		0.3315
5	<.0001	<.0001	<.0001	0.3315	



The GLM Procedure

Class Level Information		
Class	Levels	Values
title_status	6	clean lien missing parts onl rebuilt salvage

Number of Observations Read	1493278
Number of Observations Used	1493278

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	12562.2469	2512.4494	5914.73	<.0001
Error	1.49E6	634309.8416	0.4248		
Corrected Total	1.49E6	646872.0885			

R-Square	Coeff Var	Root MSE	log_age Mean
0.019420	26.63373	0.651750	2.447086

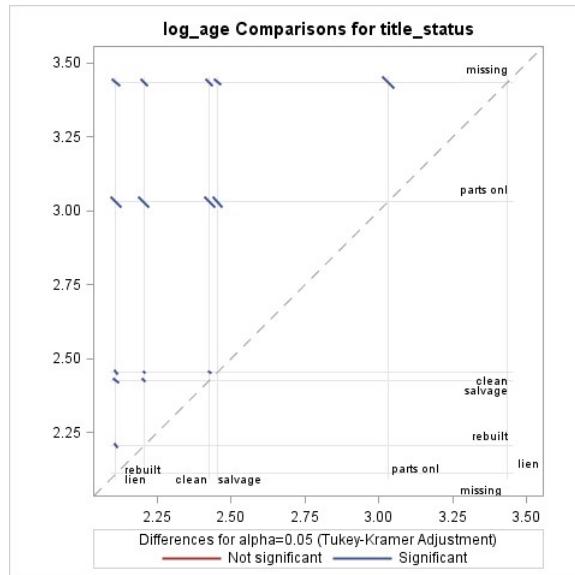
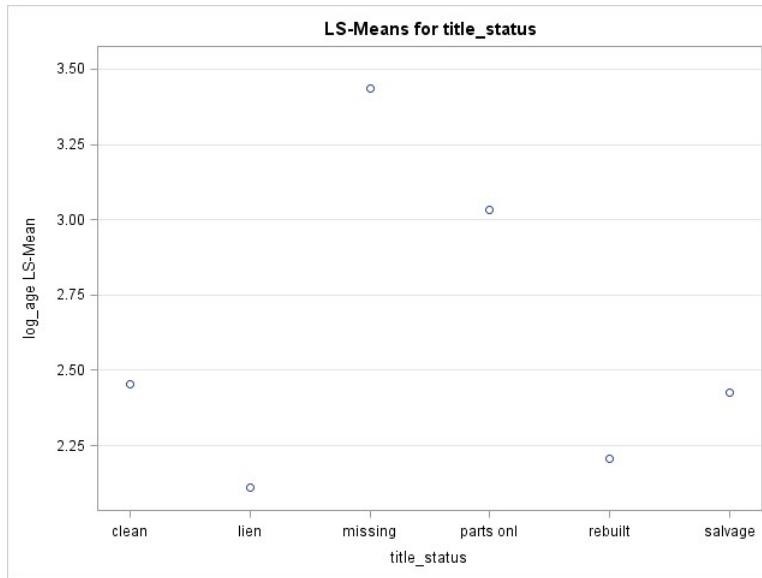
Source	DF	Type I SS	Mean Square	F Value	Pr > F
title_status	5	12562.24688	2512.44938	5914.73	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
title_status	5	12562.24688	2512.44938	5914.73	<.0001

The GLM Procedure
 Least Squares Means
 Adjustment for Multiple Comparisons: Tukey-Kramer

title_status	log_age LSMEAN	LSMEAN Number
clean	2.45379181	1
lien	2.10996123	2
missing	3.43577407	3
parts onl	3.03109515	4
rebuilt	2.20452100	5
salvage	2.42647507	6

Least Squares Means for effect title_status						
Pr > t for H0: LSMean(i)=LSMean(j)						
Dependent Variable: log_age						
i/j	1	2	3	4	5	6
1		<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	<.0001
4	<.0001	<.0001	<.0001		<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001	



The GLM Procedure

Class Level Information		
Class	Levels	Values
transmission	3	automatic manual other

Number of Observations Read	1493278
Number of Observations Used	1493278

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	34783.6218	17391.8109	42429.7	<.0001
Error	1.49E6	612088.4667	0.4099		
Corrected Total	1.49E6	646872.0885			

R-Square	Coeff Var	Root MSE	log_age Mean
0.053772	26.16303	0.640232	2.447086

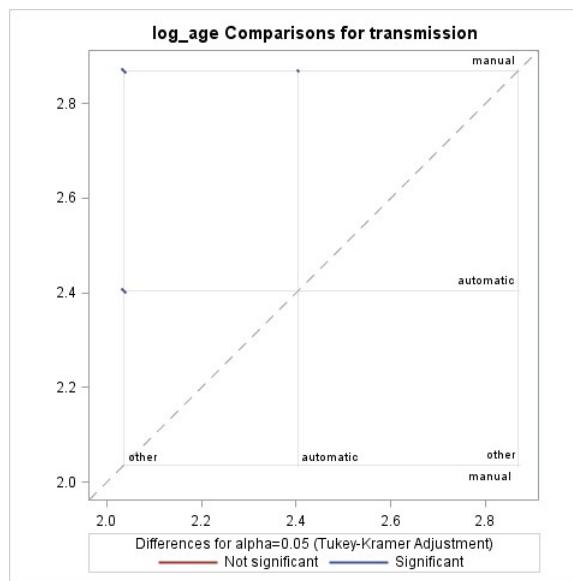
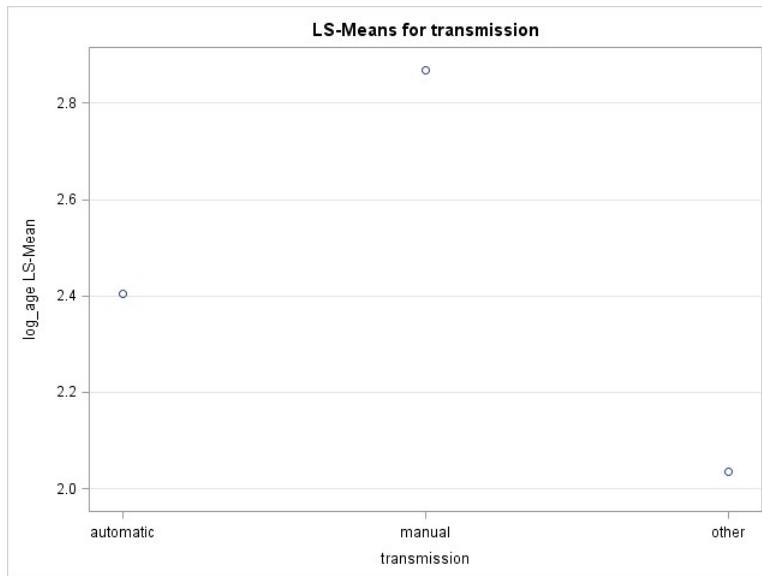
Source	DF	Type I SS	Mean Square	F Value	Pr > F
transmission	2	34783.62185	17391.81092	42429.7	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
transmission	2	34783.62185	17391.81092	42429.7	<.0001

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

transmission	log_age LSMEAN	LSMEAN Number
automatic	2.40374044	1
manual	2.86872344	2
other	2.03636612	3

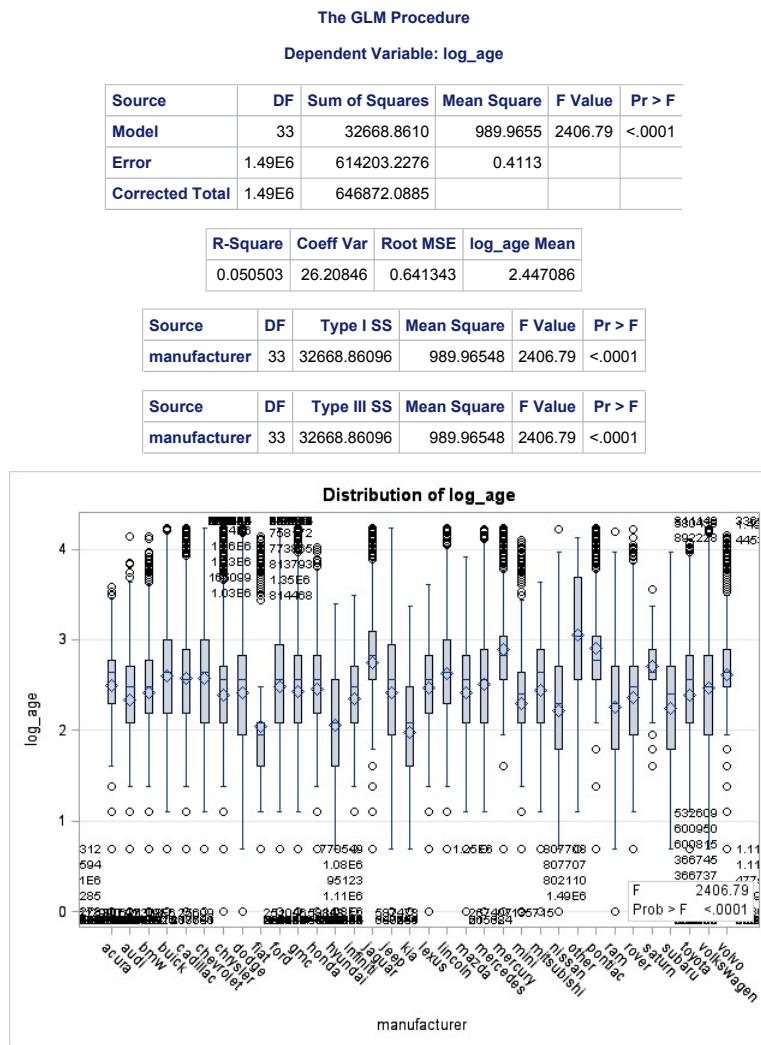
Least Squares Means for effect transmission			
Pr > t for H0: LSMean(i)=LSMean(j)			
Dependent Variable: log_age			
i/j	1	2	3
1		<.0001	<.0001
2	<.0001		<.0001
3	<.0001	<.0001	



The GLM Procedure

Class Level Information		
Class	Levels	Values
manufacturer	34	acura audi bmw buick cadillac chevrolet chrysler dodge fiat ford gmc honda hyundai infiniti jaguar jeep kia lexus lincoln mazda mercedes mercury mini mitsubishi nissan other pontiac ram rover saturn subaru toyota volkswagen volvo

Number of Observations Read	1493278
Number of Observations Used	1493278



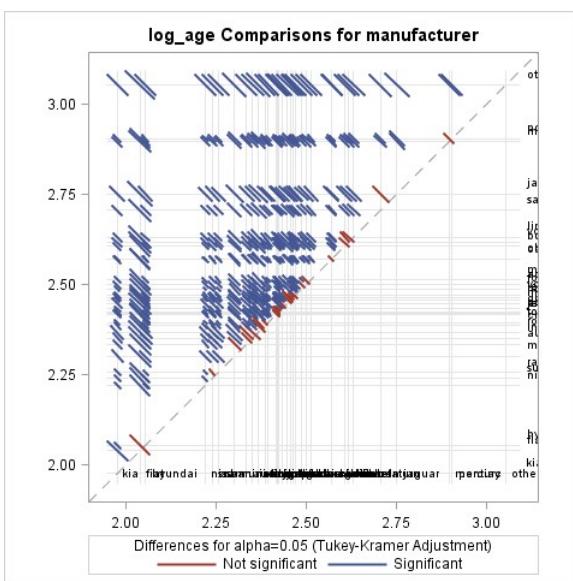
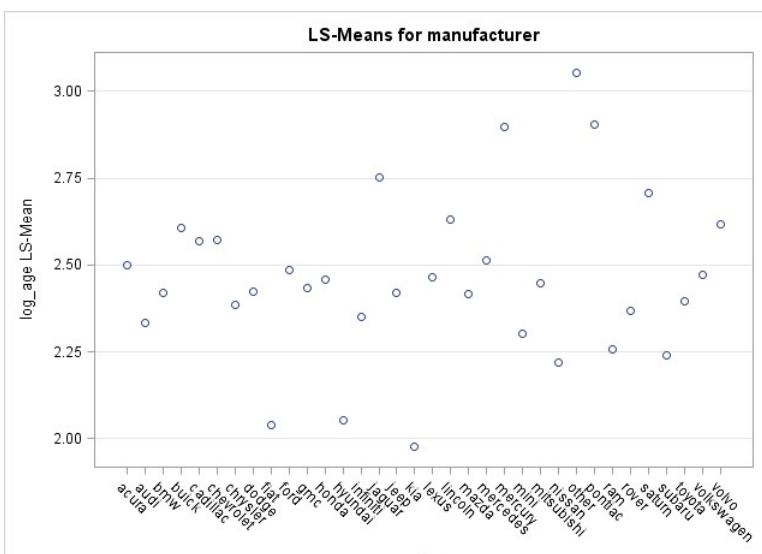
The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

manufacturer	log_age LSMEAN	LSMEAN Number
acura	2.49969801	1
audi	2.33443291	2
bmw	2.41919516	3
buick	2.60631282	4
cadillac	2.56925391	5
chevrolet	2.57185015	6
chrysler	2.38555961	7
dodge	2.42298831	8
fiat	2.03904381	9
ford	2.48745098	10
gmc	2.43234650	11
honda	2.45817279	12
hyundai	2.05453063	13
infiniti	2.35181365	14
jaguar	2.75163488	15
jeep	2.41965391	16
kia	1.97646757	17
lexus	2.46506935	18
lincoln	2.63157714	19
mazda	2.41765477	20
mercedes	2.51342468	21
mercury	2.89659636	22
mini	2.30116133	23
mitsubishi	2.44779143	24
nissan	2.21995218	25
other	3.05327619	26
pontiac	2.90488004	27
ram	2.25638443	28
rover	2.36747572	29
saturn	2.70619776	30
subaru	2.23964700	31
toyota	2.39459431	32
volkswagen	2.47110437	33
volvo	2.61716492	34

Least Squares Means for effect manufacturer
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: log_age

i\j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.8709	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.9500	
2	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.9610	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
3	<.0001	<.0001		<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	0.2493	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	1.0000	<.0001	
4	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.1410	<.0001	
5	<.0001	<.0001	<.0001	<.0001		1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
6	<.0001	<.0001	<.0001	<.0001	1.0000		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
7	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0006	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
8	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001		<.0001	<.0001	0.6867	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	1.0000	<.0001	<.0001	
9	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	0.0208	<.0001	<.0001	<.0001	<.0001	
10	0.8709	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0013	<.0001	<.0001	<.0001	
11	<.0001	<.0001	0.2493	<.0001	<.0001	<.0001	0.6867	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	0.0811	<.0001	<.0001	0.4711	<.0001	<.0001	
12	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	
13	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
14	<.0001	0.9610	<.0001	<.0001	<.0001	<.0001	0.0006	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
15	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
16	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	0.0811	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	

17	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0208	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
18	0.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0013	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
19	<.0001	<.0001	0.1410	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
20	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	0.4711	<.0001	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
21	0.9500	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
22	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
23	<.0001	0.1758	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0003	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
24	<.0001	<.0001	0.0212	<.0001	<.0001	<.0001	<.0001	0.0800	<.0001	<.0001	0.9378	0.9999	<.0001	<.0001	<.0001	0.0136	<.0001	0.9560	<.0001	0.0308	<.0001	<.0001	<.0001	
25	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
26	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
27	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
28	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
29	<.0001	0.3239	<.0001	<.0001	<.0001	<.0001	<.0001	0.9966	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0003	<.0001	
30	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0700	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
31	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	
32	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.9660	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0005	<.0001
33	0.0007	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0025	<.0001	0.2840	<.0001	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	
34	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.9994	<.0001	<.0001	



The GLM Procedure

Class Level Information		
Class	Levels	Values
condition	6	excellent fair good like new new salvage

Number of Observations Read	441328
Number of Observations Used	441328

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	29575.3017	5915.0603	20421.4	<.0001
Error	441322	127828.8143	0.2896		
Corrected Total	441327	157404.1160			

R-Square	Coeff Var	Root MSE	log_age Mean
0.187894	22.06048	0.538191	2.439617

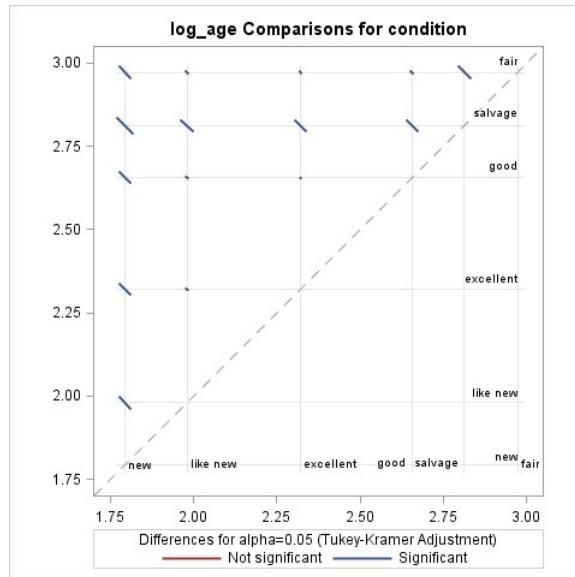
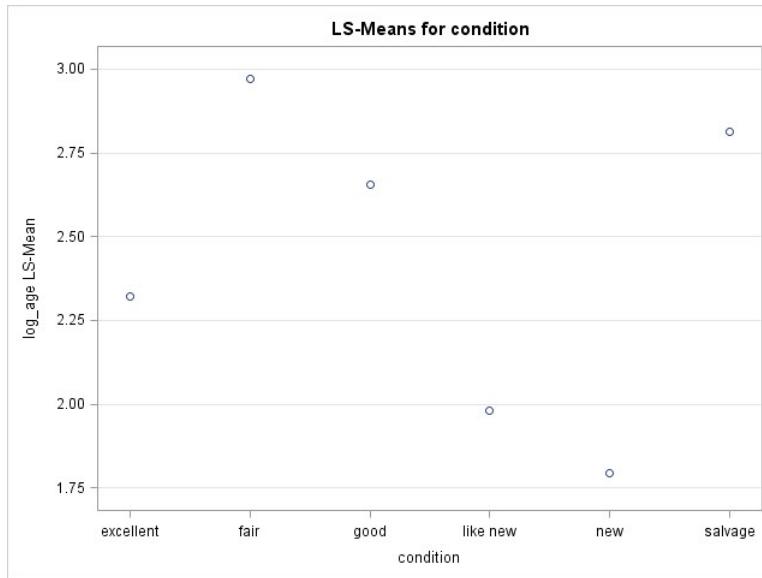
Source	DF	Type I SS	Mean Square	F Value	Pr > F
condition	5	29575.30171	5915.06034	20421.4	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
condition	5	29575.30171	5915.06034	20421.4	<.0001

The GLM Procedure
 Least Squares Means
 Adjustment for Multiple Comparisons: Tukey-Kramer

condition	log_age LSMEAN	LSMEAN Number
excellent	2.32068316	1
fair	2.97271131	2
good	2.65597653	3
like new	1.97983764	4
new	1.79294201	5
salvage	2.81289864	6

Least Squares Means for effect condition						
Pr > t for H0: LSMean(i)=LSMean(j)						
Dependent Variable: log_age						
i/j	1	2	3	4	5	6
1		<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	<.0001
4	<.0001	<.0001	<.0001		<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001	



The GLM Procedure

Class Level Information

Class	Levels	Values
cylinders	8	10 cylinders 12 cylinders 3 cylinders 4 cylinders 5 cylinders 6 cylinders 8 cylinders other

Number of Observations Read	441328
------------------------------------	--------

Number of Observations Used	441328
------------------------------------	--------

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	9055.7878	1293.6840	3848.57	<.0001
Error	441320	148348.3282	0.3361		
Corrected Total	441327	157404.1160			

R-Square	Coeff Var	Root MSE	log_age Mean
0.057532	23.76528	0.579782	2.439617

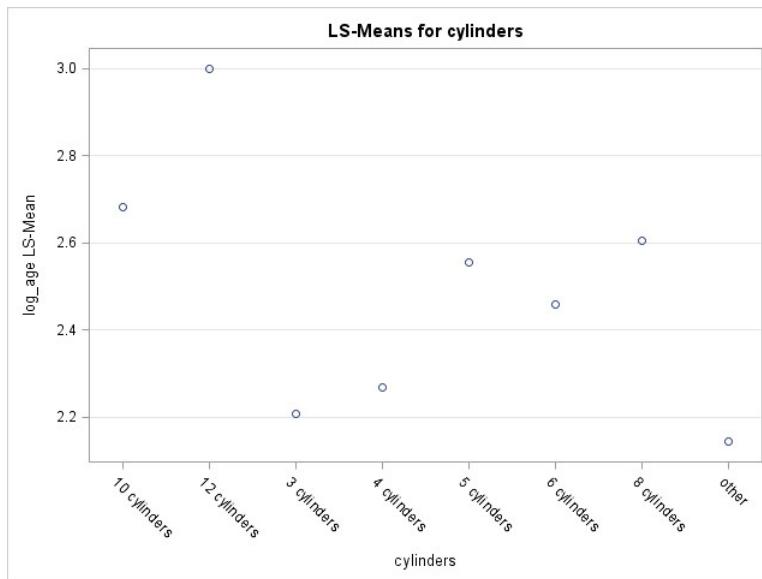
Source	DF	Type I SS	Mean Square	F Value	Pr > F
cylinders	7	9055.787797	1293.683971	3848.57	<.0001

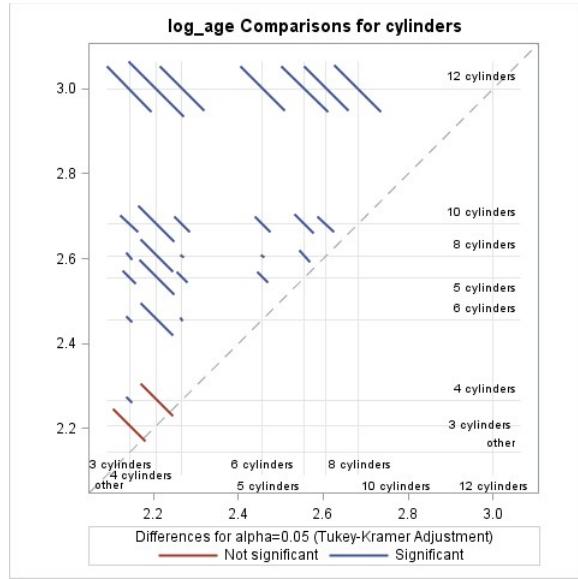
Source	DF	Type III SS	Mean Square	F Value	Pr > F
cylinders	7	9055.787797	1293.683971	3848.57	<.0001

The GLM Procedure
 Least Squares Means
 Adjustment for Multiple Comparisons: Tukey-Kramer

cylinders	log_age LSMEAN	LSMEAN Number
10 cylinders	2.68120737	1
12 cylinders	3.00000399	2
3 cylinders	2.20731222	3
4 cylinders	2.26681171	4
5 cylinders	2.55603747	5
6 cylinders	2.45700713	6
8 cylinders	2.60585431	7
other	2.14282596	8

Least Squares Means for effect cylinders								
Pr > t for H0: LSMean(i)=LSMean(j)								
Dependent Variable: log_age								
i\j	1	2	3	4	5	6	7	8
1		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		0.2625	<.0001	<.0001	<.0001	0.1907
4	<.0001	<.0001	0.2625		<.0001	<.0001	<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
7	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
8	<.0001	<.0001	0.1907	<.0001	<.0001	<.0001	<.0001	





The GLM Procedure

Class Level Information		
Class	Levels	Values
drive	3	4wd fwd rwd

Number of Observations Read	441328
Number of Observations Used	441328

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	8556.1667	4278.0834	12684.3	<.0001
Error	441325	148847.9493	0.3373		
Corrected Total	441327	157404.1160			

R-Square	Coeff Var	Root MSE	log_age Mean
0.054358	23.80513	0.580754	2.439617

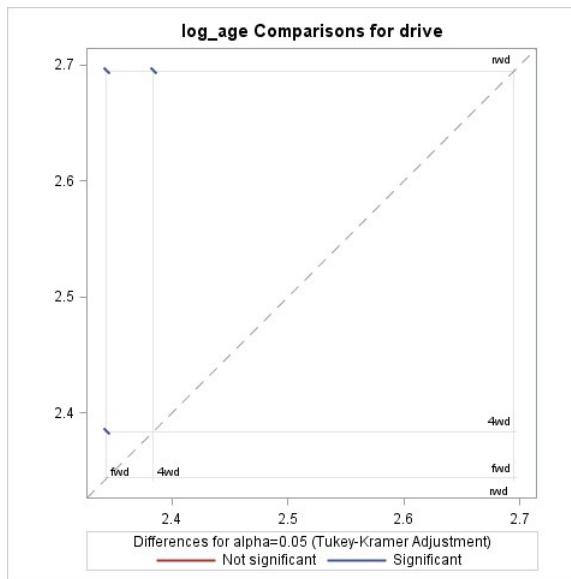
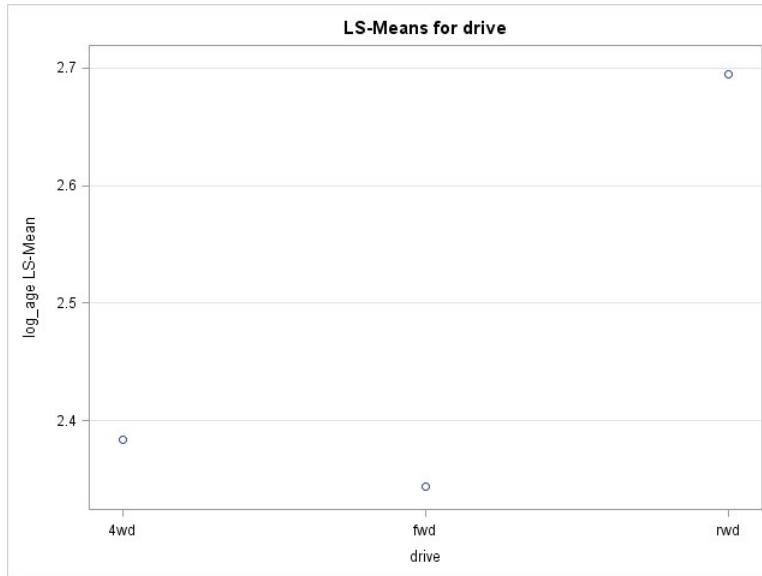
Source	DF	Type I SS	Mean Square	F Value	Pr > F
drive	2	8556.166704	4278.083352	12684.3	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
drive	2	8556.166704	4278.083352	12684.3	<.0001

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

drive	log_age LSMEAN	LSMEAN Number
4wd	2.38398796	1
fwd	2.34356757	2
rwd	2.69497697	3

Least Squares Means for effect drive			
Pr > t for H0: LSMean(i)=LSMean(j)			
Dependent Variable: log_age			
i\j	1	2	3
1		<.0001	<.0001
2	<.0001		<.0001
3	<.0001	<.0001	



The GLM Procedure

Class Level Information		
Class	Levels	Values
type	13	SUV bus converti coupe hatchbac mini-van offroad other pickup sedan truck van wagon
Number of Observations Read		441328
Number of Observations Used		441328

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	6407.8752	533.9896	1560.69	<.0001
Error	441315	150996.2409	0.3422		
Corrected Total	441327	157404.1160			

R-Square	Coeff Var	Root MSE	log_age Mean
0.040710	23.97657	0.584937	2.439617

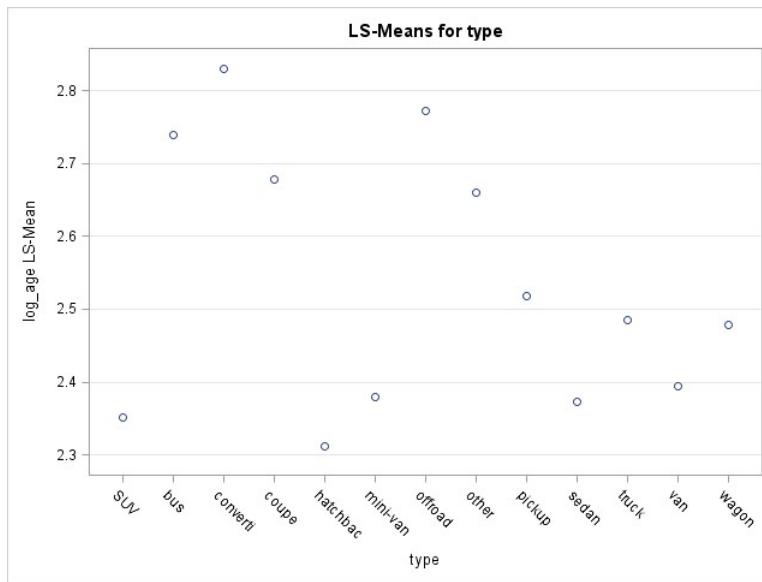
Source	DF	Type I SS	Mean Square	F Value	Pr > F
type	12	6407.875171	533.989598	1560.69	<.0001

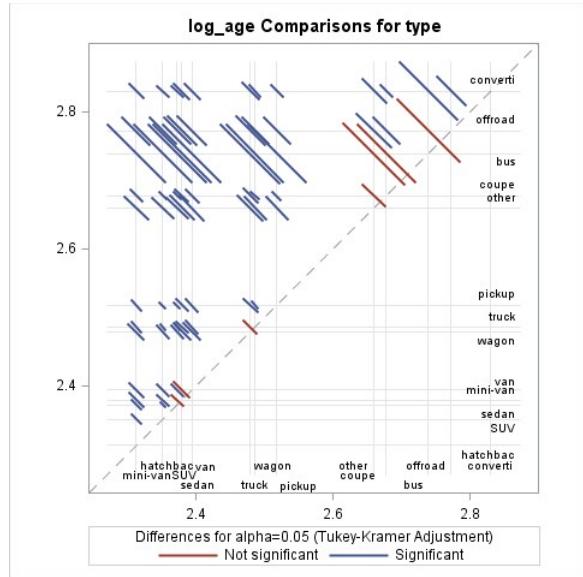
Source	DF	Type III SS	Mean Square	F Value	Pr > F
type	12	6407.875171	533.989598	1560.69	<.0001

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

type	log_age LSMEAN	LSMEAN Number
SUV	2.35114013	1
bus	2.73886174	2
converti	2.82998057	3
coupe	2.67739945	4
hatchbac	2.31263375	5
mini-van	2.37896140	6
offroad	2.77228895	7
other	2.65929381	8
pickup	2.51741844	9
sedan	2.37219899	10
truck	2.48561779	11
van	2.39363273	12
wagon	2.47901463	13

Least Squares Means for effect type Pr > t for H0: LSMean(i)=LSMean(j) Dependent Variable: log_age													
i\j	1	2	3	4	5	6	7	8	9	10	11	12	13
1		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		0.0297	0.4658	<.0001	<.0001	0.9942	0.1636	<.0001	<.0001	<.0001	<.0001	<.0001
3	<.0001	0.0297		<.0001	<.0001	<.0001	0.0006	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
4	<.0001	0.4658	<.0001		<.0001	<.0001	0.8783	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	0.9906	<.0001	0.7676	<.0001
7	<.0001	0.9942	0.0006	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
8	<.0001	0.1636	<.0001	0.8783	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
9	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001
10	<.0001	<.0001	<.0001	<.0001	<.0001	0.9906	<.0001	<.0001	<.0001		<.0001	0.0118	<.0001
11	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	0.9984
12	<.0001	<.0001	<.0001	<.0001	<.0001	0.7676	<.0001	<.0001	<.0001	0.0118	<.0001		<.0001
13	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.9984		<.0001





The GLM Procedure

Class Level Information		
Class	Levels	Values
paint_color	12	black blue brown custom green grey orange purple red silver white yellow
Number of Observations Read		441328
Number of Observations Used		441328

The GLM Procedure

Dependent Variable: log_age

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	7961.5169	723.7743	2137.36	<.0001
Error	441316	149442.5991	0.3386		
Corrected Total	441327	157404.1160			

R-Square	Coeff Var	Root MSE	log_age Mean
0.050580	23.85287	0.581919	2.439617

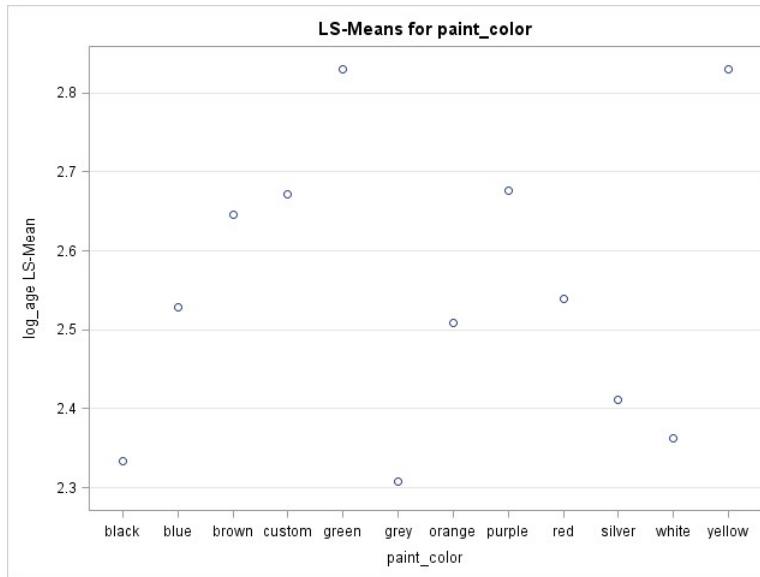
Source	DF	Type I SS	Mean Square	F Value	Pr > F
paint_color	11	7961.516914	723.774265	2137.36	<.0001

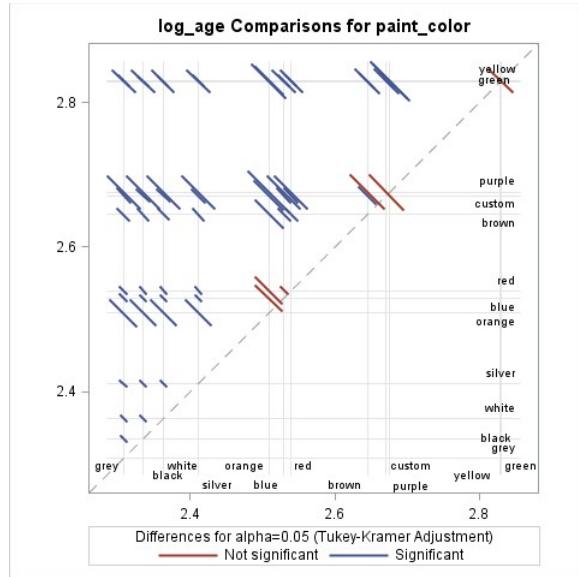
Source	DF	Type III SS	Mean Square	F Value	Pr > F
paint_color	11	7961.516914	723.774265	2137.36	<.0001

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

paint_color	log_age LSMEAN	LSMEAN Number
black	2.33438464	1
blue	2.52935677	2
brown	2.64545887	3
custom	2.67146728	4
green	2.83034080	5
grey	2.30741280	6
orange	2.50884163	7
purple	2.67609105	8
red	2.53973672	9
silver	2.41101608	10
white	2.36250553	11
yellow	2.82964146	12

Least Squares Means for effect paint_color Pr > t for H0: LSMean(i)=LSMean(j) Dependent Variable: log_age												
i\j	1	2	3	4	5	6	7	8	9	10	11	12
1	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001	0.8344	<.0001	0.1675	<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		0.0292	<.0001	<.0001	0.6531	<.0001	<.0001	<.0001	<.0001	<.0001
4	<.0001	<.0001	0.0292		<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	1.0000
6	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
7	<.0001	0.8344	<.0001	<.0001	<.0001	<.0001		0.2436	<.0001	<.0001	<.0001	<.0001
8	<.0001	<.0001	0.6531	1.0000	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001
9	<.0001	0.1675	<.0001	<.0001	<.0001	<.0001	0.2436	<.0001		<.0001	<.0001	<.0001
10	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
11	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
12	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	





The GLM Procedure

Class Level Information		
Class	Levels	Values
condition	6	excellent fair good like new new salvage
Number of Observations Read		441328
Number of Observations Used		441328

The GLM Procedure

Dependent Variable: log_odometer

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	41734.7018	8346.9404	17889.3	<.0001
Error	441322	205915.8507	0.4666		
Corrected Total	441327	247650.5524			

R-Square	Coeff Var	Root MSE	log_odometer Mean
0.168523	5.934619	0.683073	11.50997

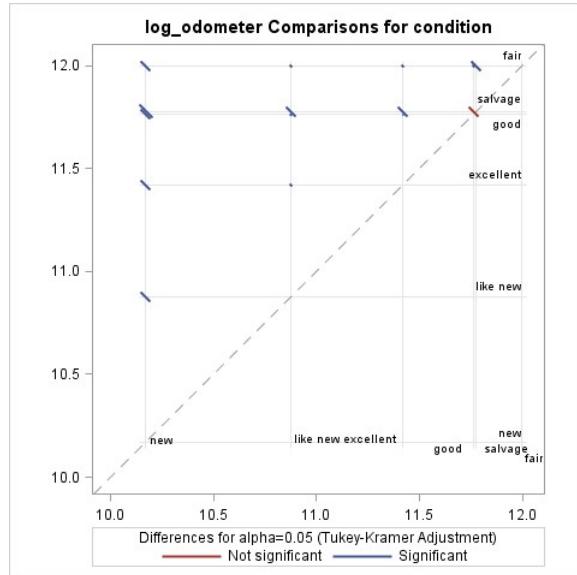
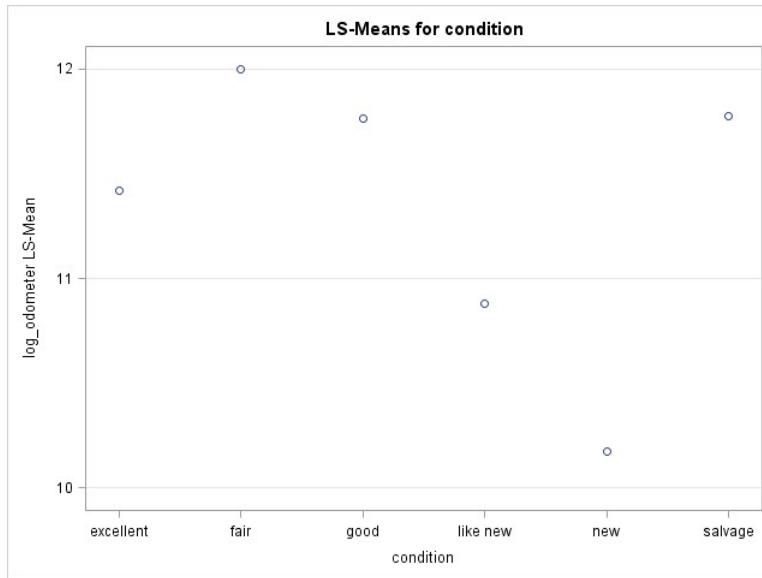
Source	DF	Type I SS	Mean Square	F Value	Pr > F
condition	5	41734.70179	8346.94036	17889.3	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
condition	5	41734.70179	8346.94036	17889.3	<.0001

The GLM Procedure
 Least Squares Means
 Adjustment for Multiple Comparisons: Tukey-Kramer

condition	log_odometer LSMEAN	LSMEAN Number
excellent	11.4206148	1
fair	11.9966477	2
good	11.7614603	3
like new	10.8772583	4
new	10.1718037	5
salvage	11.7764288	6

Least Squares Means for effect condition						
Pr > t for H0: LSMean(i)=LSMean(j)						
Dependent Variable: log_odometer						
i/j	1	2	3	4	5	6
1		<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	0.9457
4	<.0001	<.0001	<.0001		<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001
6	<.0001	<.0001	0.9457	<.0001	<.0001	



The GLM Procedure

Class Level Information

Class	Levels	Values
cylinders	8	10 cylinders 12 cylinders 3 cylinders 4 cylinders 5 cylinders 6 cylinders 8 cylinders other

Number of Observations Read	441328
Number of Observations Used	441328

The GLM Procedure

Dependent Variable: log_odometer

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	4004.7456	572.1065	1036.27	<.0001
Error	441320	243645.8069	0.5521		
Corrected Total	441327	247650.5524			

R-Square	Coeff Var	Root MSE	log_odometer Mean
0.016171	6.455479	0.743024	11.50997

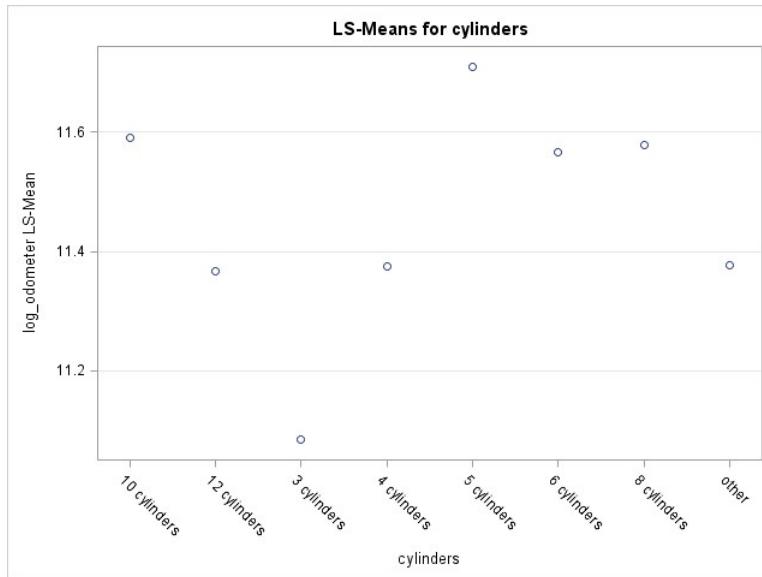
Source	DF	Type I SS	Mean Square	F Value	Pr > F
cylinders	7	4004.745567	572.106510	1036.27	<.0001

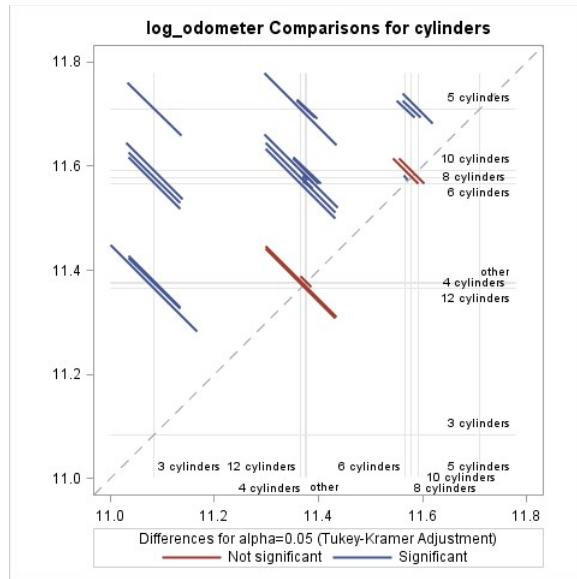
Source	DF	Type III SS	Mean Square	F Value	Pr > F
cylinders	7	4004.745567	572.106510	1036.27	<.0001

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

cylinders	log_odometer LSMEAN	LSMEAN Number
10 cylinders	11.5911121	1
12 cylinders	11.3662656	2
3 cylinders	11.0845124	3
4 cylinders	11.3752447	4
5 cylinders	11.7099553	5
6 cylinders	11.5672556	6
8 cylinders	11.5783664	7
other	11.3780070	8

Least Squares Means for effect cylinders								
Pr > t for H0: LSMean(i)=LSMean(j)								
Dependent Variable: log_odometer								
i\j	1	2	3	4	5	6	7	8
1		<.0001	<.0001	<.0001	<.0001	0.8015	0.9928	<.0001
2	<.0001		<.0001	1.0000	<.0001	0.0002	<.0001	1.0000
3	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
4	<.0001	1.0000	<.0001		<.0001	<.0001	<.0001	0.9999
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001
6	0.8015	0.0002	<.0001	<.0001	<.0001		0.0015	<.0001
7	0.9928	<.0001	<.0001	<.0001	<.0001	0.0015		<.0001
8	<.0001	1.0000	<.0001	0.9999	<.0001	<.0001	<.0001	





The GLM Procedure

Class Level Information		
Class	Levels	Values
drive	3	4wd fwd rwd

Number of Observations Read	441328
Number of Observations Used	441328

The GLM Procedure

Dependent Variable: log_odometer

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1076.2094	538.1047	963.11	<.0001
Error	441325	246574.3431	0.5587		
Corrected Total	441327	247650.5524			

R-Square	Coeff Var	Root MSE	log_odometer Mean
0.004346	6.494122	0.747472	11.50997

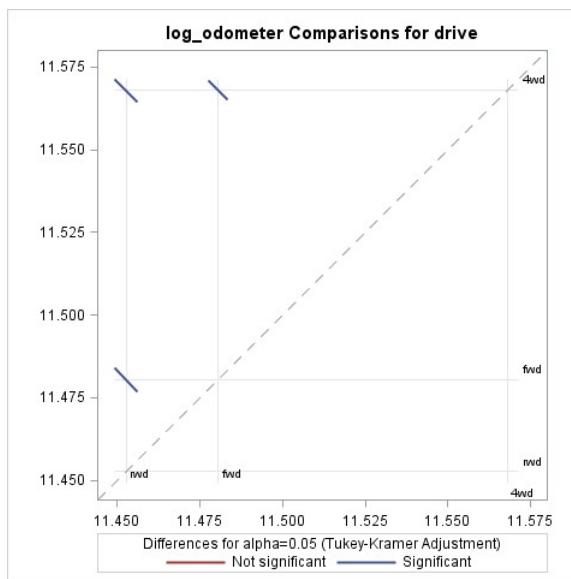
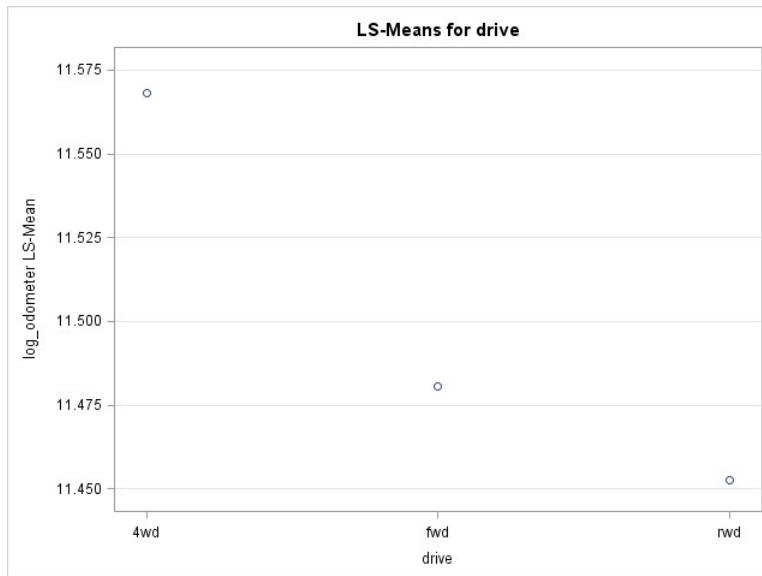
Source	DF	Type I SS	Mean Square	F Value	Pr > F
drive	2	1076.209379	538.104689	963.11	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
drive	2	1076.209379	538.104689	963.11	<.0001

The GLM Procedure
 Least Squares Means
 Adjustment for Multiple Comparisons: Tukey-Kramer

drive	log_odometer LSMEAN	LSMEAN Number
4wd	11.5680122	1
fwd	11.4804815	2
rwd	11.4526376	3

Least Squares Means for effect drive			
Pr > t for H0: LSMean(i)=LSMean(j)			
Dependent Variable: log_odometer			
i\j	1	2	3
1		<.0001	<.0001
2	<.0001		<.0001
3	<.0001	<.0001	



The GLM Procedure

Class Level Information		
Class	Levels	Values
type	13	SUV bus converti coupe hatchbac mini-van offroad other pickup sedan truck van wagon
Number of Observations Read		441328
Number of Observations Used		441328

The GLM Procedure

Dependent Variable: log_odometer

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	5151.9242	429.3270	781.32	<.0001
Error	441315	242498.6283	0.5495		
Corrected Total	441327	247650.5524			

R-Square	Coeff Var	Root MSE	log_odometer Mean
0.020803	6.440300	0.741277	11.50997

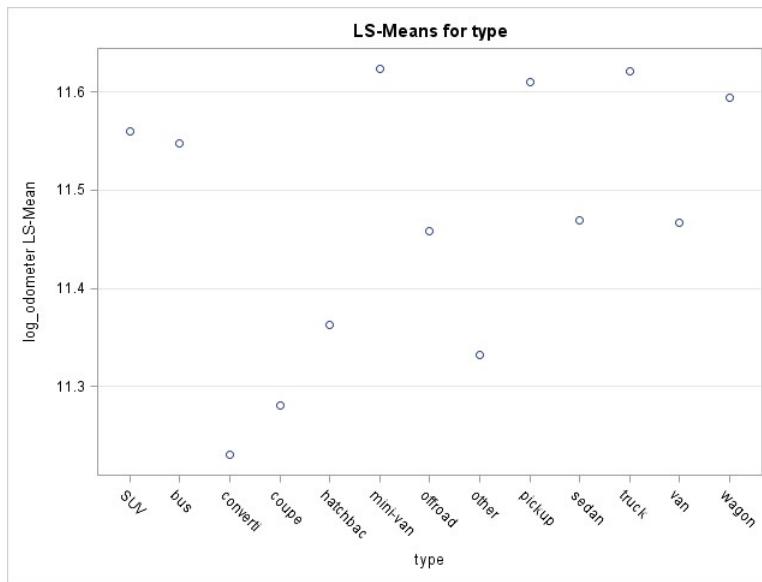
Source	DF	Type I SS	Mean Square	F Value	Pr > F
type	12	5151.924153	429.327013	781.32	<.0001

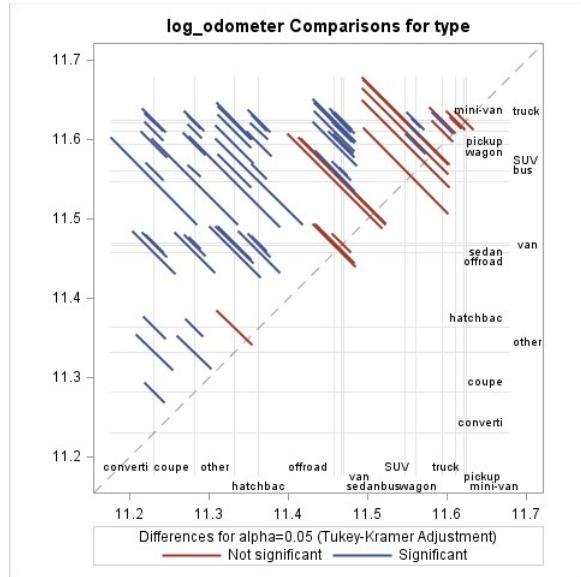
Source	DF	Type III SS	Mean Square	F Value	Pr > F
type	12	5151.924153	429.327013	781.32	<.0001

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

type	log_odometer LSMEAN	LSMEAN Number
SUV	11.5601979	1
bus	11.5472499	2
converti	11.2302263	3
coupe	11.2810839	4
hatchbac	11.3625969	5
mini-van	11.6237393	6
offroad	11.4577764	7
other	11.3316429	8
pickup	11.6103281	9
sedan	11.4691208	10
truck	11.6215251	11
van	11.4667378	12
wagon	11.5941745	13

Least Squares Means for effect type Pr > t for H0: LSMean(i)=LSMean(j) Dependent Variable: log_odometer													
i\j	1	2	3	4	5	6	7	8	9	10	11	12	13
1		1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0012
2	1.0000		<.0001	<.0001	<.0001	0.5167	0.3811	<.0001	0.7814	0.4500	0.5397	0.4328	0.9743
3	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
4	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	0.0082	<.0001	<.0001	<.0001	<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.5449	<.0001	<.0001	<.0001	<.0001	<.0001
6	<.0001	0.5167	<.0001	<.0001	<.0001		<.0001	<.0001	0.8437	<.0001	1.0000	<.0001	0.1399
7	<.0001	0.3811	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.9999	<.0001	1.0000	<.0001
8	<.0001	<.0001	<.0001	0.0082	0.5449	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
9	<.0001	0.7814	<.0001	<.0001	<.0001	0.8437	<.0001	<.0001		<.0001	0.3837	<.0001	0.7706
10	<.0001	0.4500	<.0001	<.0001	<.0001	<.0001	0.9999	<.0001	<.0001		<.0001	1.0000	<.0001
11	<.0001	0.5397	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	0.3837	<.0001		<.0001	0.0421
12	<.0001	0.4328	<.0001	<.0001	<.0001	<.0001	1.0000	<.0001	<.0001	1.0000	<.0001		<.0001
13	0.0012	0.9743	<.0001	<.0001	<.0001	0.1399	<.0001	<.0001	0.7706	<.0001	0.0421	<.0001	





The GLM Procedure

Class Level Information		
Class	Levels	Values
paint_color	12	black blue brown custom green grey orange purple red silver white yellow
Number of Observations Read		441328
Number of Observations Used		441328

The GLM Procedure

Dependent Variable: log_odometer

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	2433.7737	221.2522	398.19	<.0001
Error	441316	245216.7788	0.5556		
Corrected Total	441327	247650.5524			

R-Square	Coeff Var	Root MSE	log_odometer Mean
0.009827	6.476286	0.745419	11.50997

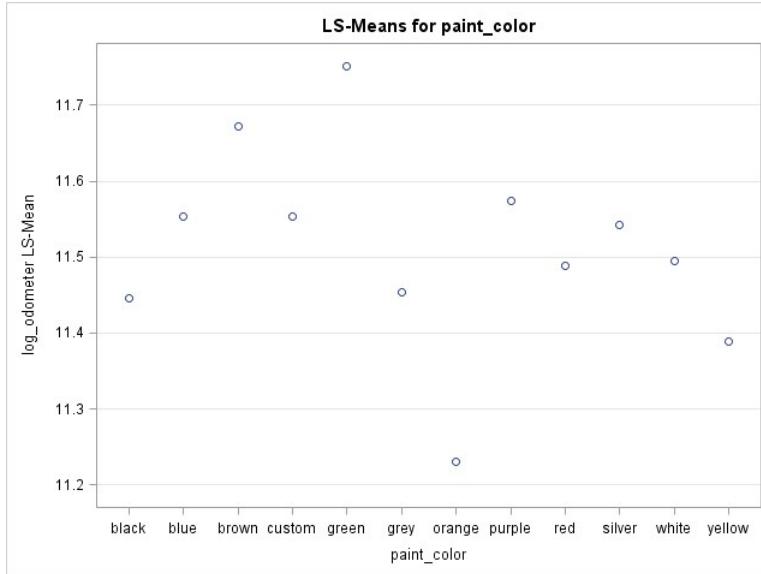
Source	DF	Type I SS	Mean Square	F Value	Pr > F
paint_color	11	2433.773670	221.252152	398.19	<.0001

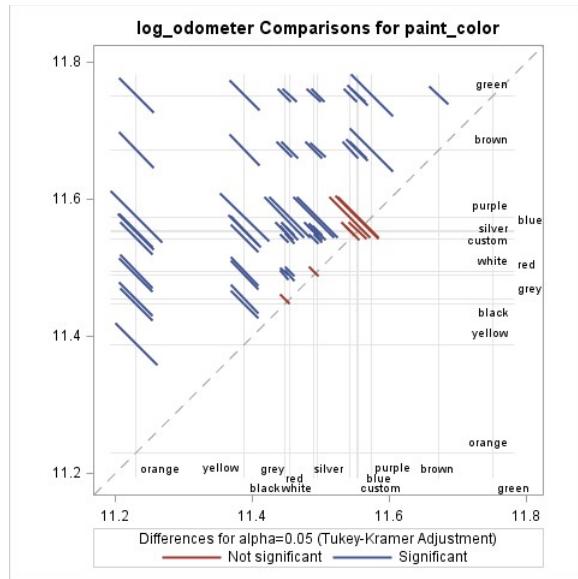
Source	DF	Type III SS	Mean Square	F Value	Pr > F
paint_color	11	2433.773670	221.252152	398.19	<.0001

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

paint_color	log_odometer LSMEAN	LSMEAN Number
black	11.4462007	1
blue	11.5541385	2
brown	11.6718375	3
custom	11.5529754	4
green	11.7515647	5
grey	11.4540691	6
orange	11.2297919	7
purple	11.5735919	8
red	11.4888015	9
silver	11.5424597	10
white	11.4945138	11
yellow	11.3880944	12

Least Squares Means for effect paint_color Pr > t for H0: LSMean(i)=LSMean(j) Dependent Variable: log_odometer												
i\j	1	2	3	4	5	6	7	8	9	10	11	12
1	<.0001	<.0001	<.0001	<.0001	0.7901	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0002
2	<.0001		<.0001	1.0000	<.0001	<.0001	<.0001	0.9960	<.0001	0.2520	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
4	<.0001	1.0000	<.0001		<.0001	<.0001	<.0001	0.9960	<.0001	0.9722	<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
6	0.7901	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
7	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
8	<.0001	0.9960	<.0001	0.9960	<.0001	<.0001	<.0001		0.0002	0.8625	0.0008	<.0001
9	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0002		<.0001	0.9667	<.0001
10	<.0001	0.2520	<.0001	0.9722	<.0001	<.0001	<.0001	0.8625	<.0001		<.0001	<.0001
11	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0008	0.9667	<.0001		<.0001
12	0.0002	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	





The GLM Procedure

Class Level Information		
Class	Levels	Values
fuel	5	diesel electric gas hybrid other
title_status	6	clean lien missing rebuilt salvage parts onl
transmission	3	automatic manual other

Number of Observations Read	1493278
Number of Observations Used	1493278

The GLM Procedure					
Dependent Variable: log_price					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	615985.178	51332.098	77749.1	<.0001
Error	1.49E6	985894.560	0.660		
Corrected Total	1.49E6	1601879.738			
R-Square	Coeff Var	Root MSE	log_price Mean		
0.384539	9.197130	0.812544	8.834755		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
log_age	1	521096.7980	521096.7980	789269	<.0001
fuel	4	74424.0306	18606.0076	28181.2	<.0001
title_status	5	17138.0637	3427.6127	5191.56	<.0001
transmission	2	3326.2856	1663.1428	2519.05	<.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
log_age	1	481193.7959	481193.7959	728830	<.0001
fuel	4	71929.7263	17982.4316	27236.7	<.0001
title_status	5	17401.7723	3480.3545	5271.45	<.0001
transmission	2	3326.2856	1663.1428	2519.05	<.0001
Parameter	Estimate		Standard Error	t Value	Pr > t
Intercept	9.612392773		B 0.01664976	577.33	<.0001
log_age	-0.898533306		0.00105250	-853.72	<.0001
fuel diesel	0.641082792		B 0.00498403	128.63	<.0001
fuel electric	-0.316091499		B 0.02366759	-13.36	<.0001
fuel gas	-0.244395652		B 0.00429852	-56.86	<.0001
fuel hybrid	-0.252637438		B 0.00912867	-27.68	<.0001
fuel other	0.000000000		B .	.	.
title_status clean	1.652348549		B 0.01556788	106.14	<.0001
title_status lien	1.934548839		B 0.01663475	116.30	<.0001
title_status missing	0.977651556		B 0.01834601	53.29	<.0001
title_status rebuilt	1.481138926		B 0.01604784	92.30	<.0001
title_status salvage	1.255873236		B 0.01630047	77.05	<.0001
title_status parts onl	0.000000000		B .	.	.
transmission automatic	-0.052636951		B 0.00528566	-9.96	<.0001
transmission manual	0.103643411		B 0.00567423	18.27	<.0001
transmission other	0.000000000		B .	.	.

Note: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

The GLM Procedure

Class Level Information		
Class	Levels	Values
manufacturer	34	acura audi bmw buick cadillac chevrolet chrysler dodge fiat ford gmc honda hyundai infiniti jaguar jeep kia lexus lincoln mazda mercedes mercury mini mitsubishi nissan pontiac ram rover saturn subaru toyota volkswagen volvo other

Number of Observations Read	1493278
Number of Observations Used	1493278

The GLM Procedure					
Dependent Variable: log_price					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	34	619030.784	18206.788	27661.6	<.0001
Error	1.49E6	982848.954	0.658		
Corrected Total	1.49E6	1601879.738			
R-Square	Coeff Var	Root MSE	log_price Mean		
0.386440	9.182980	0.811294	8.834755		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
log_age	1	521096.7980	521096.7980	791703	<.0001
manufacturer	33	97933.9858	2967.6965	4508.82	<.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
log_age	1	510626.1385	510626.1385	775795	<.0001
manufacturer	33	97933.9858	2967.6965	4508.82	<.0001
Parameter	Estimate		Standard Error	t Value	Pr > t
Intercept	11.67998703		B 0.01934005	603.93	<.0001
log_age	-0.91179136		0.00103519	-880.79	<.0001
manufacturer acura	-0.86742663		B 0.02006342	-43.23	<.0001
manufacturer audi	-0.55984240		B 0.02026168	-27.63	<.0001
manufacturer bmw	-0.46365622		B 0.01950542	-23.77	<.0001
manufacturer buick	-0.89020379		B 0.01986936	-44.80	<.0001
manufacturer cadillac	-0.47306003		B 0.01979686	-23.90	<.0001
manufacturer chevrolet	-0.47596619		B 0.01915690	-24.85	<.0001
manufacturer chrysler	-1.02437091		B 0.01965864	-52.11	<.0001
manufacturer dodge	-0.78151840		B 0.01933049	-40.43	<.0001
manufacturer fiat	-0.94782378		B 0.02682164	-35.34	<.0001
manufacturer ford	-0.49559572		B 0.01915195	-25.88	<.0001
manufacturer gmc	-0.28737663		B 0.01934987	-14.85	<.0001
manufacturer honda	-0.94177631		B 0.01929223	-48.82	<.0001
manufacturer hyundai	-1.04862674		B 0.01964236	-53.39	<.0001
manufacturer infiniti	-0.57187957		B 0.02048624	-27.92	<.0001
manufacturer jaguar	-0.49632065		B 0.02303129	-21.55	<.0001
manufacturer jeep	-0.43443853		B 0.01933194	-22.47	<.0001
manufacturer kia	-1.05157563		B 0.01982138	-53.05	<.0001
manufacturer lexus	-0.39497382		B 0.01996102	-19.79	<.0001
manufacturer lincoln	-0.68051203		B 0.02044438	-33.29	<.0001
manufacturer mazda	-0.95731289		B 0.01984601	-48.24	<.0001
manufacturer mercedes	-0.32901687		B 0.01966647	-16.73	<.0001
manufacturer mercury	-1.10232235		B 0.02078932	-53.02	<.0001
manufacturer mini	-0.70779579		B 0.02171974	-32.59	<.0001
manufacturer mitsubishi	-1.07686045		B 0.02069795	-52.03	<.0001
manufacturer nissan	-0.94655933		B 0.01932818	-48.97	<.0001
manufacturer pontiac	-0.92643823		B 0.01987464	-46.61	<.0001
manufacturer ram	-0.03984956		B 0.01941631	-2.05	0.0401
manufacturer rover	-0.13645735		B 0.02223451	-6.14	<.0001
manufacturer saturn	-1.41639493		B 0.02092782	-67.68	<.0001
manufacturer subaru	-0.72266733		B 0.01970657	-36.67	<.0001
manufacturer toyota	-0.61944080		B 0.01924773	-32.18	<.0001
manufacturer volkswagen	-0.81911298		B 0.01956001	-41.88	<.0001
manufacturer volvo	-0.91483996		B 0.02067453	-44.25	<.0001
manufacturer other	0.00000000		B

Note: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

The GLM Procedure

Class Level Information		
Class	Levels	Values
condition	6	excellent fair good like new new salvage
cylinders	8	10 cylinders 12 cylinders 3 cylinders 4 cylinders 5 cylinders 6 cylinders 8 cylinders other
drive	3	fwd rwd 4wd

Number of Observations Read	441328
Number of Observations Used	441328

The GLM Procedure					
Dependent Variable: log_price					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	16	253396.0195	15837.2512	47263.2	<.0001
Error	441311	147877.4001	0.3351		
Corrected Total	441327	401273.4197			
R-Square	Coeff Var	Root MSE	log_price Mean		
0.631480	6.518141	0.578867	8.880855		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
log_age	1	145804.0439	145804.0439	435123	<.0001
log_odometer	1	18605.6277	18605.6277	55524.8	<.0001
condition	5	32753.2641	6550.6528	19549.1	<.0001
cylinders	7	40269.5719	5752.7960	17168.1	<.0001
drive	2	15963.5120	7981.7560	23820.0	<.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
log_age	1	43881.38498	43881.38498	130955	<.0001
log_odometer	1	11849.27691	11849.27691	35361.8	<.0001
condition	5	29544.41793	5908.88359	17633.9	<.0001
cylinders	7	11866.83760	1695.26251	5059.18	<.0001
drive	2	15963.51200	7981.75600	23820.0	<.0001
Parameter	Estimate		Standard Error	t Value	Pr > t
Intercept	12.82102336		B 0.02155361	594.84	<.0001
log_age	-0.70050183		0.00193574	-361.88	<.0001
log_odometer	-0.27698836		0.00147297	-188.05	<.0001
condition excellent	1.46748612		B 0.01405023	104.45	<.0001
condition fair	0.39144406		B 0.01441897	27.15	<.0001
condition good	1.16247241		B 0.01404423	82.77	<.0001
condition like new	1.44691340		B 0.01427400	101.37	<.0001
condition new	1.12326102		B 0.01975394	56.86	<.0001
condition salvage	0.00000000		B .	.	.
cylinders 10 cylinders	0.31015242		B 0.01329205	23.33	<.0001
cylinders 12 cylinders	0.38009359		B 0.03495810	10.87	<.0001
cylinders 3 cylinders	-0.61063134		B 0.02568106	-23.78	<.0001
cylinders 4 cylinders	-0.32594400		B 0.00540527	-60.30	<.0001
cylinders 5 cylinders	-0.26402011		B 0.01001218	-26.37	<.0001
cylinders 6 cylinders	-0.18016605		B 0.00533035	-33.80	<.0001
cylinders 8 cylinders	0.15128080		B 0.00544797	27.77	<.0001
cylinders other	0.00000000		B .	.	.
drive fwd	-0.49930729		B 0.00229314	-217.74	<.0001
drive rwd	-0.13494144		B 0.00238495	-56.58	<.0001
drive 4wd	0.00000000		B .	.	.

Note: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

The GLM Procedure

Class Level Information		
Class	Levels	Values
type	13	SUV bus converti coupe hatchbac mini-van offroad pickup sedan truck van wagon other
paint_color	12	black blue brown green grey orange purple red silver white yellow custom
Number of Observations Read		441328
Number of Observations Used		441328

The GLM Procedure					
Dependent Variable: log_price					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	210097.4023	8403.8961	19399.2	<.0001
Error	441302	191176.0173	0.4332		
Corrected Total	441327	401273.4197			
R-Square	Coeff Var	Root MSE	log_price Mean		
0.523577	7.411293	0.658186	8.880855		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
log_age	1	145804.0439	145804.0439	336567	<.0001
log_odometer	1	18605.6277	18605.6277	42948.4	<.0001
type	12	43969.8856	3664.1571	8458.17	<.0001
paint_color	11	1717.8451	156.1677	360.49	<.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
log_age	1	56460.50868	56460.50868	130331	<.0001
log_odometer	1	17356.60834	17356.60834	40065.2	<.0001
type	12	42323.63850	3526.96987	8141.50	<.0001
paint_color	11	1717.84510	156.16774	360.49	<.0001
Parameter	Estimate		Standard Error	t Value	Pr > t
Intercept	14.68913103		B	0.02042256	719.26 <.0001
log_age	-0.77123760		B	0.00213631	-361.01 <.0001
log_odometer	-0.33074449		B	0.00165238	-200.16 <.0001
type SUV	-0.01912623		B	0.01131892	-1.69 0.0911
type bus	0.37551053		B	0.03102759	12.10 <.0001
type converti	0.26779246		B	0.01251803	21.39 <.0001
type coupe	-0.03582146		B	0.01171610	-3.06 0.0022
type hatchbac	-0.43384369		B	0.01218340	-35.61 <.0001
type mini-van	-0.37052350		B	0.01257901	-29.46 <.0001
type offroad	0.37621092		B	0.01742038	21.60 <.0001
type pickup	0.33788954		B	0.01153014	29.30 <.0001
type sedan	-0.42728121		B	0.01129058	-37.84 <.0001
type truck	0.40331506		B	0.01142219	35.31 <.0001
type van	-0.10845854		B	0.01273539	-8.52 <.0001
type wagon	-0.29423580		B	0.01300717	-22.62 <.0001
type other	0.00000000		B	.	.
paint_color black	0.00435113		B	0.00681937	0.64 0.5234
paint_color blue	-0.11978805		B	0.00703158	-17.04 <.0001
paint_color brown	-0.13654849		B	0.00856618	-15.94 <.0001
paint_color green	-0.24409450		B	0.00799297	-30.54 <.0001
paint_color grey	-0.04522703		B	0.00706922	-6.40 <.0001
paint_color orange	0.00812714		B	0.01430908	0.57 0.5701
paint_color purple	-0.23734662		B	0.01707889	-13.90 <.0001
paint_color red	-0.12058141		B	0.00704129	-17.12 <.0001
paint_color silver	-0.08364177		B	0.00689901	-12.12 <.0001
paint_color white	-0.02182426		B	0.00677128	-3.22 0.0013
paint_color yellow	0.08264198		B	0.01258758	6.57 <.0001
paint_color custom	0.00000000		B	.	.

Note: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.