

The SAS System

The CONTENTS Procedure

Data Set Name	PROJECT.USEDCARSPROJECT	Observations	509577
Member Type	DATA	Variables	25
Engine	V9	Indexes	0
Created	04/23/2020 16:40:19	Observation Length	9576
Last Modified	04/23/2020 16:40:19	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	262144
Number of Data Set Pages	18874
First Data Page	1
Max Obs per Page	27
Obs in First Data Page	26
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	E:\Users\bsp190005\Documents\My SAS Files\Codes\Project\usedcarsproject.sas7bdat
Release Created	9.0401M4
Host Created	X64_SR12R2
Owner Name	CAMPUS\bsp190005
File Size	5GB
File Size (bytes)	4947968000

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
9	condition	Char	9	\$9.	\$9.
22	county	Char	1	\$1.	\$1.
10	cylinders	Char	11	\$11.	\$11.
21	description	Char	9212	\$9212.	\$9212.
16	drive	Char	3	\$3.	\$3.
11	fuel	Char	6	\$6.	\$6.
1	id	Num	8	BEST12.	BEST32.
20	image_url	Char	59	\$59.	\$59.
24	lat	Num	8	BEST12.	BEST32.
25	long	Num	8	BEST12.	BEST32.
7	manufacturer	Char	13	\$13.	\$13.
8	model	Char	11	\$11.	\$11.
12	odometer	Num	8	BEST12.	BEST32.
19	paint_color	Char	6	\$6.	\$6.

5	price	Num	8	BEST12.	BEST32.
3	region	Char	14	\$14.	\$14.
4	region_url	Char	35	\$35.	\$35.
17	size	Char	9	\$9.	\$9.
23	state	Char	2	\$2.	\$2.
13	title_status	Char	5	\$5.	\$5.
14	transmission	Char	9	\$9.	\$9.
18	type	Char	9	\$9.	\$9.
2	url	Char	97	\$97.	\$97.
15	vin	Char	17	\$17.	\$17.
6	year	Num	8	BEST12.	BEST32.

Statistics for Numeric columns**The MEANS Procedure**

Variable	N	Mean	Median	Std Dev	Lower Quartile	Upper Quartile	Quartile Range	Minimum	Maximum
price	509577	54796.84	9377.00	9575025.12	3995.00	17955.00	13960.00	0.00	3600028900
odometer	417253	101729.96	94894.00	107378.99	49488.00	138778.00	89290.00	0.00	10000000.0

Content description after outliers removed and columns dropped
The CONTENTS Procedure

Data Set Name	PROJECT.USEDCARSPROJECT3	Observations	411825
Member Type	DATA	Variables	16
Engine	V9	Indexes	0
Created	04/23/2020 16:44:03	Observation Length	136
Last Modified	04/23/2020 16:44:03	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	857
First Data Page	1
Max Obs per Page	481
Obs in First Data Page	461
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	E:\Users\bsp190005\Documents\My SAS Files\Codes\Projectusedcarsproject3.sas7bdat
Release Created	9.0401M4
Host Created	X64_SR12R2
Owner Name	CAMPUS\bsp190005
File Size	54MB
File Size (bytes)	56229888

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
6	condition	Char	9	\$9.	\$9.
7	cylinders	Char	11	\$11.	\$11.
12	drive	Char	3	\$3.	\$3.
8	fuel	Char	6	\$6.	\$6.
16	logPrice	Num	8		
4	manufacturer	Char	13	\$13.	\$13.
5	model	Char	11	\$11.	\$11.
9	odometer	Num	8	BEST12.	BEST32.
14	paint_color	Char	6	\$6.	\$6.
2	price	Num	8	BEST12.	BEST32.
1	region	Char	14	\$14.	\$14.
15	state	Char	2	\$2.	\$2.
10	title_status	Char	5	\$5.	\$5.

11	transmission	Char	9	\$9.	\$9.
13	type	Char	9	\$9.	\$9.
3	year	Num	8	BEST12.	BEST32.

Content description after outliers removed and columns dropped
The CONTENTS Procedure

Data Set Name	PROJECT.USEDCARSPROJECT4	Observations	411825
Member Type	DATA	Variables	17
Engine	V9	Indexes	0
Created	04/23/2020 17:01:46	Observation Length	144
Last Modified	04/23/2020 17:01:46	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	908
First Data Page	1
Max Obs per Page	454
Obs in First Data Page	435
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	E:\Users\bsp190005\Documents\My SAS Files\Codes\Projectusedcarsproject4.sas7bdat
Release Created	9.0401M4
Host Created	X64_SR12R2
Owner Name	CAMPUS\bsp190005
File Size	57MB
File Size (bytes)	59572224

Alphabetic List of Variables and Attributes						
#	Variable	Type	Len	Format	Informat	
17	Age	Num	8			
6	condition	Char	9	\$9.	\$9.	
7	cylinders	Char	11	\$11.	\$11.	
12	drive	Char	3	\$3.	\$3.	
8	fuel	Char	6	\$6.	\$6.	
16	logPrice	Num	8			
4	manufacturer	Char	13	\$13.	\$13.	
5	model	Char	11	\$11.	\$11.	
9	odometer	Num	8	BEST12.	BEST32.	
14	paint_color	Char	6	\$6.	\$6.	
2	price	Num	8	BEST12.	BEST32.	
1	region	Char	14	\$14.	\$14.	
15	state	Char	2	\$2.	\$2.	

10	title_status	Char	5	\$5.	\$5.
11	transmission	Char	9	\$9.	\$9.
13	type	Char	9	\$9.	\$9.
3	year	Num	8	BEST12.	BEST32.

Content description after outliers removed and columns dropped
The UNIVARIATE Procedure
Variable: price

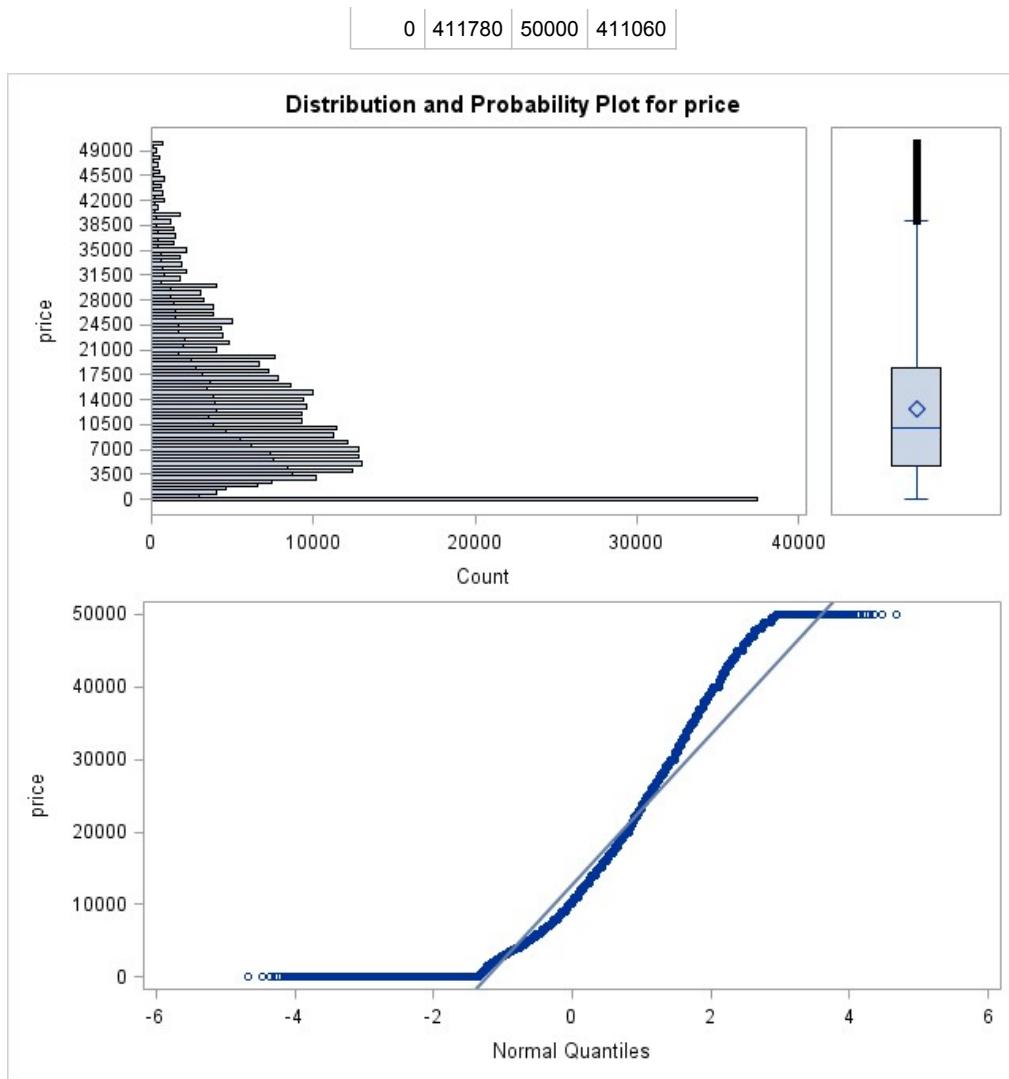
Moments			
N	411825	Sum Weights	411825
Mean	12714.4285	Sum Observations	5236119507
Std Deviation	10393.8516	Variance	108032152
Skewness	0.99900146	Kurtosis	0.61972872
Uncorrected SS	1.11064E14	Corrected SS	4.44902E13
Coeff Variation	81.7484769	Std Error Mean	16.1964621

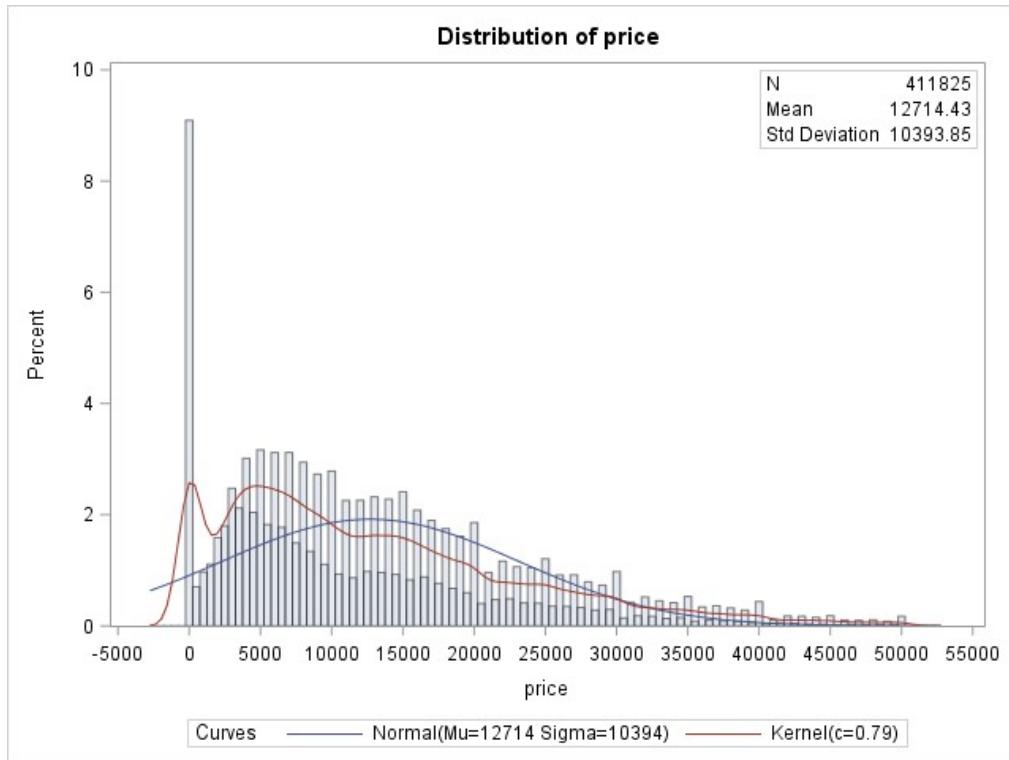
Basic Statistical Measures			
Location		Variability	
Mean	12714.43	Std Deviation	10394
Median	10000.00	Variance	108032152
Mode	0.00	Range	50000
		Interquartile Range	13719

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	785.0127	Pr > t 	<.0001
Sign	M	188931	Pr >= M 	<.0001
Signed Rank	S	3.57E10	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	50000
99%	43988
95%	33500
90%	27888
75% Q3	18499
50% Median	10000
25% Q1	4780
10%	900
5%	0
1%	0
0% Min	0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	411806	50000	392363
0	411805	50000	399082
0	411804	50000	401456
0	411788	50000	408967



Distribution Analysis - Continous Variables**The UNIVARIATE Procedure**

Distribution Analysis - Continous Variables

The UNIVARIATE Procedure
Fitted Normal Distribution for price

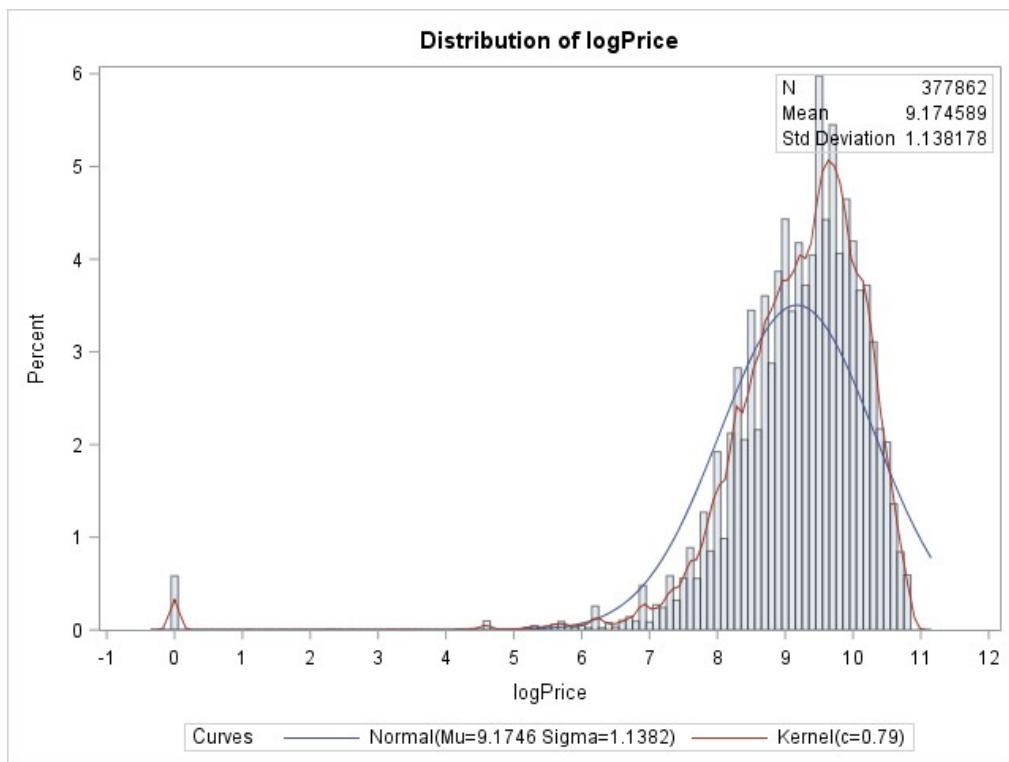
Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	12714.43
Std Dev	Sigma	10393.85

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic	p Value		
Kolmogorov-Smirnov	D	0.11061	Pr > D	<0.010
Cramer-von Mises	W-Sq	1359.11675	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	8654.45110	Pr > A-Sq	<0.005

Percent	Quantile	
	Observed	Estimated
1.0	0.0	-11465.286
5.0	0.0	-4381.936
10.0	900.0	-605.828
25.0	4780.0	5703.882
50.0	10000.0	12714.428
75.0	18499.0	19724.975
90.0	27888.0	26034.685
95.0	33500.0	29810.793
99.0	43988.0	36894.143

Distribution Analysis - Continous Variables

The UNIVARIATE Procedure



Distribution Analysis - Continous Variables

The UNIVARIATE Procedure
Fitted Normal Distribution for logPrice

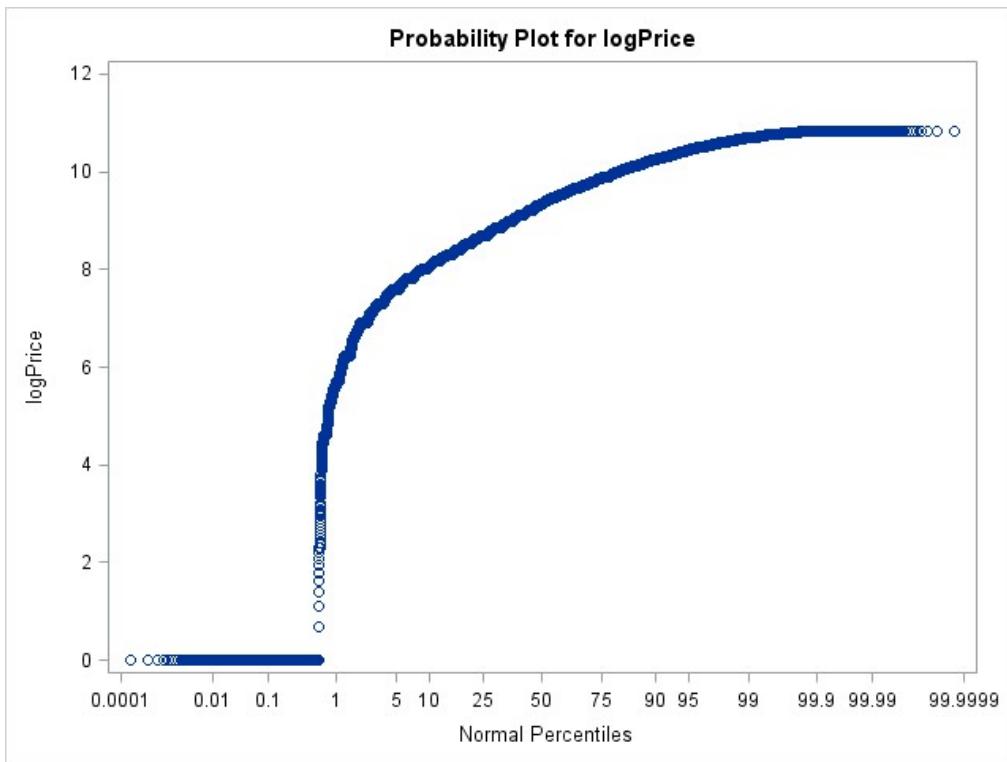
Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	9.174589
Std Dev	Sigma	1.138178

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.09222	Pr > D	<0.010
Cramer-von Mises	W-Sq	1444.72973	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	9921.53892	Pr > A-Sq	<0.005

Percent	Quantile	
	Observed	Estimated
1.0	5.66643	6.52679
5.0	7.60090	7.30245
10.0	8.05516	7.71596
25.0	8.69785	8.40690
50.0	9.35010	9.17459
75.0	9.87298	9.94228
90.0	10.25748	10.63322
95.0	10.43412	11.04673
99.0	10.69194	11.82239

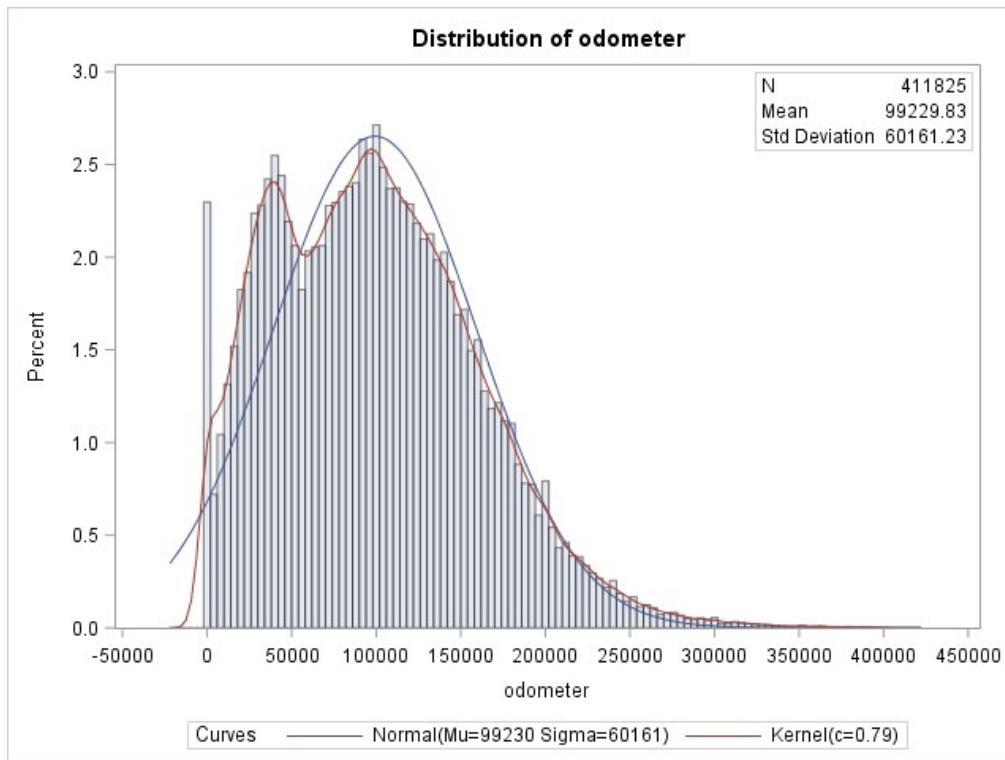
Distribution Analysis - Continous Variables

The UNIVARIATE Procedure



Distribution Analysis - Continous Variables

The UNIVARIATE Procedure



Distribution Analysis - Continous Variables

The UNIVARIATE Procedure
Fitted Normal Distribution for odometer

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	99229.83
Std Dev	Sigma	60161.23

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic	p Value		
Kolmogorov-Smirnov	D	0.04953	Pr > D	<0.010
Cramer-von Mises	W-Sq	233.27683	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	1858.47284	Pr > A-Sq	<0.005

Percent	Quantile	
	Observed	Estimated
1.0	89.0000	-40726.121
5.0	12900.0000	273.413
10.0	24692.0000	22130.113
25.0	50346.0000	58651.700
50.0	95171.0000	99229.834
75.0	138700.0000	139807.969
90.0	178805.0000	176329.556
95.0	204000.0000	198186.256
99.0	260000.0000	239185.789

Distribution analysis for price

The UNIVARIATE Procedure
Variable: price

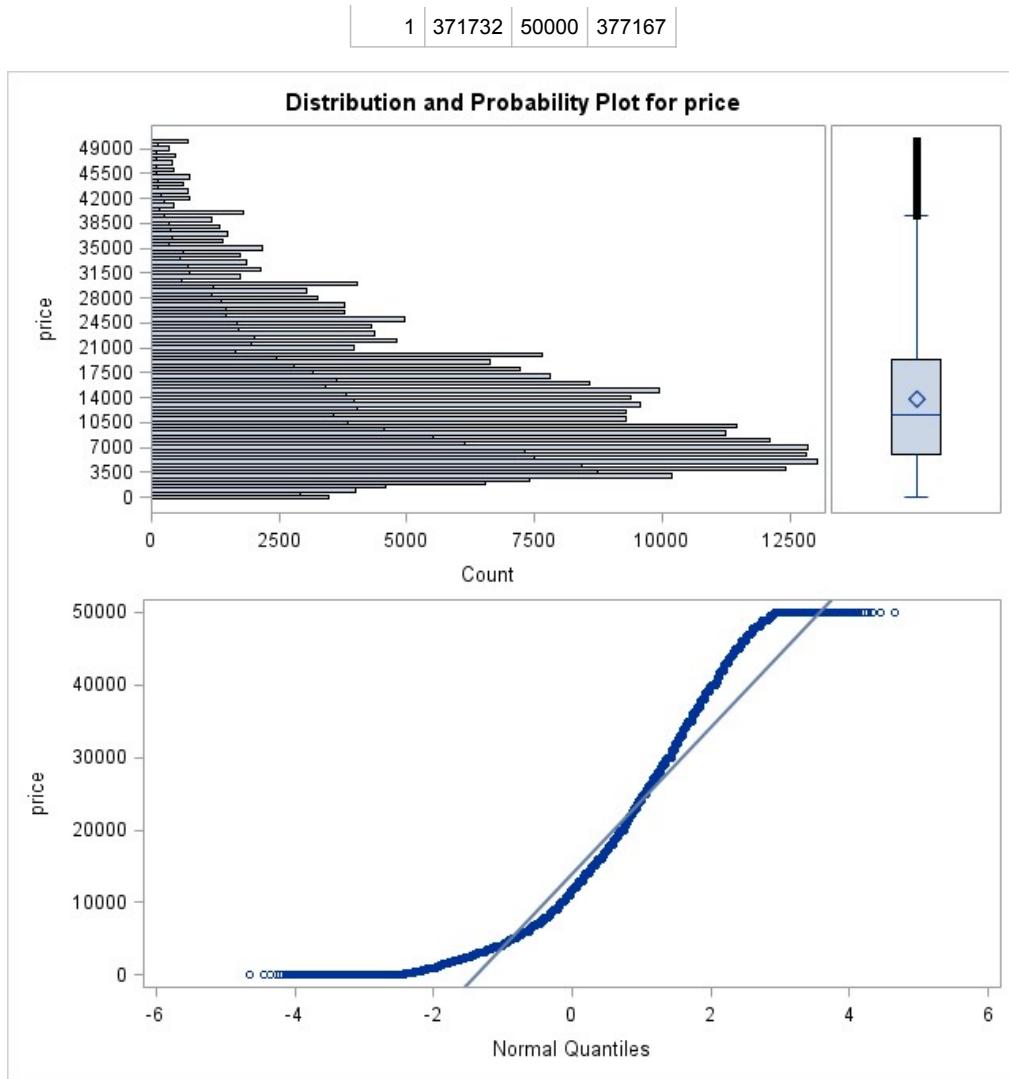
Moments			
N	377862	Sum Weights	377862
Mean	13857.227	Sum Observations	5236119507
Std Deviation	10094.8633	Variance	101906265
Skewness	1.02694404	Kurtosis	0.66536831
Uncorrected SS	1.11064E14	Corrected SS	3.85064E13
Coeff Variation	72.8490866	Std Error Mean	16.4222944

Basic Statistical Measures			
Location		Variability	
Mean	13857.23	Std Deviation	10095
Median	11500.00	Variance	101906265
Mode	6995.00	Range	49999
		Interquartile Range	13409

Tests for Location: Mu0=0				
Test		Statistic	p Value	
Student's t	t	843.8058	Pr > t 	<.0001
Sign	M	188931	Pr >= M 	<.0001
Signed Rank	S	3.57E10	Pr >= S 	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	50000
99%	44000
95%	34000
90%	28495
75% Q3	19399
50% Median	11500
25% Q1	5990
10%	3150
5%	2000
1%	289
0% Min	1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	376164	50000	359875
1	375915	50000	366179
1	375381	50000	368247
1	374977	50000	375179



Count Missing values for all variables**The FREQ Procedure**

region	Frequency
Not Missing	509577

manufacturer	Frequency
Missing	22764
Not Missing	486813

model	Frequency
Missing	8001
Not Missing	501576

condition	Frequency
Missing	231934
Not Missing	277643

cylinders	Frequency
Missing	199683
Not Missing	309894

fuel	Frequency
Missing	3985
Not Missing	505592

title_status	Frequency
Missing	3062
Not Missing	506515

transmission	Frequency
Missing	3719
Not Missing	505858

drive	Frequency
Missing	144143
Not Missing	365434

type	Frequency
Missing	141531
Not Missing	368046

paint_color	Frequency
Missing	164706
Not Missing	344871

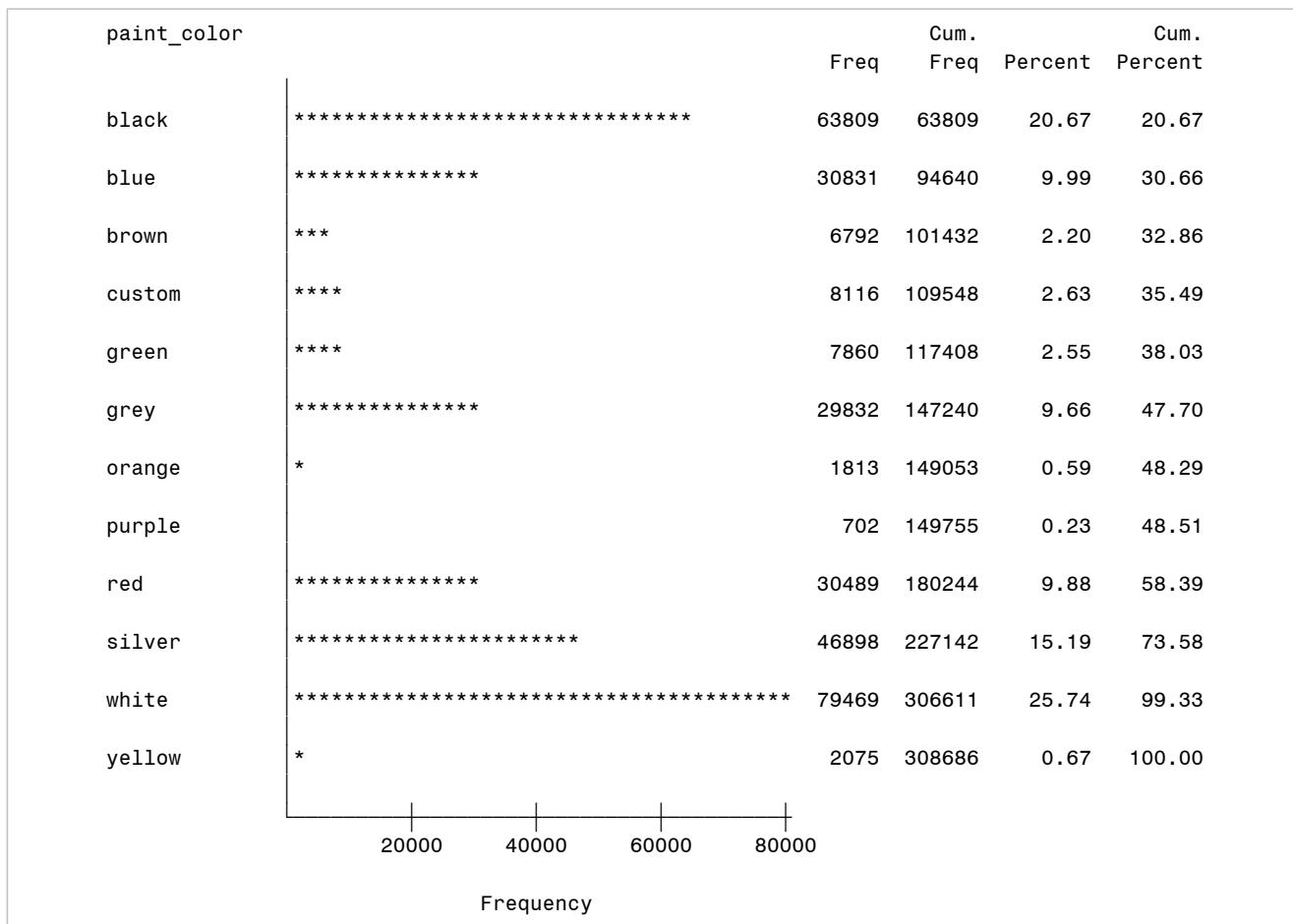
state	Frequency
Not Missing	509577

price	Frequency
Not Missing	509577

year	Frequency
Missing	1527
Not Missing	508050

odometer	Frequency
Missing	92324
Not Missing	417253

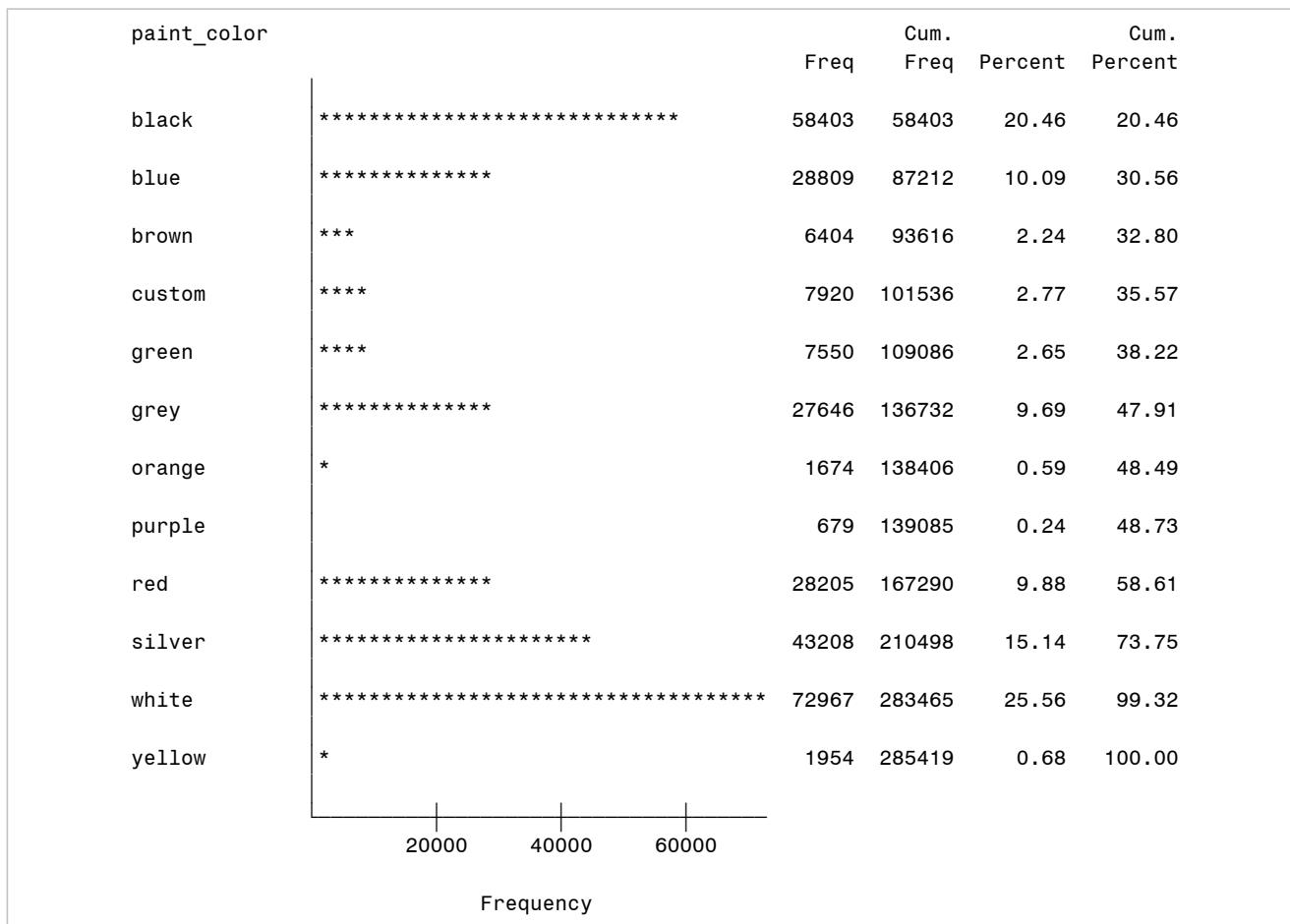
Frequency distribution for paint_color of the car



price stats as per the car color

The MEANS Procedure

Analysis Variable : price						
paint_color	N Obs	N	Mean	Median	Maximum	Minimum
black	63809	63809	13950.82	11995.00	50000.00	0
blue	30831	30831	10983.29	8500.00	50000.00	0
brown	6792	6792	10812.48	7950.00	50000.00	0
custom	8116	8116	12959.08	10952.00	50000.00	0
green	7860	7860	9039.69	5995.00	49999.00	0
grey	29832	29832	11592.92	8997.00	50000.00	0
orange	1813	1813	13534.53	12490.00	49980.00	0
purple	702	702	9890.35	6988.00	47991.00	0
red	30489	30489	12137.34	9688.00	49995.00	0
silver	46898	46898	11195.86	8900.00	50000.00	0
white	79469	79469	14616.28	12995.00	50000.00	0
yellow	2075	2075	12564.81	10000.00	50000.00	0

Frequency distribution for paint_color of the car

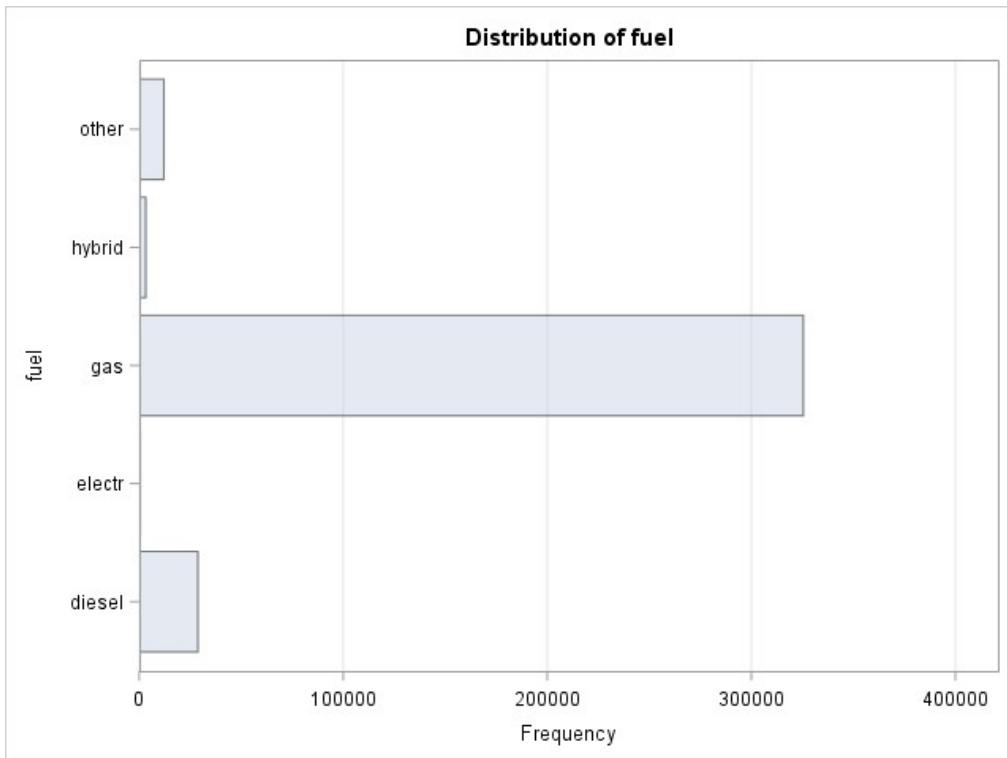
price stats as per the car color**The MEANS Procedure**

Analysis Variable : price						
paint_color	N Obs	N	Mean	Median	Maximum	Minimum
black	58403	58403	15242.16	12999.00	50000.00	1.0000000
blue	28809	28809	11754.17	9000.00	50000.00	1.0000000
brown	6404	6404	11467.57	8500.00	50000.00	1.0000000
custom	7920	7920	13279.78	10995.00	50000.00	1.0000000
green	7550	7550	9410.85	6223.50	49999.00	1.0000000
grey	27646	27646	12509.58	9995.00	50000.00	1.0000000
orange	1674	1674	14658.36	13525.00	49980.00	1.0000000
purple	679	679	10225.37	6995.00	47991.00	295.0000000
red	28205	28205	13120.20	10635.00	49995.00	1.0000000
silver	43208	43208	12151.99	9800.00	50000.00	1.0000000
white	72967	72967	15918.72	13995.00	50000.00	1.0000000
yellow	1954	1954	13342.88	11000.00	50000.00	1.0000000

Frequency Analysis - Fuel

The FREQ Procedure

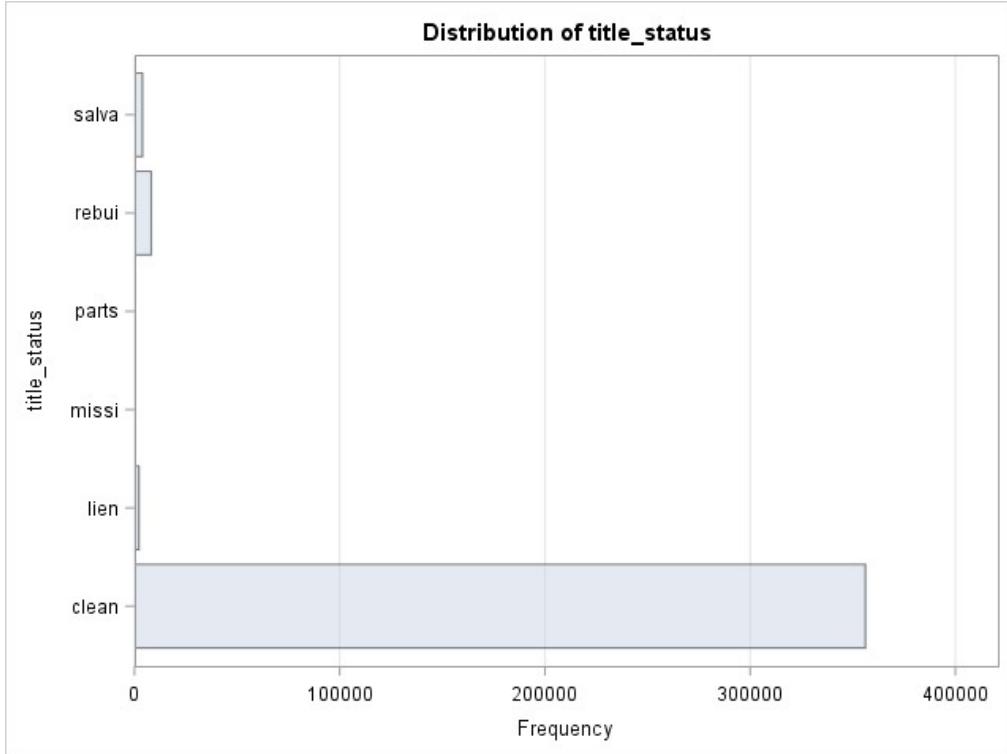
fuel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
diesel	28732	7.77	28732	7.77
electr	476	0.13	29208	7.90
gas	325429	87.98	354637	95.88
hybrid	3174	0.86	357811	96.74
other	12064	3.26	369875	100.00
Frequency Missing = 2969				

Distribution of fuel

Frequency Analysis - title

The FREQ Procedure

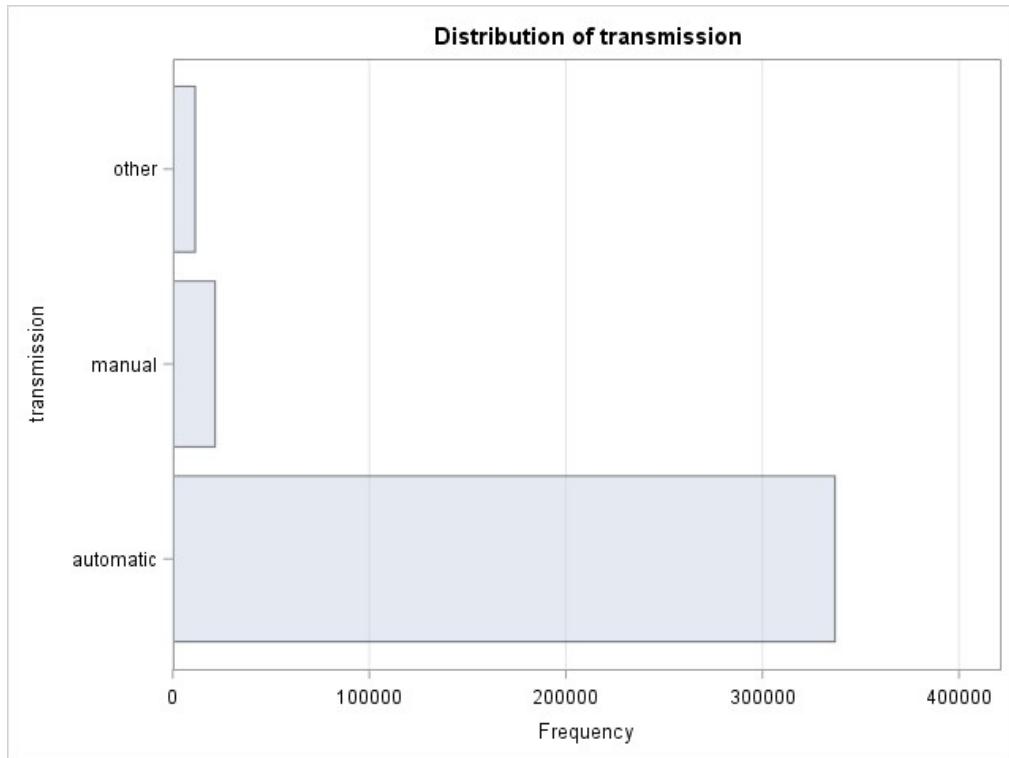
title_status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
clean	356133	96.15	356133	96.15
lien	2147	0.58	358280	96.73
missi	99	0.03	358379	96.76
parts	39	0.01	358418	96.77
rebui	8110	2.19	366528	98.96
salva	3856	1.04	370384	100.00
Frequency Missing = 2460				

Distribution of title_status

Frequency Analysis - Transmission

The FREQ Procedure

transmission	Frequency	Percent	Cumulative Frequency	Cumulative Percent
automatic	336906	91.10	336906	91.10
manual	21528	5.82	358434	96.92
other	11378	3.08	369812	100.00
Frequency Missing = 3032				

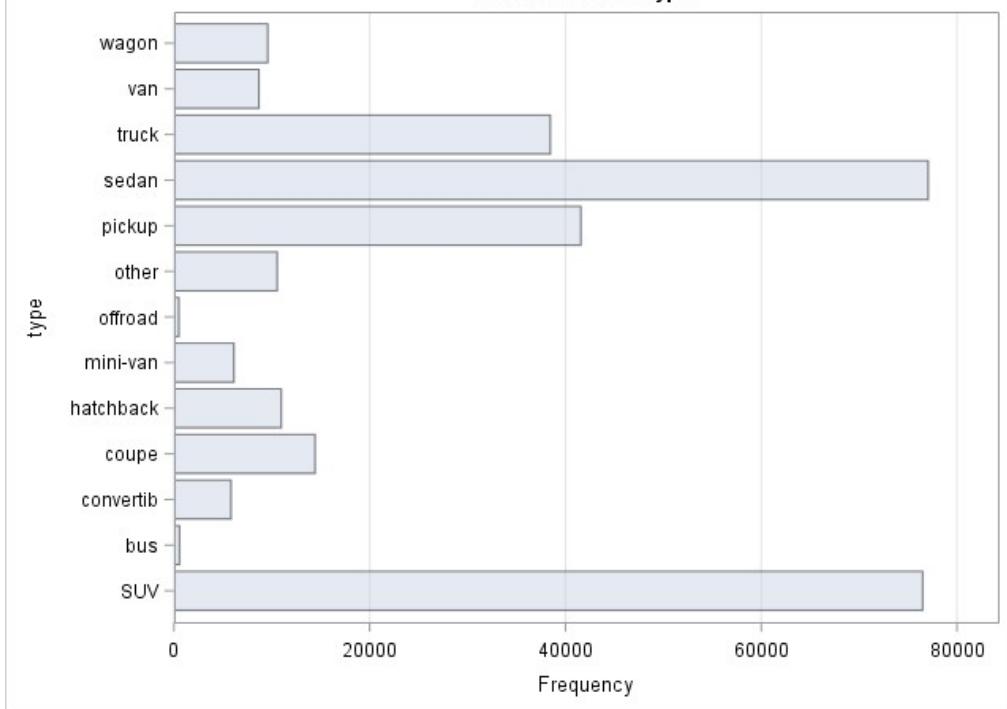


Frequency Analysis - Type

The FREQ Procedure

type	Frequency	Percent	Cumulative Frequency	Cumulative Percent
SUV	76492	25.44	76492	25.44
bus	560	0.19	77052	25.62
convertib	5825	1.94	82877	27.56
coupe	14410	4.79	97287	32.35
hatchback	10959	3.64	108246	35.99
mini-van	6124	2.04	114370	38.03
offroad	495	0.16	114865	38.20
other	10539	3.50	125404	41.70
pickup	41567	13.82	166971	55.52
sedan	77056	25.62	244027	81.15
truck	38445	12.78	282472	93.93
van	8676	2.89	291148	96.82
wagon	9578	3.18	300726	100.00

Frequency Missing = 72118

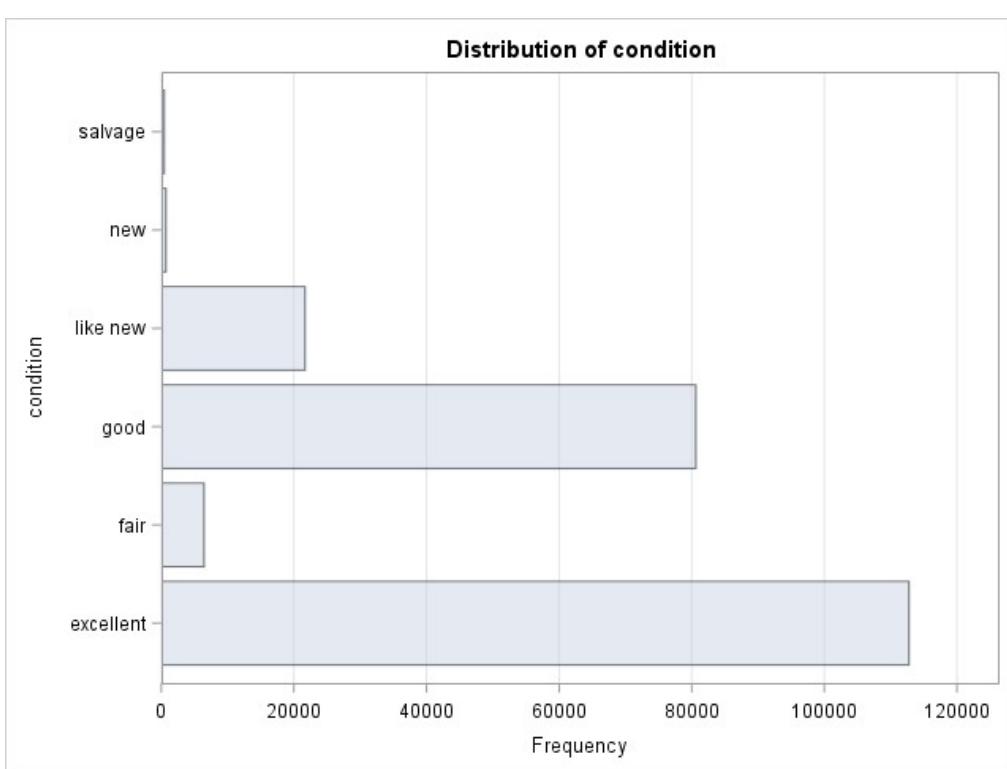
Distribution of type

Frequency Analysis - Condition

The FREQ Procedure

condition	Frequency	Percent	Cumulative Frequency	Cumulative Percent
excellent	112750	50.66	112750	50.66
fair	6437	2.89	119187	53.55
good	80607	36.21	199794	89.76
like new	21668	9.73	221462	99.50
new	686	0.31	222148	99.80
salvage	436	0.20	222584	100.00

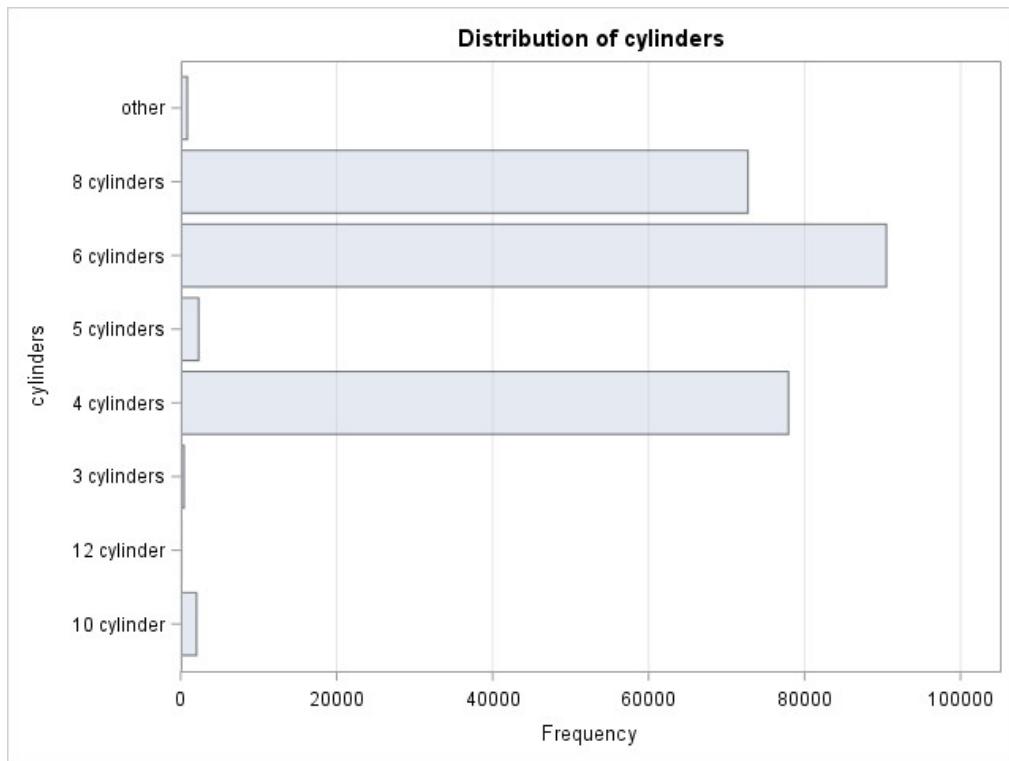
Frequency Missing = 150260

Distribution of condition

Frequency Analysis - Cylinders

The FREQ Procedure

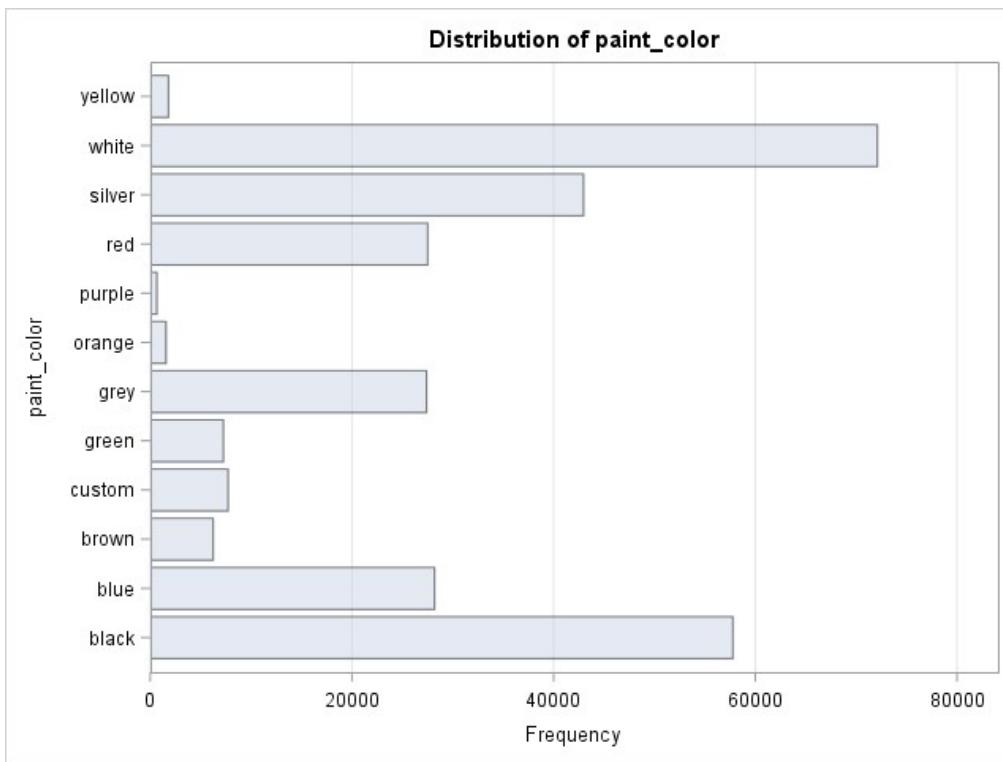
cylinders	Frequency	Percent	Cumulative Frequency	Cumulative Percent
10 cylinder	2025	0.82	2025	0.82
12 cylinder	102	0.04	2127	0.86
3 cylinders	423	0.17	2550	1.03
4 cylinders	77930	31.57	80480	32.60
5 cylinders	2299	0.93	82779	33.54
6 cylinders	90480	36.66	173259	70.19
8 cylinders	72737	29.47	245996	99.66
other	845	0.34	246841	100.00
Frequency Missing = 126003				



Frequency Analysis - Color

The FREQ Procedure

paint_color	Frequency	Percent	Cumulative Frequency	Cumulative Percent
black	57767	20.55	57767	20.55
blue	28188	10.03	85955	30.58
brown	6234	2.22	92189	32.80
custom	7711	2.74	99900	35.54
green	7247	2.58	107147	38.12
grey	27393	9.74	134540	47.86
orange	1557	0.55	136097	48.42
purple	647	0.23	136744	48.65
red	27507	9.79	164251	58.43
silver	42960	15.28	207211	73.71
white	72098	25.65	279309	99.36
yellow	1794	0.64	281103	100.00
Frequency Missing = 91741				

Distribution of paint_color

Frequency Analysis - Region

The FREQ Procedure

region	Frequency	Percent	Cumulative Frequency	Cumulative Percent
SF bay area	2445	0.66	2445	0.66
abilene	349	0.09	2794	0.75
akron / canton	2480	0.67	5274	1.41
albany	2044	0.55	7318	1.96
albuquerque	1853	0.50	9171	2.46
altoona-johnst	161	0.04	9332	2.50
amarillo	867	0.23	10199	2.74
ames	198	0.05	10397	2.79
anchorage / ma	2417	0.65	12814	3.44
ann arbor	792	0.21	13606	3.65
annapolis	268	0.07	13874	3.72
appleton-oshko	1902	0.51	15776	4.23
asheville	1572	0.42	17348	4.65
ashtabula	167	0.04	17515	4.70
athens	511	0.14	18026	4.83
atlanta	2061	0.55	20087	5.39
auburn	89	0.02	20176	5.41
augusta	421	0.11	20597	5.52
austin	2034	0.55	22631	6.07
bakersfield	2231	0.60	24862	6.67
baltimore	1974	0.53	26836	7.20
baton rouge	652	0.17	27488	7.37
battle creek	487	0.13	27975	7.50
beaumont / por	56	0.02	28031	7.52
bellingham	2714	0.73	30745	8.25
bemidji	473	0.13	31218	8.37
bend	2173	0.58	33391	8.96
billings	1602	0.43	34993	9.39
binghamton	585	0.16	35578	9.54
birmingham	2363	0.63	37941	10.18
bismarck	143	0.04	38084	10.21
bloomington	137	0.04	38221	10.25
bloomington-no	308	0.08	38529	10.33
boise	2531	0.68	41060	11.01
boone	148	0.04	41208	11.05
boston	2426	0.65	43634	11.70
boulder	908	0.24	44542	11.95
bowling green	297	0.08	44839	12.03
bozeman	1302	0.35	46141	12.38

brainerd	588	0.16	46729	12.53
brownsville	166	0.04	46895	12.58
brunswick	89	0.02	46984	12.60
buffalo	1821	0.49	48805	13.09
butte	251	0.07	49056	13.16
cape cod / isl	562	0.15	49618	13.31
catskills	64	0.02	49682	13.33
cedar rapids	970	0.26	50652	13.59
central NJ	2128	0.57	52780	14.16
central louisi	60	0.02	52840	14.17
central michig	317	0.09	53157	14.26
champaign urba	215	0.06	53372	14.31
charleston	1897	0.51	55269	14.82
charlotte	1979	0.53	57248	15.35
charlottesville	676	0.18	57924	15.54
chattanooga	1204	0.32	59128	15.86
chautauqua	109	0.03	59237	15.89
chicago	2098	0.56	61335	16.45
chico	1372	0.37	62707	16.82
chillicothe	133	0.04	62840	16.85
cincinnati	2105	0.56	64945	17.42
clarksville	482	0.13	65427	17.55
cleveland	2490	0.67	67917	18.22
clovis / porta	42	0.01	67959	18.23
college statio	241	0.06	68200	18.29
colorado sprin	2524	0.68	70724	18.97
columbia	1255	0.34	71979	19.31
columbia / jef	1326	0.36	73305	19.66
columbus	2727	0.73	76032	20.39
cookeville	87	0.02	76119	20.42
corpus christi	2057	0.55	78176	20.97
corvallis/alba	1272	0.34	79448	21.31
cumberland val	87	0.02	79535	21.33
dallas / fort	2116	0.57	81651	21.90
danville	190	0.05	81841	21.95
dayton / sprin	1599	0.43	83440	22.38
daytona beach	1945	0.52	85385	22.90
decatur	99	0.03	85484	22.93
deep east texa	69	0.02	85553	22.95
del rio / eagl	47	0.01	85600	22.96
delaware	1034	0.28	86634	23.24
denver	2441	0.65	89075	23.89
des moines	2305	0.62	91380	24.51
detroit metro	2141	0.57	93521	25.08
dothan				

	229	0.06	93750	25.14
dubuque	334	0.09	94084	25.23
duluth / super	884	0.24	94968	25.47
east idaho	778	0.21	95746	25.68
east oregon	306	0.08	96052	25.76
eastern CO	114	0.03	96166	25.79
eastern CT	560	0.15	96726	25.94
eastern NC	835	0.22	97561	26.17
eastern kentuc	123	0.03	97684	26.20
eastern montan	66	0.02	97750	26.22
eastern panhan	77	0.02	97827	26.24
eastern shore	414	0.11	98241	26.35
eau claire	935	0.25	99176	26.60
el paso	1898	0.51	101074	27.11
elko	56	0.02	101130	27.12
elmira-corning	211	0.06	101341	27.18
erie	757	0.20	102098	27.38
eugene	2238	0.60	104336	27.98
evansville	319	0.09	104655	28.07
fairbanks	365	0.10	105020	28.17
fargo / moorhe	1391	0.37	106411	28.54
farmington	82	0.02	106493	28.56
fayetteville	3160	0.85	109653	29.41
finger lakes	211	0.06	109864	29.47
flagstaff / se	603	0.16	110467	29.63
flint	1493	0.40	111960	30.03
florence	244	0.07	112204	30.09
florence / mus	260	0.07	112464	30.16
florida keys	86	0.02	112550	30.19
fort collins /	2565	0.69	115115	30.87
fort dodge	67	0.02	115182	30.89
fort smith	482	0.13	115664	31.02
fort wayne	786	0.21	116450	31.23
frederick	274	0.07	116724	31.31
fredericksburg	1252	0.34	117976	31.64
fresno / mader	1904	0.51	119880	32.15
ft myers / SW	2338	0.63	122218	32.78
gadsden-annist	145	0.04	122363	32.82
gainesville	739	0.20	123102	33.02
galveston	68	0.02	123170	33.04
glens falls	214	0.06	123384	33.09
gold country	545	0.15	123929	33.24
grand forks	293	0.08	124222	33.32
grand island	392	0.11	124614	33.42
grand rapids				

	2595	0.70	127209	34.12
great falls	538	0.14	127747	34.26
green bay	1521	0.41	129268	34.67
greensboro	1882	0.50	131150	35.18
greenville / u	1993	0.53	133143	35.71
gulfport / bil	510	0.14	133653	35.85
hanford-corcor	61	0.02	133714	35.86
harrisburg	1279	0.34	134993	36.21
harrisonburg	483	0.13	135476	36.34
hartford	2468	0.66	137944	37.00
hattiesburg	248	0.07	138192	37.06
hawaii	2122	0.57	140314	37.63
heartland flor	100	0.03	140414	37.66
helena	293	0.08	140707	37.74
hickory / leno	389	0.10	141096	37.84
high rockies	213	0.06	141309	37.90
hilton head	89	0.02	141398	37.92
holland	434	0.12	141832	38.04
houma	65	0.02	141897	38.06
houston	1998	0.54	143895	38.59
hudson valley	1951	0.52	145846	39.12
humboldt count	313	0.08	146159	39.20
huntington-ash	194	0.05	146353	39.25
huntsville / d	1487	0.40	147840	39.65
imperial count	182	0.05	148022	39.70
indianapolis	2212	0.59	150234	40.29
inland empire	1674	0.45	151908	40.74
iowa city	493	0.13	152401	40.88
ithaca	210	0.06	152611	40.93
jackson	1219	0.33	153830	41.26
jacksonville	2419	0.65	156249	41.91
janesville	148	0.04	156397	41.95
jersey shore	1090	0.29	157487	42.24
jonesboro	353	0.09	157840	42.33
joplin	454	0.12	158294	42.46
kalamazoo	1264	0.34	159558	42.79
kalispell	1133	0.30	160691	43.10
kansas city	3	0.00	160694	43.10
kansas city, M	2234	0.60	162928	43.70
kenai pensinsul	266	0.07	163194	43.77
kennewick-pasc	2259	0.61	165453	44.38
kenosha-racine	889	0.24	166342	44.61
killeen / temp	725	0.19	167067	44.81
kirksville	50	0.01	167117	44.82
klamath falls				

	212	0.06	167329	44.88
knoxville	2113	0.57	169442	45.45
kokomo	87	0.02	169529	45.47
la crosse	521	0.14	170050	45.61
la salle co	49	0.01	170099	45.62
lafayette	153	0.04	170252	45.66
lafayette / we	226	0.06	170478	45.72
lake charles	39	0.01	170517	45.73
lake of the oz	378	0.10	170895	45.84
lakeland	1290	0.35	172185	46.18
lancaster	1265	0.34	173450	46.52
lansing	1624	0.44	175074	46.96
laredo	116	0.03	175190	46.99
las cruces	187	0.05	175377	47.04
las vegas	2128	0.57	177505	47.61
lawrence	86	0.02	177591	47.63
lawton	91	0.02	177682	47.66
lehigh valley	1179	0.32	178861	47.97
lewiston / cla	1119	0.30	179980	48.27
lexington	1468	0.39	181448	48.67
lima / findlay	328	0.09	181776	48.75
lincoln	1407	0.38	183183	49.13
little rock	1383	0.37	184566	49.50
logan	48	0.01	184614	49.52
long island	1920	0.51	186534	50.03
los angeles	2032	0.55	188566	50.58
louisville	2224	0.60	190790	51.17
lubbock	928	0.25	191718	51.42
lynchburg	457	0.12	192175	51.54
macon / warner	441	0.12	192616	51.66
madison	2408	0.65	195024	52.31
maine	2174	0.58	197198	52.89
manhattan	106	0.03	197304	52.92
mankato	311	0.08	197615	53.00
mansfield	341	0.09	197956	53.09
mason city	168	0.05	198124	53.14
mattoon-charle	66	0.02	198190	53.16
mcallen / edin	1026	0.28	199216	53.43
meadville	20	0.01	199236	53.44
medford-ashlan	2128	0.57	201364	54.01
memphis	2411	0.65	203775	54.65
mendocino coun	115	0.03	203890	54.69
merced	237	0.06	204127	54.75
meridian	31	0.01	204158	54.76
milwaukee				

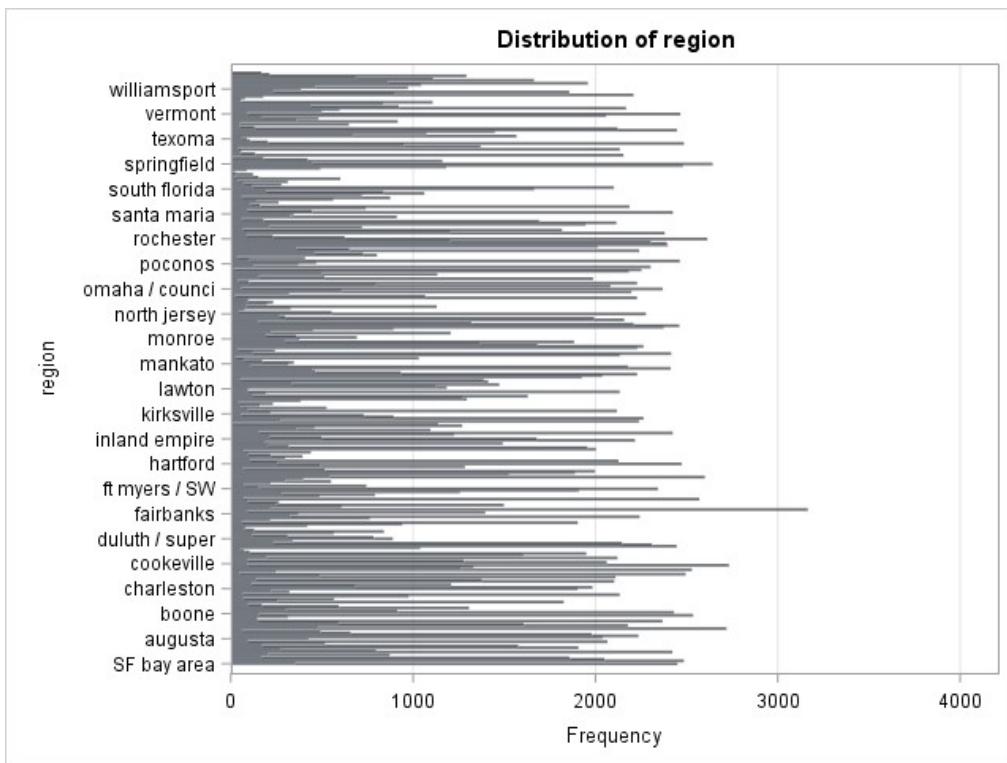
	2224	0.60	206382	55.35
minneapolis /	2259	0.61	208641	55.96
missoula	1677	0.45	210318	56.41
mobile	1358	0.36	211676	56.77
modesto	1878	0.50	213554	57.28
mohave county	290	0.08	213844	57.35
monroe	371	0.10	214215	57.45
monterey bay	687	0.18	214902	57.64
montgomery	355	0.10	215257	57.73
morgantown	186	0.05	215443	57.78
moses lake	1201	0.32	216644	58.11
muncie / ander	217	0.06	216861	58.16
muskegon	446	0.12	217307	58.28
myrtle beach	887	0.24	218194	58.52
nashville	2369	0.64	220563	59.16
new hampshire	2455	0.66	223018	59.82
new haven	2205	0.59	225223	60.41
new orleans	1315	0.35	226538	60.76
new river vall	143	0.04	226681	60.80
new york city	2153	0.58	228834	61.38
norfolk / hamp	1988	0.53	230822	61.91
north central	294	0.08	231116	61.99
north dakota	262	0.07	231378	62.06
north jersey	2272	0.61	233650	62.67
north mississi	549	0.15	234199	62.81
north platte	38	0.01	234237	62.82
northeast SD	70	0.02	234307	62.84
northern WI	323	0.09	234630	62.93
northern michi	1125	0.30	235755	63.23
northern panha	78	0.02	235833	63.25
northwest CT	194	0.05	236027	63.30
northwest GA	228	0.06	236255	63.37
northwest KS	85	0.02	236340	63.39
northwest OK	94	0.03	236434	63.41
ocala	2223	0.60	238657	64.01
odessa / midla	1062	0.28	239719	64.29
ogden-clearfie	18	0.00	239737	64.30
okaloosa / wal	315	0.08	240052	64.38
oklahoma city	2192	0.59	242244	64.97
olympic penins	602	0.16	242846	65.13
omaha / counci	2364	0.63	245210	65.77
oneonta	53	0.01	245263	65.78
orange county	2078	0.56	247341	66.34
oregon coast	791	0.21	248132	66.55
orlando				

	2224	0.60	250356	67.15
outer banks	94	0.03	250450	67.17
owensboro	32	0.01	250482	67.18
palm springs	1982	0.53	252464	67.71
panama city	510	0.14	252974	67.85
parkersburg-ma	150	0.04	253124	67.89
pensacola	1128	0.30	254252	68.19
peoria	494	0.13	254746	68.33
philadelphia	2178	0.58	256924	68.91
phoenix	2247	0.60	259171	69.51
pierre / centr	13	0.00	259184	69.52
pittsburgh	2298	0.62	261482	70.13
plattsburgh-ad	118	0.03	261600	70.16
poconos	366	0.10	261966	70.26
port huron	464	0.12	262430	70.39
portland	2458	0.66	264888	71.05
potsdam-canton	97	0.03	264985	71.07
prescott	403	0.11	265388	71.18
provo / orem	30	0.01	265418	71.19
pueblo	797	0.21	266215	71.40
pullman / mosc	721	0.19	266936	71.59
quad cities, i	455	0.12	267391	71.72
raleigh / durh	2235	0.60	269626	72.32
rapid city / w	645	0.17	270271	72.49
reading	354	0.09	270625	72.58
redding	2007	0.54	272632	73.12
reno / tahoe	2392	0.64	275024	73.76
rhode island	2386	0.64	277410	74.40
richmond	2298	0.62	279708	75.02
roanoke	1198	0.32	280906	75.34
rochester	2609	0.70	283515	76.04
rockford	620	0.17	284135	76.21
roseburg	226	0.06	284361	76.27
roswell / carl	81	0.02	284442	76.29
sacramento	2375	0.64	286817	76.93
saginaw-midlan	1198	0.32	288015	77.25
salem	1812	0.49	289827	77.73
salina	60	0.02	289887	77.75
salt lake city	715	0.19	290602	77.94
san angelo	207	0.06	290809	78.00
san antonio	1940	0.52	292749	78.52
san diego	2110	0.57	294859	79.08
san luis obisp	1687	0.45	296546	79.54
san marcos	174	0.05	296720	79.58
sandusky				

	53	0.01	296773	79.60
santa barbara	906	0.24	297679	79.84
santa fe / tao	318	0.09	297997	79.93
santa maria	340	0.09	298337	80.02
sarasota-brade	2420	0.65	300757	80.67
savannah / hin	441	0.12	301198	80.78
scottsbluff /	84	0.02	301282	80.81
scranton / wil	735	0.20	302017	81.00
seattle-tacoma	2182	0.59	304199	81.59
sheboygan	155	0.04	304354	81.63
show low	100	0.03	304454	81.66
shreveport	257	0.07	304711	81.73
sierra vista	135	0.04	304846	81.76
sioux city	556	0.15	305402	81.91
sioux falls /	870	0.23	306272	82.14
siskiyou count	55	0.01	306327	82.16
skagit / islan	714	0.19	307041	82.35
south bend / m	1057	0.28	308098	82.63
south coast	833	0.22	308931	82.86
south dakota	188	0.05	309119	82.91
south florida	1659	0.44	310778	83.35
south jersey	2096	0.56	312874	83.92
southeast IA	112	0.03	312986	83.95
southeast KS	273	0.07	313259	84.02
southeast alas	62	0.02	313321	84.04
southeast miss	308	0.08	313629	84.12
southern WV	32	0.01	313661	84.13
southern illin	596	0.16	314257	84.29
southern maryl	147	0.04	314404	84.33
southwest KS	16	0.00	314420	84.33
southwest MN	117	0.03	314537	84.36
southwest MS	21	0.01	314558	84.37
southwest TX	4	0.00	314562	84.37
southwest VA	83	0.02	314645	84.39
southwest mich	486	0.13	315131	84.52
space coast	1176	0.32	316307	84.84
spokane / coeu	2474	0.66	318781	85.50
springfield	2638	0.71	321419	86.21
st augustine	437	0.12	321856	86.32
st cloud	1156	0.31	323012	86.63
st george	417	0.11	323429	86.75
st joseph	174	0.05	323603	86.79
st louis	3	0.00	323606	86.79
st louis, MO	2149	0.58	325755	87.37
state college				

	129	0.03	325884	87.40
statesboro	46	0.01	325930	87.42
stillwater	52	0.01	325982	87.43
stockton	2129	0.57	328111	88.00
susanville	36	0.01	328147	88.01
syracuse	1365	0.37	329512	88.38
tallahassee	945	0.25	330457	88.63
tampa bay area	2480	0.67	332937	89.30
terre haute	197	0.05	333134	89.35
texarkana	100	0.03	333234	89.38
texoma	84	0.02	333318	89.40
the thumb	56	0.02	333374	89.41
toledo	1562	0.42	334936	89.83
topeka	664	0.18	335600	90.01
treasure coast	1068	0.29	336668	90.30
tri-cities	1443	0.39	338111	90.68
tucson	2442	0.65	340553	91.34
tulsa	2117	0.57	342670	91.91
tuscaloosa	125	0.03	342795	91.94
tuscarawas co	44	0.01	342839	91.95
twin falls	643	0.17	343482	92.12
twin tiers NY/	45	0.01	343527	92.14
tyler / east T	911	0.24	344438	92.38
upper pensul	354	0.09	344792	92.48
utica-rome-one	476	0.13	345268	92.60
valdosta	165	0.04	345433	92.65
ventura county	2053	0.55	347486	93.20
vermont	2461	0.66	349947	93.86
victoria	86	0.02	350033	93.88
visalia-tulare	494	0.13	350527	94.01
waco	591	0.16	351118	94.17
washington, DC	2163	0.58	353281	94.75
waterloo / ced	917	0.25	354198	95.00
watertown	431	0.12	354629	95.11
wausau	830	0.22	355459	95.34
wenatchee	1102	0.30	356561	95.63
west virginia	49	0.01	356610	95.65
western IL	52	0.01	356662	95.66
western KY	73	0.02	356735	95.68
western maryla	173	0.05	356908	95.73
western massac	2204	0.59	359112	96.32
western slope	894	0.24	360006	96.56
wichita	1852	0.50	361858	97.05
wichita falls	229	0.06	362087	97.11
williamsport				

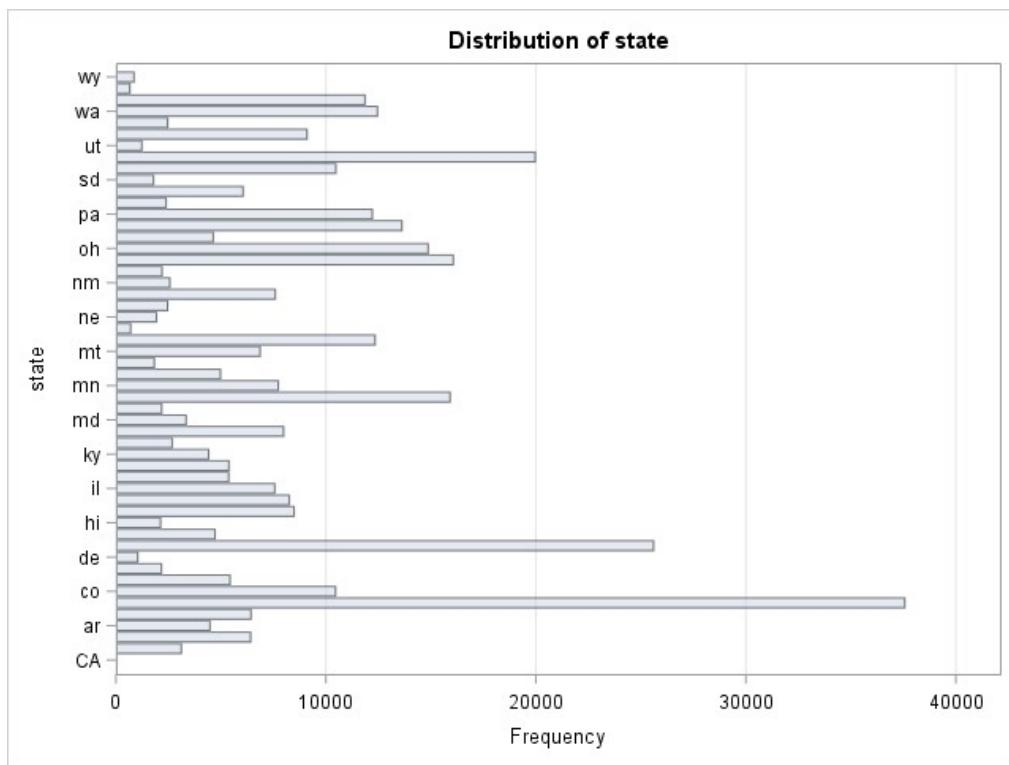
	380	0.10	362467	97.22
wilmington	967	0.26	363434	97.48
winchester	458	0.12	363892	97.60
winston-salem	1039	0.28	364931	97.88
worcester / ce	1954	0.52	366885	98.40
wyoming	856	0.23	367741	98.63
yakima	1659	0.44	369400	99.08
york	1105	0.30	370505	99.37
youngstown	680	0.18	371185	99.56
yuba-sutter	1288	0.35	372473	99.90
yuma	208	0.06	372681	99.96
zanesville / c	163	0.04	372844	100.00



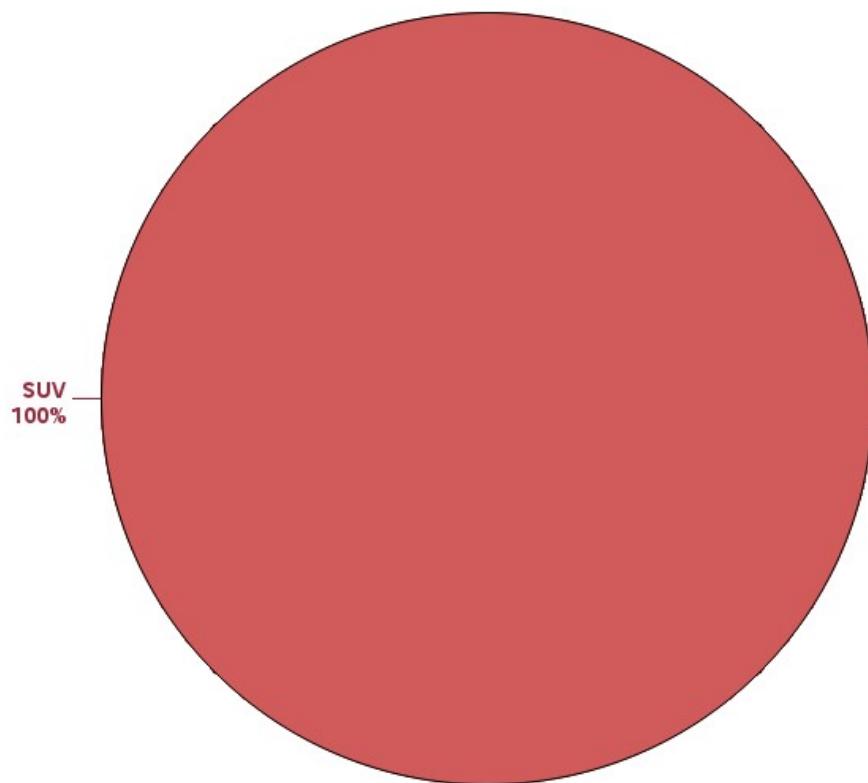
Frequency Analysis - State**The FREQ Procedure**

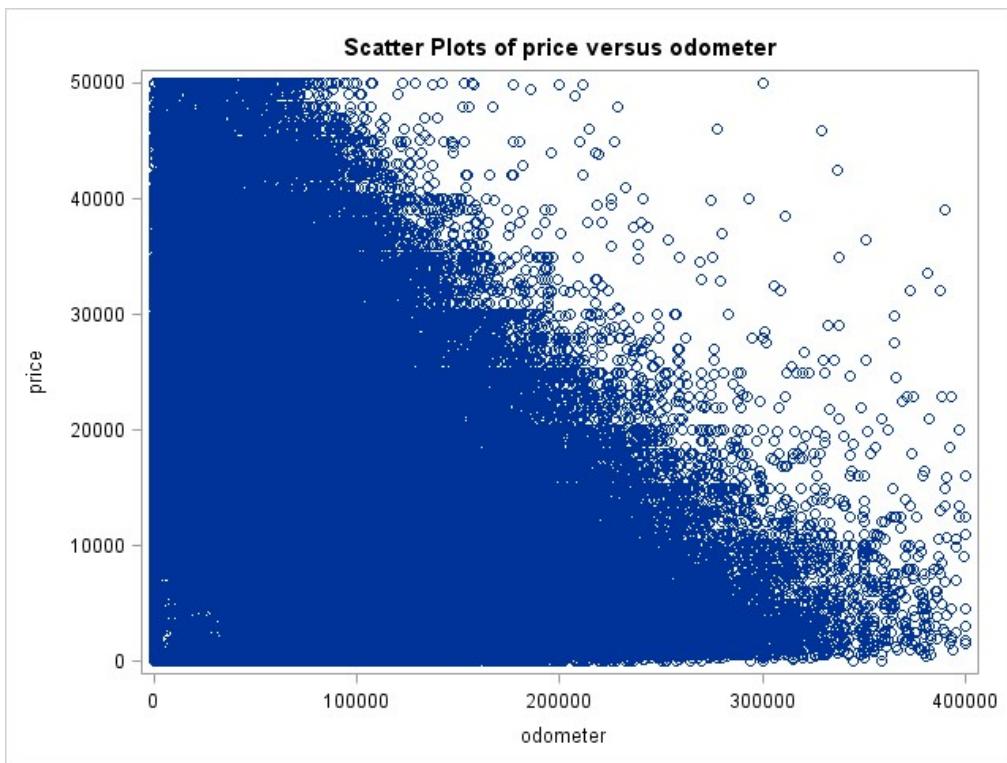
state	Frequency	Percent	Cumulative Frequency	Cumulative Percent
CA	36	0.01	36	0.01
ak	3110	0.83	3146	0.84
al	6411	1.72	9557	2.56
ar	4475	1.20	14032	3.76
az	6428	1.72	20460	5.49
ca	37565	10.08	58025	15.56
co	10456	2.80	68481	18.37
ct	5427	1.46	73908	19.82
dc	2163	0.58	76071	20.40
de	1034	0.28	77105	20.68
fl	25600	6.87	102705	27.55
ga	4715	1.26	107420	28.81
hi	2122	0.57	109542	29.38
ia	8481	2.27	118023	31.65
id	8247	2.21	126270	33.87
il	7561	2.03	133831	35.89
in	5368	1.44	139199	37.33
ks	5376	1.44	144575	38.78
ky	4411	1.18	148986	39.96
la	2674	0.72	151660	40.68
ma	7979	2.14	159639	42.82
md	3337	0.90	162976	43.71
me	2174	0.58	165150	44.29
mi	15909	4.27	181059	48.56
mn	7726	2.07	188785	50.63
mo	4974	1.33	193759	51.97
ms	1824	0.49	195583	52.46
mt	6862	1.84	202445	54.30
nc	12339	3.31	214784	57.61
nd	698	0.19	215482	57.79
ne	1924	0.52	217406	58.31
nh	2455	0.66	219861	58.97
nj	7586	2.03	227447	61.00
nm	2563	0.69	230010	61.69
nv	2191	0.59	232201	62.28
ny	16063	4.31	248264	66.59
oh	14864	3.99	263128	70.57
ok	4630	1.24	267758	71.82
or	13616	3.65	281374	75.47

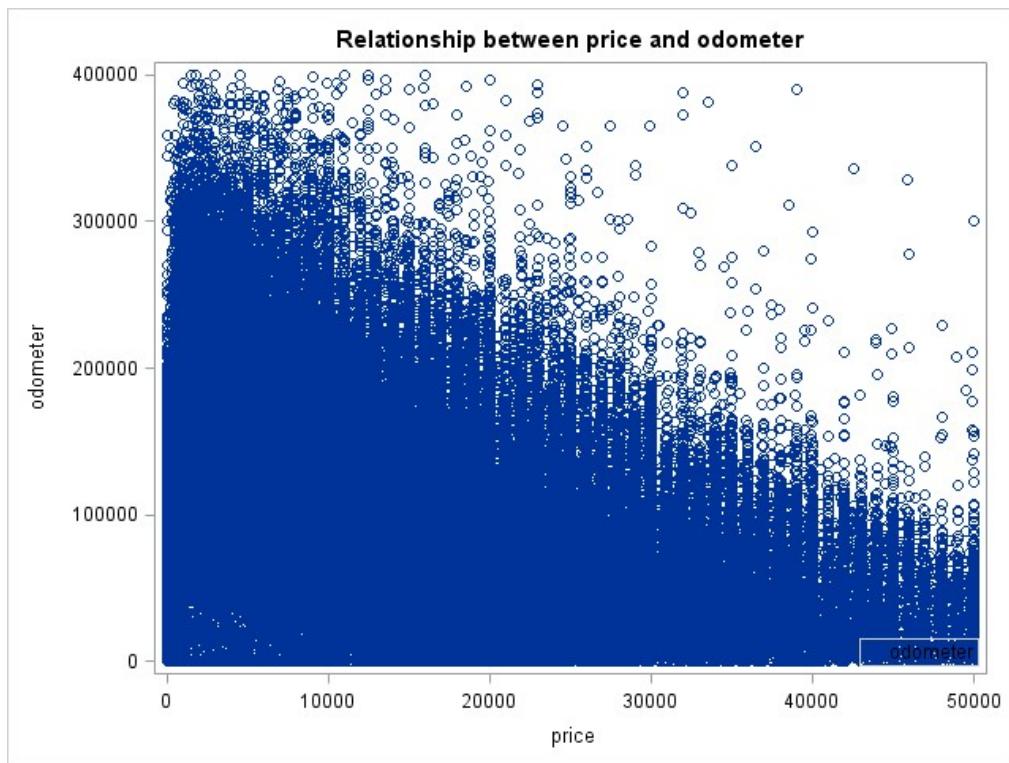
pa	12206	3.27	293580	78.74
ri	2386	0.64	295966	79.38
sc	6051	1.62	302017	81.00
sd	1786	0.48	303803	81.48
tn	10468	2.81	314271	84.29
tx	19965	5.35	334236	89.64
ut	1228	0.33	335464	89.97
va	9096	2.44	344560	92.41
vt	2461	0.66	347021	93.07
wa	12452	3.34	359473	96.41
wi	11857	3.18	371330	99.59
wv	658	0.18	371988	99.77
wy	856	0.23	372844	100.00

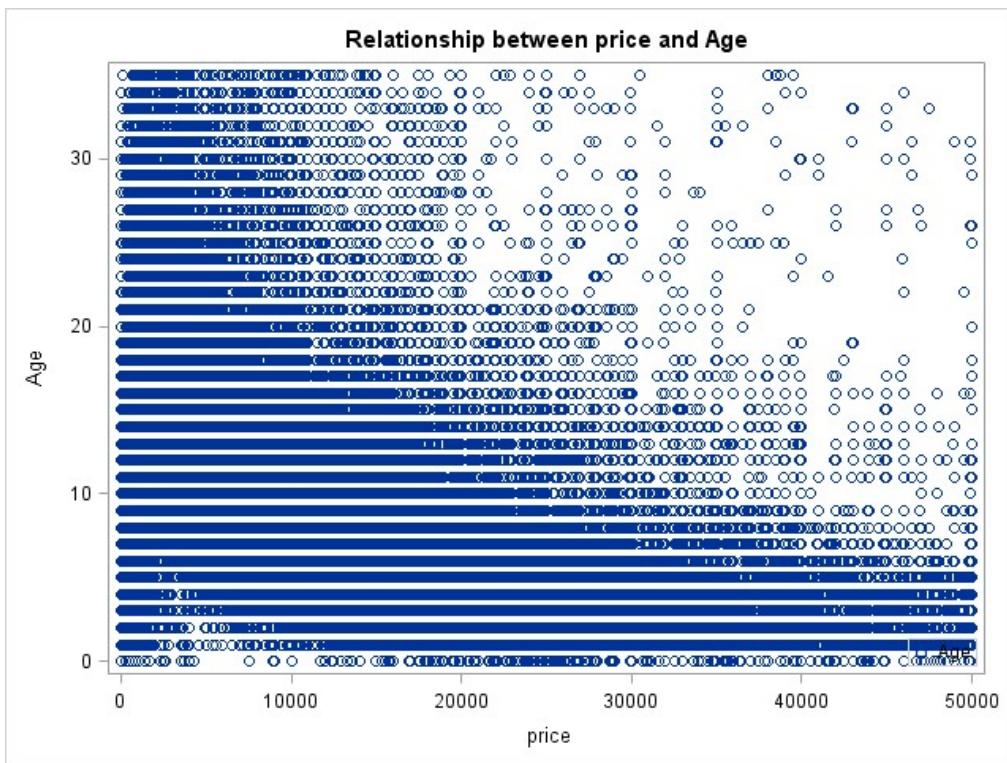


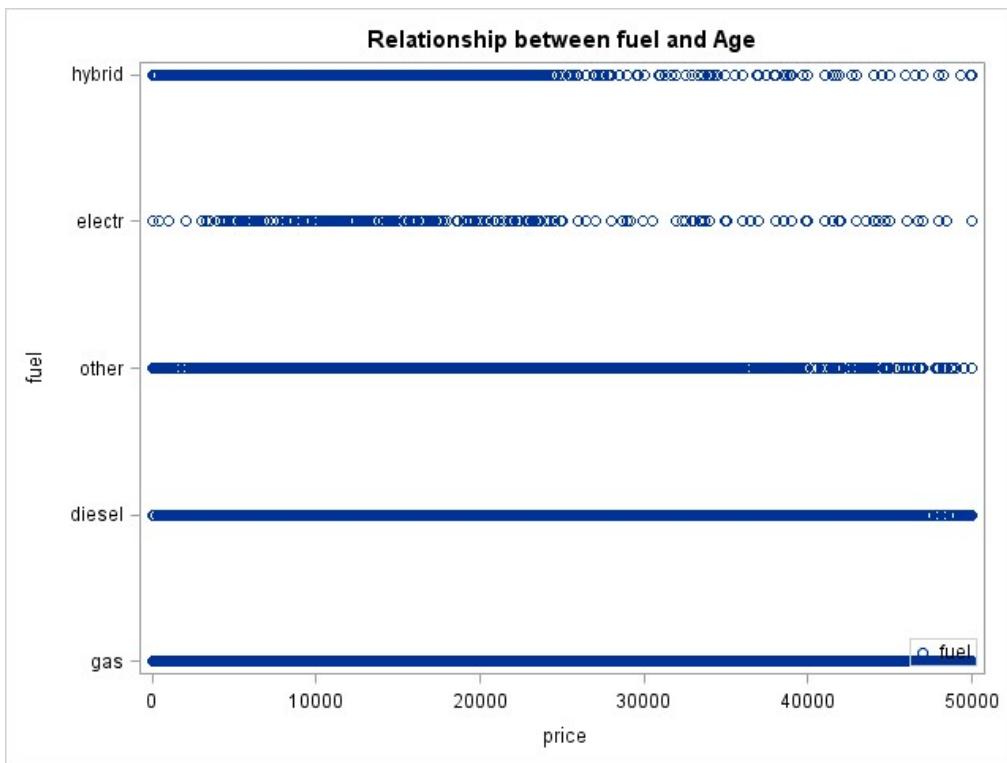
Frequency Analysis - State

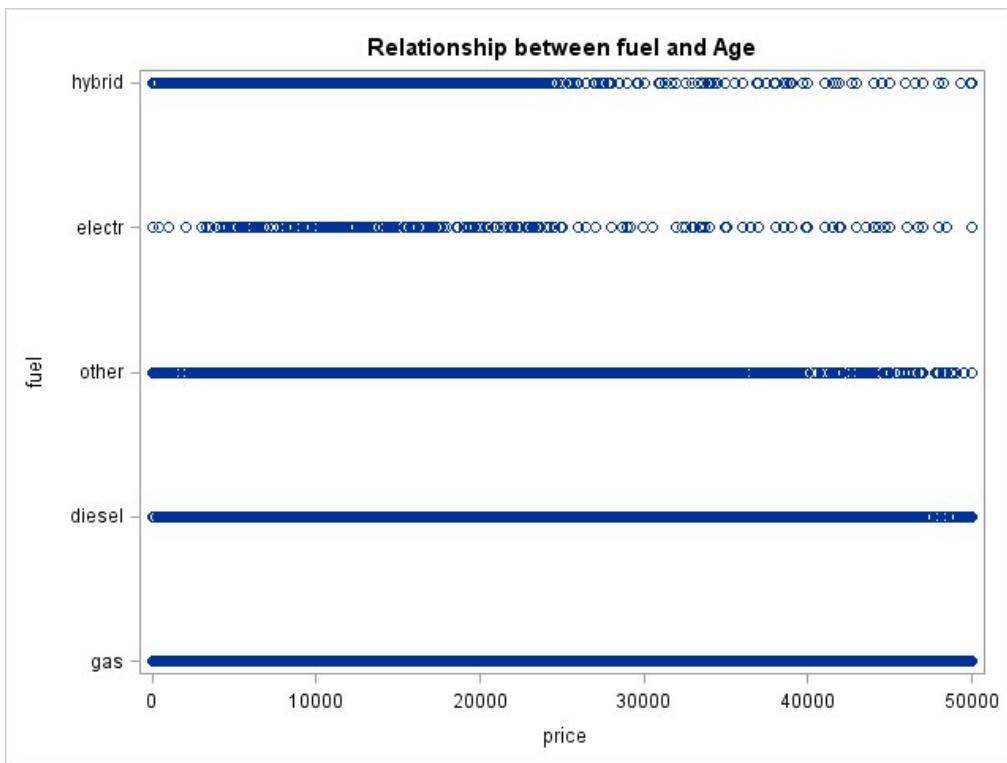


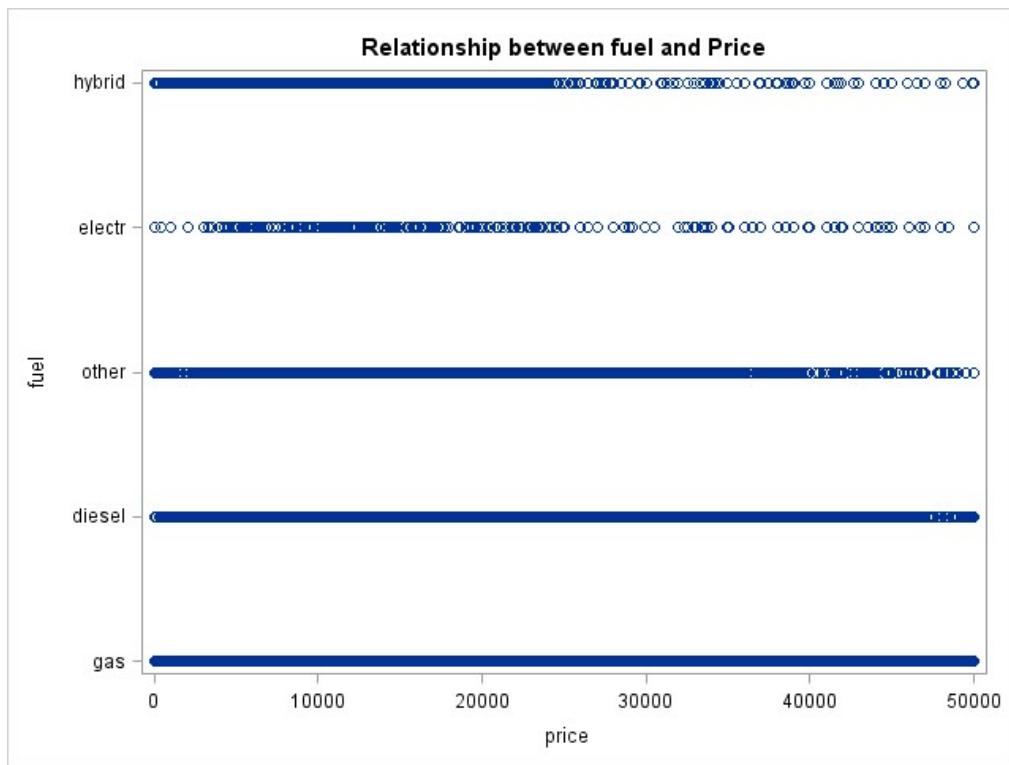


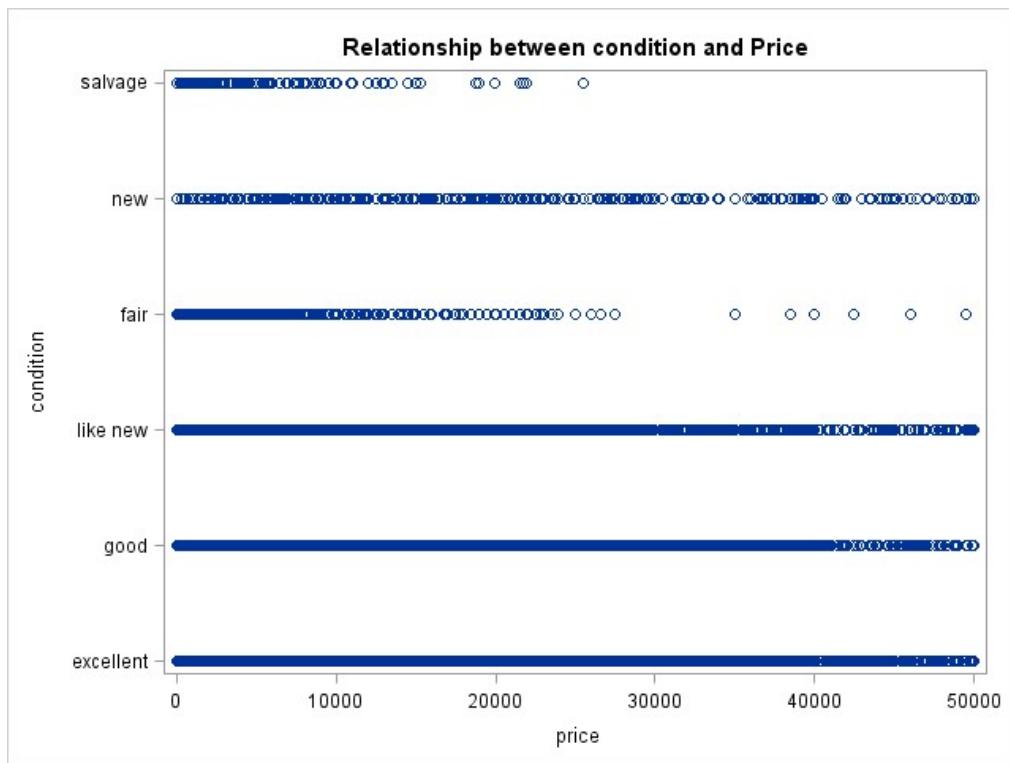


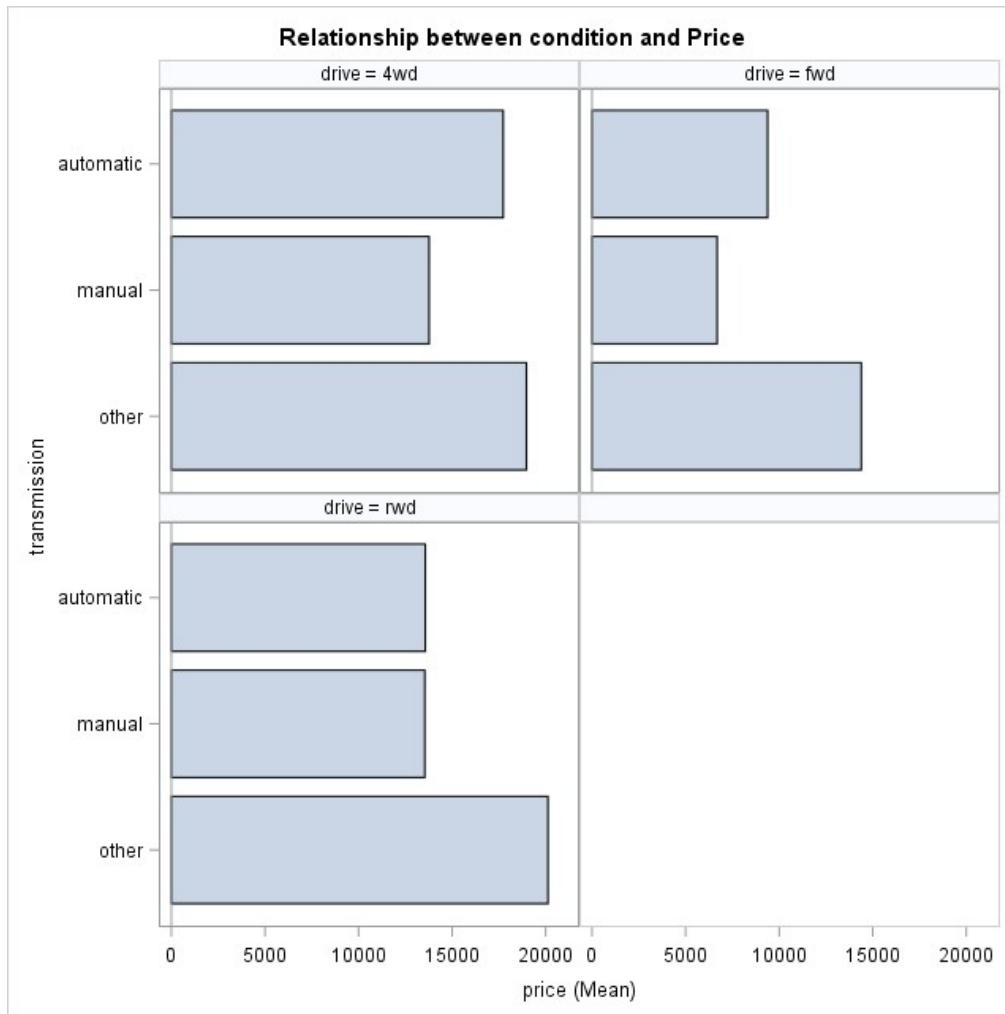


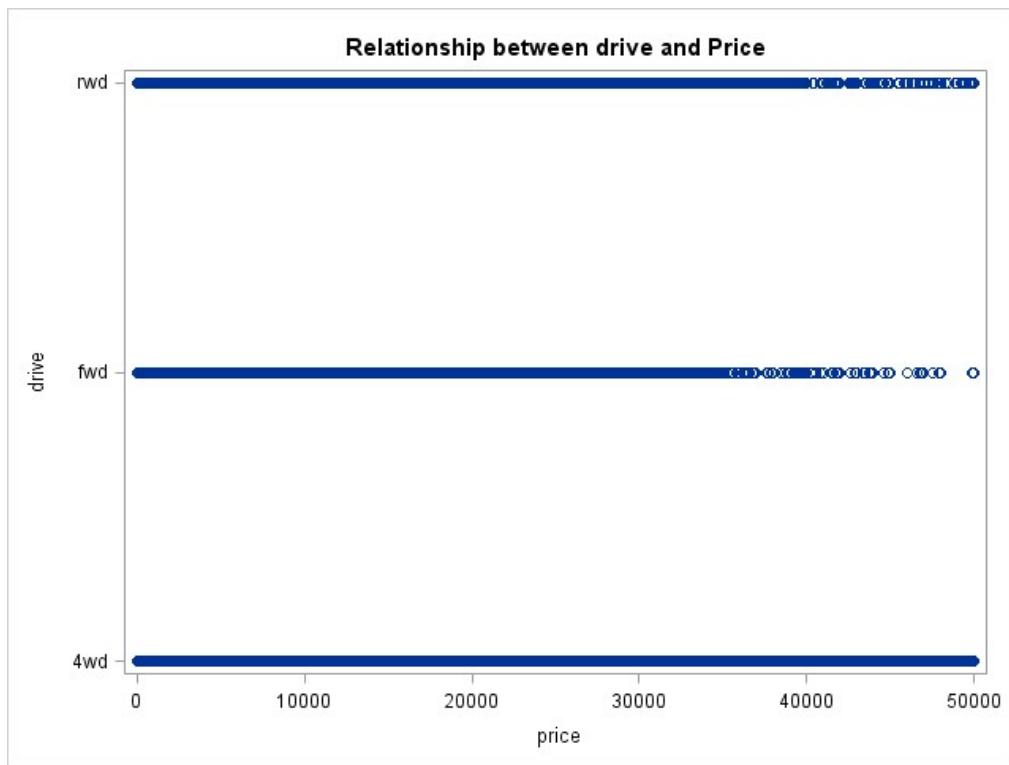


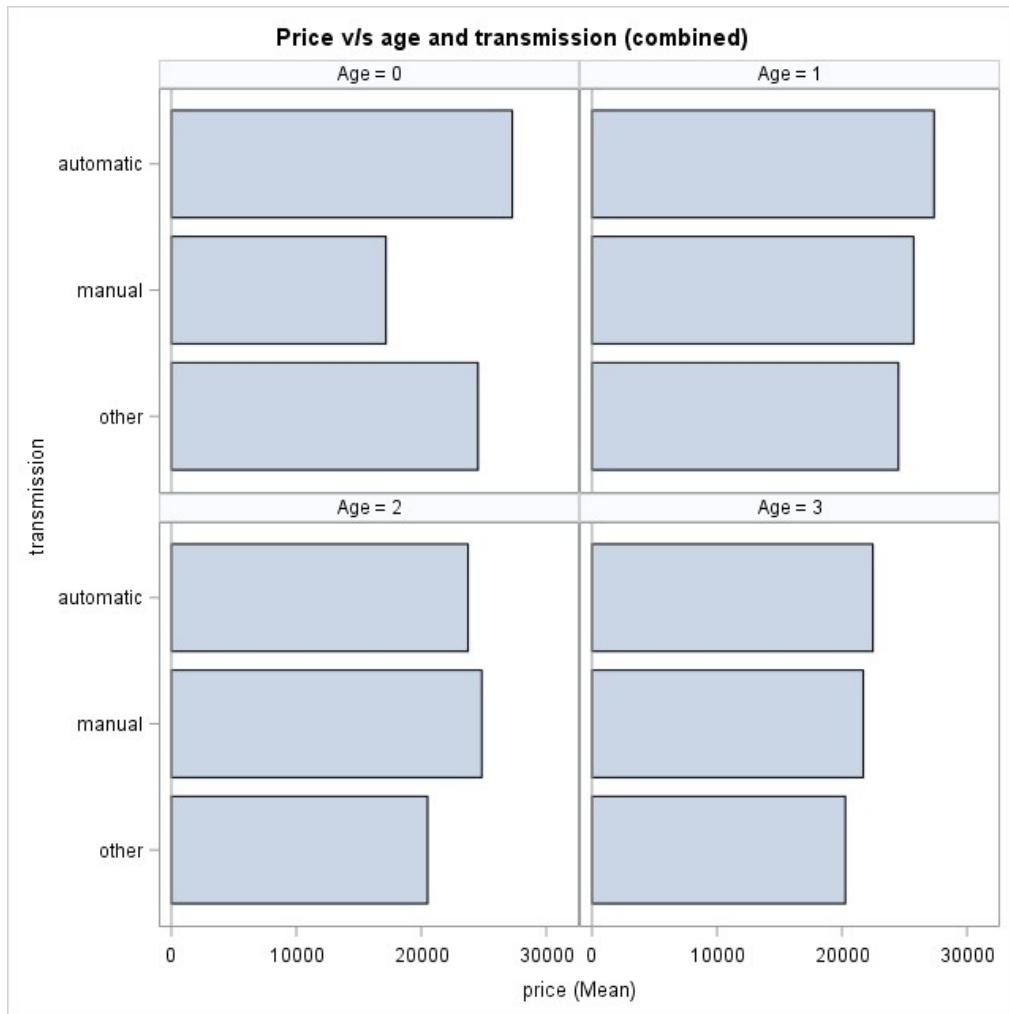


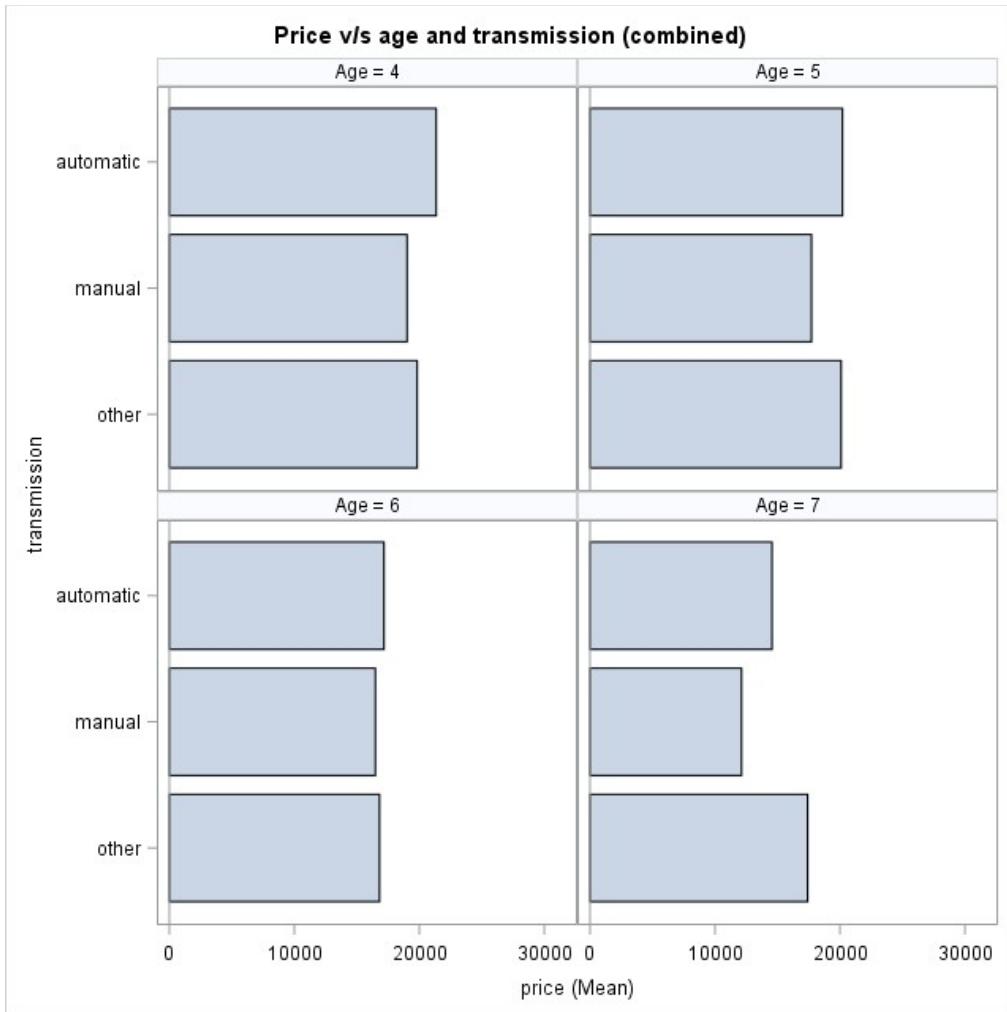


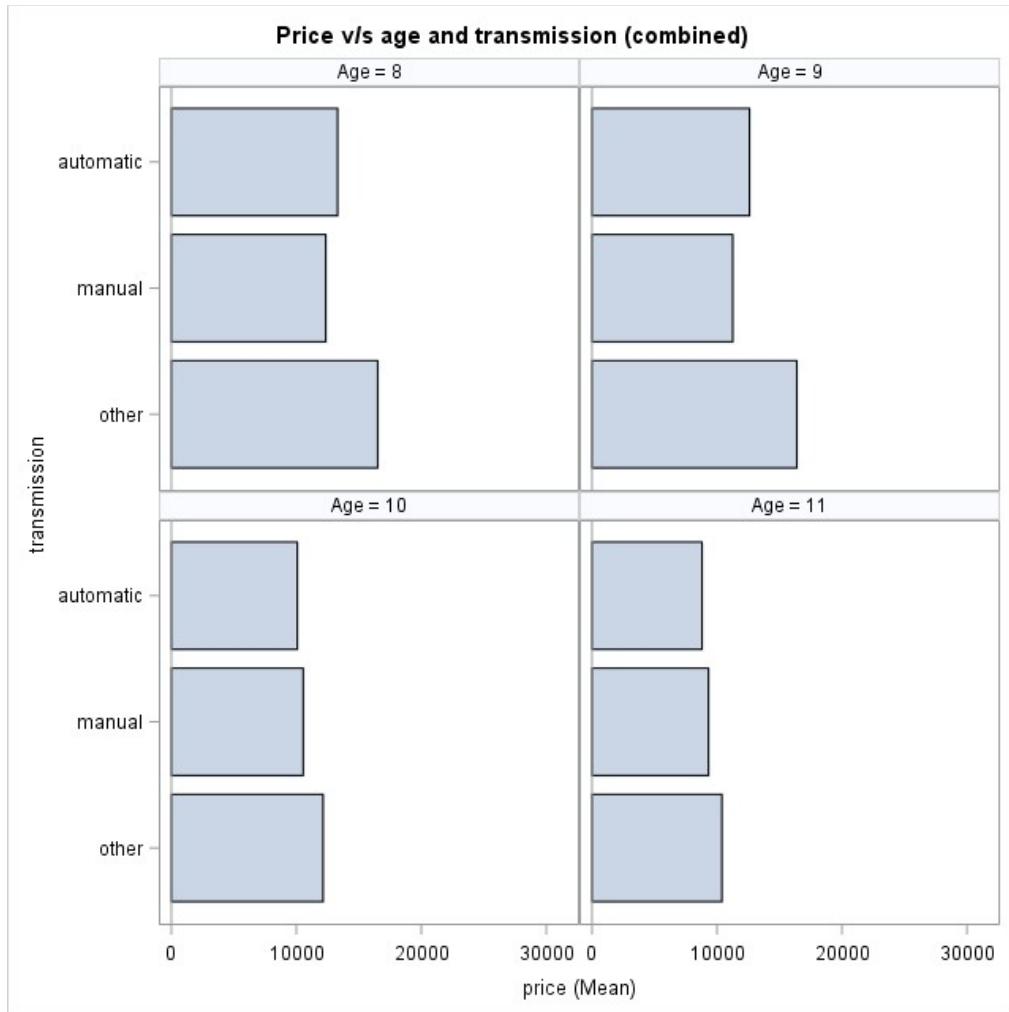


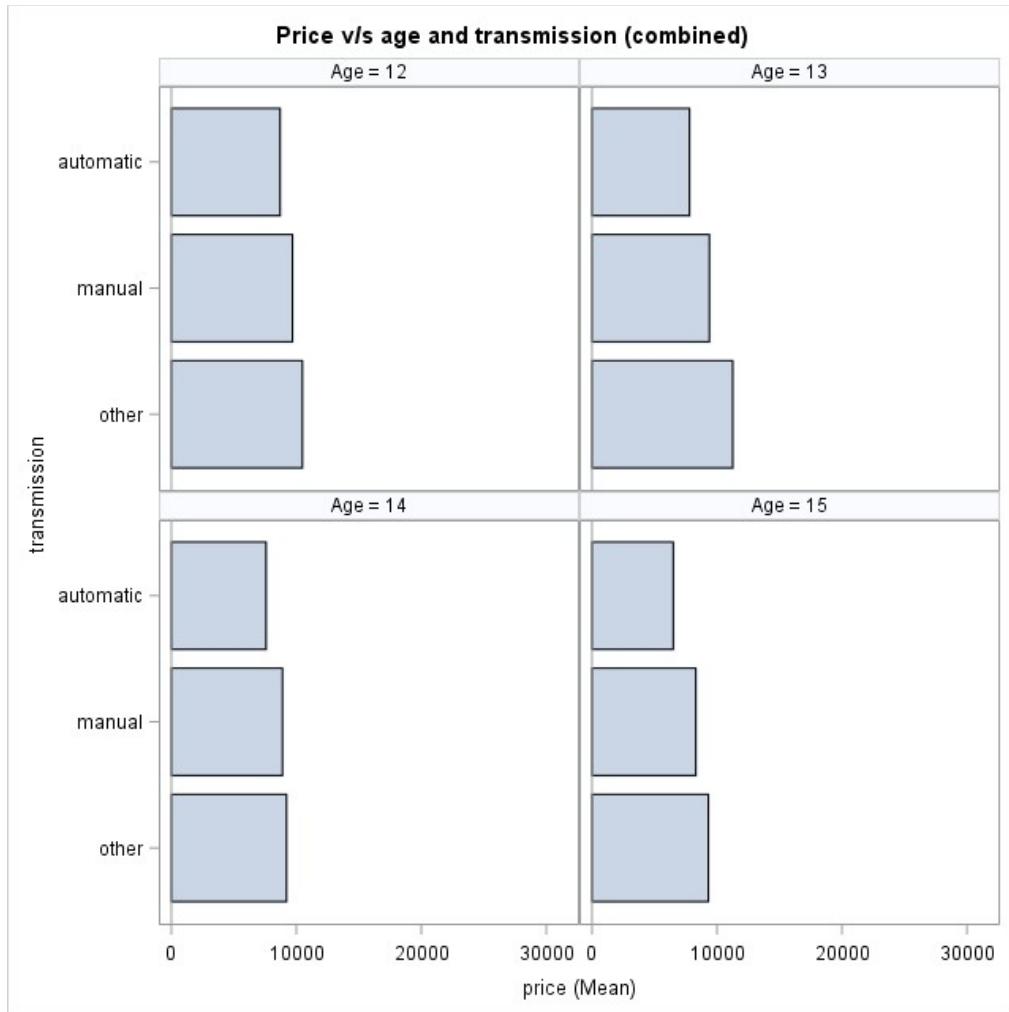


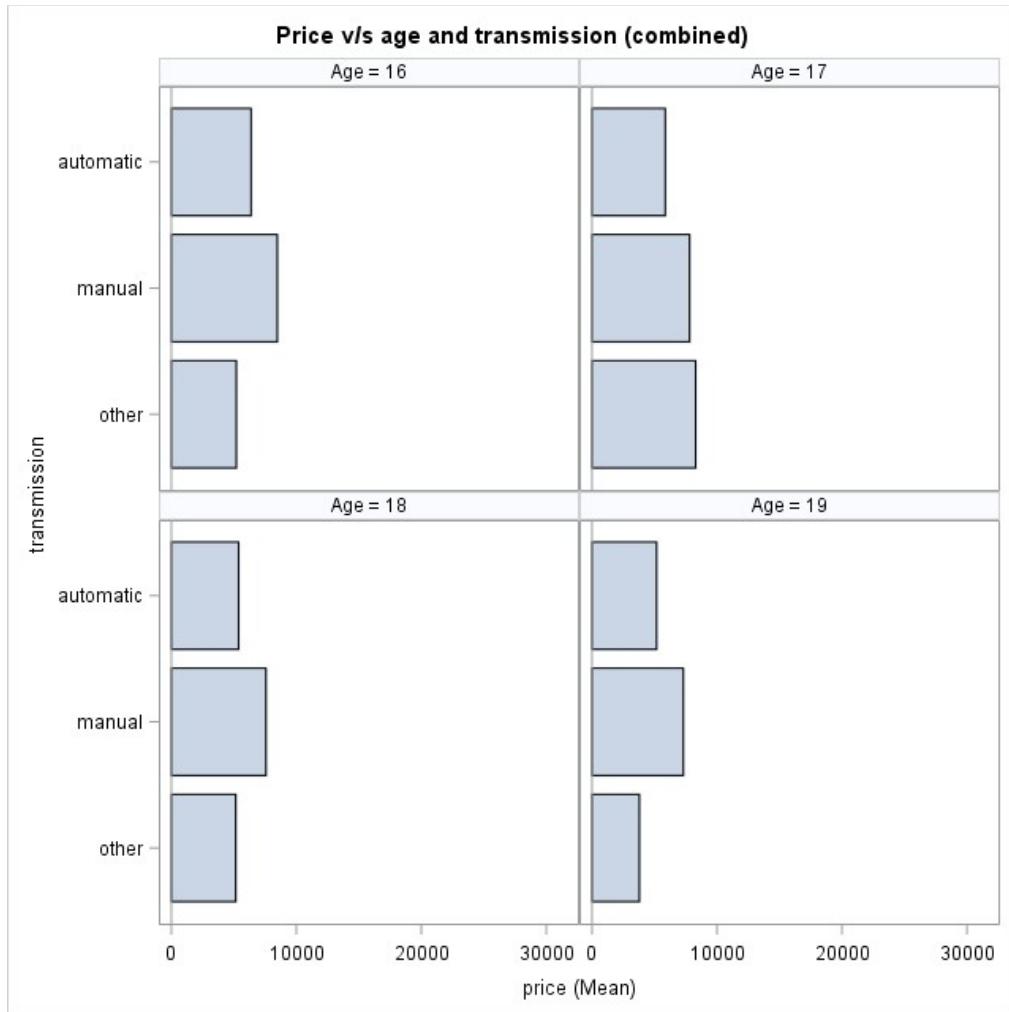


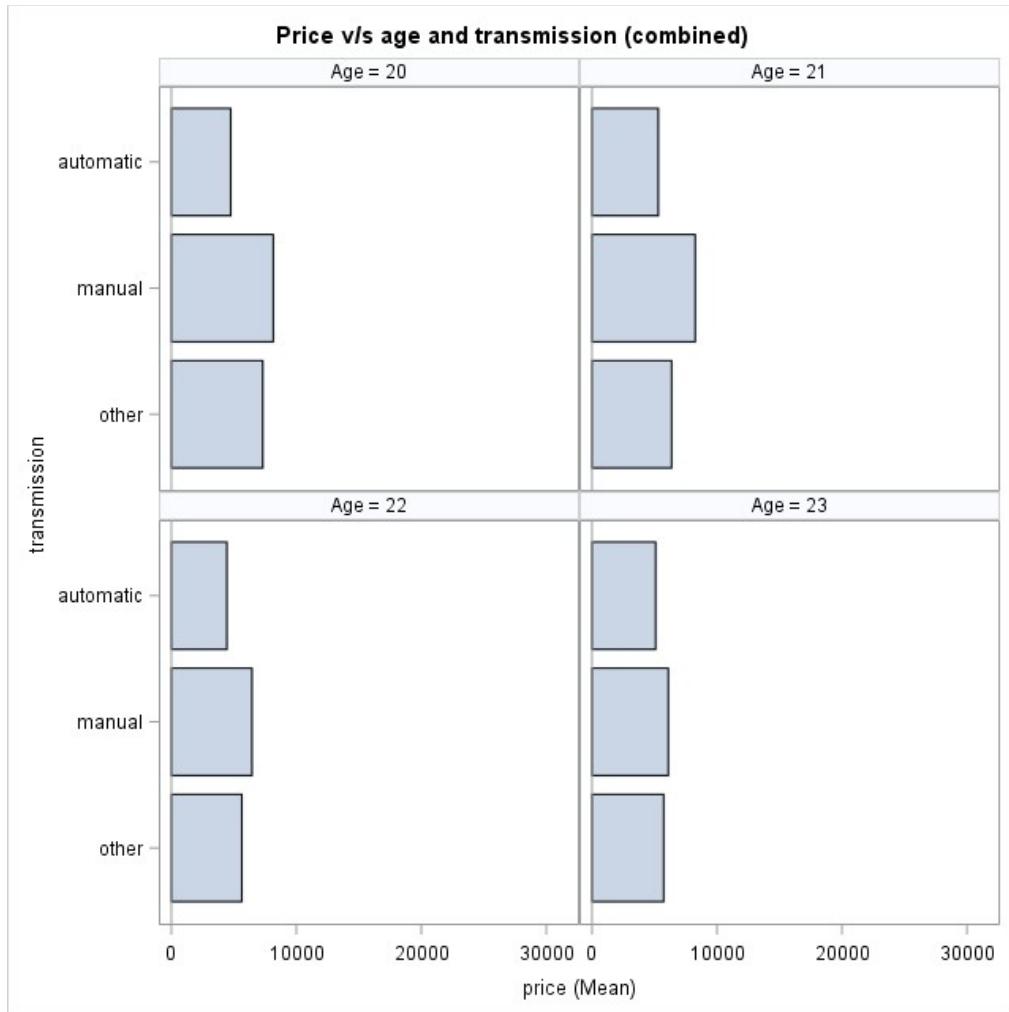


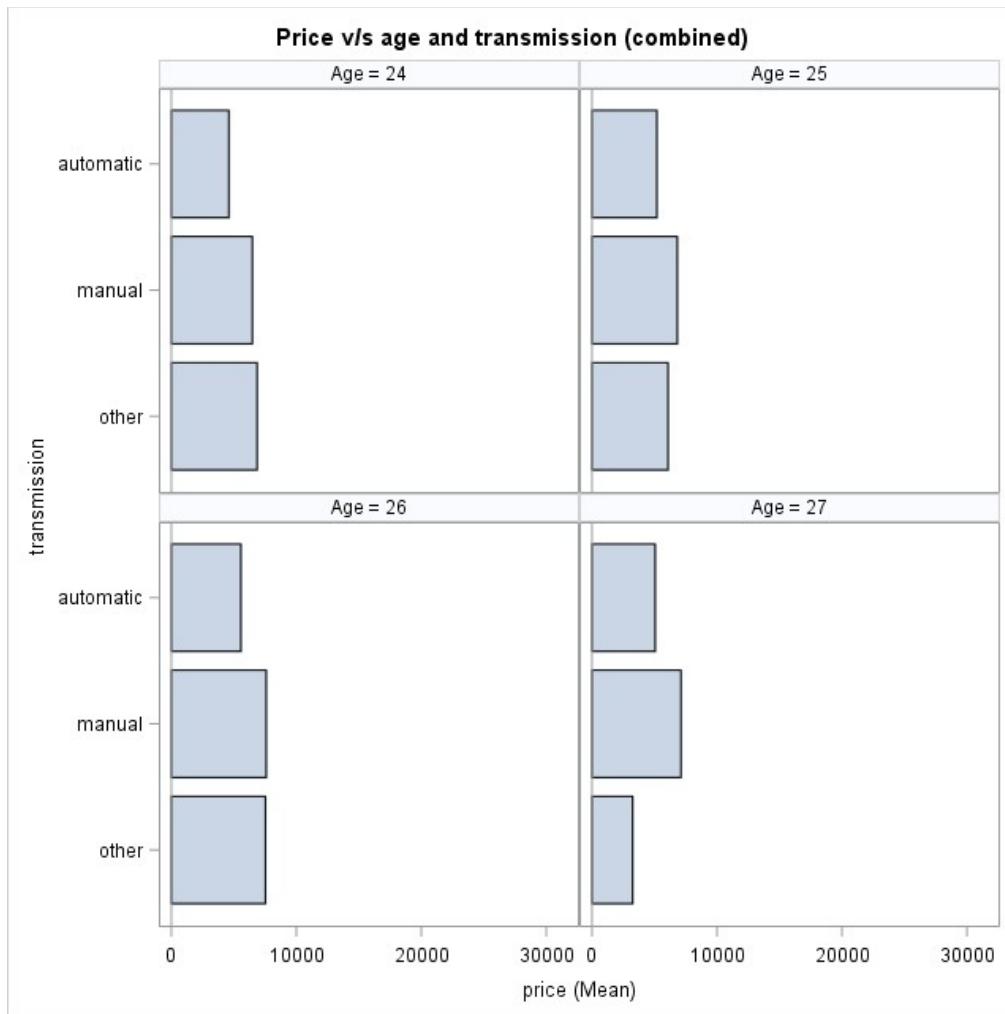


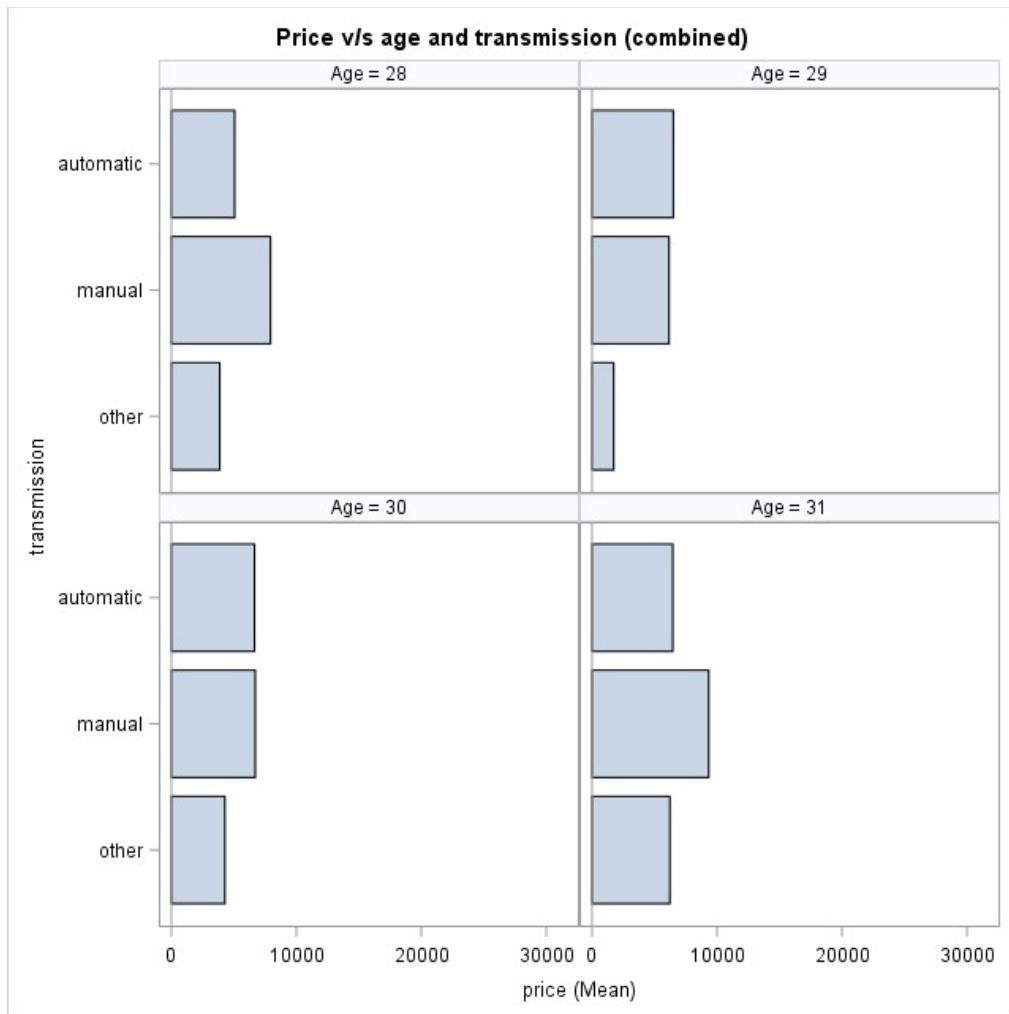


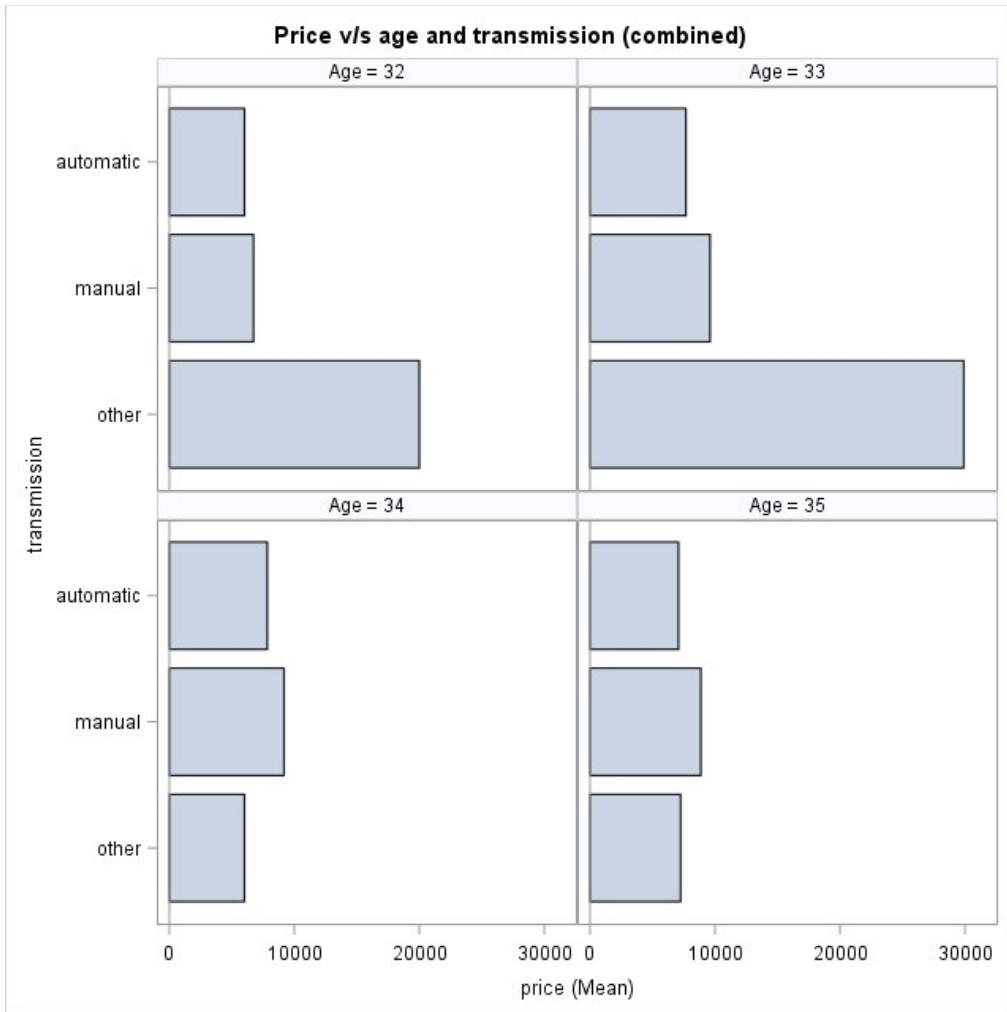


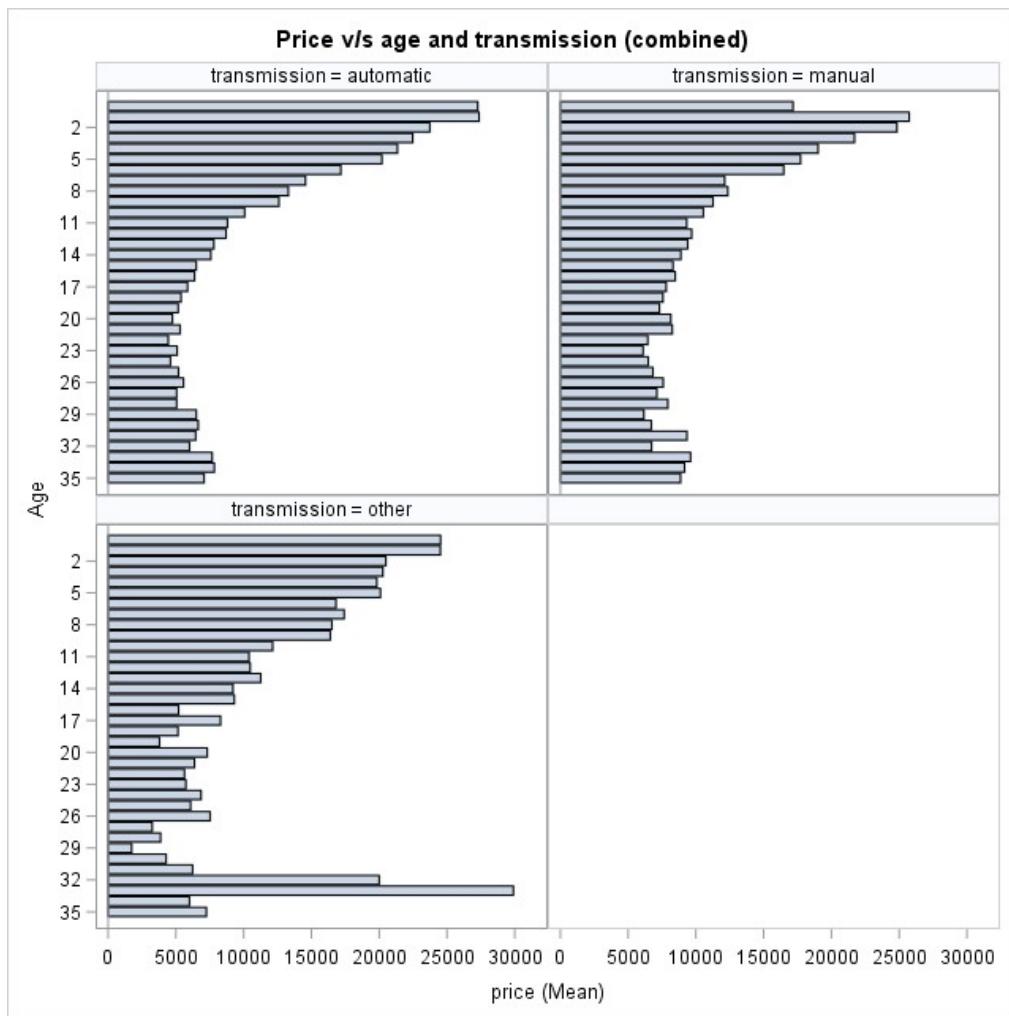


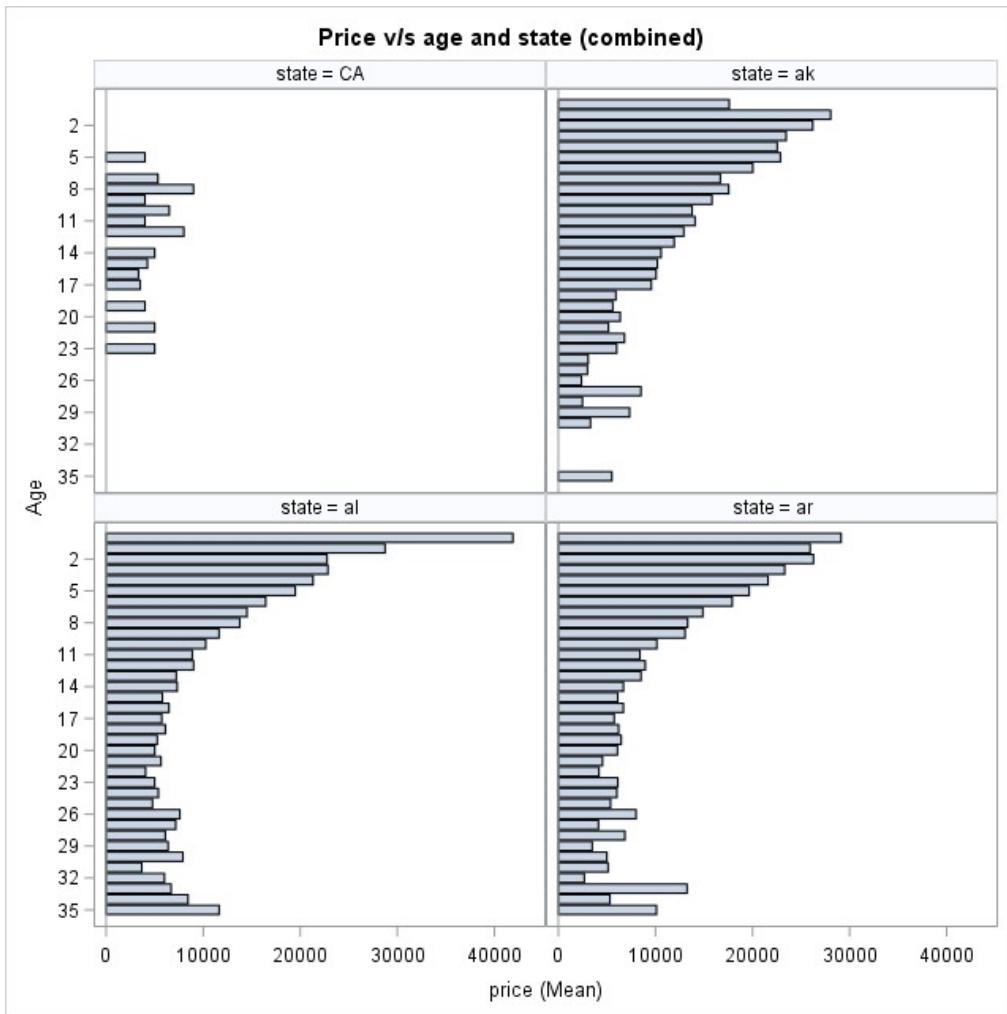


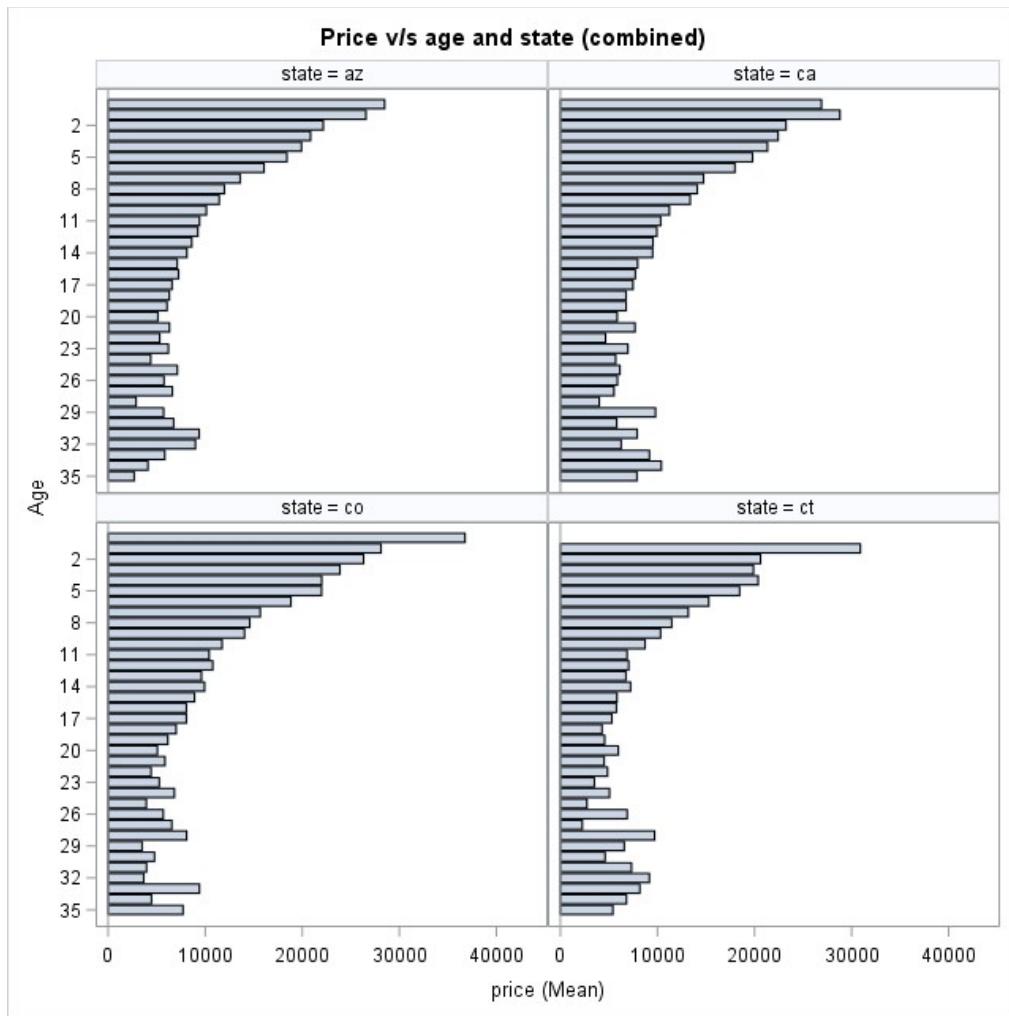


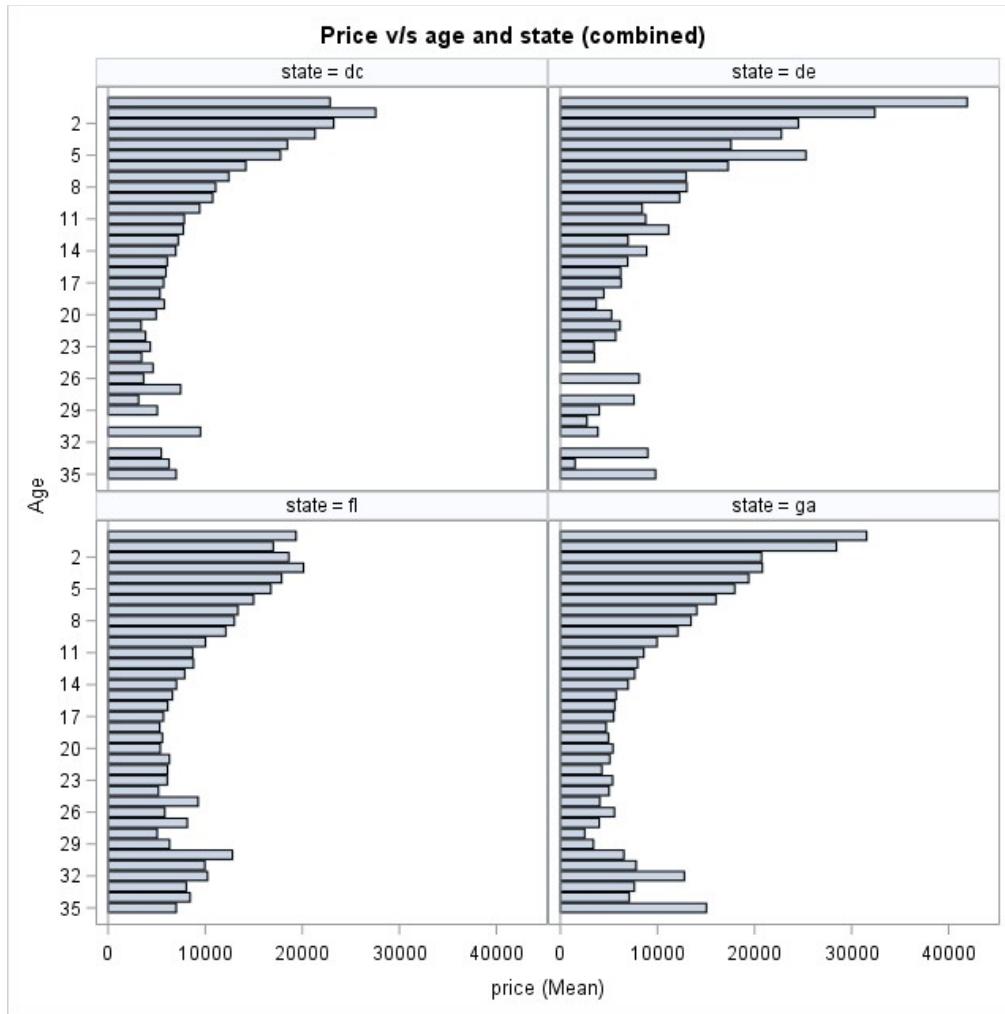


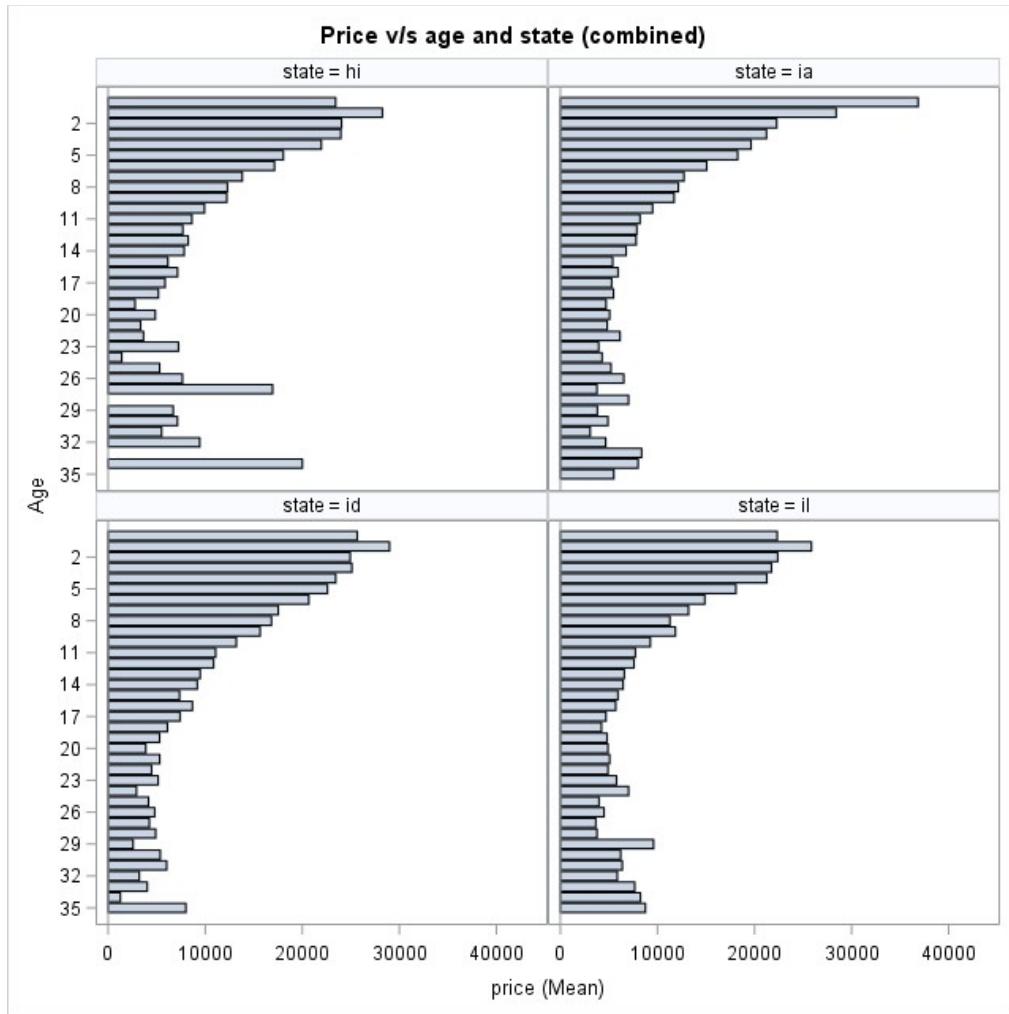


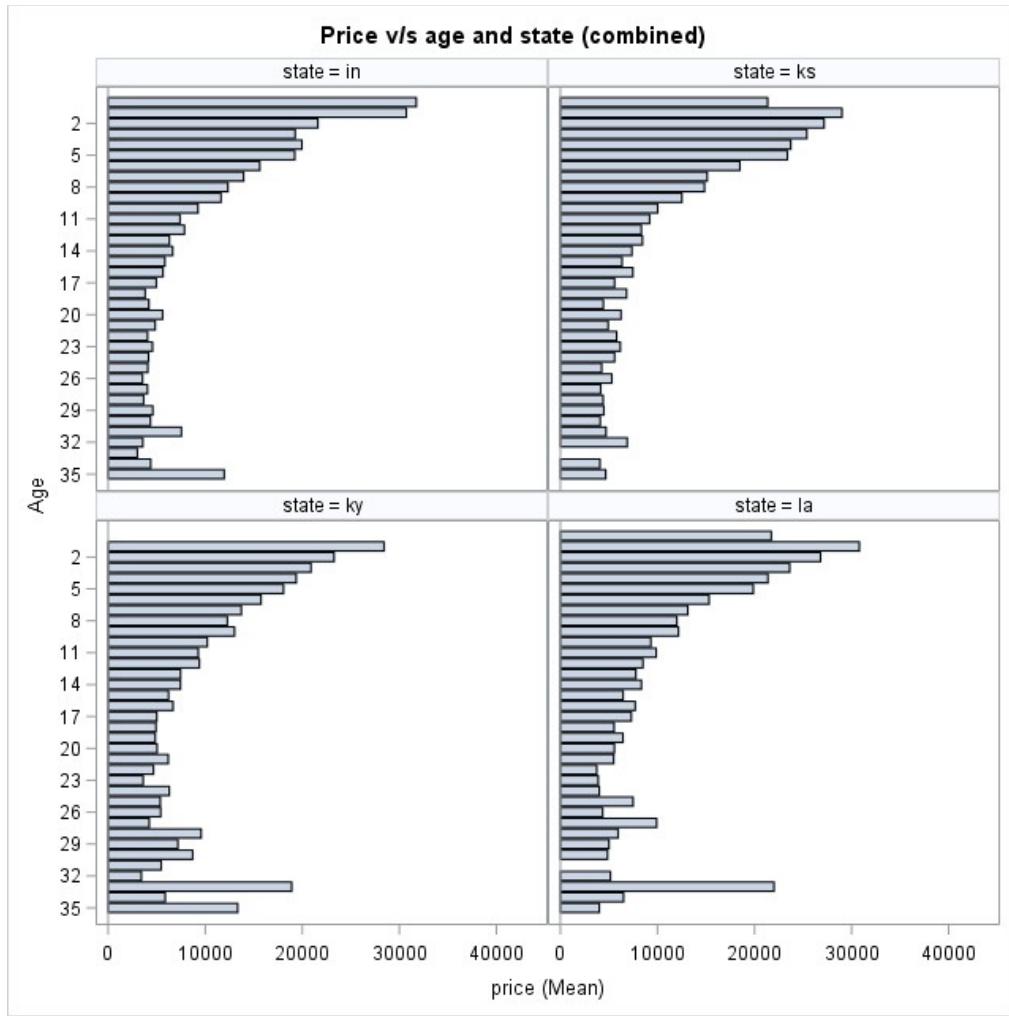


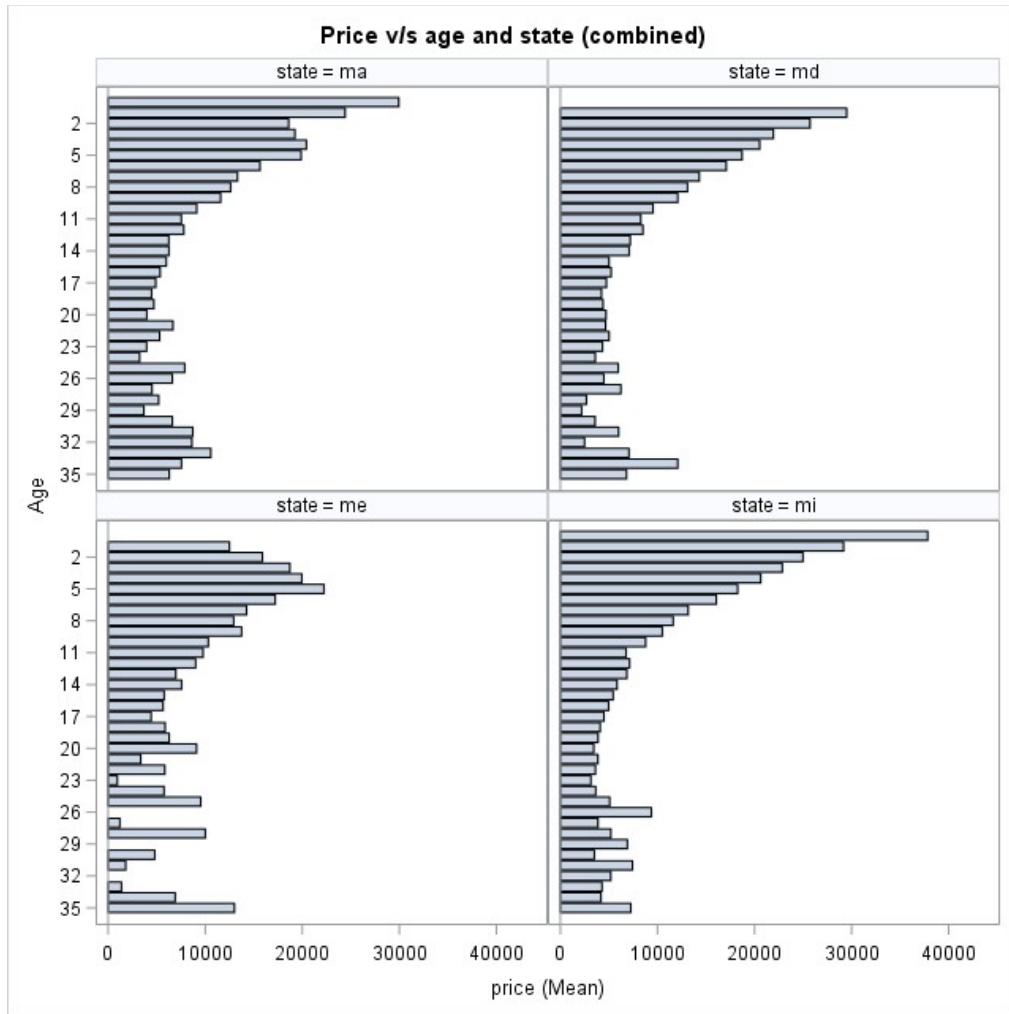


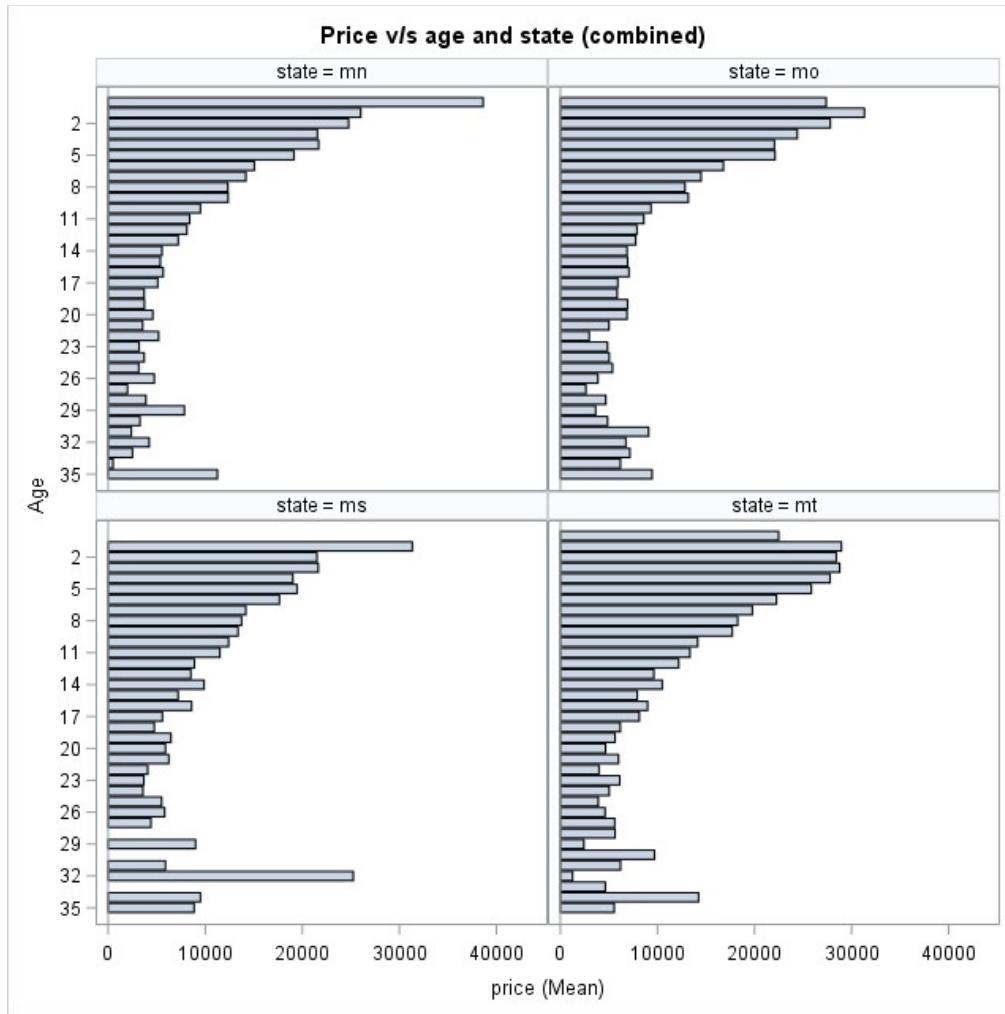


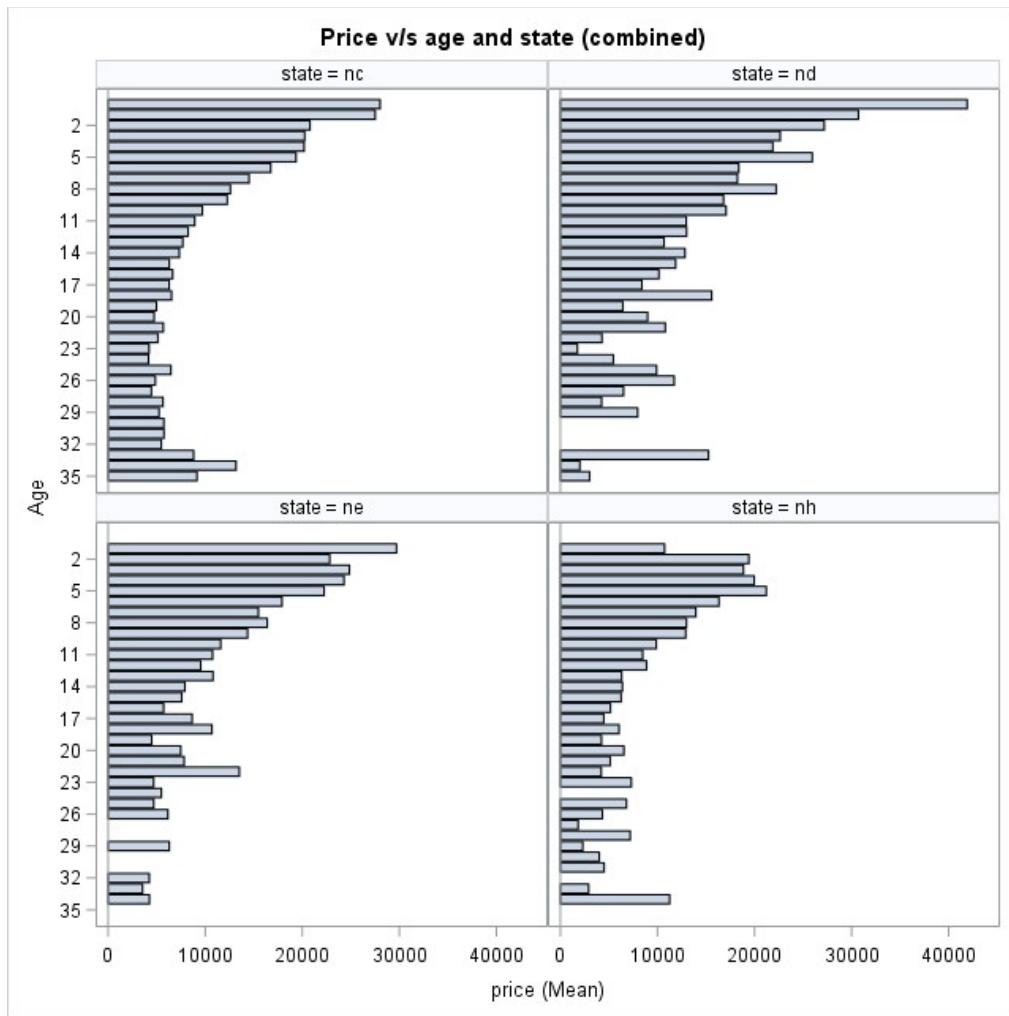


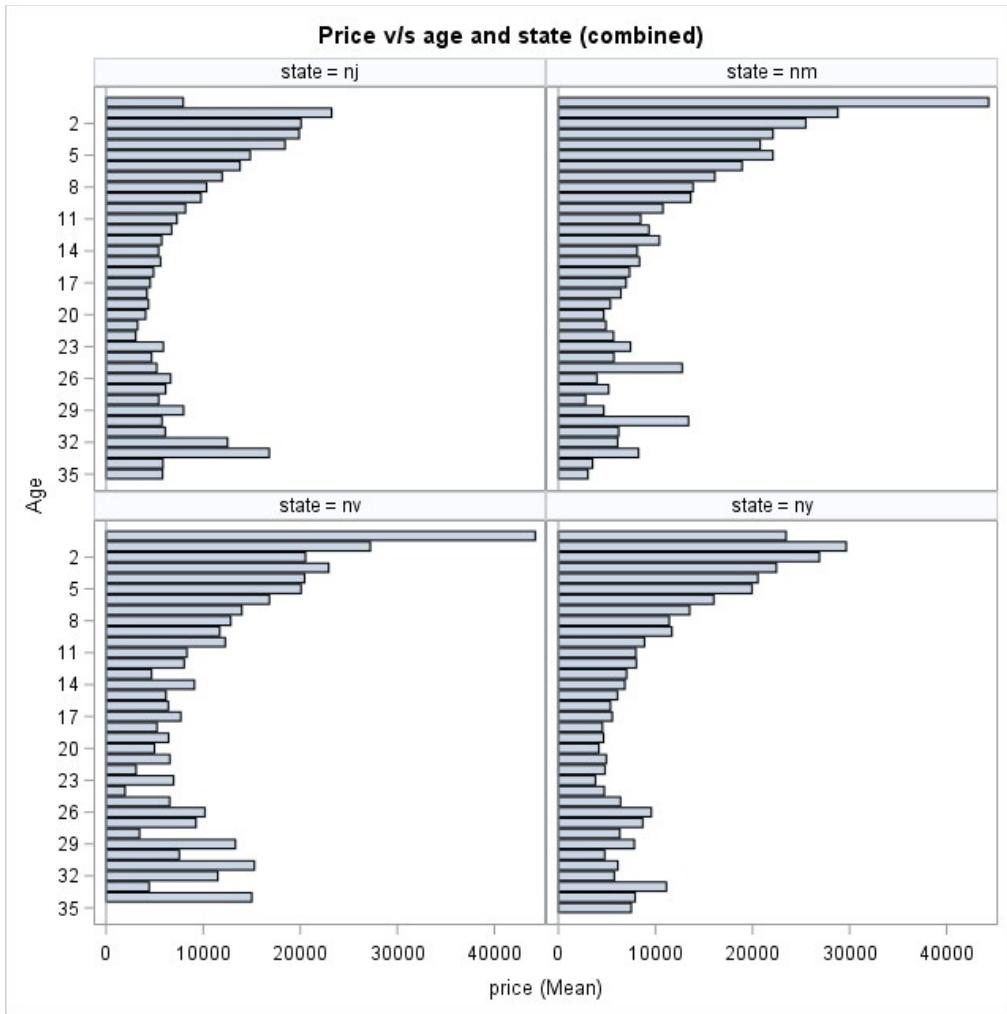


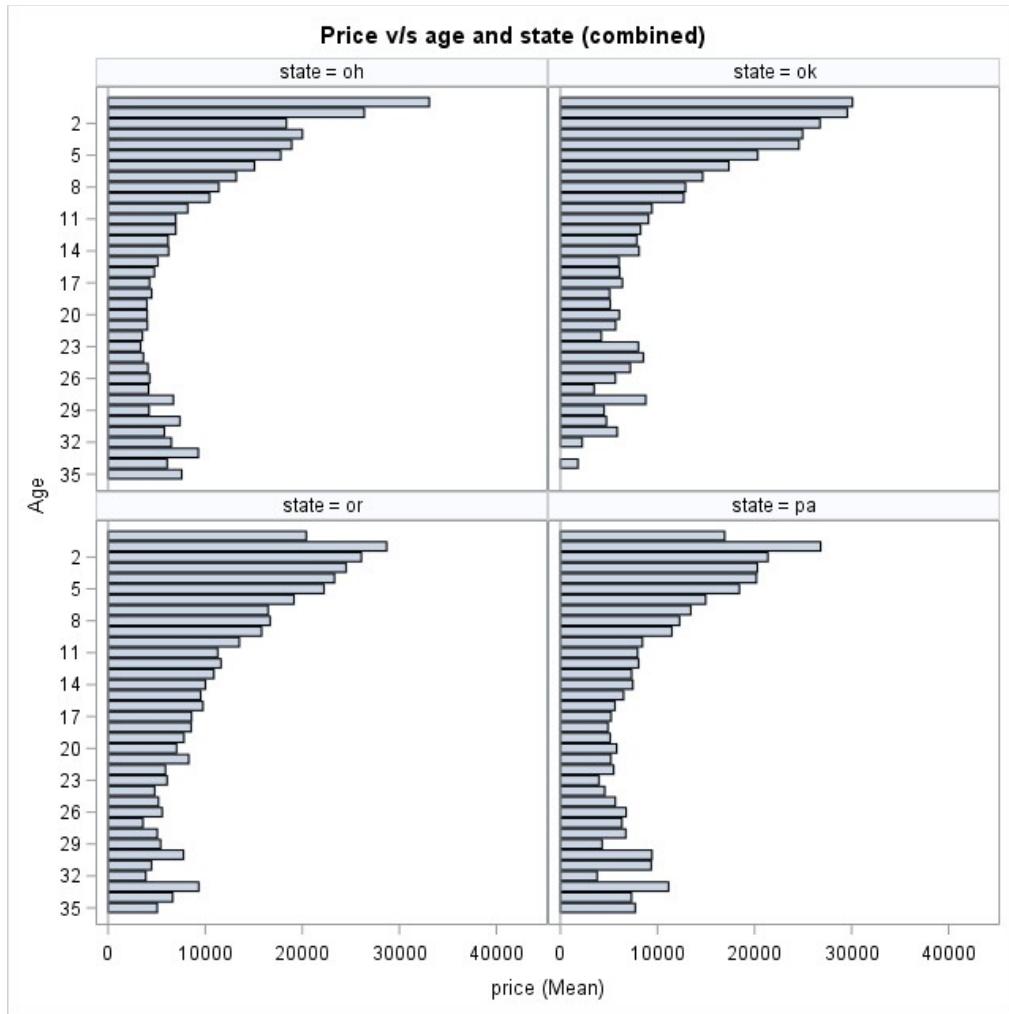


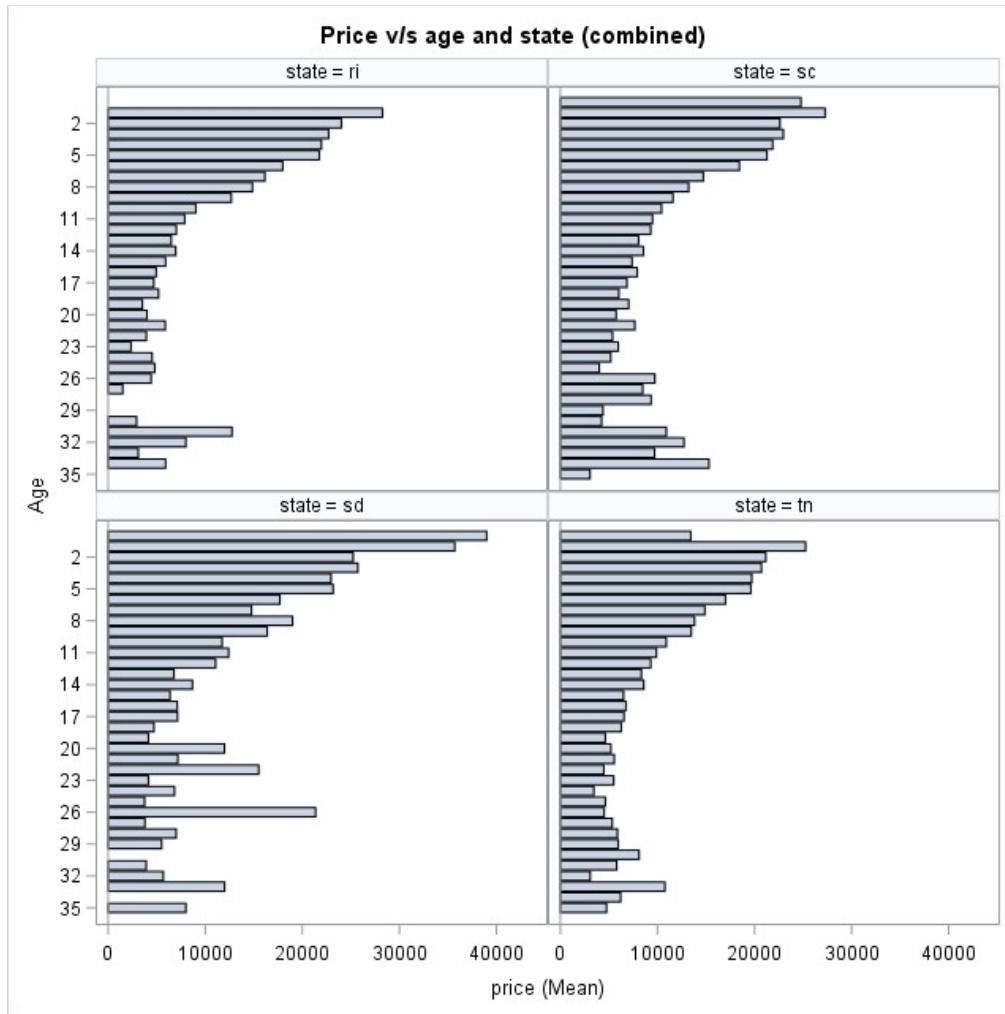


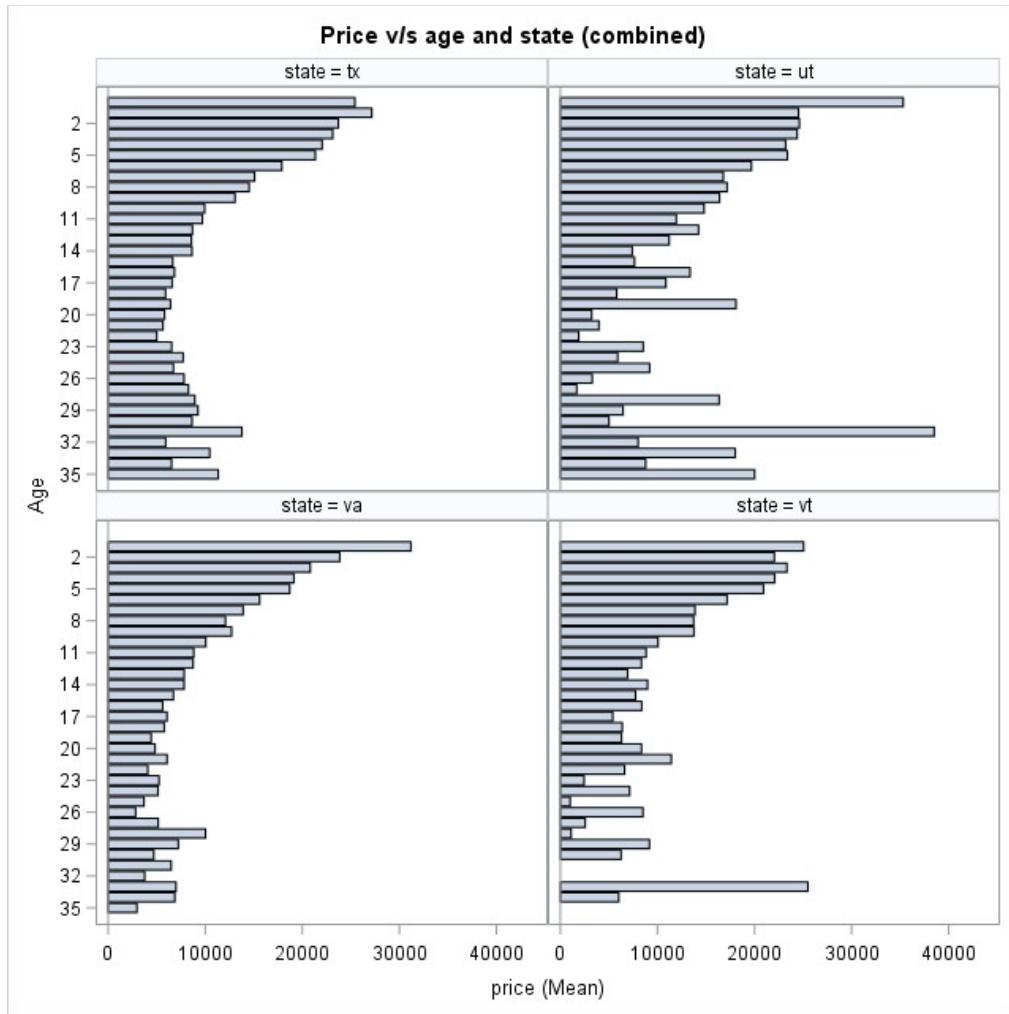


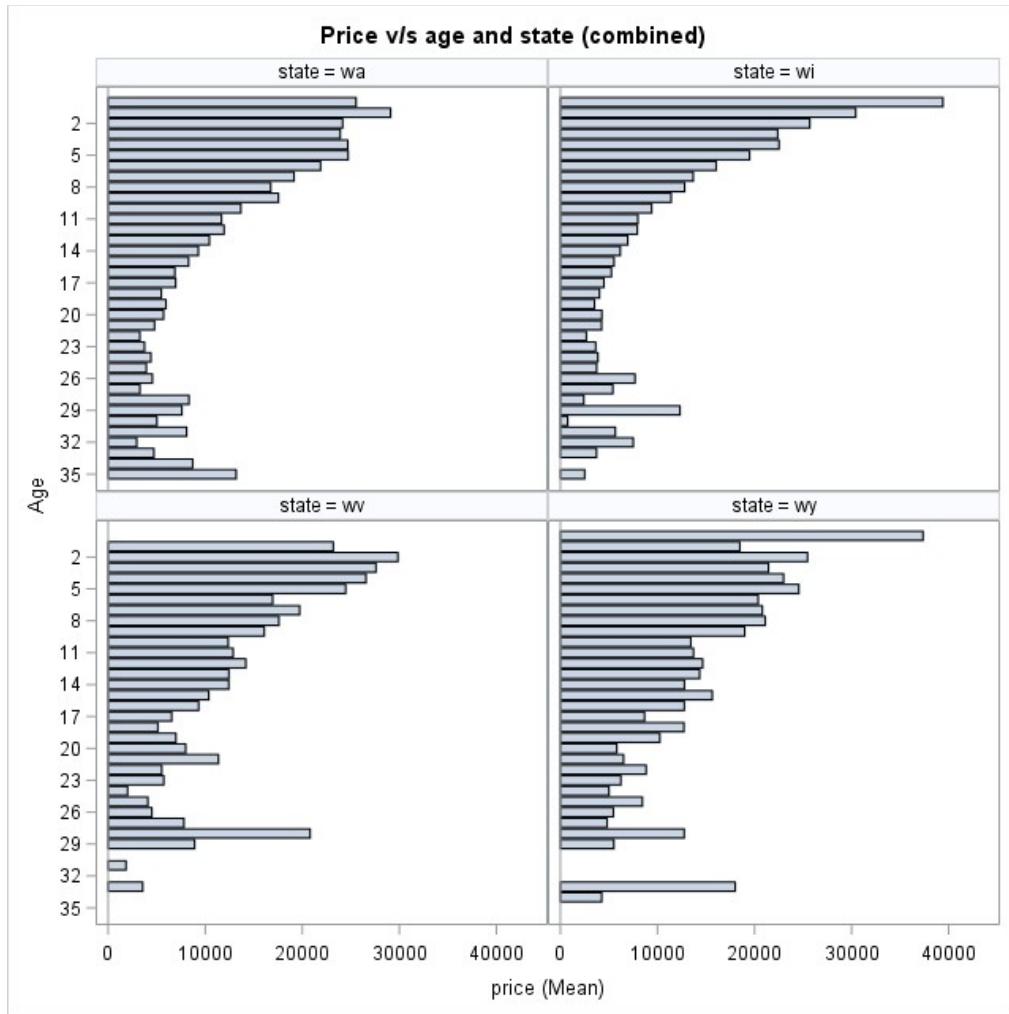


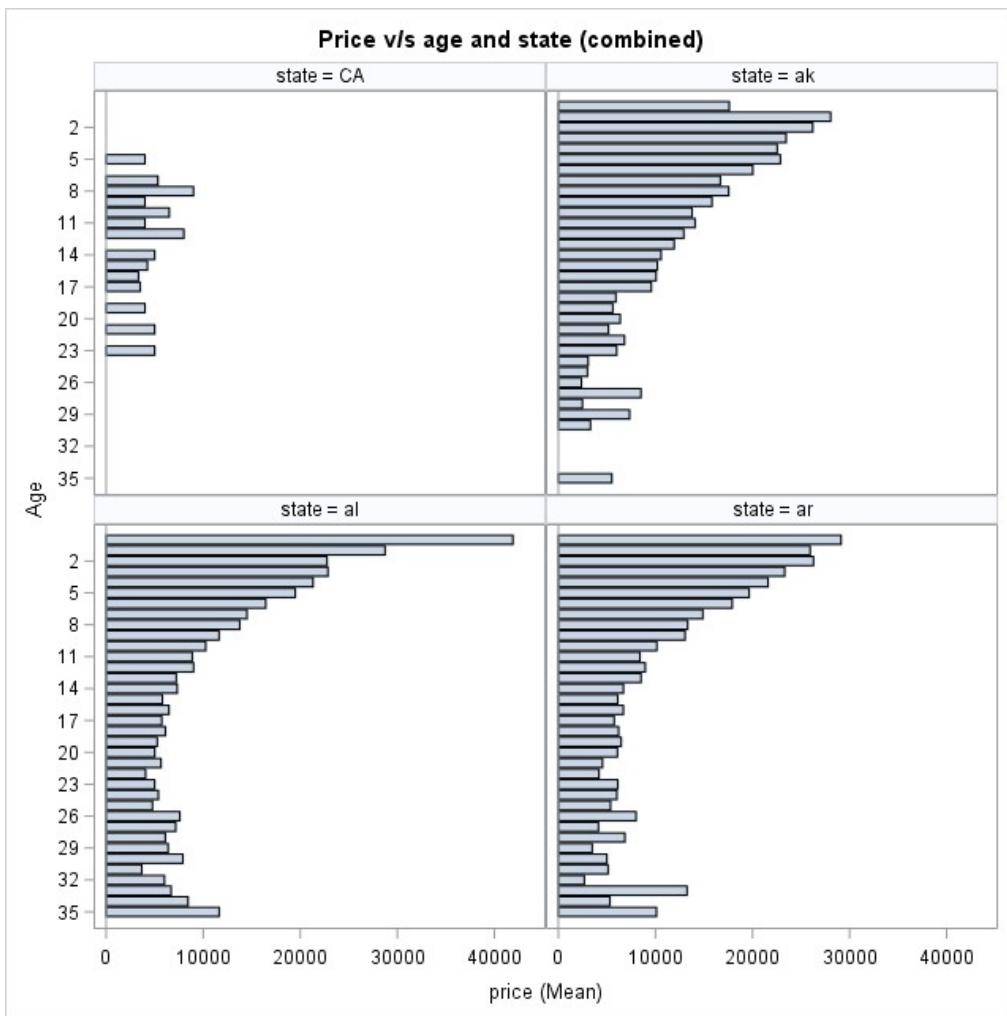


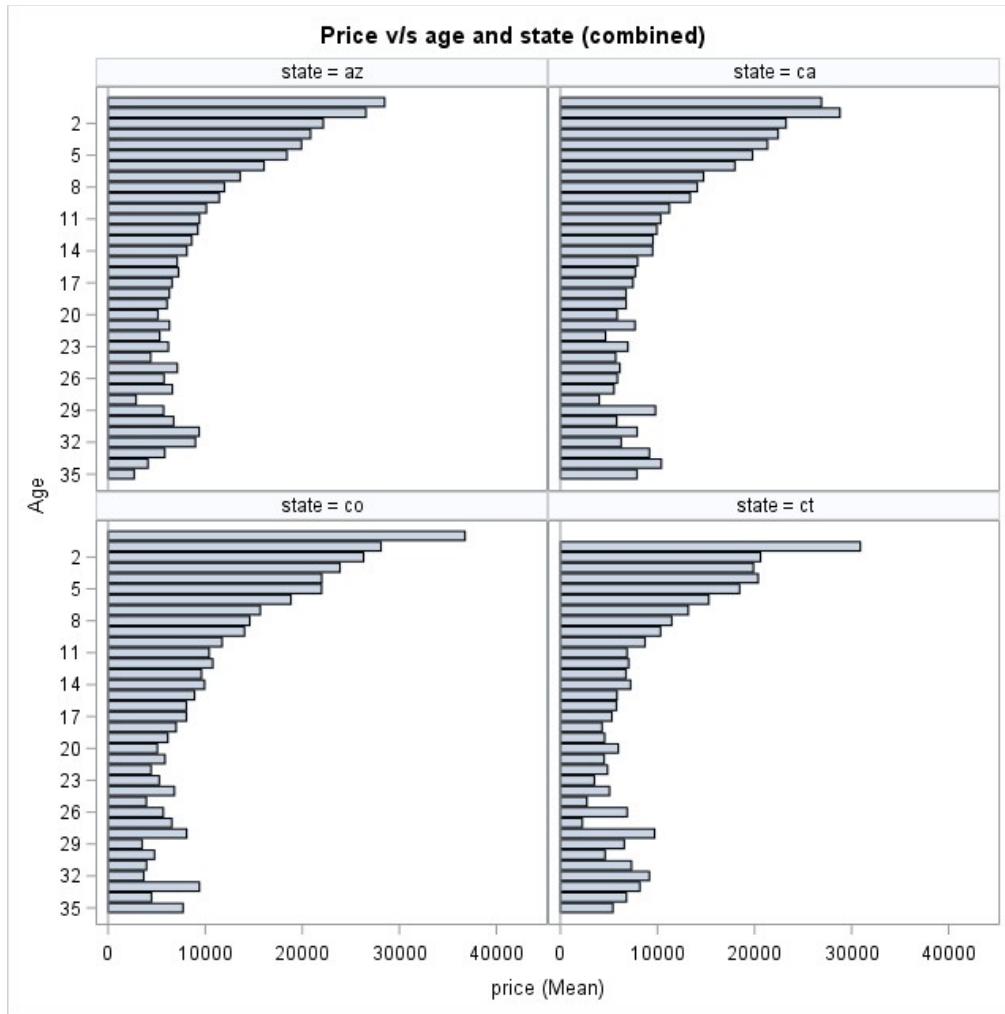


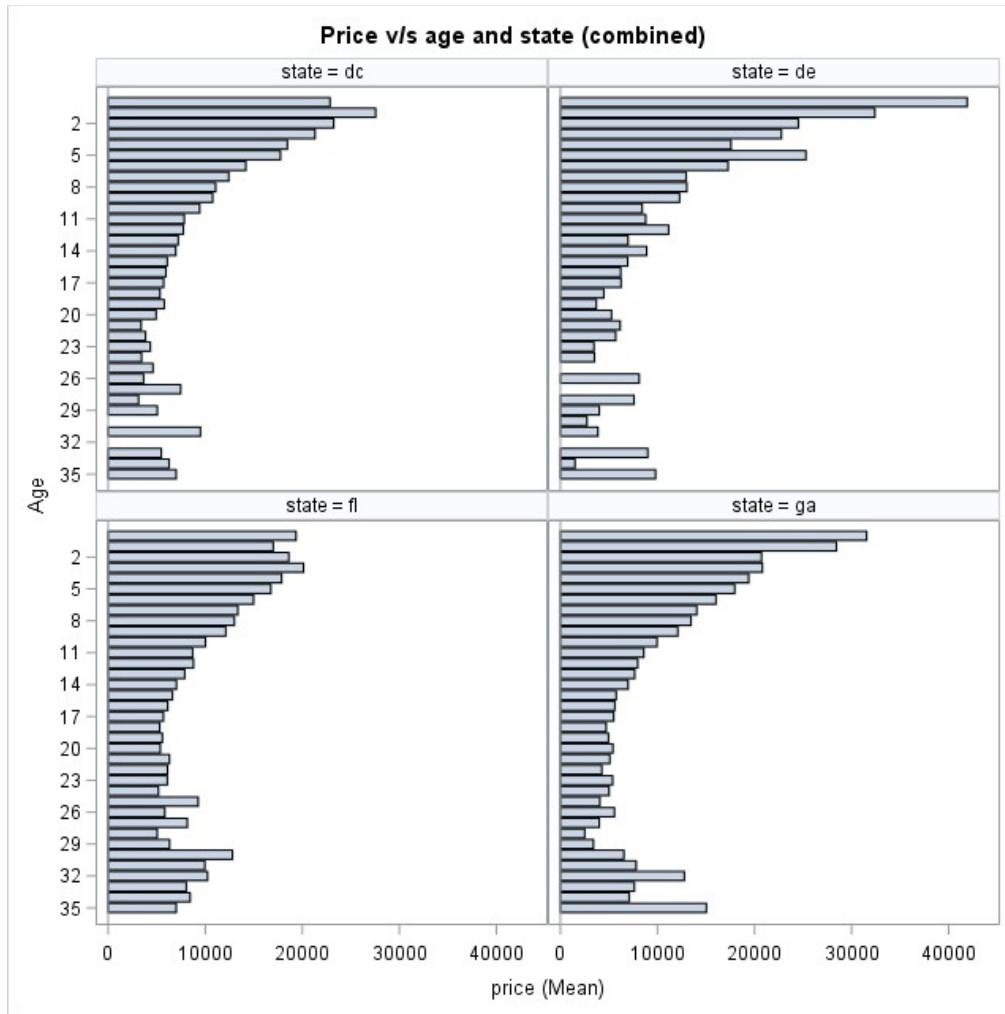


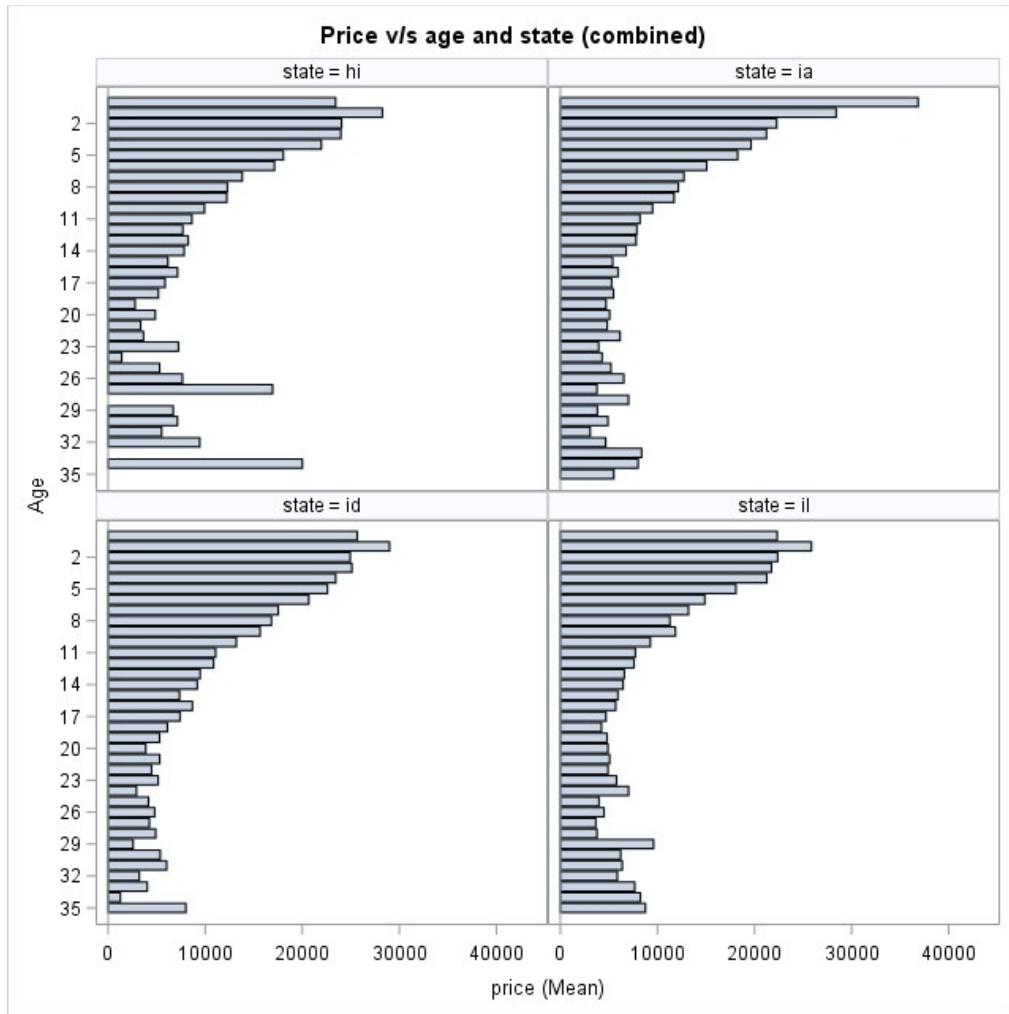


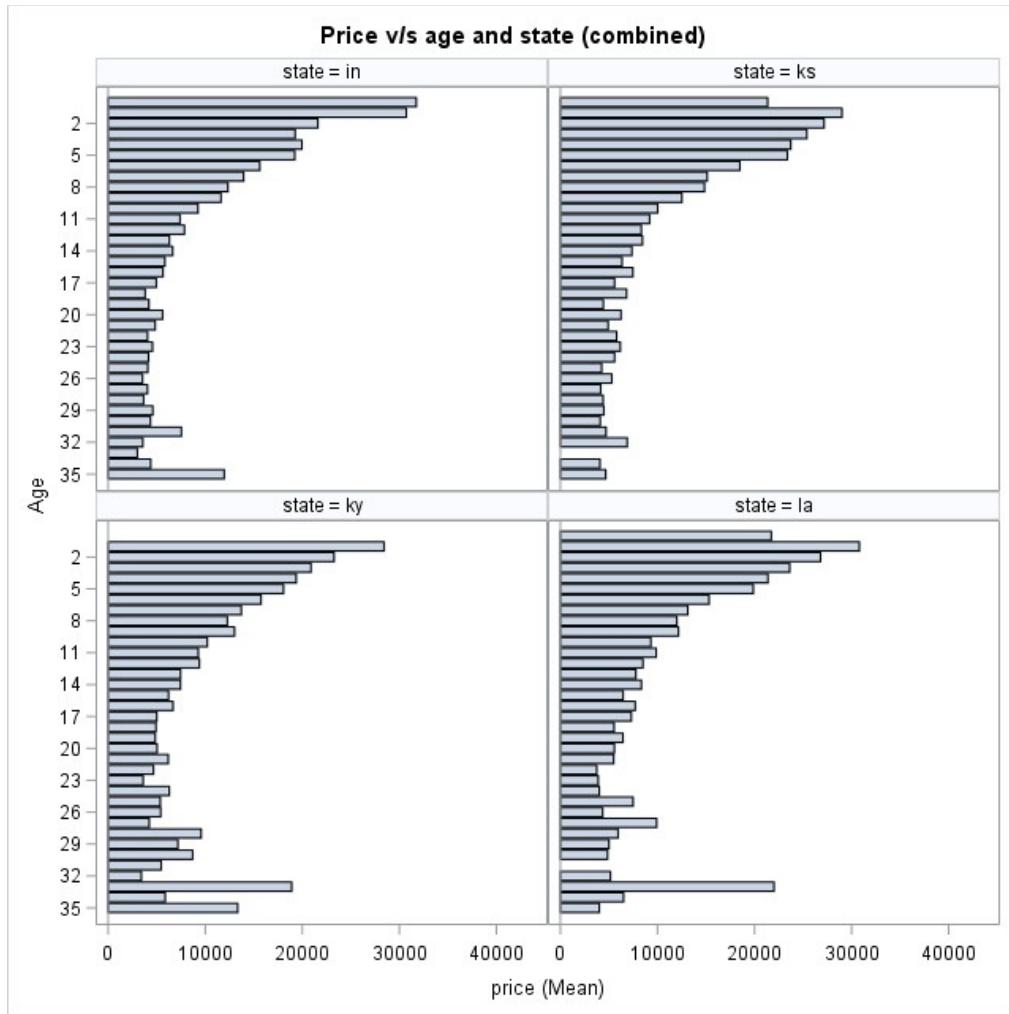


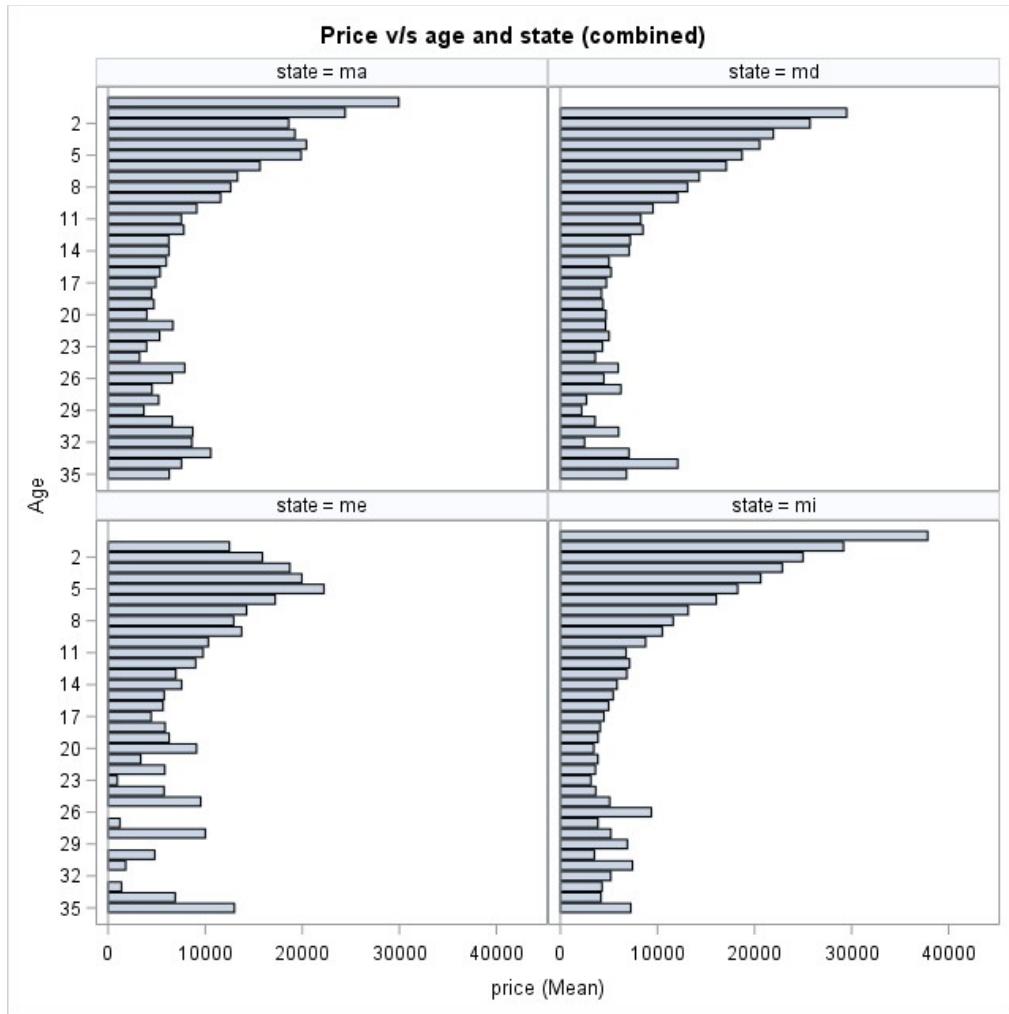


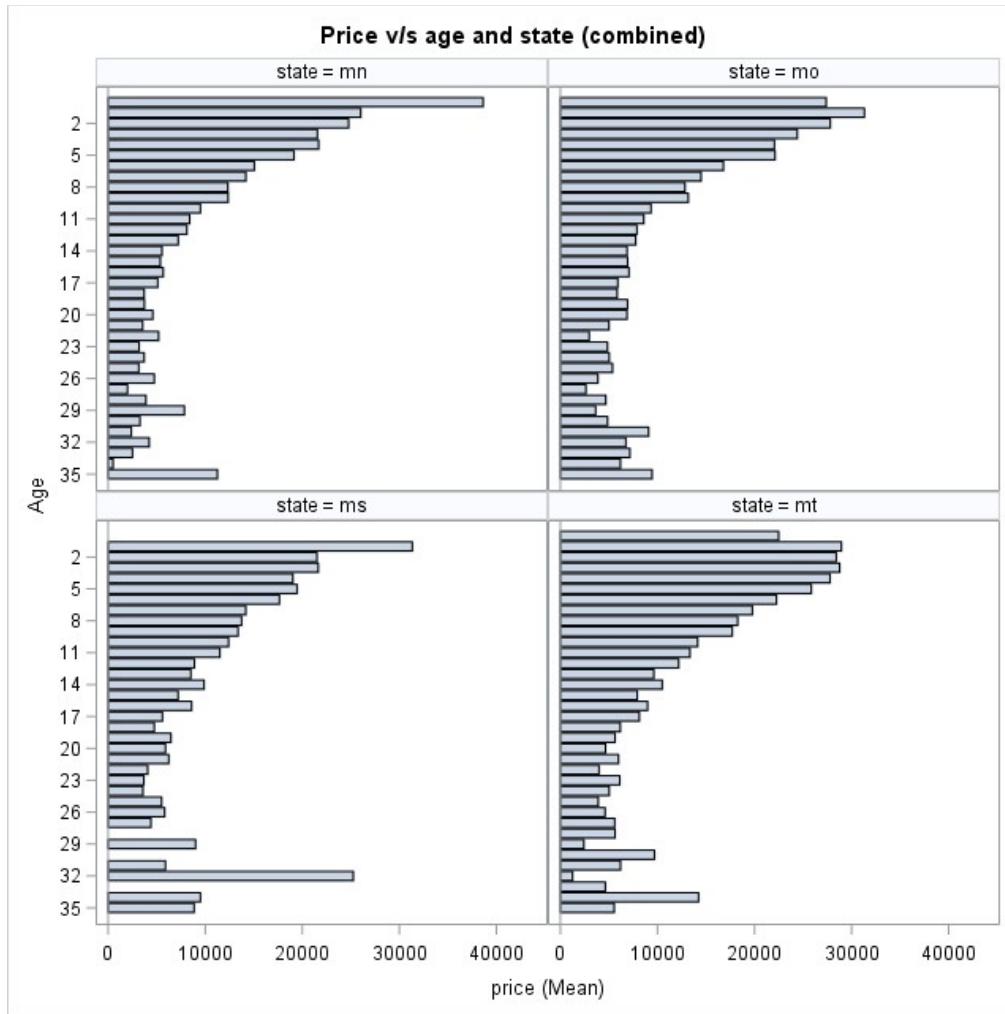


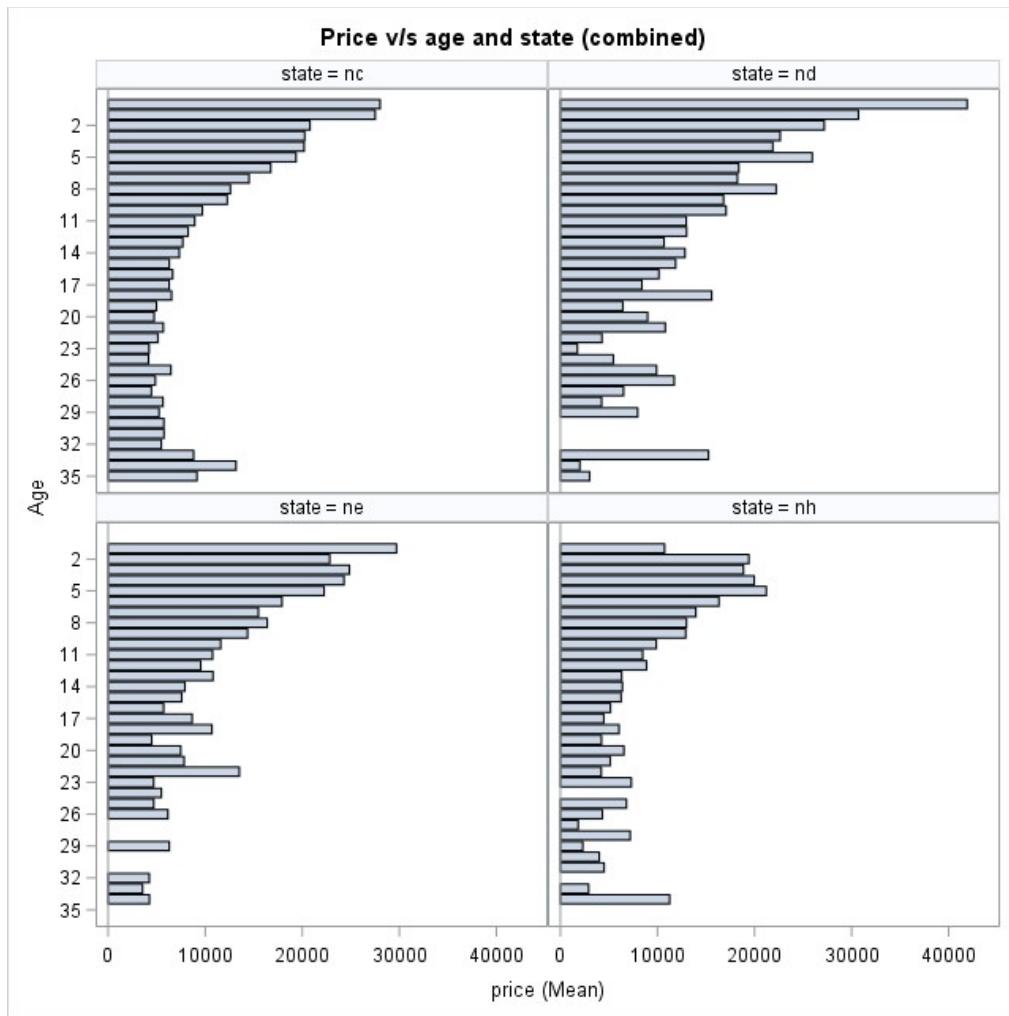


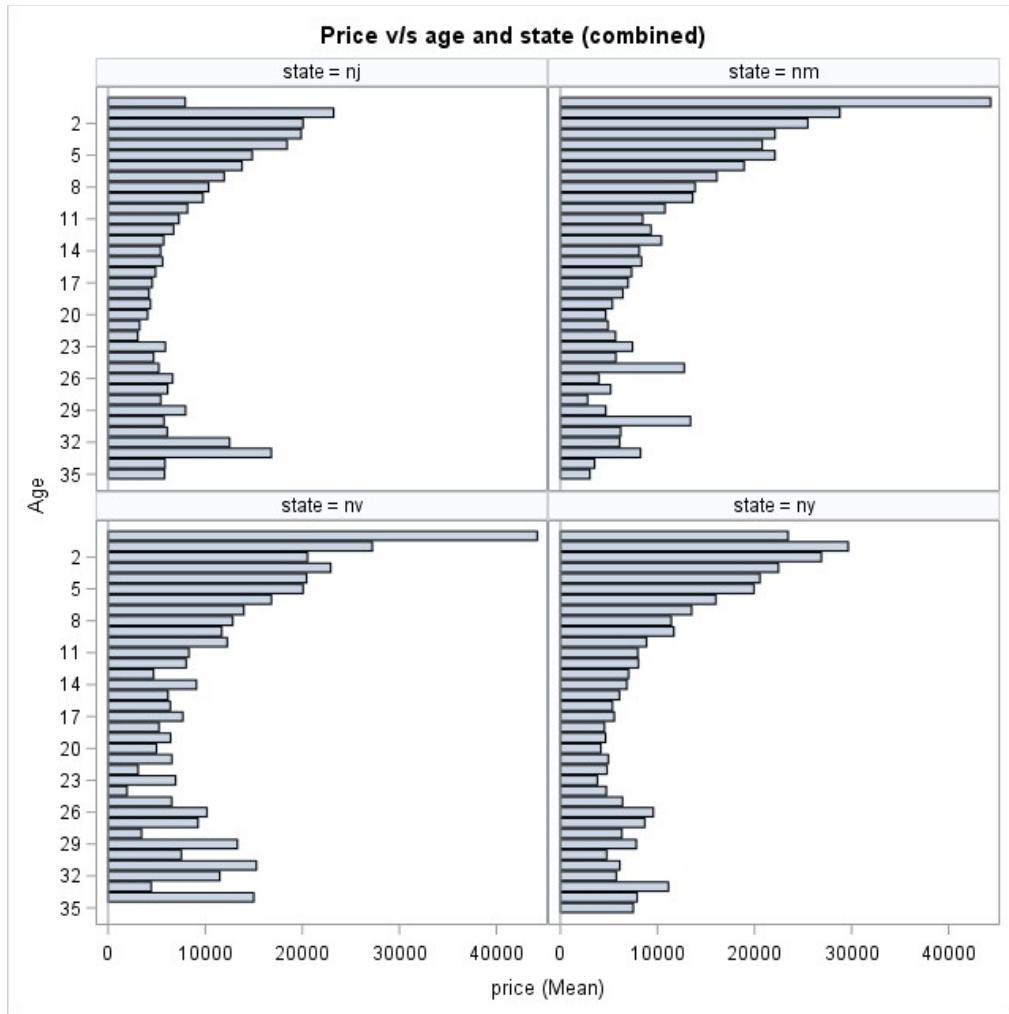


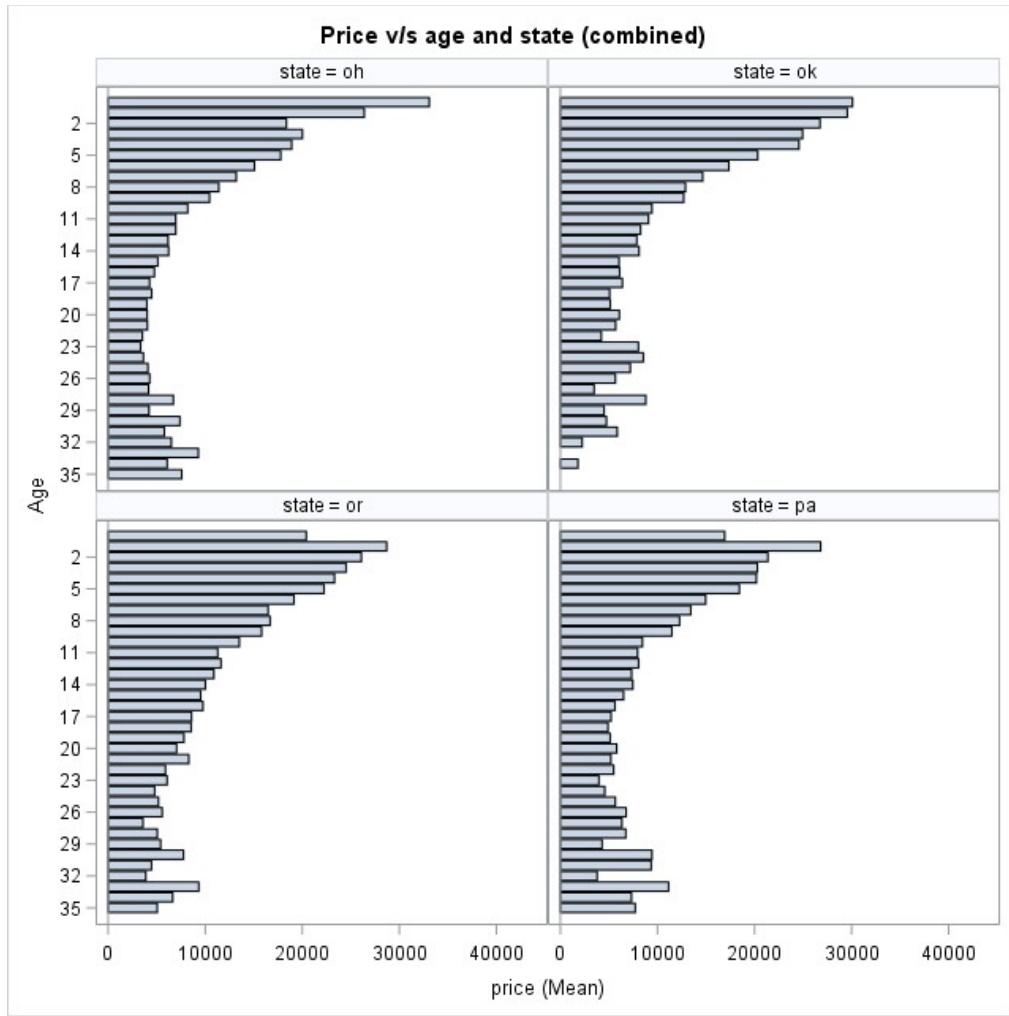


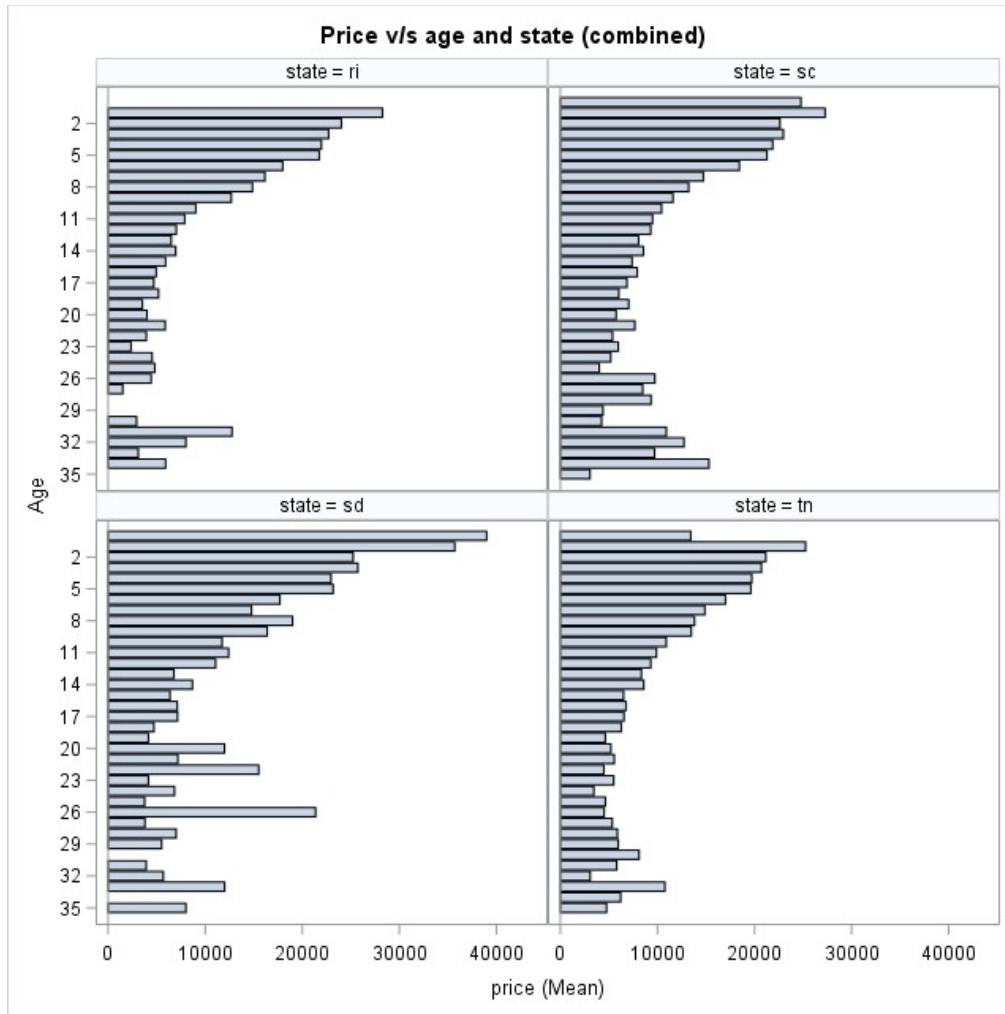


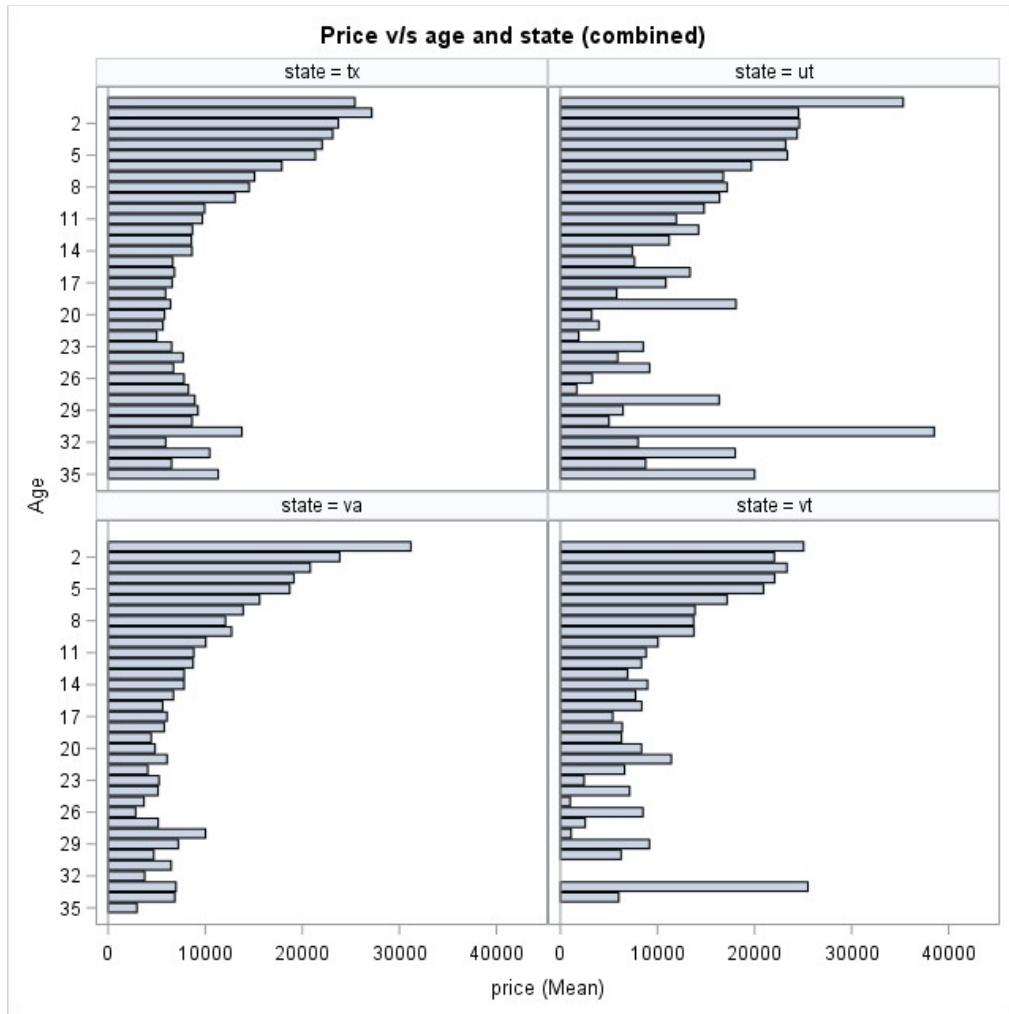


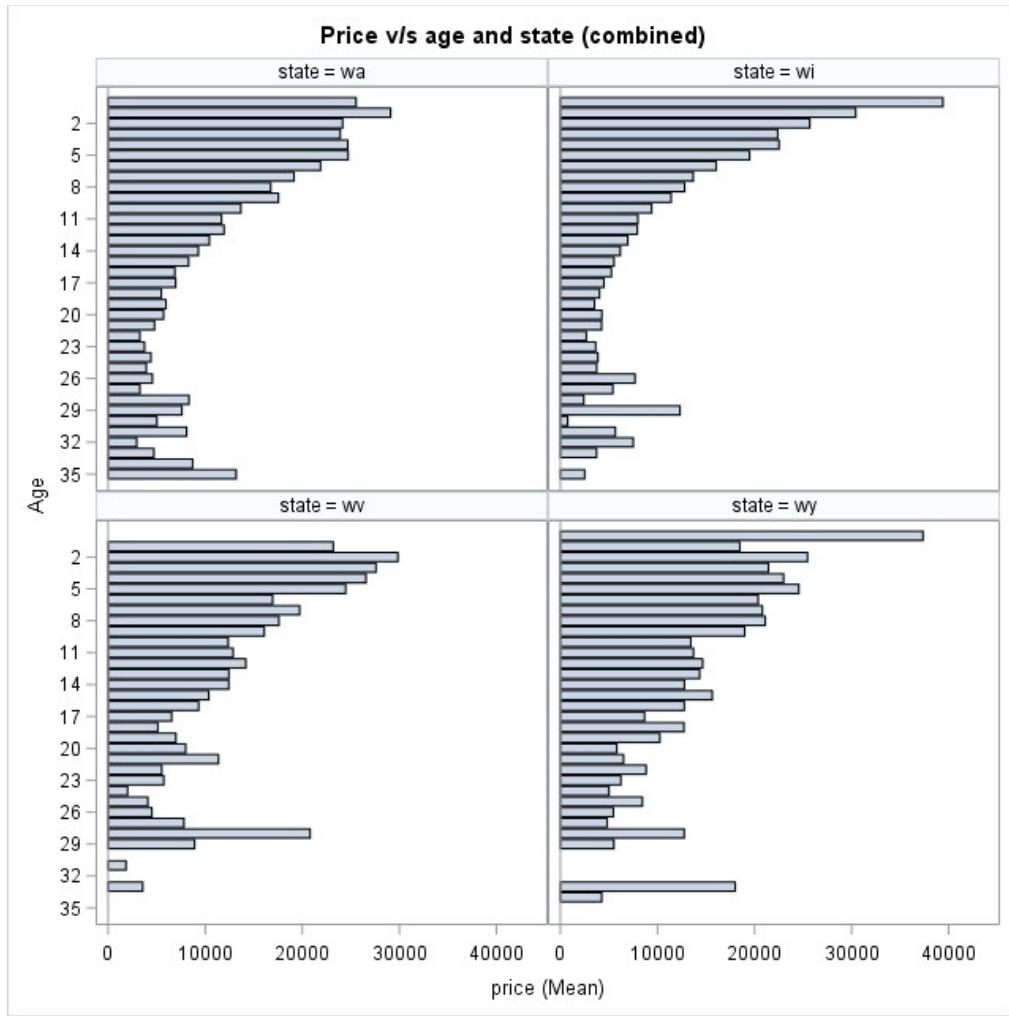


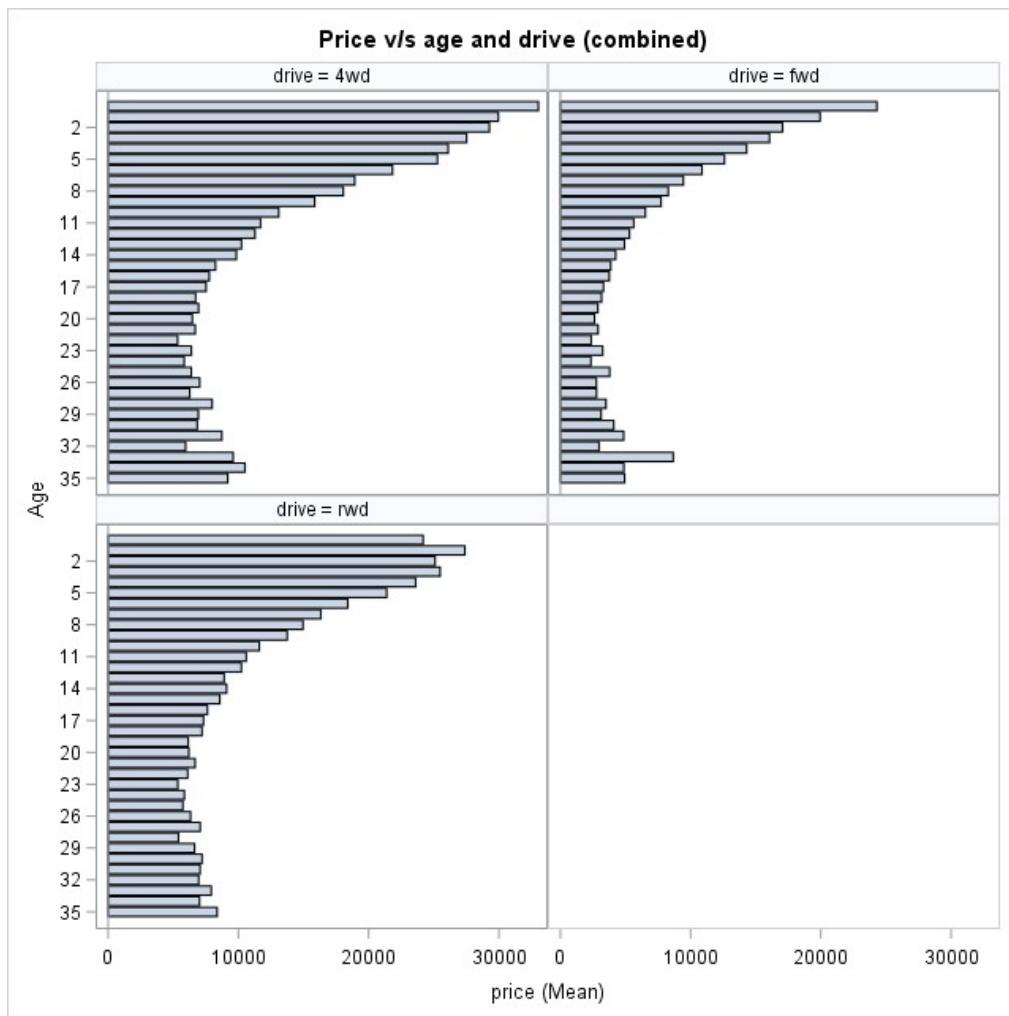


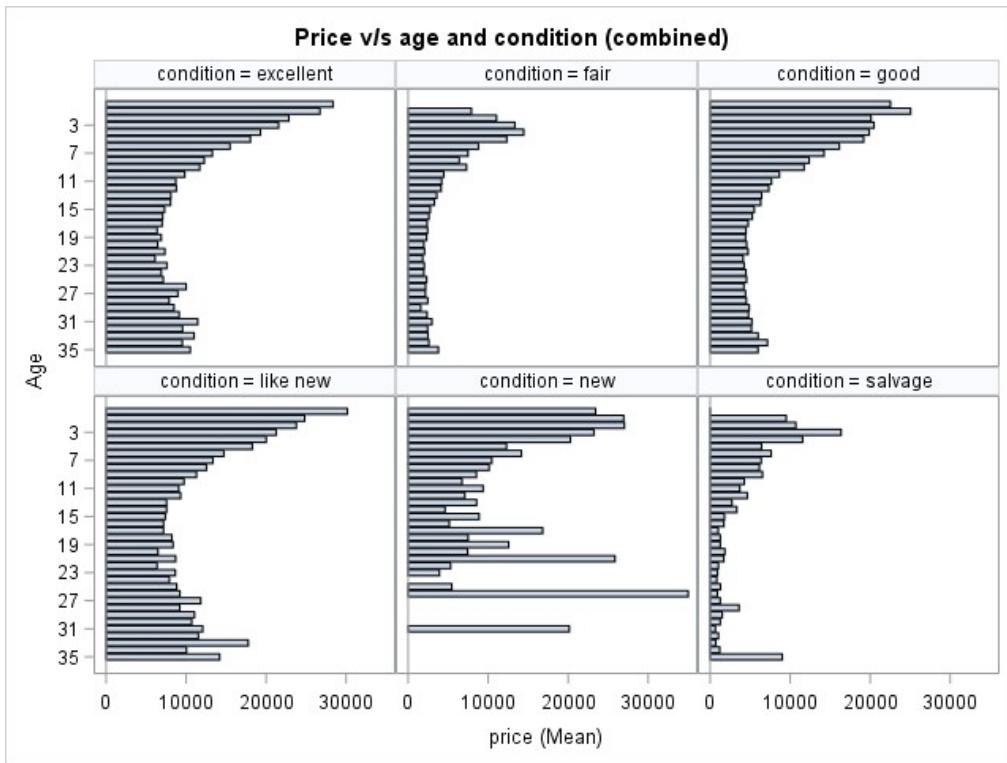


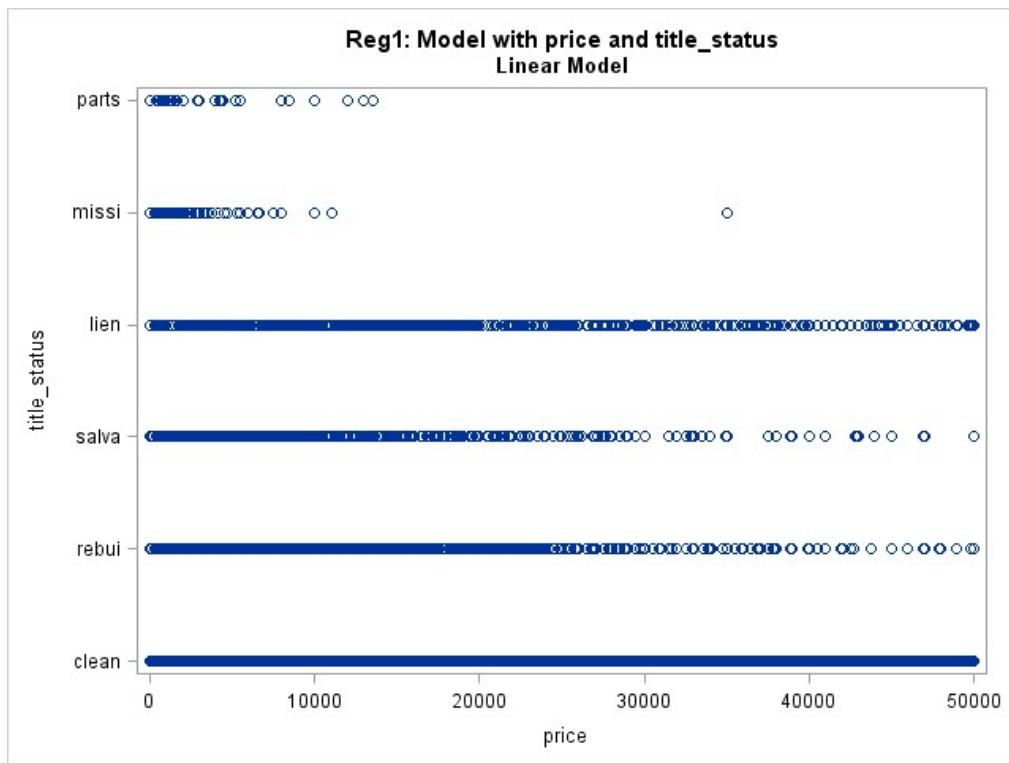


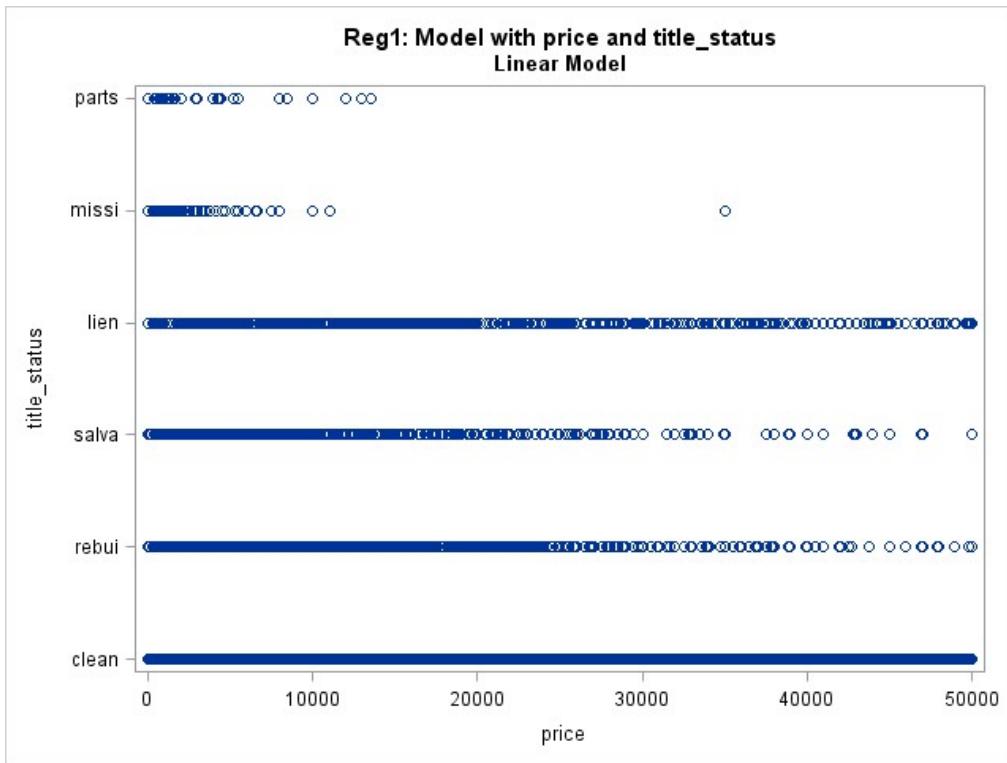


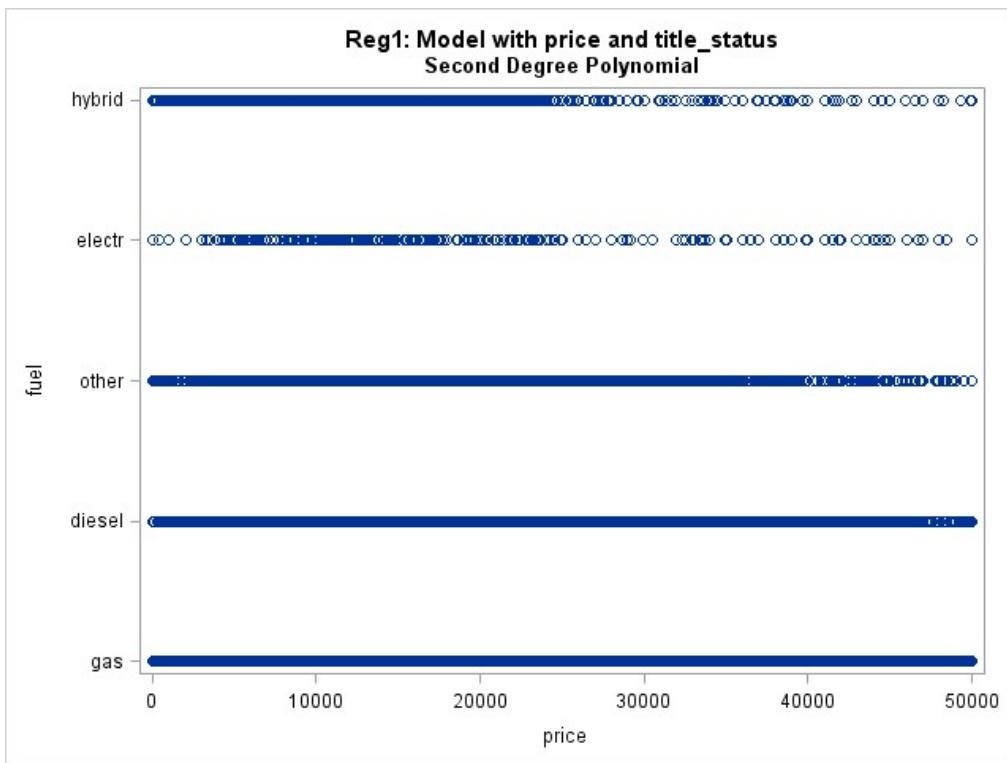












Reg1: Model with price and odometer

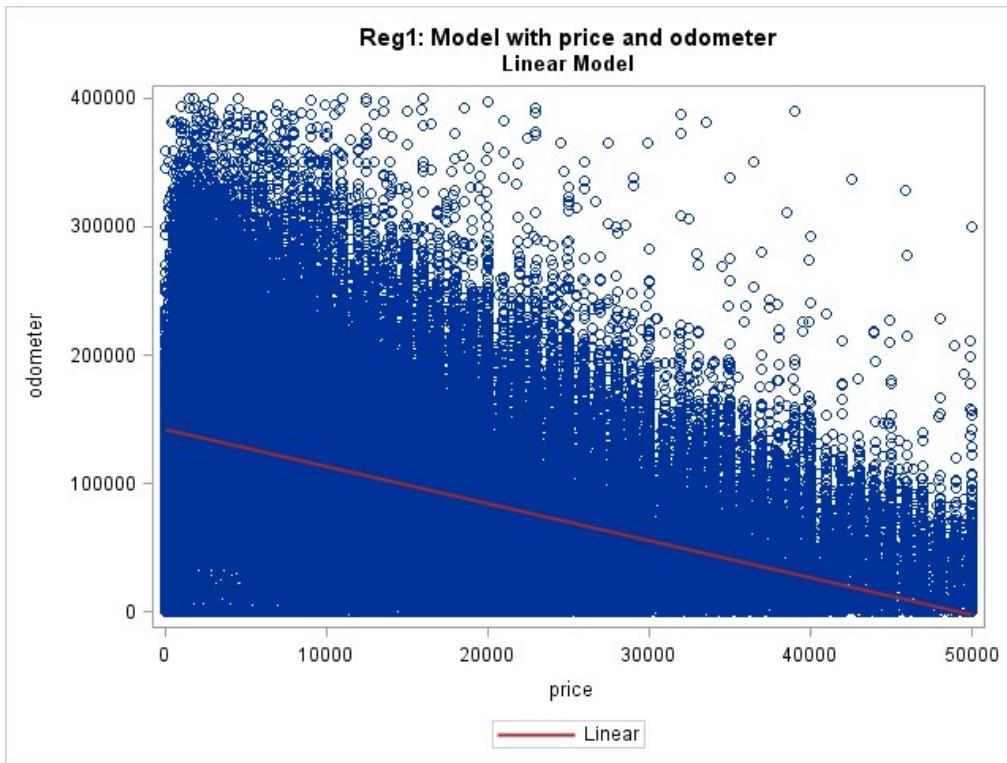
The REG Procedure
Model: MODEL1
Dependent Variable: price

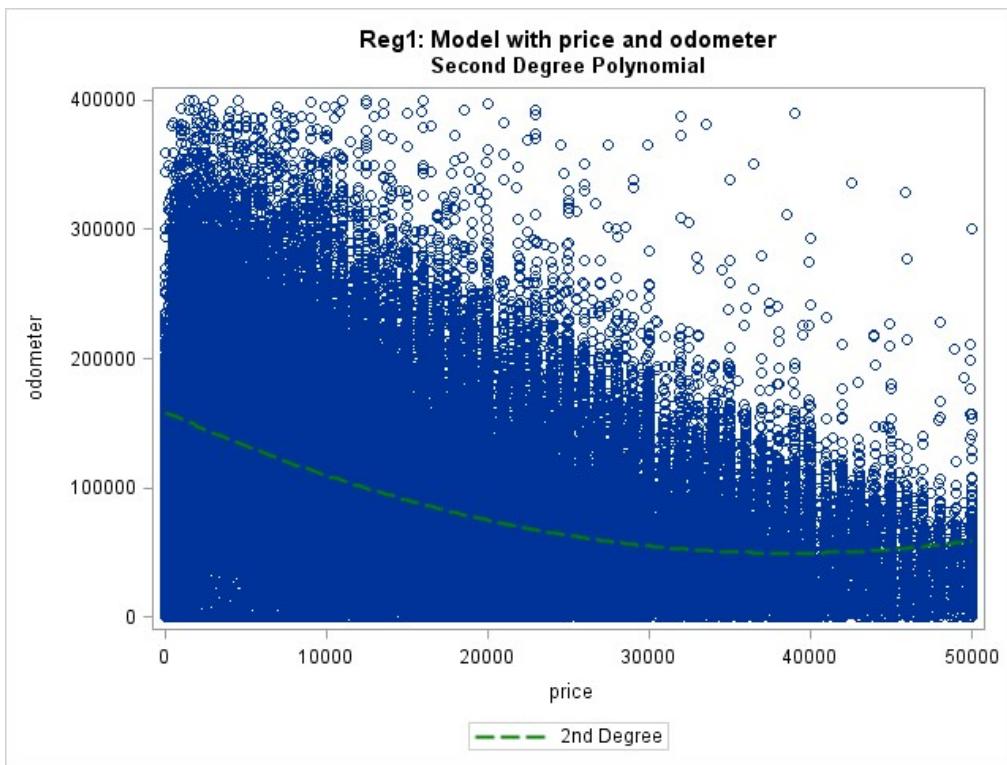
Number of Observations Read	372844
Number of Observations Used	372844

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	8.69911E12	8.69911E12	112365	<.0001
Error	372842	2.886471E13	77418083		
Corrected Total	372843	3.756382E13			

Root MSE	8798.75464	R-Square	0.2316
Dependent Mean	13803	Adj R-Sq	0.2316
Coeff Var	63.74378		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	21995	28.36978	775.30	<.0001
odometer	1	-0.08020	0.00023926	-335.21	<.0001





4th Degree Polynomial using Usedcars Dataset**The GLMSELECT Procedure**

Data Set	PROJECT.USEDCARSPROJECT5
Dependent Variable	price
Selection Method	None

Number of Observations Read	372844
Number of Observations Used	372844

Dimensions	
Number of Effects	5
Number of Parameters	5

4th Degree Polynomial using Usedcars Dataset**The GLMSELECT Procedure**

Least Squares Summary			
Step	Effect Entered	Number Effects In	SBC
0	Intercept	1	6870836.93
1	odometer	2	6772634.60
2	odometer^2	3	6764571.87
3	odometer^3	4	6762923.88
4	odometer^4	5	6758396.08*

* Optimal Value of Criterion

4th Degree Polynomial using Usedcars Dataset

The GLMSELECT Procedure
Least Squares Model (No Selection)

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	9.783508E12	2.445877E12	32826.1	<.0001
Error	372839	2.778031E13	74510216		
Corrected Total	372843	3.756382E13			

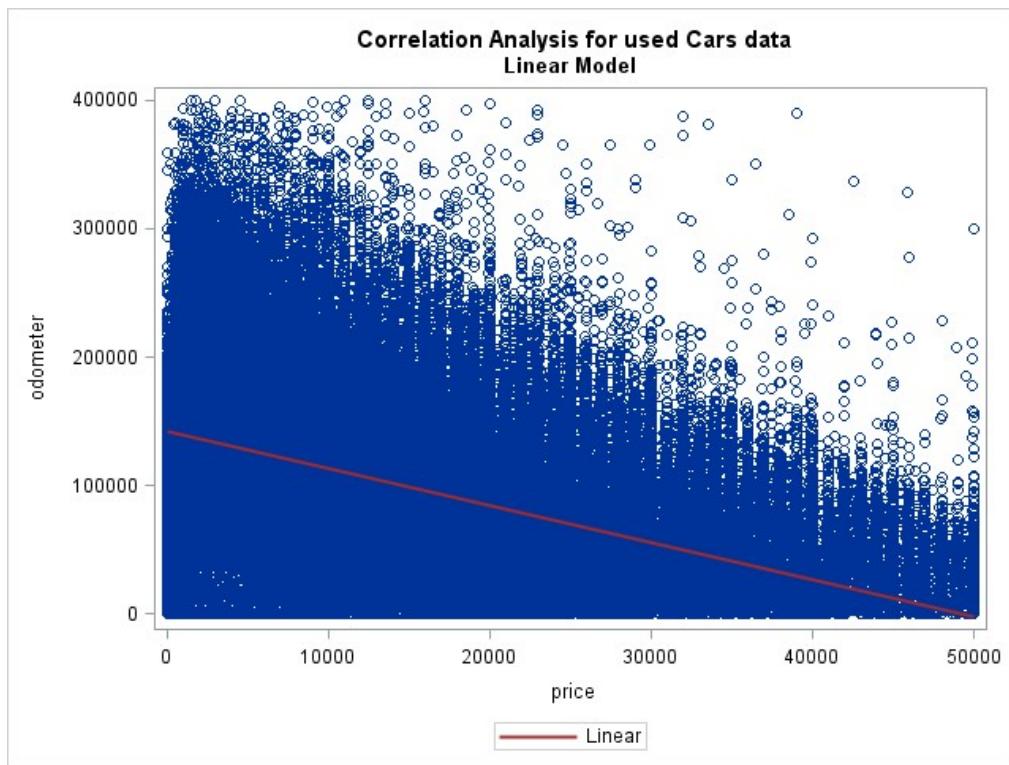
Root MSE	8631.93002
Dependent Mean	13803
R-Square	0.2605
Adj R-Sq	0.2604
AIC	7131188
AICC	7131188
SBC	6758396

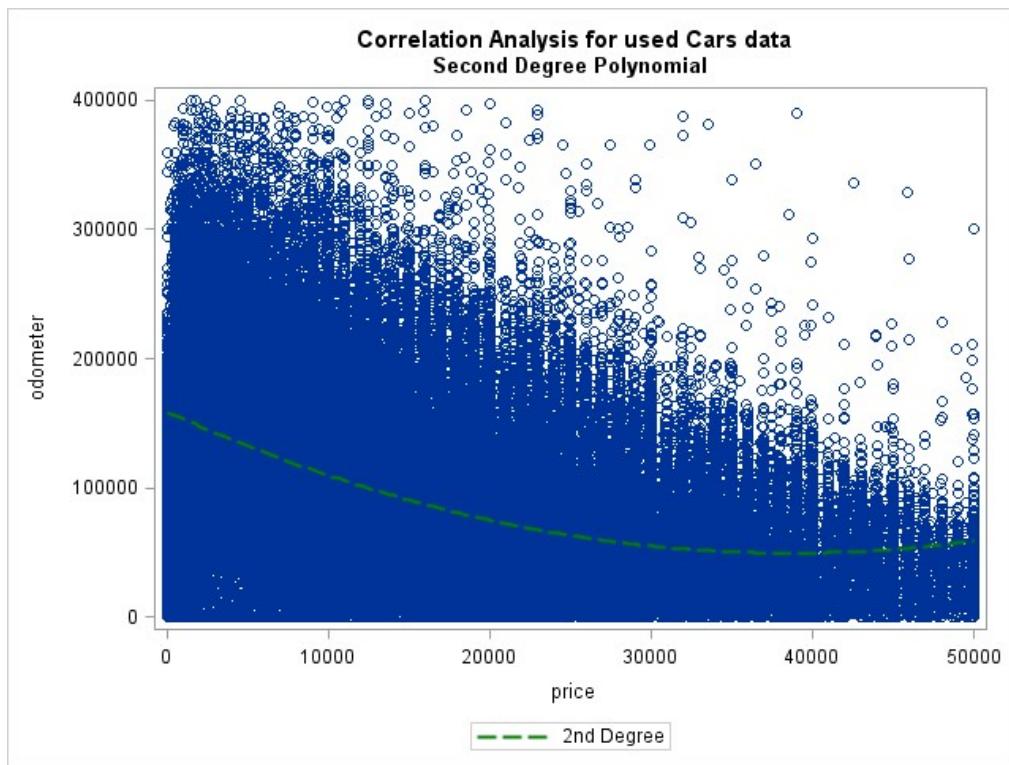
Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Pr > t
Intercept	1	20869	62.334968	334.79	<.0001
odometer	1	0.061693	0.002655	23.23	<.0001
odometer^2	1	-0.000002418	3.4620963E-8	-69.85	<.0001
odometer^3	1	1.22936E-11	1.674602E-13	73.41	<.0001
odometer^4	1	-1.77865E-17	2.631542E-19	-67.59	<.0001

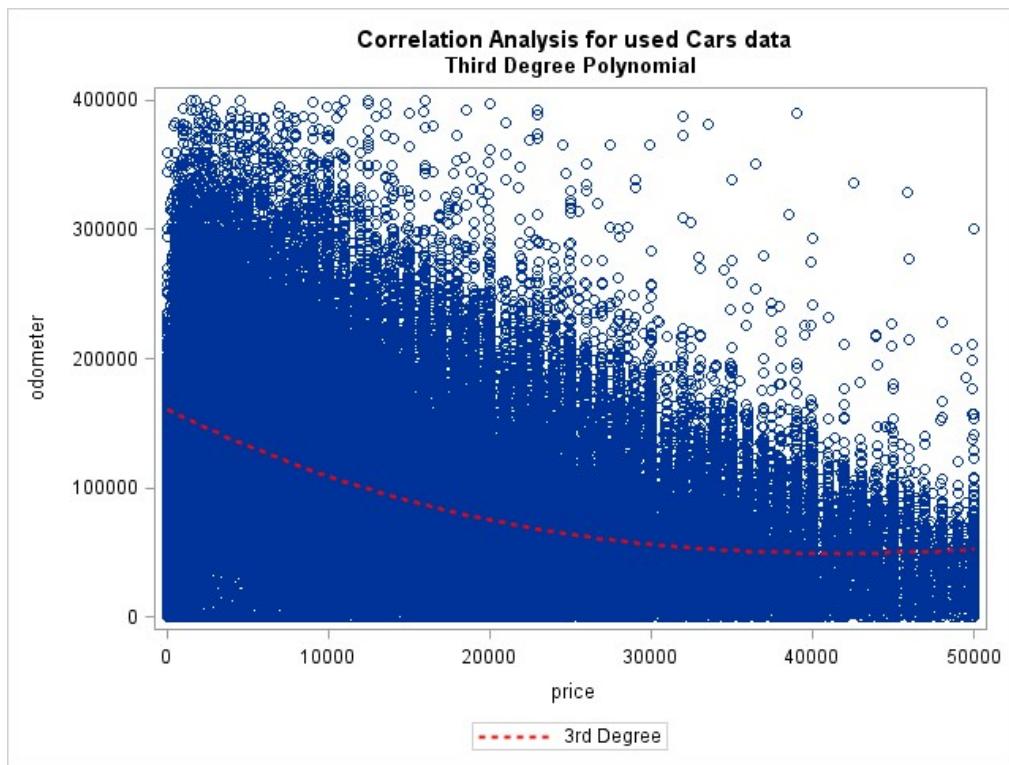
Correlation Analysis for used Cars data**The CORR Procedure**

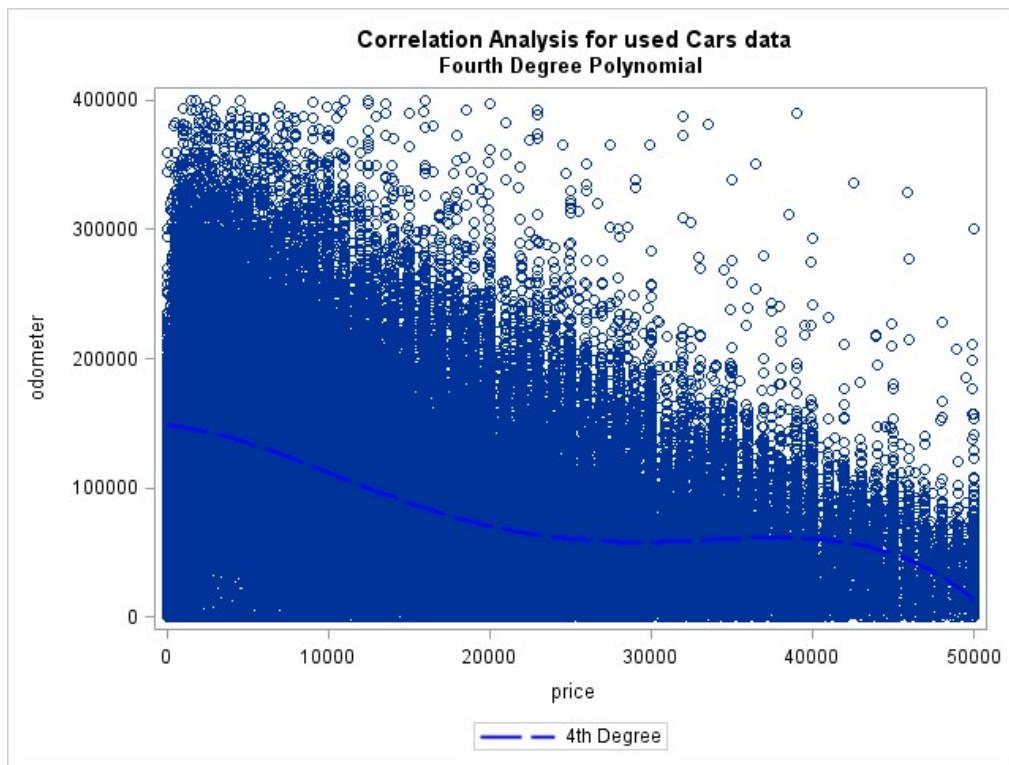
3 Variables:	price	odometer	Age
---------------------	-------	----------	-----

Pearson Correlation Coefficients, N = 372844			
	Prob > r under H0: Rho=0		
	price	odometer	Age
price	1.00000	-0.48123 <.0001	-0.56260 <.0001
odometer	-0.48123 <.0001	1.00000	0.57867 <.0001
Age	-0.56260 <.0001	0.57867 <.0001	1.00000









3rd Degree Polynomial: Paper Data

The REG Procedure
 Model: Cubic_Model
 Dependent Variable: price

Number of Observations Read	372844
Number of Observations Used	372844

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	9.783508E12	2.445877E12	32826.1	<.0001
Error	372839	2.778031E13	74510216		
Lack of Fit	111377	1.743137E13	156507784	3.95	<.0001
Pure Error	261462	1.034895E13	39581075		
Corrected Total	372843	3.756382E13			

Root MSE	8631.93002	R-Square	0.2605
Dependent Mean	13803	Adj R-Sq	0.2604
Coeff Var	62.53520		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	20869	62.33497	334.79	<.0001
odometer	odometer	1	0.06169	0.00266	23.23	<.0001
odometer_2	odometer^2	1	-0.00000242	3.462096E-8	-69.85	<.0001
odometer_3	odometer^3	1	1.22936E-11	1.6746E-13	73.41	<.0001
odometer_4	odometer^4	1	-1.7786E-17	2.63154E-19	-67.59	<.0001

Quadratic Polynomial using UsedCars Dataset

The GLMSELECT Procedure

Data Set	PROJECT.USEDCARSPROJECT5
Dependent Variable	price
Selection Method	None

Number of Observations Read	372844
Number of Observations Used	372844

Dimensions	
Number of Effects	3
Number of Parameters	3

Quadratic Polynomial using UsedCars Dataset

The GLMSELECT Procedure

Least Squares Summary			
Step	Effect Entered	Number Effects In	SBC
0	Intercept	1	6870836.93
1	odometer	2	6772634.60
2	odometer^2	3	6764571.87*

* Optimal Value of Criterion

Quadratic Polynomial using UsedCars Dataset**The GLMSELECT Procedure
Least Squares Model (No Selection)**

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	9.317579E12	4.658789E12	61494.5	<.0001
Error	372841	2.824624E13	75759491		
Corrected Total	372843	3.756382E13			

Root MSE	8703.99280
Dependent Mean	13803
R-Square	0.2480
Adj R-Sq	0.2480
AIC	7137385
AICC	7137385
SBC	6764572

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Pr > t
Intercept	1	24591	40.160284	612.31	<.0001
odometer	1	-0.140733	0.000711	-198.07	<.0001
odometer^2	1	0.000000255	2.823598E-9	90.35	<.0001

Diagnostic plots: UsedCars Data

The REG Procedure
Model: Cubic_Model
Dependent Variable: price

Number of Observations Read	372844
Number of Observations Used	372844

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	9.317579E12	4.658789E12	61494.5	<.0001
Error	372841	2.824624E13	75759491		
Lack of Fit	111379	1.78973E13	160688256	4.06	<.0001
Pure Error	261462	1.034895E13	39581075		
Corrected Total	372843	3.756382E13			

Root MSE	8703.99280	R-Square	0.2480
Dependent Mean	13803	Adj R-Sq	0.2480
Coeff Var	63.05727		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	24591	40.16028	612.31	<.0001
odometer	odometer	1	-0.14073	0.00071051	-198.07	<.0001
odometer_2	odometer^2	1	2.551192E-7	2.823598E-9	90.35	<.0001

Diagnostic plots: UsedCars Data

The CONTENTS Procedure

Data Set Name	PROJECT.USEDCARSPROJECT5	Observations	372844
Member Type	DATA	Variables	18
Engine	V9	Indexes	0
Created	04/23/2020 17:58:59	Observation Length	152
Last Modified	04/23/2020 17:58:59	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	868
First Data Page	1
Max Obs per Page	430
Obs in First Data Page	411
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	E:\Users\bsp190005\Documents\My SAS Files\Codes\Projectusedcarsproject5.sas7bdat
Release Created	9.0401M4
Host Created	X64_SR12R2
Owner Name	CAMPUS\bsp190005
File Size	54MB
File Size (bytes)	56950784

Alphabetic List of Variables and Attributes						
#	Variable	Type	Len	Format	Informat	
17	Age	Num	8			
6	condition	Char	9	\$9.	\$9.	
7	cylinders	Char	11	\$11.	\$11.	
12	drive	Char	3	\$3.	\$3.	
8	fuel	Char	6	\$6.	\$6.	
18	logAge	Num	8			
16	logPrice	Num	8			
4	manufacturer	Char	13	\$13.	\$13.	
5	model	Char	11	\$11.	\$11.	
9	odometer	Num	8	BEST12.	BEST32.	
14	paint_color	Char	6	\$6.	\$6.	
2	price	Num	8	BEST12.	BEST32.	
1	region	Char	14	\$14.	\$14.	

15	state	Char	2	\$2.	\$2.
10	title_status	Char	5	\$5.	\$5.
11	transmission	Char	9	\$9.	\$9.
13	type	Char	9	\$9.	\$9.
3	year	Num	8	BEST12.	BEST32.

Price Model - Generate Diagnostic Plots

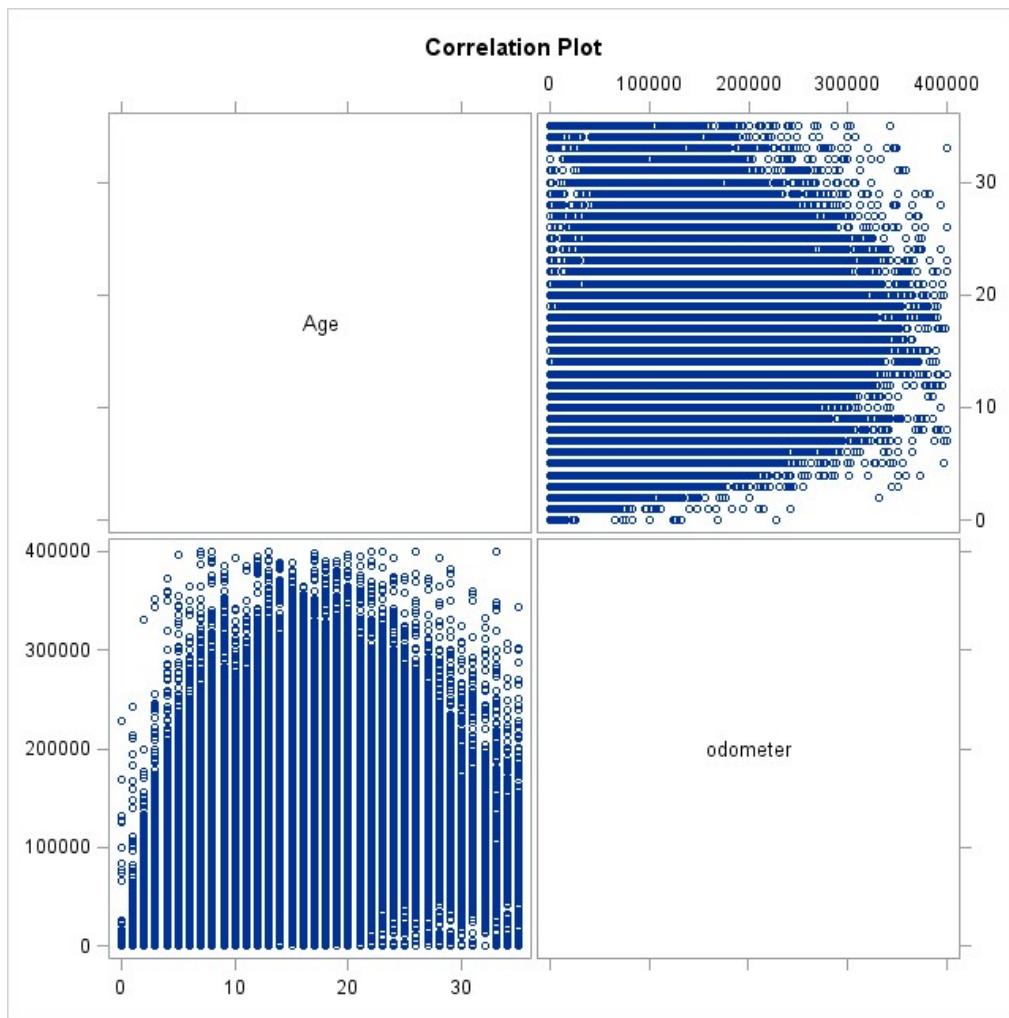
The REG Procedure
Model: continuous
Dependent Variable: price

Number of Observations Read	372844
Number of Observations Used	372844

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1.325811E13	6.629053E12	101687	<.0001
Error	372841	2.430572E13	65190567		
Corrected Total	372843	3.756382E13			

Root MSE	8074.06759	R-Square	0.3529
Dependent Mean	13803	Adj R-Sq	0.3529
Coeff Var	58.49369		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	24887	28.23637	881.37	<.0001
Age	1	-759.54497	2.87218	-264.45	<.0001
odometer	1	-0.03901	0.00026921	-144.89	<.0001



Collinearity Diagnostics

The REG Procedure
Model: wSCORE
Dependent Variable: price

Number of Observations Read	372844
Number of Observations Used	372844

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1.325811E13	6.629053E12	101687	<.0001
Error	372841	2.430572E13	65190567		
Corrected Total	372843	3.756382E13			

Root MSE	8074.06759	R-Square	0.3529
Dependent Mean	13803	Adj R-Sq	0.3529
Coeff Var	58.49369		

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
Intercept	1	24887	28.23637	881.37	<.0001	0
Age	1	-759.54497	2.87218	-264.45	<.0001	1.50343
odometer	1	-0.03901	0.00026921	-144.89	<.0001	1.50343

Quadratic Polynomial using Cars Dataset

The GLMSELECT Procedure

Data Set	PROJECT.USEDCARSPROJECT5
Dependent Variable	price
Selection Method	None

Number of Observations Read	372844
Number of Observations Used	372844

Dimensions	
Number of Effects	5
Number of Parameters	5

Quadratic Polynomial using Cars Dataset**The GLMSELECT Procedure**

Least Squares Summary			
Step	Effect Entered	Number Effects In	SBC
0	Intercept	1	6870836.93
1	odometer	2	6772634.60
2	odometer^2	3	6764571.87
3	odometer^3	4	6762923.88
4	odometer^4	5	6758396.08*

* Optimal Value of Criterion

Quadratic Polynomial using Cars Dataset**The GLMSELECT Procedure
Least Squares Model (No Selection)**

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	9.783508E12	2.445877E12	32826.1	<.0001
Error	372839	2.778031E13	74510216		
Corrected Total	372843	3.756382E13			

Root MSE	8631.93002
Dependent Mean	13803
R-Square	0.2605
Adj R-Sq	0.2604
AIC	7131188
AICC	7131188
SBC	6758396

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Pr > t
Intercept	1	20869	62.334968	334.79	<.0001
odometer	1	0.061693	0.002655	23.23	<.0001
odometer^2	1	-0.000002418	3.4620963E-8	-69.85	<.0001
odometer^3	1	1.22936E-11	1.674602E-13	73.41	<.0001
odometer^4	1	-1.77865E-17	2.631542E-19	-67.59	<.0001

Diagnostic plots: Cars Data

The REG Procedure
 Model: Cubic_Model
 Dependent Variable: price

Number of Observations Read	372844
Number of Observations Used	372844

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	9.783508E12	2.445877E12	32826.1	<.0001
Error	372839	2.778031E13	74510216		
Lack of Fit	111377	1.743137E13	156507784	3.95	<.0001
Pure Error	261462	1.034895E13	39581075		
Corrected Total	372843	3.756382E13			

Root MSE	8631.93002	R-Square	0.2605
Dependent Mean	13803	Adj R-Sq	0.2604
Coeff Var	62.53520		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	20869	62.33497	334.79	<.0001
odometer	odometer	1	0.06169	0.00266	23.23	<.0001
odometer_2	odometer^2	1	-0.00000242	3.462096E-8	-69.85	<.0001
odometer_3	odometer^3	1	1.22936E-11	1.6746E-13	73.41	<.0001
odometer_4	odometer^4	1	-1.7786E-17	2.63154E-19	-67.59	<.0001

Quadratic Polynomial using Cars Dataset

The GLMSELECT Procedure

Data Set	PROJECT.USEDCARSPROJECT5
Dependent Variable	price
Selection Method	None

Number of Observations Read	372844
Number of Observations Used	372844

Dimensions	
Number of Effects	5
Number of Parameters	5

Quadratic Polynomial using Cars Dataset**The GLMSELECT Procedure**

Least Squares Summary			
Step	Effect Entered	Number Effects In	SBC
0	Intercept	1	6870836.93
1	odometer	2	6772634.60
2	odometer^2	3	6764571.87
3	odometer^3	4	6762923.88
4	odometer^4	5	6758396.08*

* Optimal Value of Criterion

Quadratic Polynomial using Cars Dataset**The GLMSELECT Procedure
Least Squares Model (No Selection)**

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	9.783508E12	2.445877E12	32826.1	<.0001
Error	372839	2.778031E13	74510216		
Corrected Total	372843	3.756382E13			

Root MSE	8631.93002
Dependent Mean	13803
R-Square	0.2605
Adj R-Sq	0.2604
AIC	7131188
AICC	7131188
SBC	6758396

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Pr > t
Intercept	1	20869	62.334968	334.79	<.0001
odometer	1	0.061693	0.002655	23.23	<.0001
odometer^2	1	-0.000002418	3.4620963E-8	-69.85	<.0001
odometer^3	1	1.22936E-11	1.674602E-13	73.41	<.0001
odometer^4	1	-1.77865E-17	2.631542E-19	-67.59	<.0001

Diagnostic plots: Cars Data

The REG Procedure
 Model: Cubic_Model
 Dependent Variable: price

Number of Observations Read	372844
Number of Observations Used	372844

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	9.783508E12	2.445877E12	32826.1	<.0001
Error	372839	2.778031E13	74510216		
Lack of Fit	111377	1.743137E13	156507784	3.95	<.0001
Pure Error	261462	1.034895E13	39581075		
Corrected Total	372843	3.756382E13			

Root MSE	8631.93002	R-Square	0.2605
Dependent Mean	13803	Adj R-Sq	0.2604
Coeff Var	62.53520		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	20869	62.33497	334.79	<.0001
odometer	odometer	1	0.06169	0.00266	23.23	<.0001
odometer_2	odometer^2	1	-0.00000242	3.462096E-8	-69.85	<.0001
odometer_3	odometer^3	1	1.22936E-11	1.6746E-13	73.41	<.0001
odometer_4	odometer^4	1	-1.7786E-17	2.63154E-19	-67.59	<.0001