

# Exploratory Data Analysis of:

`/Applications/Stata/ado/base/a/auto.dta`

billy

October 29, 2015

# Contents

<b>1</b>	<b>Graphs</b>	<b>4</b>
1.1	Univariate Distributions . . . . .	5
1.1.1	Categorical Variables . . . . .	6
1.1.2	Continuous Variables . . . . .	7
1.2	Bivariate Distributions . . . . .	8
1.2.1	Box Plots . . . . .	9
	Groups of rep78 . . . . .	9
	Groups of headroom . . . . .	17
	Groups of foreign . . . . .	25
<b>2</b>	<b>Descriptive Statistics</b>	<b>33</b>
2.1	Categorical Variables . . . . .	34
2.2	Continuous Variables . . . . .	41
2.3	Conditional Descriptive Statistics . . . . .	43
2.3.1	Tables by groups of rep78 . . . . .	44
2.3.2	Tables by groups of headroom . . . . .	47
2.3.3	Tables by groups of foreign . . . . .	50

# List of Figures

1.1	Box Plot of Price by Repair Record 1978 . . . . .	9
1.2	Box Plot of Mileage (mpg) by Repair Record 1978 . . . . .	10
1.3	Box Plot of Trunk space (cu. ft.) by Repair Record 1978 . . . . .	11
1.4	Box Plot of Weight (lbs.) by Repair Record 1978 . . . . .	12
1.5	Box Plot of Length (in.) by Repair Record 1978 . . . . .	13
1.6	Box Plot of Turn Circle (ft.) by Repair Record 1978 . . . . .	14
1.7	Box Plot of Displacement (cu. in.) by Repair Record 1978 . . . . .	15
1.8	Box Plot of Gear Ratio by Repair Record 1978 . . . . .	16
1.9	Box Plot of Price by Headroom (in.) . . . . .	17
1.10	Box Plot of Mileage (mpg) by Headroom (in.) . . . . .	18
1.11	Box Plot of Trunk space (cu. ft.) by Headroom (in.) . . . . .	19
1.12	Box Plot of Weight (lbs.) by Headroom (in.) . . . . .	20
1.13	Box Plot of Length (in.) by Headroom (in.) . . . . .	21
1.14	Box Plot of Turn Circle (ft.) by Headroom (in.) . . . . .	22
1.15	Box Plot of Displacement (cu. in.) by Headroom (in.) . . . . .	23
1.16	Box Plot of Gear Ratio by Headroom (in.) . . . . .	24
1.17	Box Plot of Price by Car type . . . . .	25
1.18	Box Plot of Mileage (mpg) by Car type . . . . .	26
1.19	Box Plot of Trunk space (cu. ft.) by Car type . . . . .	27
1.20	Box Plot of Weight (lbs.) by Car type . . . . .	28
1.21	Box Plot of Length (in.) by Car type . . . . .	29
1.22	Box Plot of Turn Circle (ft.) by Car type . . . . .	30
1.23	Box Plot of Displacement (cu. in.) by Car type . . . . .	31
1.24	Box Plot of Gear Ratio by Car type . . . . .	32

# List of Tables

2.1	Distribution of Repair Record 1978 . . . . .	35
2.2	Distribution of Headroom (in.) . . . . .	36
2.3	Distribution of Car type . . . . .	37
2.4	Frequency of Headroom (in.) by Car type . . . . .	38
2.5	Overall % of Headroom (in.) by Car type . . . . .	38
2.6	Column % of Headroom (in.) by Car type . . . . .	38
2.7	Row % of Headroom (in.) by Car type . . . . .	38
2.8	Frequency of Repair Record 1978 by Car type . . . . .	39
2.9	Overall % of Repair Record 1978 by Car type . . . . .	39
2.10	Column % of Repair Record 1978 by Car type . . . . .	39
2.11	Row % of Repair Record 1978 by Car type . . . . .	39
2.12	Frequency of Repair Record 1978 by Headroom (in.) . . . . .	40
2.13	Overall % of Repair Record 1978 by Headroom (in.) . . . . .	40
2.14	Column % of Repair Record 1978 by Headroom (in.) . . . . .	40
2.15	Row % of Repair Record 1978 by Headroom (in.) . . . . .	40
2.16	Descriptive Statistics of Continuous Variables . . . . .	42
2.17	Higher Order Moment Conditions . . . . .	42
2.18	Order Statistics . . . . .	42
2.19	Averages by groups of Repair Record 1978 . . . . .	45
2.20	Standard Deviations by groups of Repair Record 1978 . . . . .	46
2.21	Averages by groups of Headroom (in.) . . . . .	48
2.22	Standard Deviations by groups of Headroom (in.) . . . . .	49
2.23	Averages by groups of Car type . . . . .	51
2.24	Standard Deviations by groups of Car type . . . . .	52

# Chapter 1.

## Graphs

## Univariate Distributions

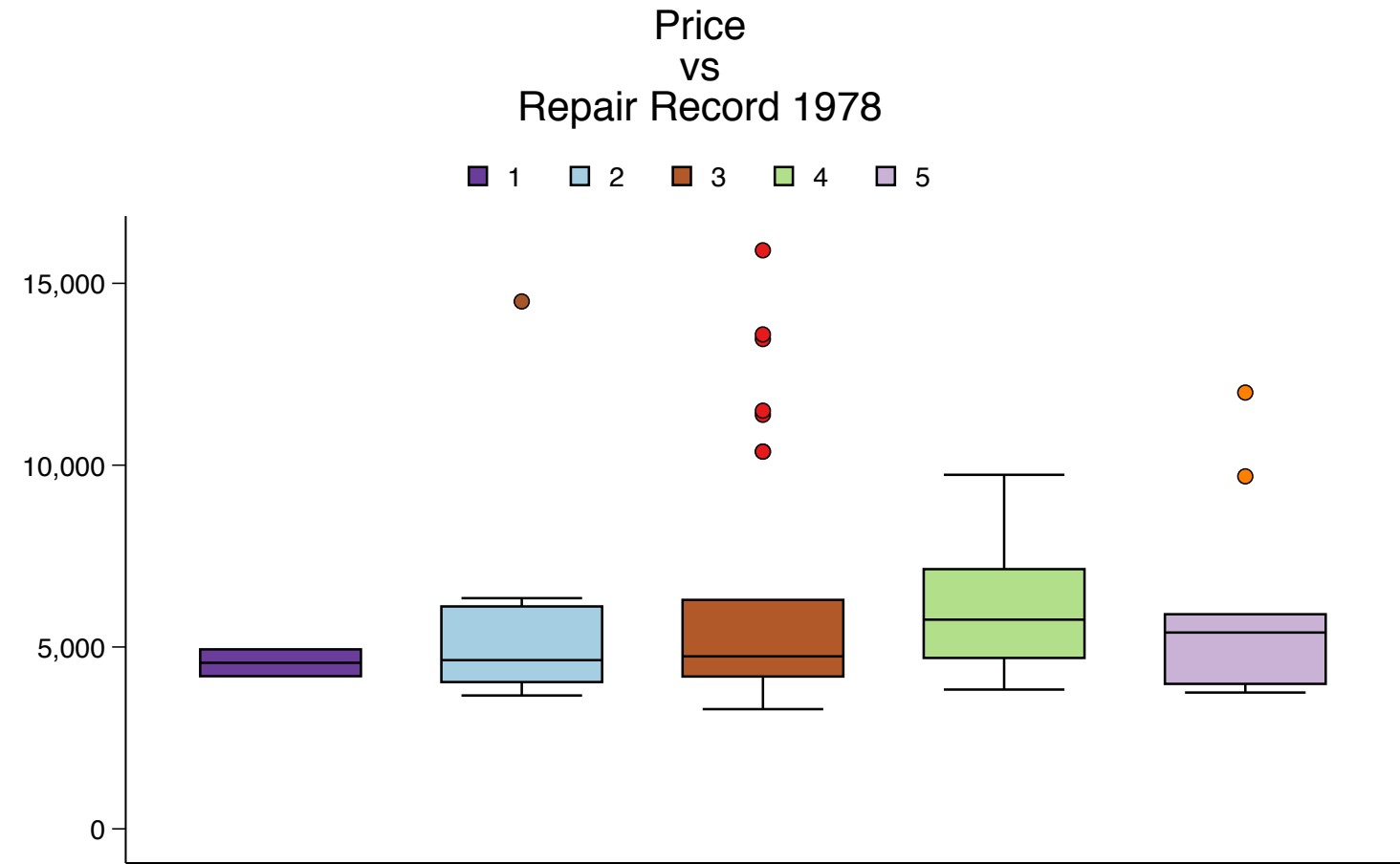
Categorical Variables

Continuous Variables



## Bivariate Distributions

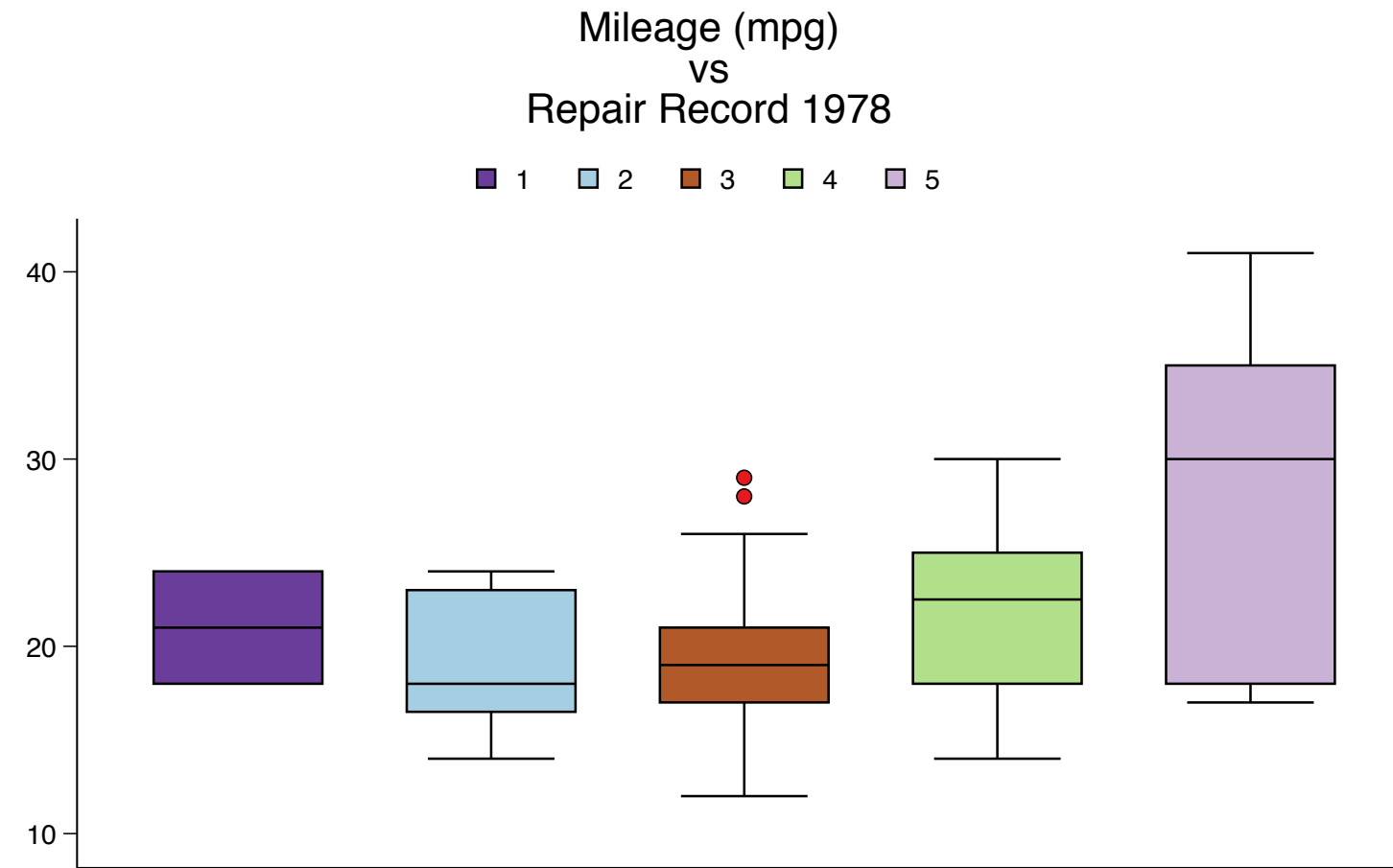
Figure 1.1: Box Plot of Price by Repair Record 1978



Created on: 29 Oct 2015 at: 15:15:34

Groups of rep78.

Figure 1.2: Box Plot of Mileage (mpg) by Repair Record 1978

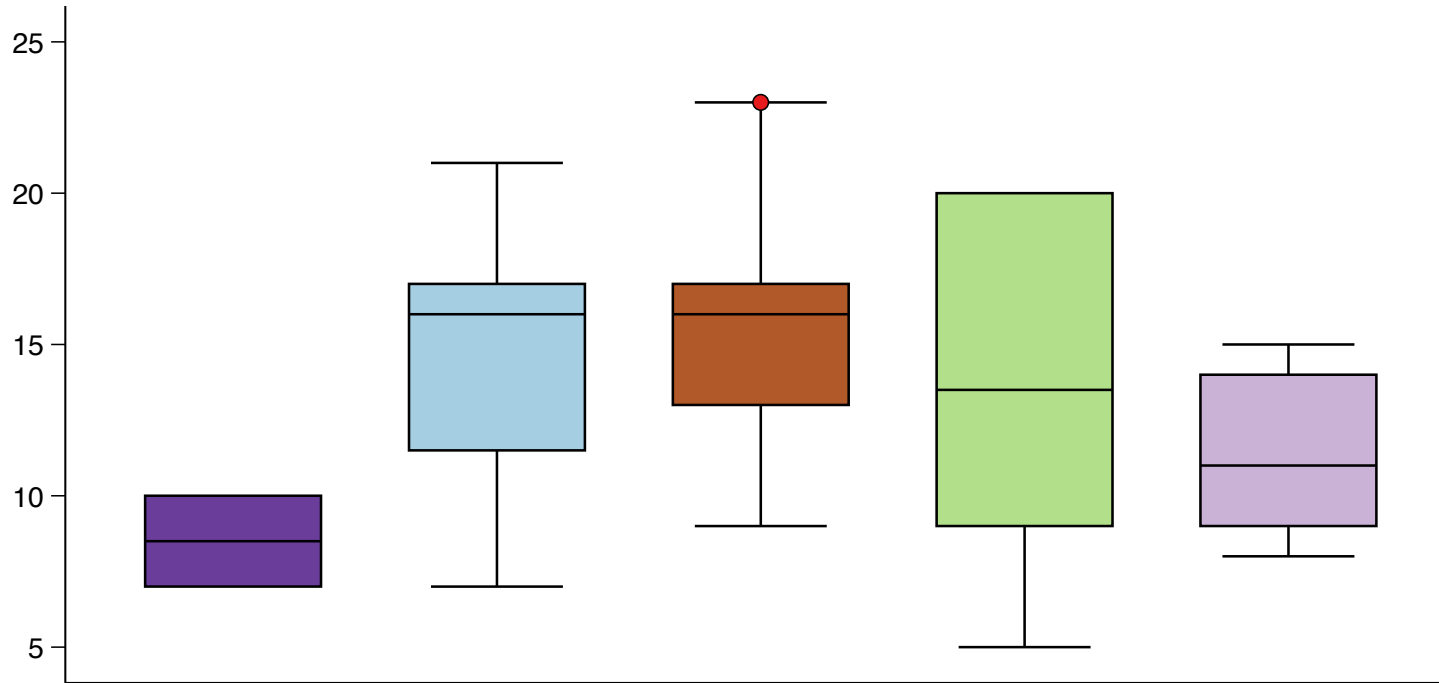


Created on: 29 Oct 2015 at: 15:15:37

Figure 1.3: Box Plot of Trunk space (cu. ft.) by Repair Record 1978

Trunk space (cu. ft.)  
vs  
Repair Record 1978

1 2 3 4 5

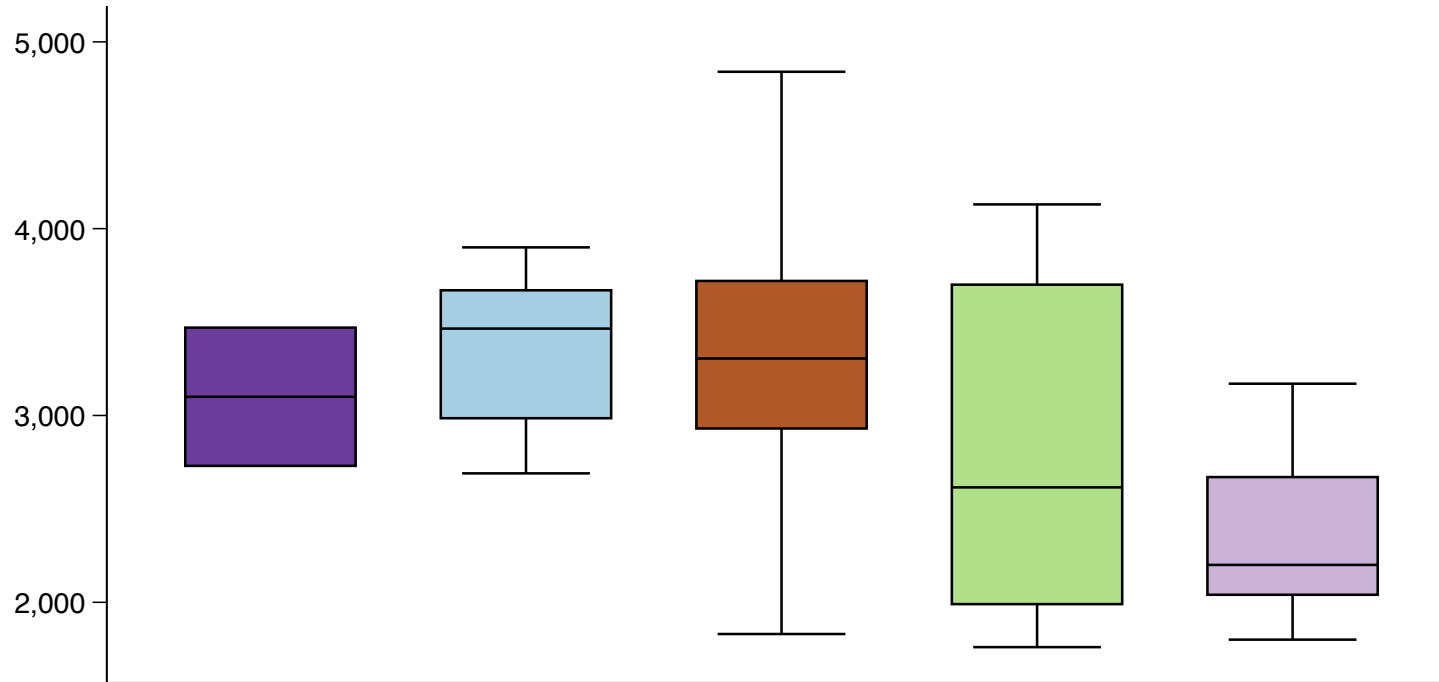


Created on: 29 Oct 2015 at: 15:15:39

Figure 1.4: Box Plot of Weight (lbs.) by Repair Record 1978

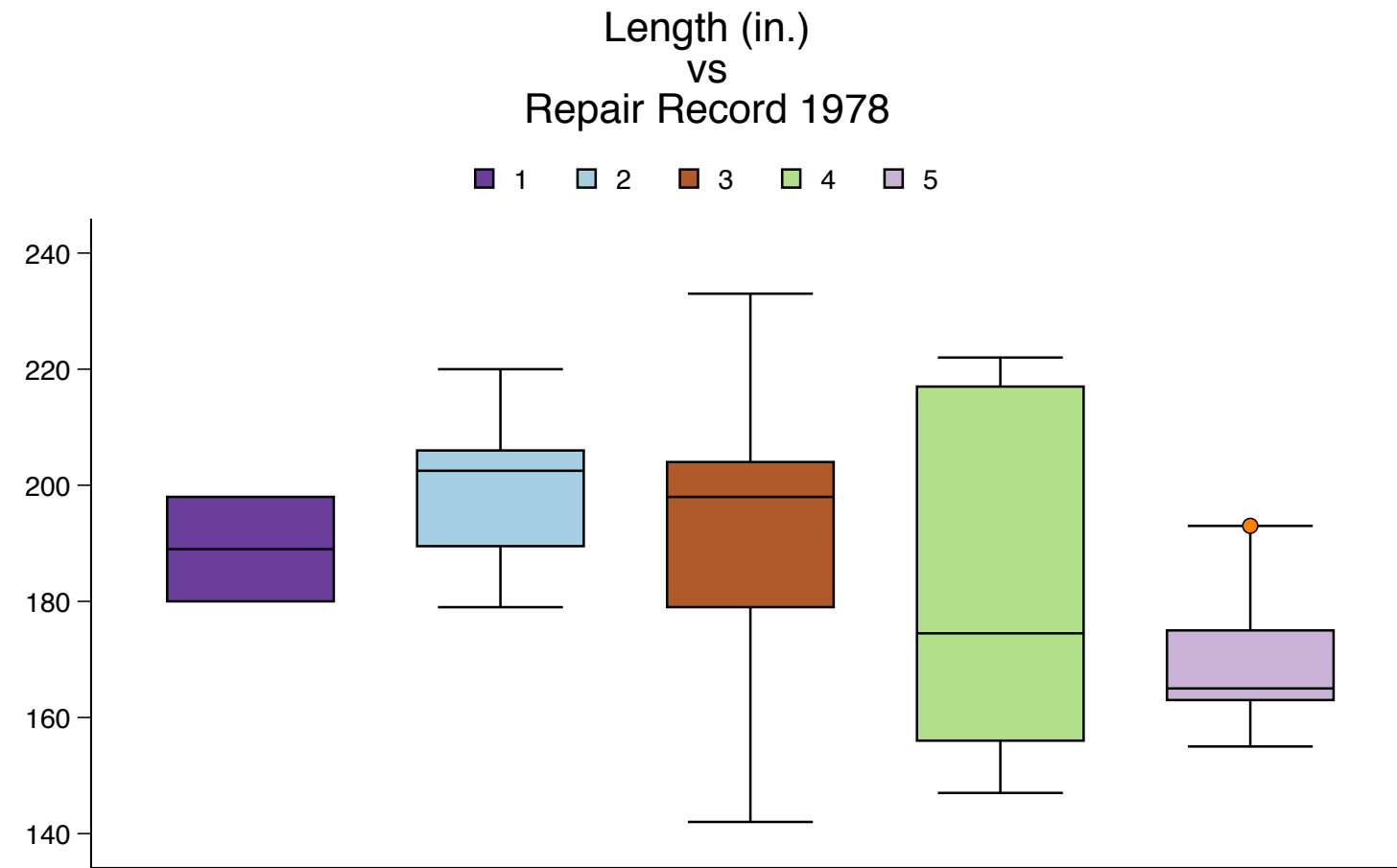
Weight (lbs.)  
vs  
Repair Record 1978

1 2 3 4 5



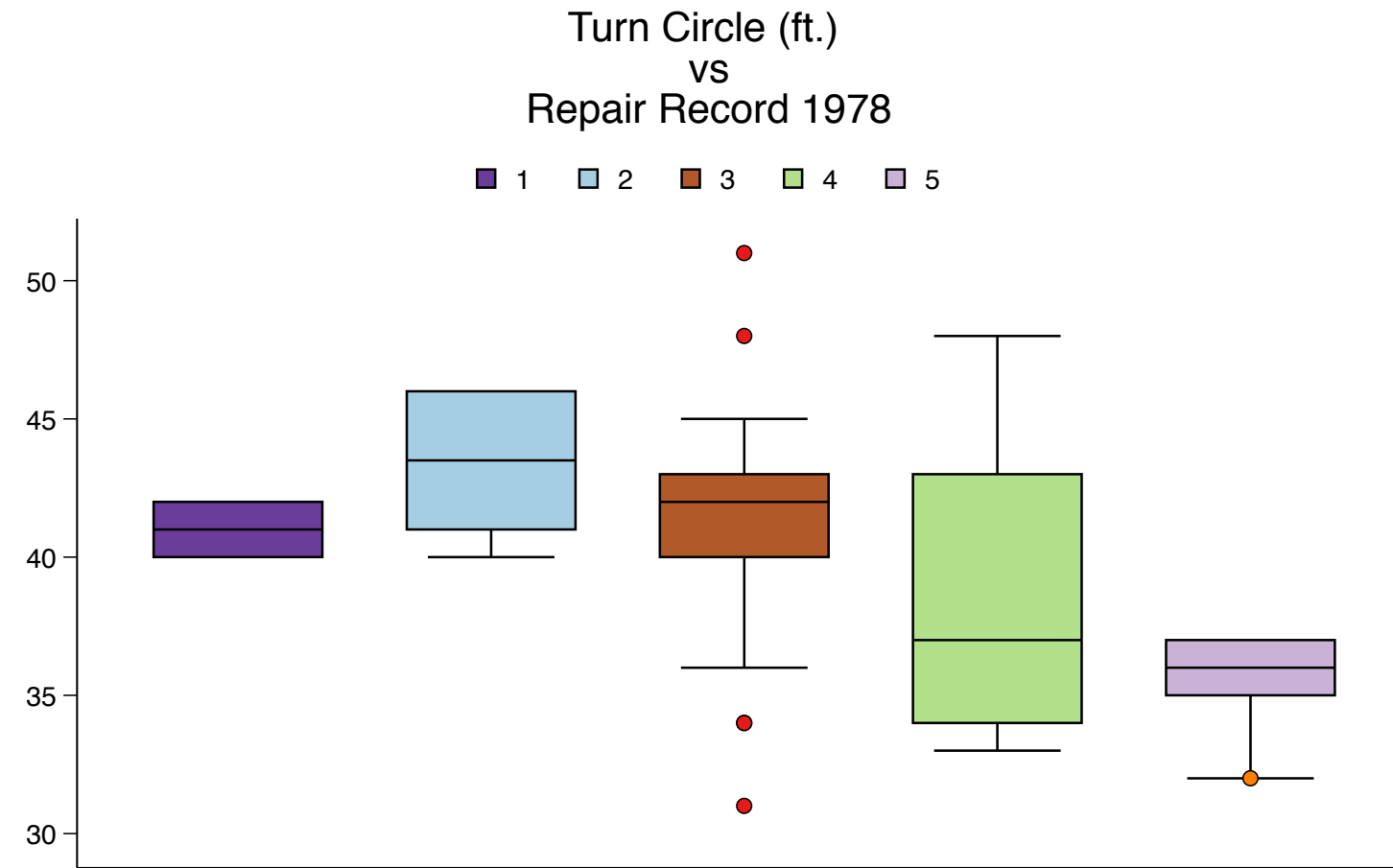
Created on: 29 Oct 2015 at: 15:15:42

Figure 1.5: Box Plot of Length (in.) by Repair Record 1978



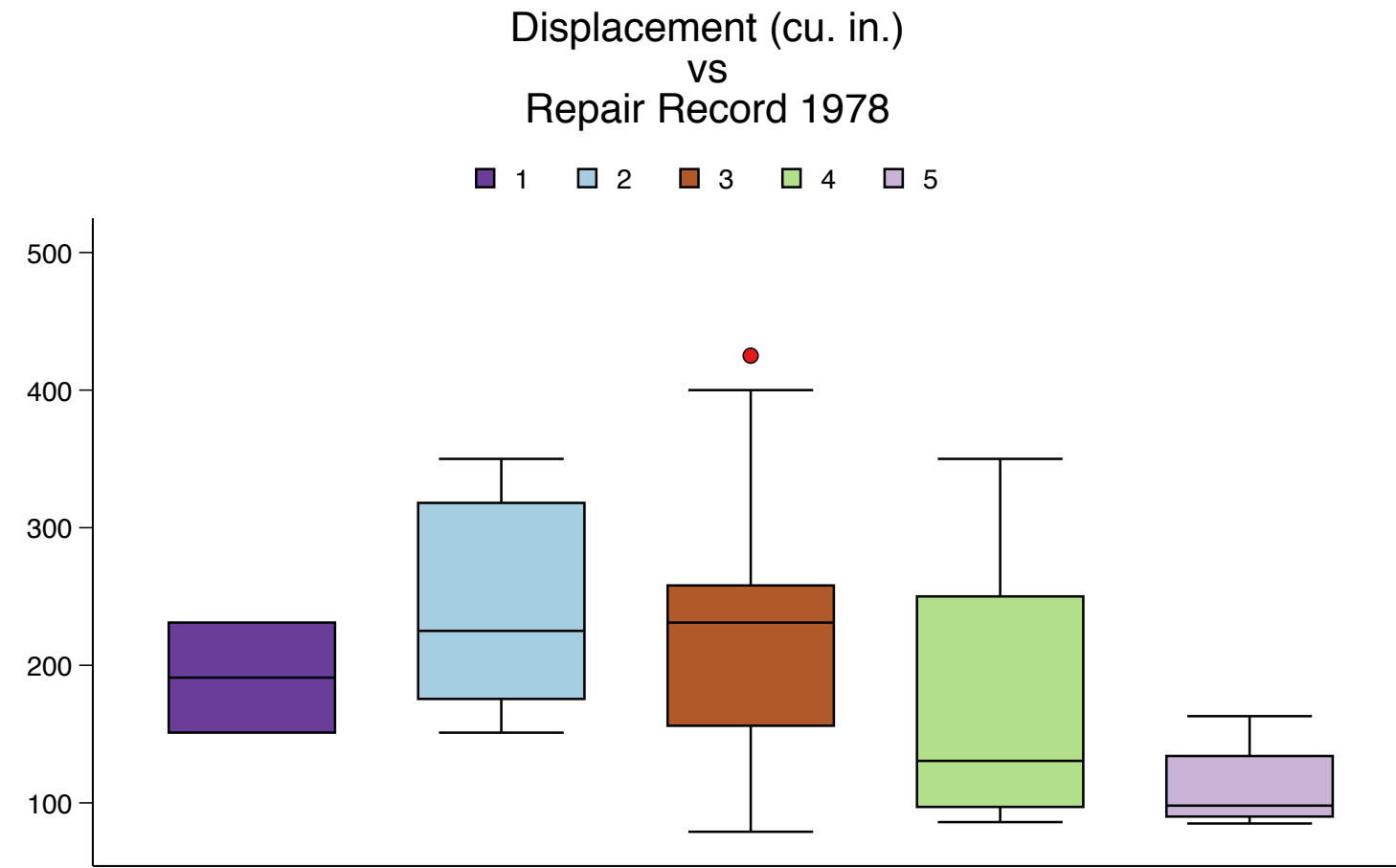
Created on: 29 Oct 2015 at: 15:15:44

Figure 1.6: Box Plot of Turn Circle (ft.) by Repair Record 1978



Created on: 29 Oct 2015 at: 15:15:47

Figure 1.7: Box Plot of Displacement (cu. in.) by Repair Record 1978



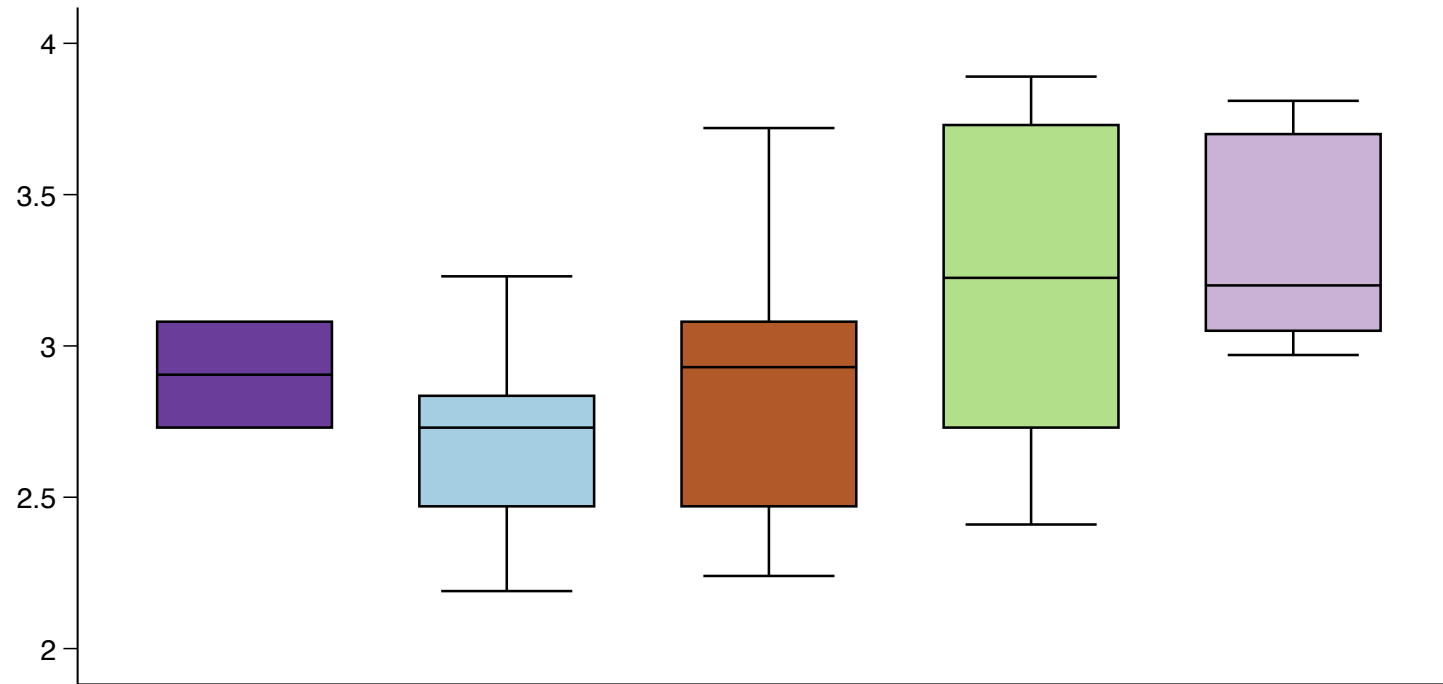
Created on: 29 Oct 2015 at: 15:15:49



Figure 1.8: Box Plot of Gear Ratio by Repair Record 1978

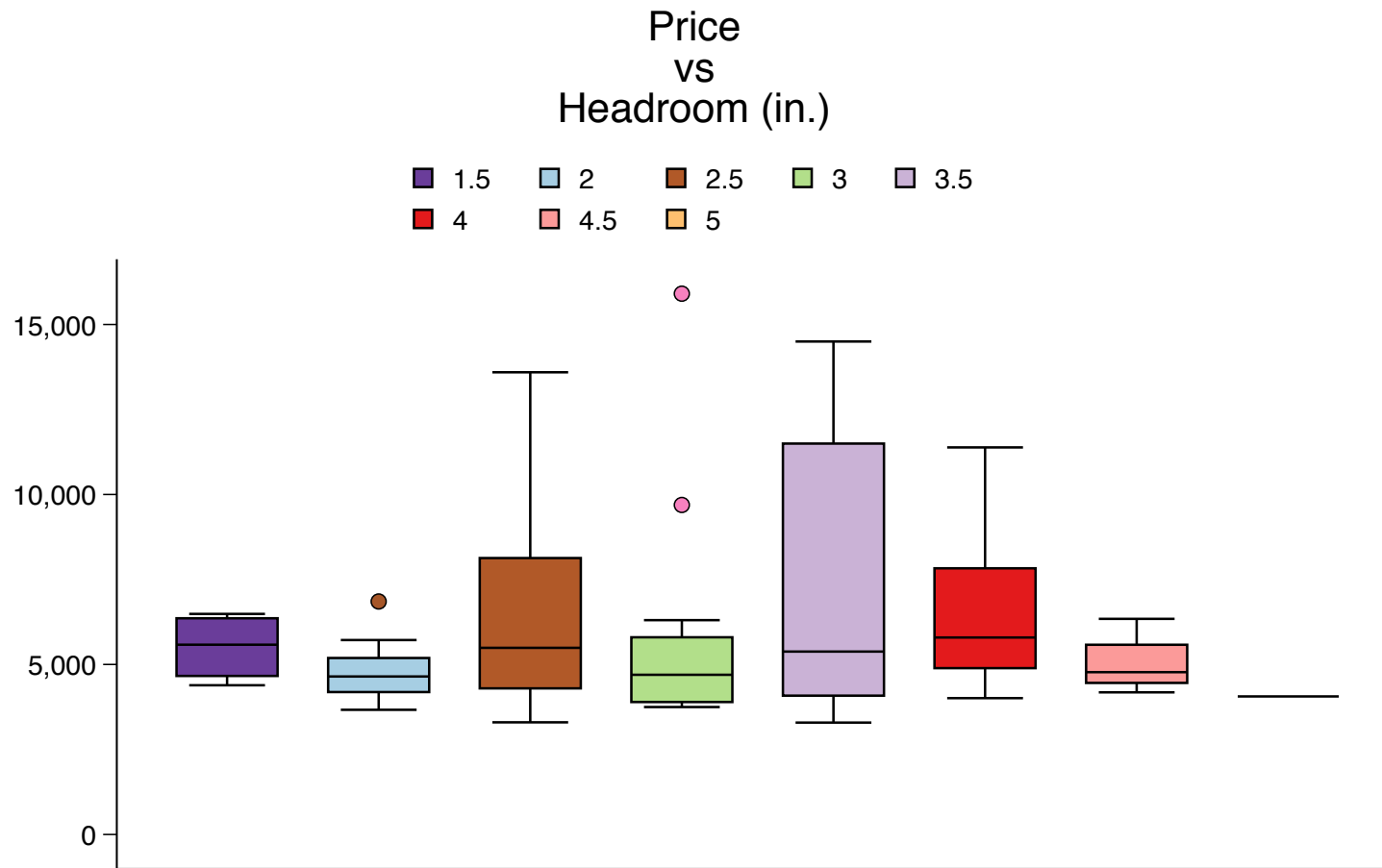
# Gear Ratio vs Repair Record 1978

1 2 3 4 5



Created on: 29 Oct 2015 at: 15:15:52

Figure 1.9: Box Plot of Price by Headroom (in.)



Created on: 29 Oct 2015 at: 15:15:54

Groups of headroom.

Figure 1.10: Box Plot of Mileage (mpg) by Headroom (in.)

Mileage (mpg)  
vs  
Headroom (in.)

1.5 2 2.5 3 3.5  
4 4.5 5

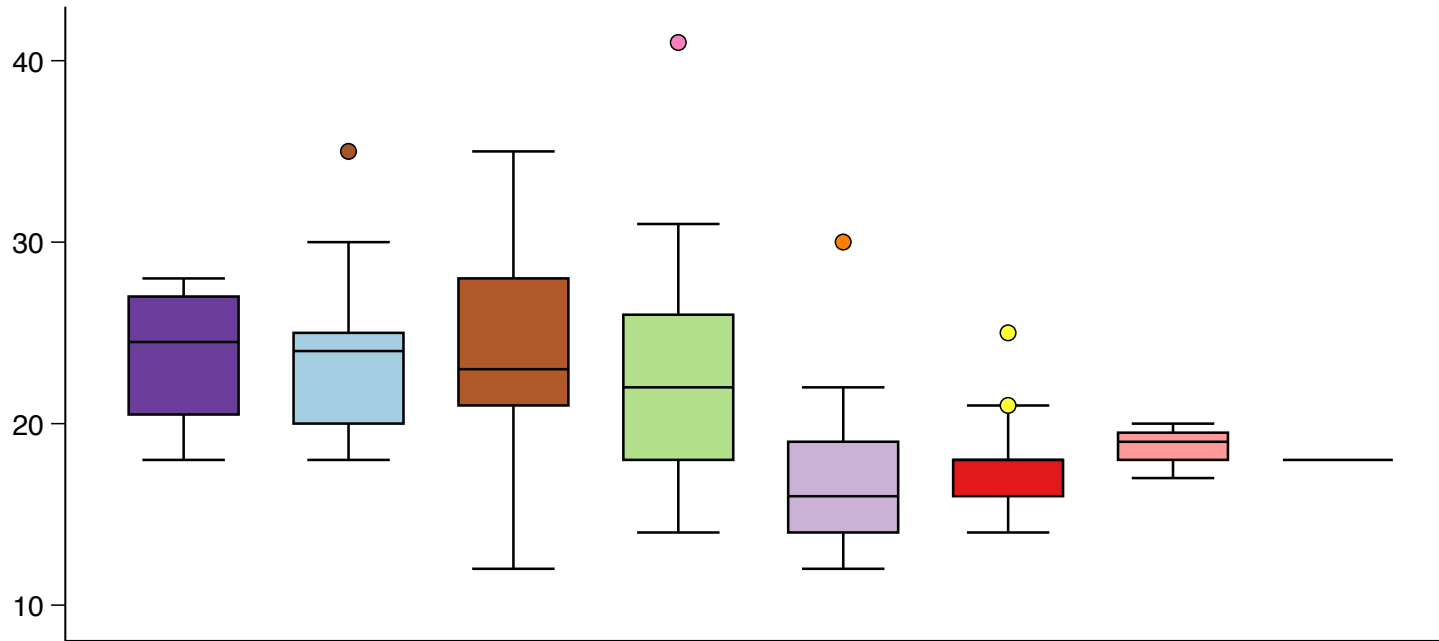
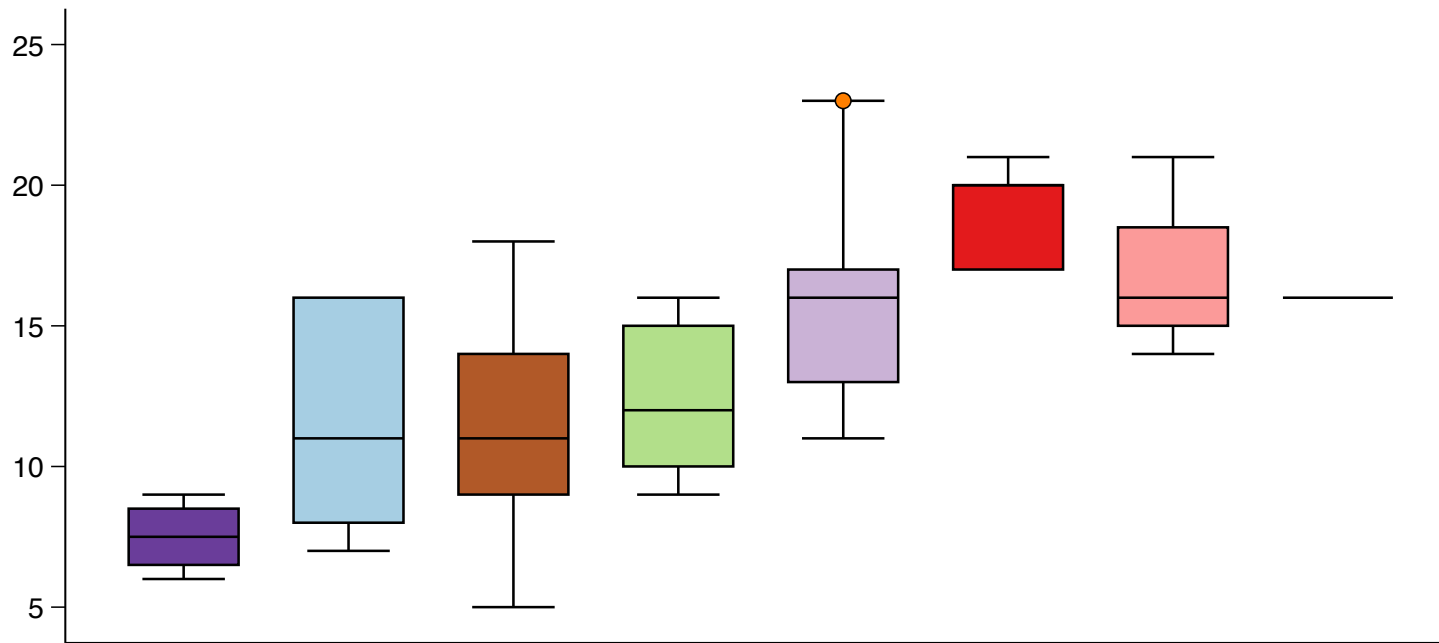


Figure 1.11: Box Plot of Trunk space (cu. ft.) by Headroom (in.)

Trunk space (cu. ft.)  
vs  
Headroom (in.)

1.5 2 2.5 3 3.5  
4 4.5 5

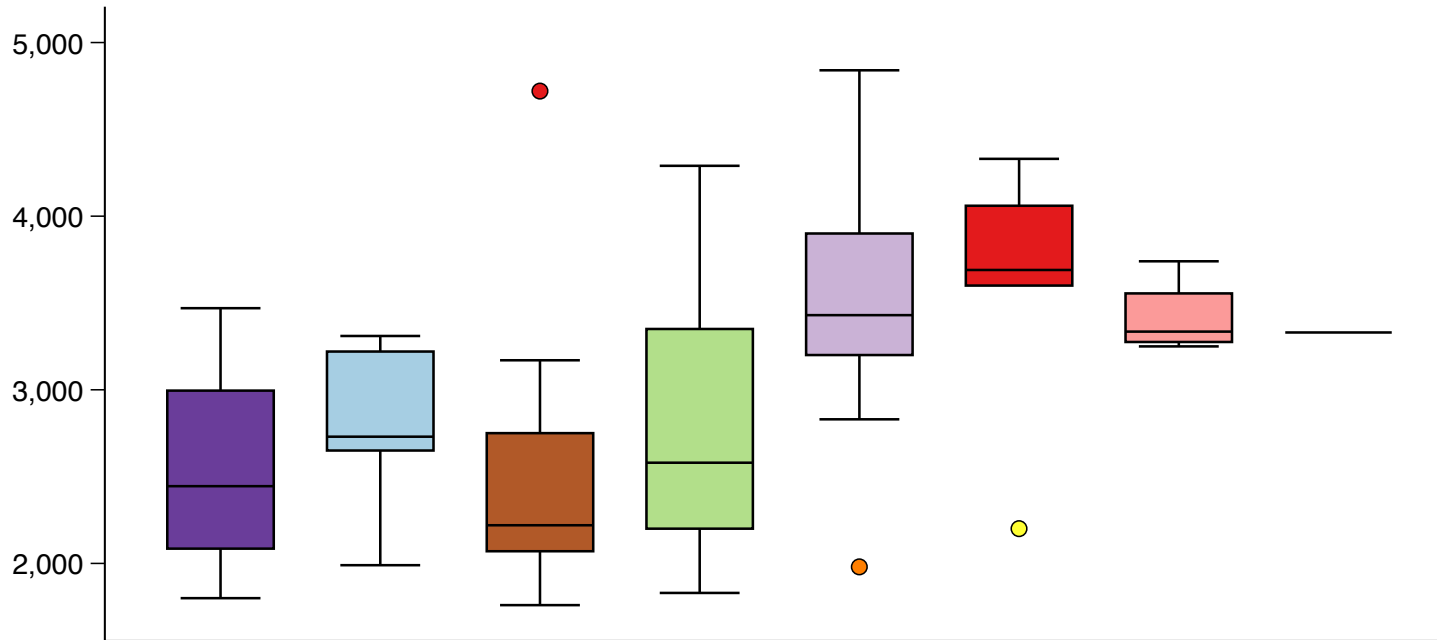


Created on: 29 Oct 2015 at: 15:16:00

Figure 1.12: Box Plot of Weight (lbs.) by Headroom (in.)

Weight (lbs.)  
vs  
Headroom (in.)

1.5 2 2.5 3 3.5  
4 4.5 5

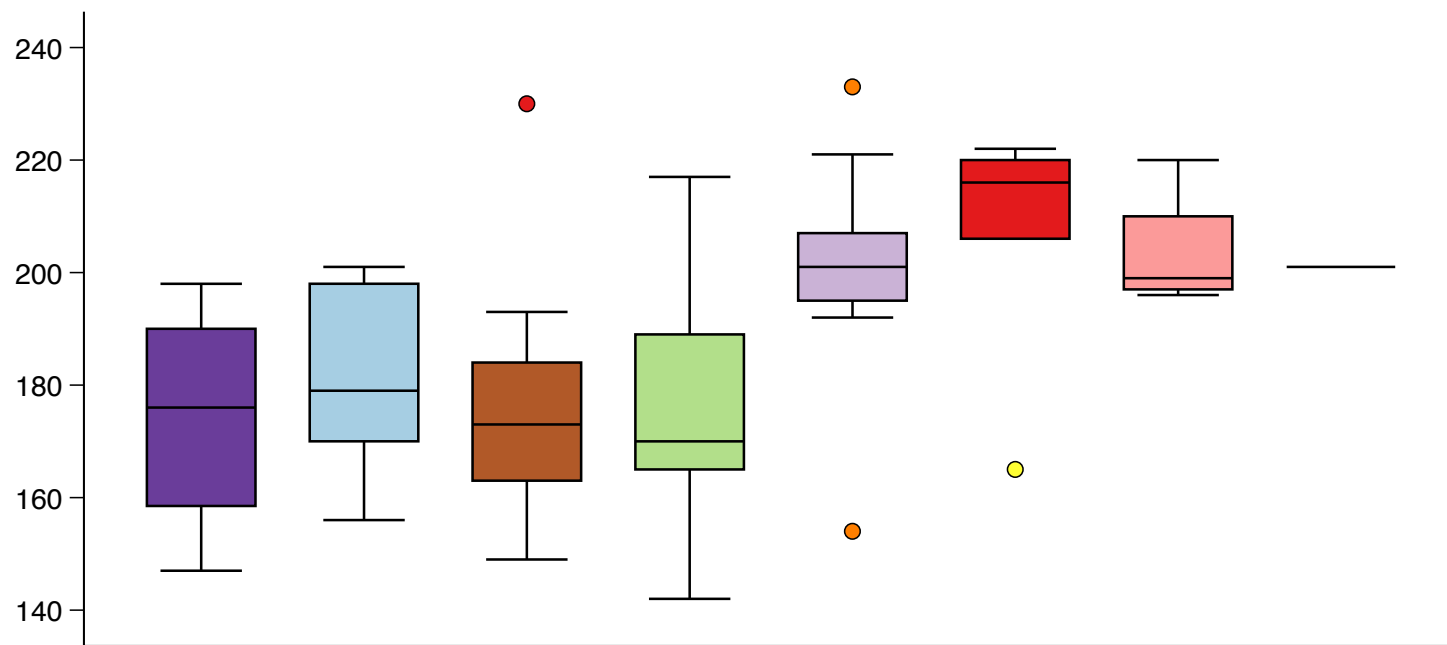


Created on: 29 Oct 2015 at: 15:16:02

Figure 1.13: Box Plot of Length (in.) by Headroom (in.)

Length (in.)  
VS  
Headroom (in.)

1.5 2 2.5 3 3.5  
4 4.5 5

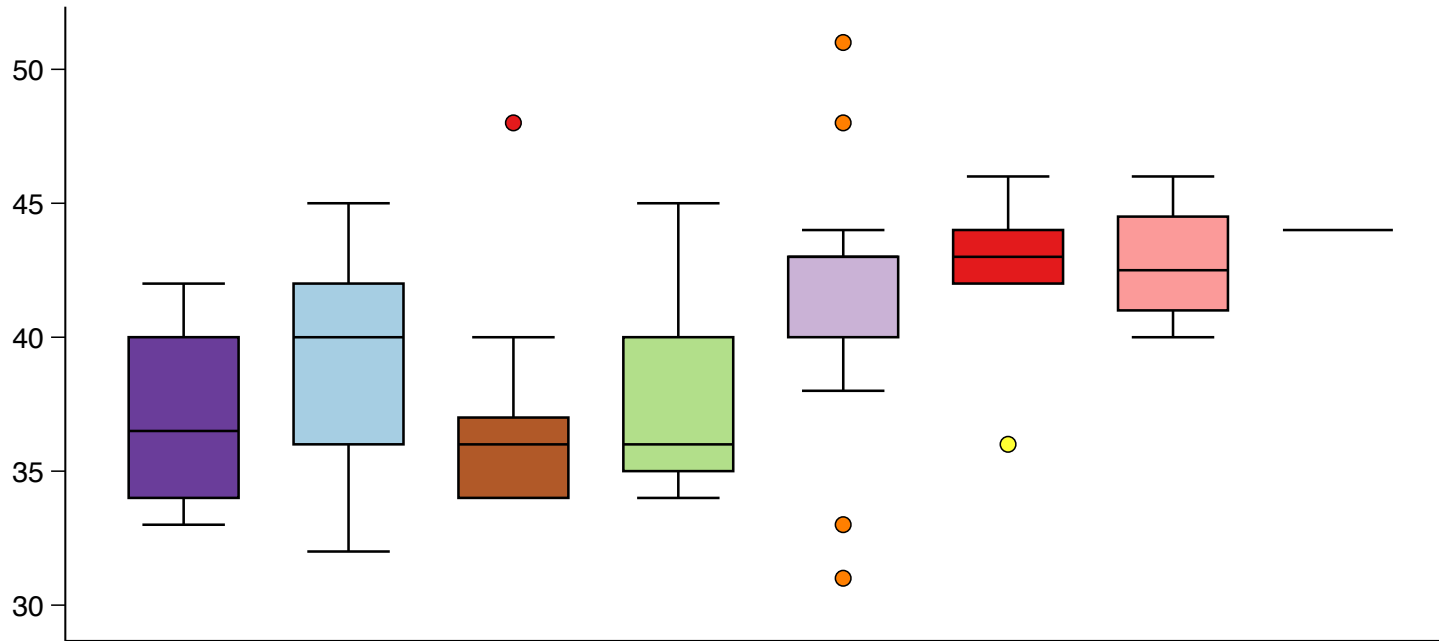


Created on: 29 Oct 2015 at: 15:16:05

Figure 1.14: Box Plot of Turn Circle (ft.) by Headroom (in.)

Turn Circle (ft.)  
vs  
Headroom (in.)

1.5 2 2.5 3 3.5  
4 4.5 5

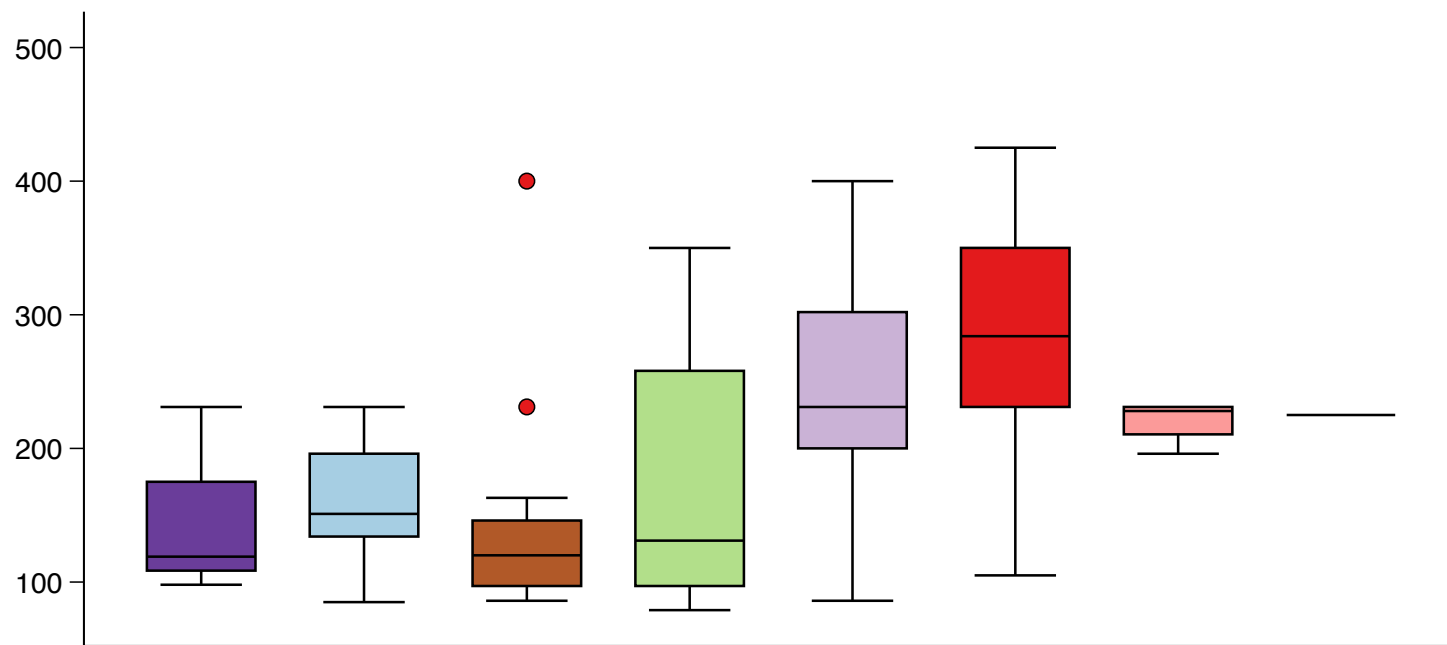


Created on: 29 Oct 2015 at: 15:16:08

Figure 1.15: Box Plot of Displacement (cu. in.) by Headroom (in.)

Displacement (cu. in.)  
vs  
Headroom (in.)

1.5 2 2.5 3 3.5  
4 4.5 5



Created on: 29 Oct 2015 at: 15:16:11



Figure 1.16: Box Plot of Gear Ratio by Headroom (in.)

# Gear Ratio vs Headroom (in.)

1.5 2 2.5 3 3.5  
4 4.5 5

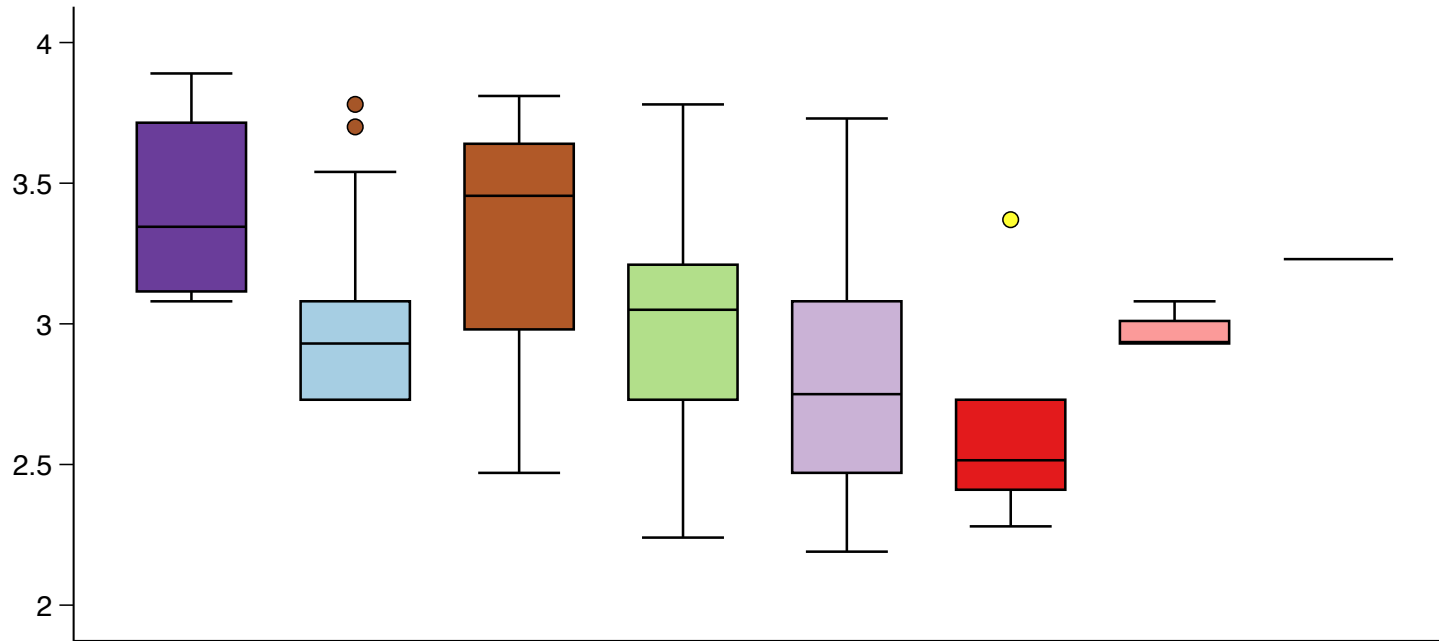
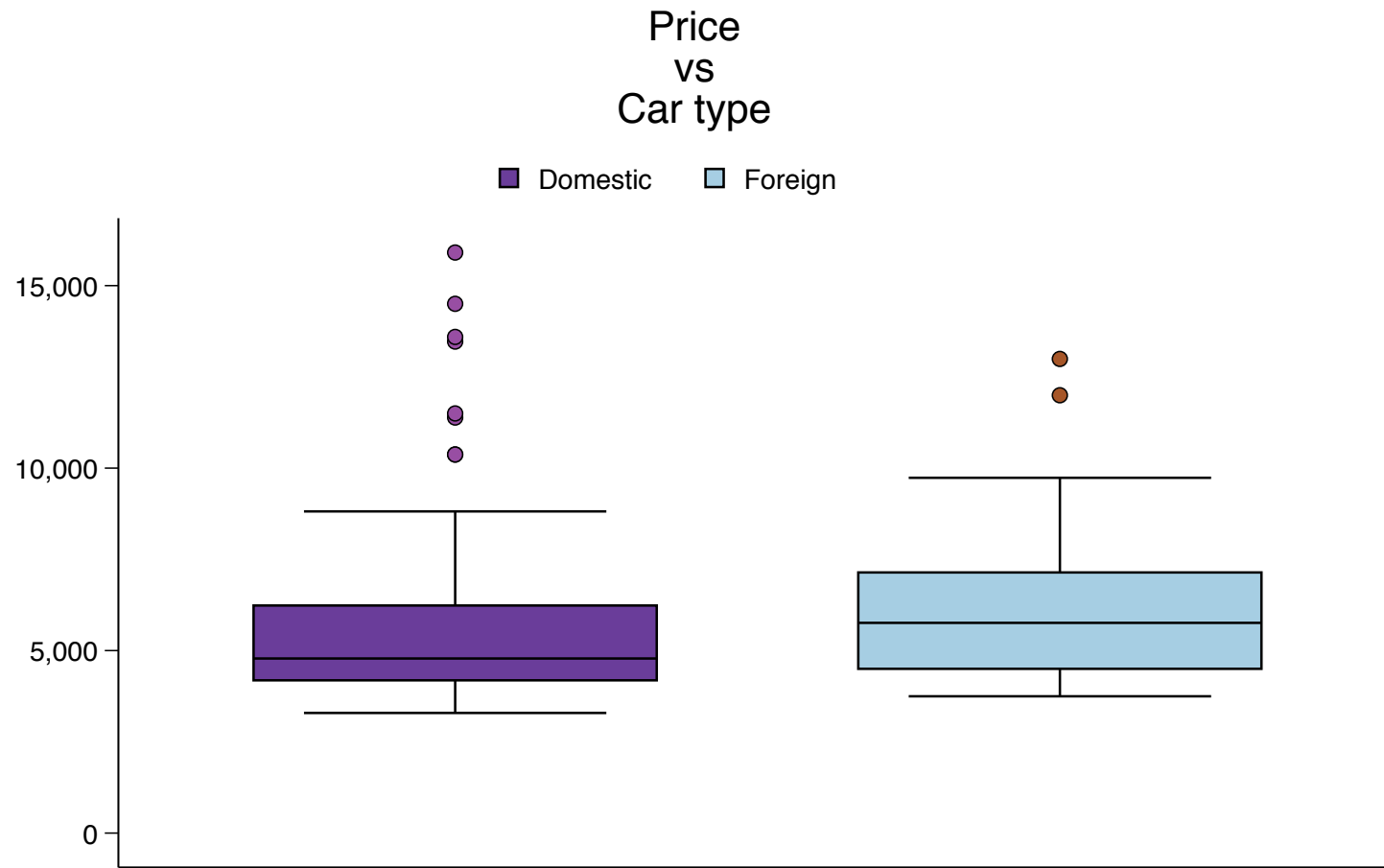


Figure 1.17: Box Plot of Price by Car type



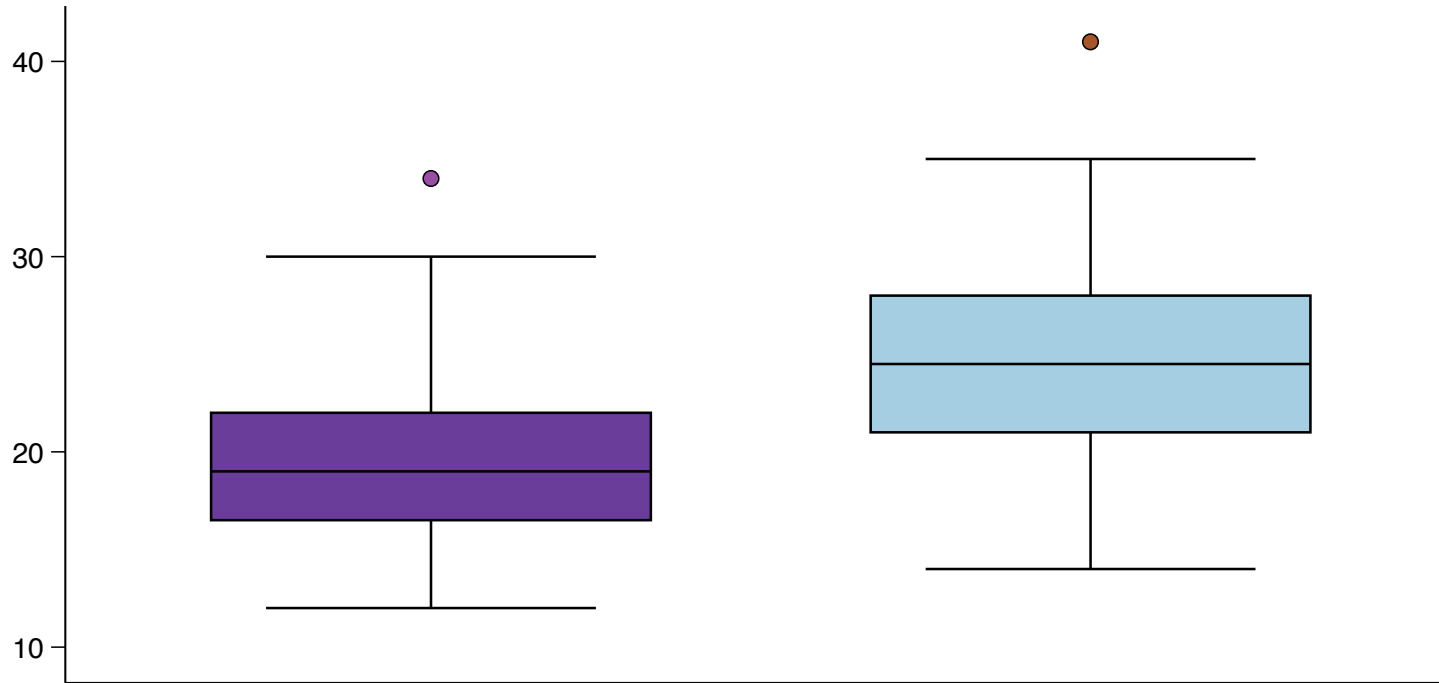
Created on: 29 Oct 2015 at: 15:16:16

Groups of foreign.

Figure 1.18: Box Plot of Mileage (mpg) by Car type

Mileage (mpg)  
vs  
Car type

■ Domestic    ■ Foreign

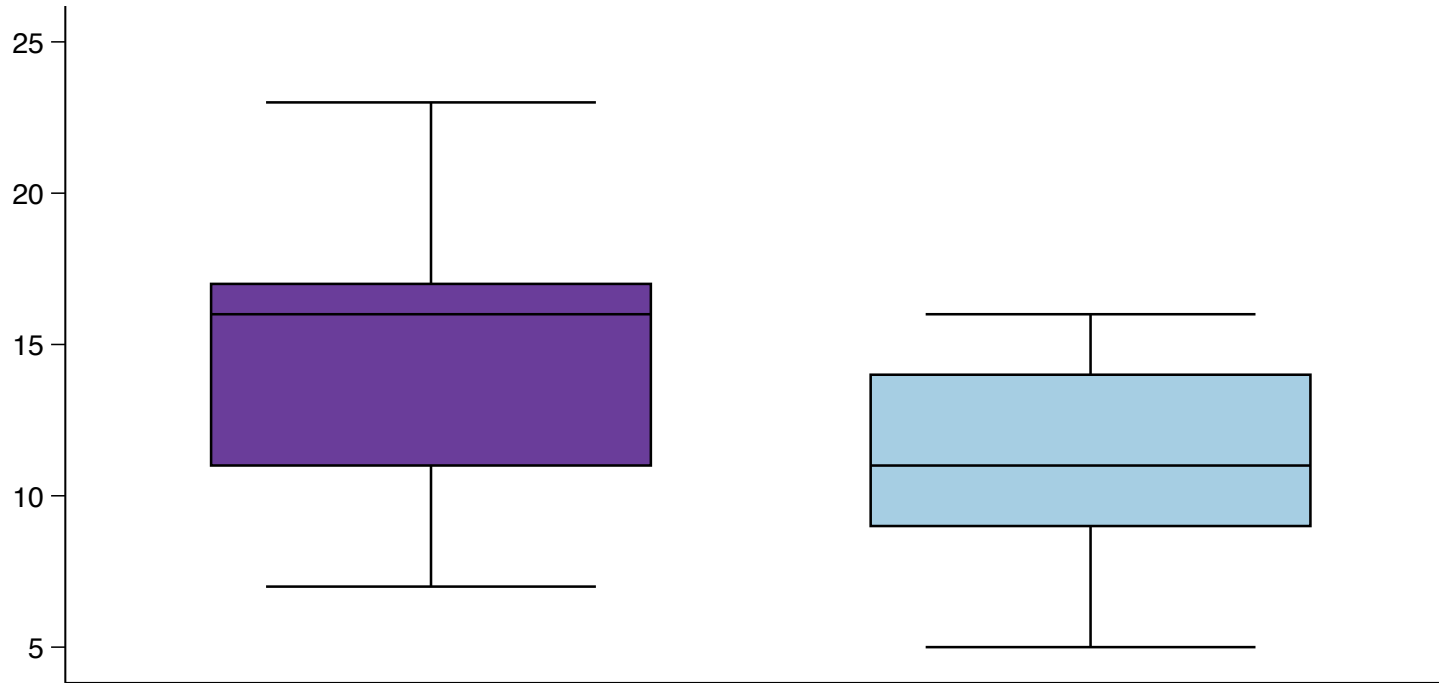


Created on: 29 Oct 2015 at: 15:16:18

Figure 1.19: Box Plot of Trunk space (cu. ft.) by Car type

Trunk space (cu. ft.)  
vs  
Car type

■ Domestic    ■ Foreign

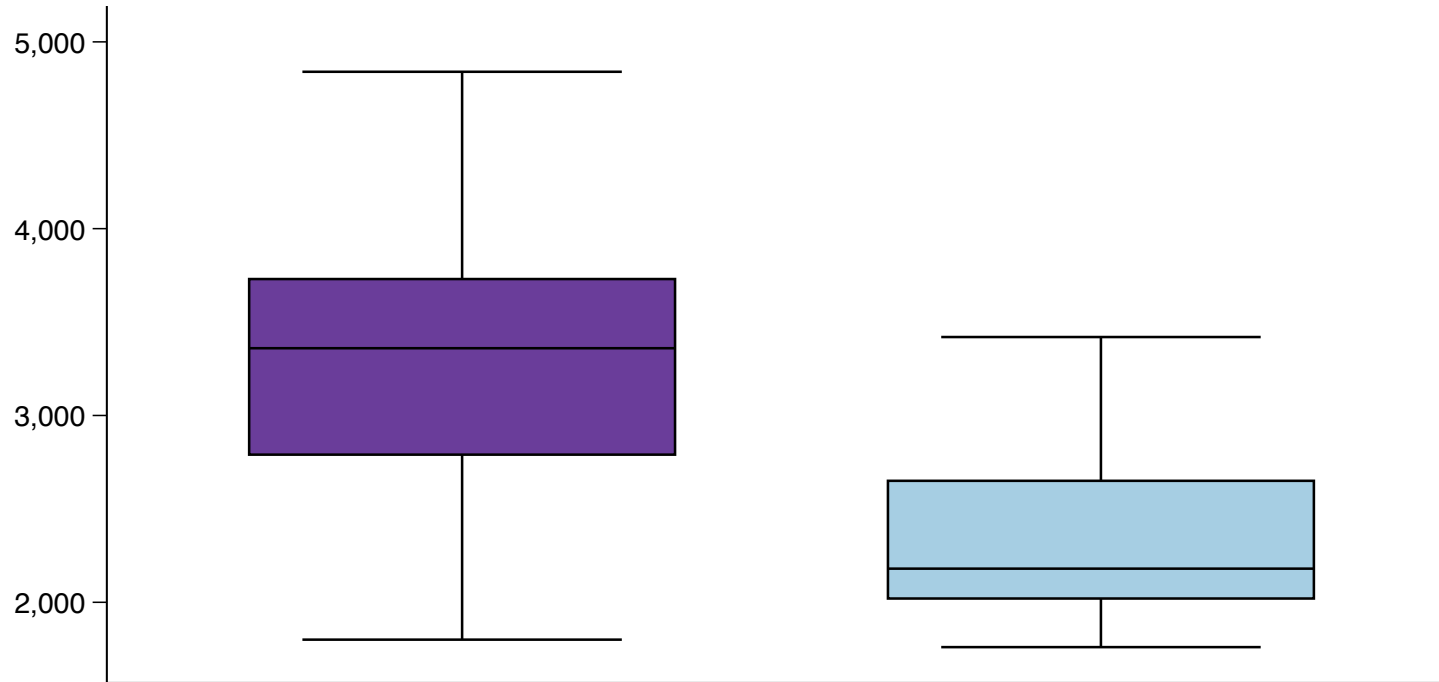


Created on: 29 Oct 2015 at: 15:16:20

Figure 1.20: Box Plot of Weight (lbs.) by Car type

Weight (lbs.)  
vs  
Car type

■ Domestic    ■ Foreign

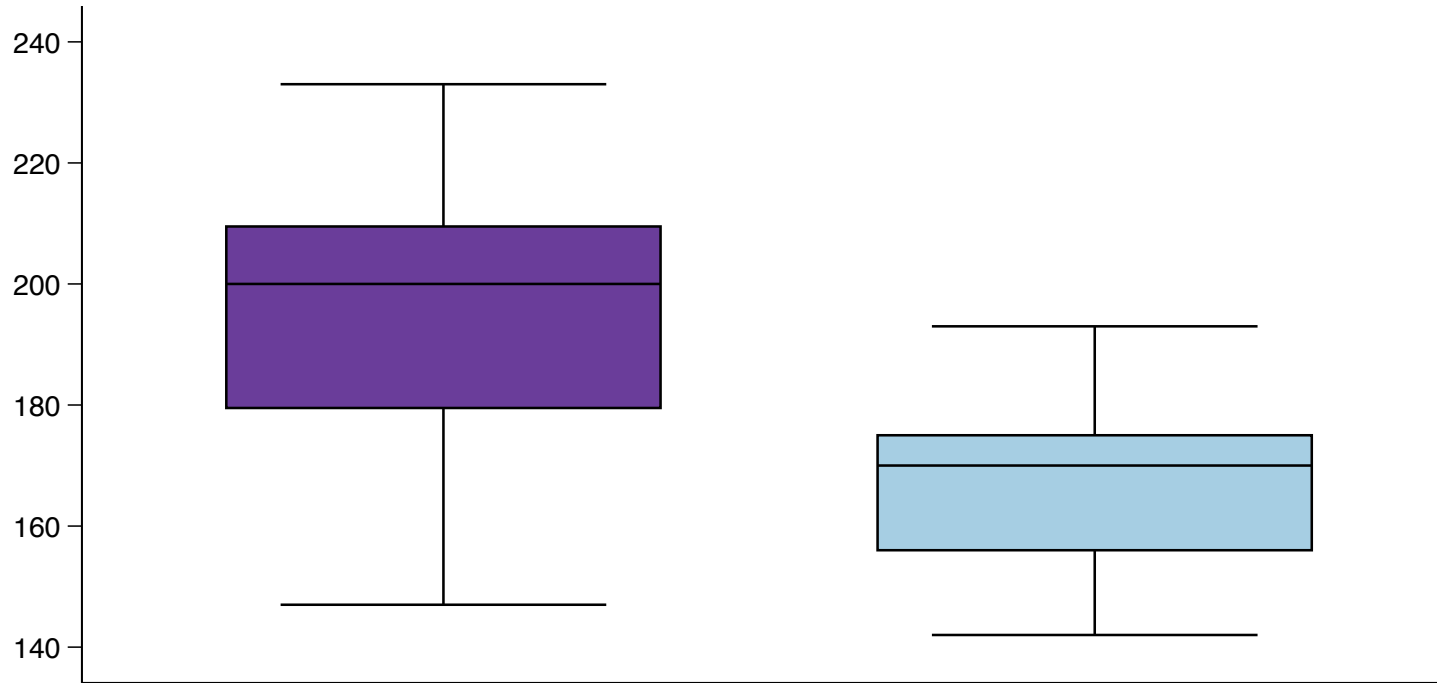


Created on: 29 Oct 2015 at: 15:16:22

Figure 1.21: Box Plot of Length (in.) by Car type

Length (in.)  
vs  
Car type

■ Domestic    ■ Foreign

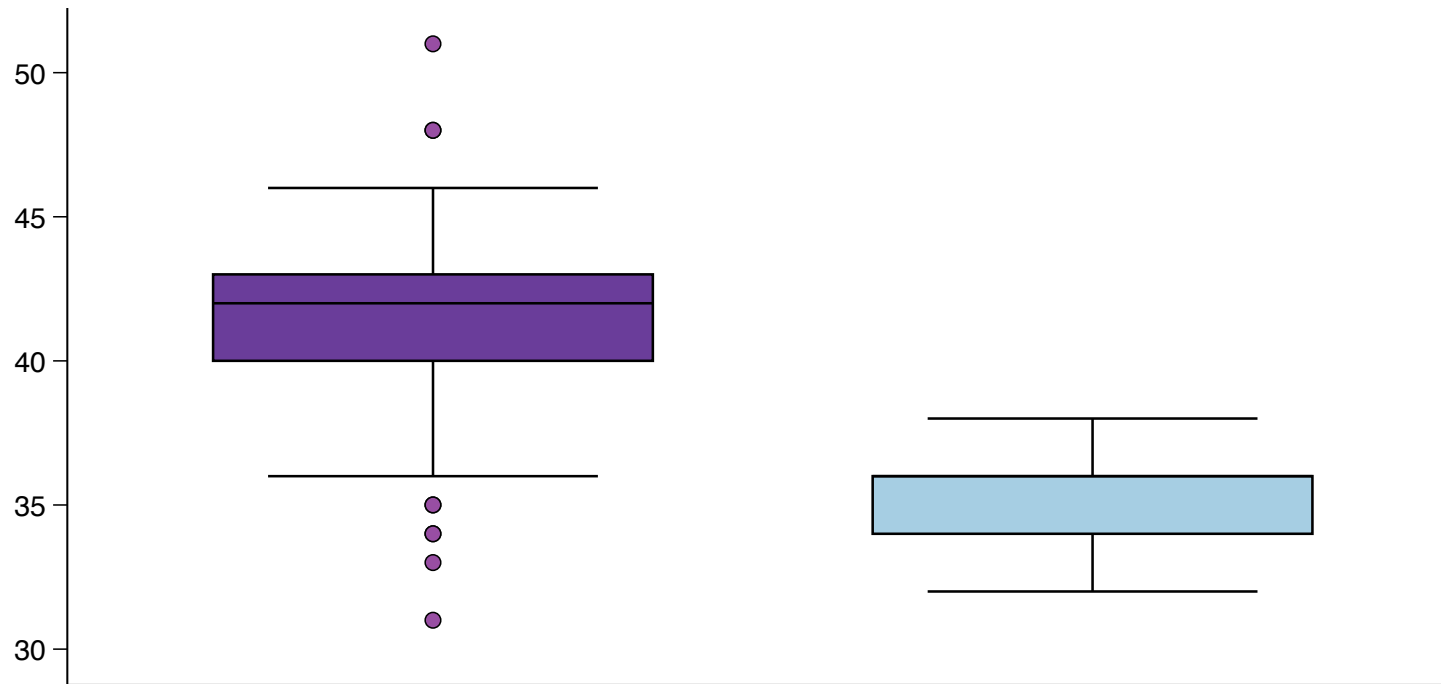


Created on: 29 Oct 2015 at: 15:16:25

Figure 1.22: Box Plot of Turn Circle (ft.) by Car type

Turn Circle (ft.)  
vs  
Car type

■ Domestic   ■ Foreign

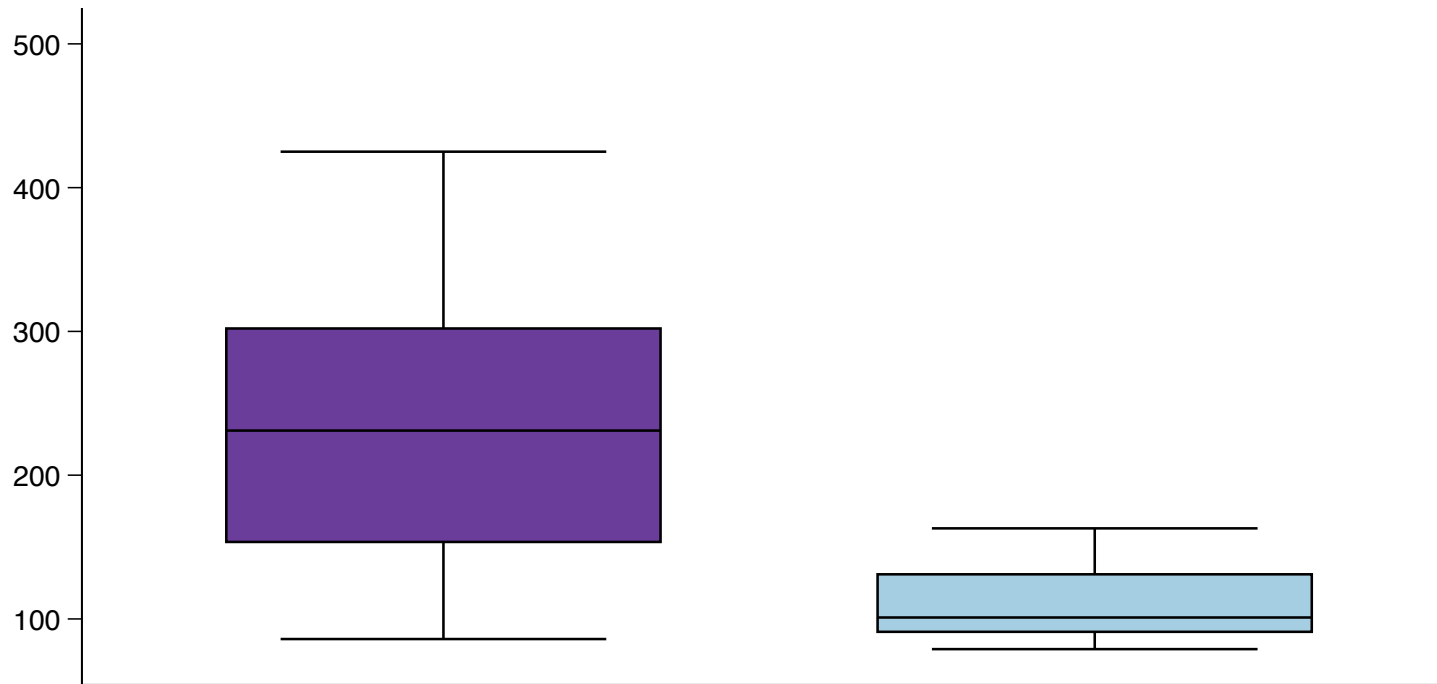


Created on: 29 Oct 2015 at: 15:16:27

Figure 1.23: Box Plot of Displacement (cu. in.) by Car type

Displacement (cu. in.)  
vs  
Car type

■ Domestic    ■ Foreign



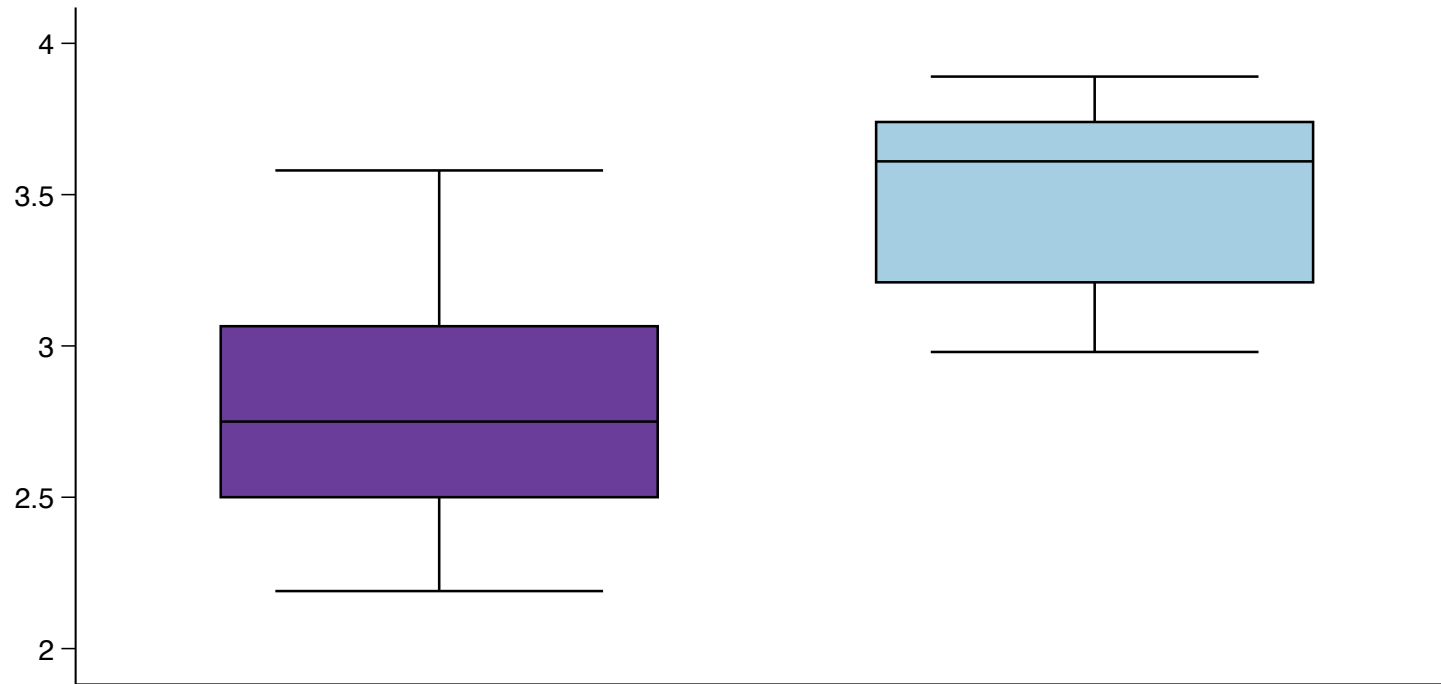
Created on: 29 Oct 2015 at: 15:16:29



Figure 1.24: Box Plot of Gear Ratio by Car type

# Gear Ratio vs Car type

■ Domestic    ■ Foreign



Created on: 29 Oct 2015 at: 15:16:31

**Chapter 2.**  
**Descriptive Statistics**

## Categorical Variables

Table 2.1: Distribution of Repair Record 1978

	Frequency	Percentage
1	2	2.703
2	8	10.81
3	30	40.54
4	18	24.32
5	11	14.86
.	5	6.757

Table 2.2: Distribution of Headroom (in.)

	Frequency	Percentage
1.5	4	5.405
2	13	17.57
2.5	14	18.92
3	13	17.57
3.5	15	20.27
4	10	13.51
4.5	4	5.405
5	1	1.351

Table 2.3: Distribution of Car type

	Frequency	Percentage
Domestic	52	70.27
Foreign	22	29.73

Table 2.4: Frequency of Headroom (in.) by Car type

	Domestic Frequency	Foreign Frequency
1.5	3	1
2	10	3
2.5	4	10
3	7	6
3.5	13	2
4	10	0
4.5	4	0
5	1	0

Table 2.5: Overall % of Headroom (in.) by Car type

	Domestic Overall%	Foreign Overall%
1.5	4.054	1.351
2	13.51	4.054
2.5	5.405	13.51
3	9.459	8.108
3.5	17.57	2.703
4	13.51	0
4.5	5.405	0
5	1.351	0

Table 2.6: Column % of Headroom (in.) by Car type

	Domestic Column%	Foreign Column%
1.5	5.769	4.545
2	19.23	13.64
2.5	7.692	45.45
3	13.46	27.27
3.5	25	9.091
4	19.23	0
4.5	7.692	0
5	1.923	0

Table 2.7: Row % of Headroom (in.) by Car type

	Domestic Row%	Foreign Row%
1.5	75	25
2	76.92	23.08
2.5	28.57	71.43
3	53.85	46.15
3.5	86.67	13.33
4	100	0
4.5	100	0
5	100	0

Table 2.8: Frequency of Repair Record 1978 by Car type

	Domestic Frequency	Foreign Frequency
1	2	0
2	8	0
3	27	3
4	9	9
5	2	9
.	4	1

Table 2.9: Overall % of Repair Record 1978 by Car type

	Domestic Overall%	Foreign Overall%
1	2.703	0
2	10.81	0
3	36.49	4.054
4	12.16	12.16
5	2.703	12.16
.	5.405	1.351

Table 2.10: Column % of Repair Record 1978 by Car type

	Domestic Column%	Foreign Column%
1	3.846	0
2	15.38	0
3	51.92	13.64
4	17.31	40.91
5	3.846	40.91
.	7.692	4.545

Table 2.11: Row % of Repair Record 1978 by Car type

	Domestic Row%	Foreign Row%
1	100	0
2	100	0
3	90	10
4	50	50
5	18.18	81.82
.	80	20



Table 2.12: Frequency of Repair Record 1978 by Headroom (in.)

	1.5	2	2.5	3	3.5	4	4.5	5
	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency
1	1	1	0	0	0	0	0	0
2	0	3	0	0	1	2	1	1
3	0	5	5	4	10	3	3	0
4	2	1	5	3	2	5	0	0
5	0	3	4	4	0	0	0	0
.	1	0	0	2	2	0	0	0

Table 2.13: Overall % of Repair Record 1978 by Headroom (in.)

	1.5	2	2.5	3	3.5	4	4.5	5
	Overall%	Overall%	Overall%	Overall%	Overall%	Overall%	Overall%	Overall%
1	1.351	1.351	0	0	0	0	0	0
2	0	4.054	0	0	1.351	2.703	1.351	1.351
3	0	6.757	6.757	5.405	13.51	4.054	4.054	0
4	2.703	1.351	6.757	4.054	2.703	6.757	0	0
5	0	4.054	5.405	5.405	0	0	0	0
.	1.351	0	0	2.703	2.703	0	0	0

Table 2.14: Column % of Repair Record 1978 by Headroom (in.)

	1.5	2	2.5	3	3.5	4	4.5	5
	Column%	Column%	Column%	Column%	Column%	Column%	Column%	Column%
1	25	7.692	0	0	0	0	0	0
2	0	23.08	0	0	6.667	20	25	100
3	0	38.46	35.71	30.77	66.67	30	75	0
4	50	7.692	35.71	23.08	13.33	50	0	0
5	0	23.08	28.57	30.77	0	0	0	0
.	25	0	0	15.38	13.33	0	0	0

Table 2.15: Row % of Repair Record 1978 by Headroom (in.)

1.5	2	2.5	3	3.5	4	4.5	5
-----	---	-----	---	-----	---	-----	---

## Continuous Variables

Table 2.16: Descriptive Statistics of Continuous Variables

	N	$\mu$	$\sigma$
Price	74	6165.257	2949.496
Mileage (mpg)	74	21.297	5.786
Trunk space (cu. ft.)	74	13.757	4.277
Weight (lbs.)	74	3019.459	777.194
Length (in.)	74	187.932	22.266
Turn Circle (ft.)	74	39.649	4.399
Displacement (cu. in.)	74	197.297	91.837
Gear Ratio	74	3.015	0.456

$\mu$  = Average  $\sigma$  = Standard Deviation

Table 2.17: Higher Order Moment Conditions

	Skewness	Kurtosis
Price	1.653	4.819
Mileage (mpg)	0.949	3.975
Trunk space (cu. ft.)	0.029	2.192
Weight (lbs.)	0.148	2.118
Length (in.)	-0.041	2.042
Turn Circle (ft.)	0.124	2.229
Displacement (cu. in.)	0.592	2.376
Gear Ratio	0.219	2.102

Table 2.18: Order Statistics

	Min.	25%ile	Median	75%ile	Max
Price	3291.000	4195.000	5006.500	6342.000	15906.000
Mileage (mpg)	12.000	18.000	20.000	25.000	41.000
Trunk space (cu. ft.)	5.000	10.000	14.000	17.000	23.000
Weight (lbs.)	1760.000	2240.000	3190.000	3600.000	4840.000
Length (in.)	142.000	170.000	192.500	204.000	233.000
Turn Circle (ft.)	31.000	36.000	40.000	43.000	51.000
Displacement (cu. in.)	79.000	119.000	196.000	250.000	425.000
Gear Ratio	2.190	2.730	2.955	3.370	3.890

This is also known as Tukey's Five Number Summary

## Conditional Descriptive Statistics



Table 2.19: Averages by groups of Repair Record 1978

	1	2	3	4	5	Total
Price	4564.5	5967.6	6429.2	6071.5	5913	6146.0
Mileage (mpg)	21	19.12	19.43	21.67	27.36	21.29
Trunk space (cu. ft.)	8.500	14.62	15.27	13.50	11.45	13.93
Weight (lbs.)	3100	3353.8	3299	2870	2322.7	3032.0
Length (in.)	189	199.4	194	184.8	170.2	188.3
Turn Circle (ft.)	41	43.38	41.07	38.50	35.64	39.80
Displacement (cu. in.)	191	242.2	230.0	178.8	111.1	198
Gear Ratio	2.905	2.686	2.871	3.176	3.305	2.999
Observations	69					

mean coefficients; *t* statistics in parentheses

Table 2.20: Standard Deviations by groups of Repair  
Record 1978

	1	2	3	4	5	Total
Price	522.6	3579.4	3525.1	1709.6	2615.8	2912.4
Mileage (mpg)	4.243	3.758	4.141	4.935	8.732	5.866
Trunk space (cu. ft.)	2.121	4.984	3.591	5.272	2.659	4.343
Weight (lbs.)	523.3	446.0	752.4	907.3	410.6	792.9
Length (in.)	12.73	13.98	20.72	28.65	12.10	22.75
Turn Circle (ft.)	1.414	2.504	4.250	4.706	1.433	4.441
Displacement (cu. in.)	56.57	77.57	91.93	96.06	25.46	93.15
Gear Ratio	0.247	0.317	0.411	0.536	0.335	0.463
Observations	69					

sd coefficients; *t* statistics in parentheses

Tables by groups of headroom



Table 2.21: Averages by groups of Headroom (in.)

	1_missing_5 1.5	2 1.5	2_missing_5 1.5	3 1.5	3_missing_5 1.5	4 1.5	4_missing_5 1.5	5 1.5	Total 1.5
Price	5509.5	4822.8	6591.6	5906.2	7580.9	6458.5	5018	4060	6165.3
Mileage (mpg)	23.75	23.77	23.71	23.46	17.53	17.90	18.75	18	21.30
Trunk space (cu. ft.)	7.500	11.69	11.43	12.31	16.07	19.20	16.75	16	13.76
Weight (lbs.)	2540	2761.5	2499.3	2743.1	3541.3	3662	3415	3330	3019.5
Length (in.)	174.2	180.4	175.3	175.2	201.1	210.2	203.5	201	187.9
Turn Circle (ft.)	37	39.23	36.86	37.77	41.73	42.80	42.75	44	39.65
Displacement (cu. in.)	141.8	155.5	143.4	178.3	251.3	280.9	220.8	225	197.3
Gear Ratio	3.415	3.070	3.331	3.028	2.817	2.616	2.970	3.230	3.015
Observations	74								

mean coefficients;  $t$  statistics in parentheses

Table 2.22: Standard Deviations by groups of Headroom  
(in.)

	1_missing_5 1.5	2 1.5	2_missing_5 1.5	3 1.5	3_missing_5 1.5	4 1.5	4_missing_5 1.5	Total 1.5
Price	1009.6	854.1	3196.5	3404.0	4086.4	2256.7	926.6	2949.5
Mileage (mpg)	4.349	4.902	6.293	7.043	4.486	3.178	1.258	5.786
Trunk space (cu. ft.)	1.291	3.772	3.390	2.626	3.240	1.549	2.986	4.277
Weight (lbs.)	693.3	484.2	761.2	812.6	644.0	570.3	222.2	777.2
Length (in.)	21.48	15.22	19.71	21.29	16.97	16.90	11.12	22.27
Turn Circle (ft.)	3.916	3.723	3.634	3.919	4.992	2.781	2.500	4.399
Displacement (cu. in.)	60.32	47.86	83.01	98.48	84.98	89.71	16.74	91.84
Gear Ratio	0.376	0.370	0.394	0.508	0.450	0.308	0.0735	0.456
Observations	74							

sd coefficients;  $t$  statistics in parentheses

Tables by groups of foreign

Table 2.23: Averages by groups of Car type

	Domestic	Foreign	Total
Price	6072.4	6384.7	6165.3
Mileage (mpg)	19.83	24.77	21.30
Trunk space (cu. ft.)	14.75	11.41	13.76
Weight (lbs.)	3317.1	2315.9	3019.5
Length (in.)	196.1	168.5	187.9
Turn Circle (ft.)	41.44	35.41	39.65
Displacement (cu. in.)	233.7	111.2	197.3
Gear Ratio	2.807	3.507	3.015
Observations	74		

mean coefficients;  $t$  statistics in parentheses

Table 2.24: Standard Deviations by groups of Car type

	Domestic	Foreign	Total
Price	3097.1	2621.9	2949.5
Mileage (mpg)	4.743	6.611	5.786
Trunk space (cu. ft.)	4.306	3.217	4.277
Weight (lbs.)	695.4	433.0	777.2
Length (in.)	20.05	13.68	22.27
Turn Circle (ft.)	3.968	1.501	4.399
Displacement (cu. in.)	85.26	24.88	91.84
Gear Ratio	0.336	0.297	0.456
Observations	74		

sd coefficients;  $t$  statistics in parentheses