Price Comparison On Various Automobile Characteristics

Sheng (Jeremy) Jiang
NUID:001725452
Course: IE 6200
Lecturer: Shivani Patel

Table of Content

Abstract

Background:

For centuries, the evolution of technology has driven the development of the automobile industry. With increasing in automobile mass production, the automobile is becoming a key component in modern life as it blurred social class distinctions, expanded markets, and stimulated the economy. Buying a car has always been a part of the whole life that everyone went through or experiencing now. But how automobile industry determined their car price in terms of car's characteristics remained ambiguous.

Aims:

The aim of this observational study is to estimate the average car price in the industry and find the basic pattern of automobile characteristics with association to its price.This study particularly focuses on the following car characteristics: drive wheel mode, aspiration and body style. The scope of this observational study will be based on Automobile dataset compiled by Jeffrey C. Schlimmer (Schlimmer).

Methods:

Sampling Strategy:

The original dataset of this study comes from UCI Machine Learning Repository. It was compiled by Jeffrey C. Schlimmer.There were some limitations on the sampling strategy since the dataset only focused on the year 1985.These resources Schlimer found may not have enough coverage for the entire target population at that period which may lead to selection bias. And there may be confounding factors like currency fluctuation and auto companies' transformation that can affect this study's key variable automobile price dramatically.

Exploratory Analysis:

This study performed plots to visually show the car price distribution in terms of body styles, drive wheels and aspiration types.From the graphs and data summary, the whole dataset is right skewed at lower price（mean=13207 median=10295). Rear wheel drive cars tend to have relative

higher car price than other types of drive wheels, while constitutes 37.3% of the dataset. Convertible (mean=17084.5) and hardtop (19687.5) are relatively expensive than other cars, while convertible hardtop have disperse price distributions.Hatchback and sedan constitute 33.8% and 46.8% of the dataset. For aspiration types, most cars are standard type, which constitutes 82.1% of dataset.

Statistical Analysis:

There are five tests in this study.

At the significant level $\alpha = 0.05$ and with a 95% confidence,Test 1 suggests the true mean car price in 1985 is between $12101.8 and $14312.46 which is greater than $7000 estimate mean price from research. Test 2 fails to prove the true proportion of rear wheel drive cars is greater than 0.3333, but between 0.3162255 and 1. Test 3 indicates that true mean price for turbo cars is greater than standard cars (CI:1257.027 to 5328.797). In the further study at Test 4 on comparison of the true proportion of turbo type between luxury cars and economy cars.Turbo cars in luxury class population has a higher proportion than turbo cars in economy class population (CI:0.07224941 to 0.30527772).Finally at Test 5, analyst finds not all true proportion of body styles are evenly distributed in whole car population.

Conclusion:

These findings on automobile price, and its characteristics in 1985 are highly based on dataset from Jeffrey C. Schlimmer on limited sources. This study firstly suggests that the average car price was greater than the estimate mean price $7000 in 1980s.Secondly, proportion of rear wheel drive cars was equal to the other two drive wheel cars. In addition, price of turbo aspiration car was significantly higher than standard aspiration cars, and turbo aspiration car took a greater percentage in luxury class population than economy class population. Finally, the availability of different body style cars was different. But this study may have limitations due to confounding factors, potential biases, and limited sources so that this topic needs to be further explored.

Introduction

For centuries, the evolution of technology has driven the development of the automobile industry. With increasing in automobile mass production, the automobile is becoming a key component in modern life as it blurred social class distinctions, expanded markets, and stimulated the economy. Buying a car has always been a part of the whole life that everyone went through or experiencing now. And automobile development over the years has always been the symbol of technology evolution that draws attention to everyone. While driving the car, people always appreciate the emergence of new car features like hybrid fuel type,electric charging and automation driving mode. But how automobile industry determined their car price in terms of car's characteristics remained ambiguous.

What features of automobile have significant influences on car total price has been an interest to car buyers and car owners. Analyst of this study is like others as outsiders of automobile industry seek to figure out how these insiders price car in terms of car's characteristics. So by having this observational study on automobile dataset for cars in 1985, analyst could help car customers by providing buying guide to select optimized car in 1985. Analyst can also find the trend of car features in 1985 to estimate what car features are most popular nowadays.

This observational study is designed to:

- Estimate the average car price in the industry in 1985
- Find the basic pattern of automobile characteristics with association to its price.

This study particularly focuses on the following car characteristics: drive wheel mode, aspiration and body style. The scope of this observational study will be based on the year 1985 where the collected Automobile dataset focused on.

The goal of this observational study is to provide estimations for these questions below:

1. What is the average car price in 1985?
2. How rear wheel drive cars distributed in whole car population in 1985?
3. Is there a difference in price between turbo and standard car in 1985?

4. How turbo aspiration cars distributed in luxury class compared to economy class in 1985?
5. Are certain body style more preferred than others in 1985?

Variable Definition:

- **Price**

  Definition: Price is the most important variable for this study. The price column from dataset states car total price for each car. Analyst converts this variable into numeric to better perform analysis in the next step.

  - **Type:** Numeric Variable in decimal; Response Variable; Quantitative Continuous Variable.
  - **Scale:** Price values range from 5118 to 45400 per car.

- **Wheel drive**

  Definition: Wheel drive is one of the car characteristics. Wheel drive has three modes: front wheel drive, rear wheel drive, and four wheel drive.Each car has to have only one of these three drive wheel modes. In order to perform statistical analysis, analyst will group front wheel drive and rear wheel drive into one category called two wheel drive.

  - **Type:** Categories Variable; Explanatory Variable; Discrete Variable.
  - **Scale:** Wheel drive has only 3 categories: fwd, rwd and 4wd and in statistical analysis we group fwd and 4wd into a single group called fwd(front and four wheel drive). So wheel drive will either be rwd or fwd..

- **Aspiration**

  Definition: Aspiration is one of the car characteristics. Aspiration has two types: turbo and standard. Each car can only has one and only aspiration type.

  - **Type:** Categories Variable; Explanatory Variable; Discrete Variable.
  - **Scale:** Aspiration has only 2 categories: turbo and standard. Car aspiration either to be turbo or standard.

- **Body style**

  Definition: Car has many body style which could influence the car price. Body style has five general categories:convertible, hardtop, hatchback, sedan, and wagon.Each car can only has one and only body style.

  - **Type:** Categories Variable; Explanatory Variable; Discrete Variable.
  - **Scale:** Body style has 5 general categories: convertible, hardtop, hatchback, sedan, and wagon. Each car must be in one of these 5 body styles.

- **Car class**

  Definition: This study created car class variable based on car price values with comparison of dataset's mean car price. Cars under price $13207 will be economy class, while cars over price $13207 will be luxury class.

  - **Type:** Categories Variable; Explanatory Variable; Discrete Variable.
  - **Scale:** Each car must be in either economy class or luxury class.

Methods

Target population for this study: The entire automobile population in 1985.

Sampling Size: 201 data points.

Sampling Strategy

This study modifies the original dataset from UCI Machine Learning Repository. The original dataset was compiled by Jeffrey C. Schlimmer from 1985 Model Import Car and Truck Specifications, 1985 Ward's Automotive Yearbook, Personal Auto Manuals, Insurance Services Office, and Insurance Collision Report, Insurance Institute for Highway Safety. In this study, “Symboling” and “normalized losses” as these were introduced for insurance rating purposes in Schlimmer’s study and are not meaningful to this study’s topic. Before preparing the dataset, there are rows with missing values and several wrong data types.

- Limitations:
    - The dataset only focused on the year 1985. From the datasetSchlimmer provided, it's not clear what year of car was produced and into the market which may lead to **information bias** in this study.
    - These resources Schlimer found may not have enough coverage for the entire target population at that period which may lead to **selection bias**.
    - In this observational study, analyst uses this dataset for price comparison on various automobile characteristics. But there may be **confounding factors** like currency fluctuation and auto companies' transformation that can affect the automobile price dramatically.

Data Preparation

1. Analyst removes the "symboling" and "normalized losses" from the dataset since these were introduced for insurance rating purposes in Schlimmer's study and are irrelevant to this study's topic.

2. Analyst removes rows with missing values and convert several data types from string into numeric data.

3. Analyst create a column called "Car class" based on car price values. By comparing each car price from dataset to average car price of dataset , analyst is able to prepare data points into two car class categories: economy and luxury. Price under $13207 will be economy class while price over $13207 will be luxury class.

Analysis

Exploratory Analysis

	4wd	fwd	rwd	Sum
Convertible	0	1	5	6
Hardtop	0	1	6	7
Hatchback	1	40	15	56
Sedan	2	48	26	76
Wagon	3	12	5	20
Sum	6	102	57	1

Table 1. Standard Type Aspiration Car Summary

	4wd	fwd	rwd	Sum
Convertible	0	0	0	0
Hardtop	0	0	1	1
Hatchback	0	9	3	12
Sedan	1	7	10	18
Wagon	1	0	4	5
Sum	2	16	18	36

Table 2. Turbo Type Aspiration Car Summary

	4wd	fwd	rwd	Sum
Convertible	0	1	5	6
Hardtop	0	1	7	8
Hatchback	1	49	18	68
Sedan	3	55	36	94
Wagon	4	12	9	25
Sum	8	118	75	201

Table3. Total Car Summary on drive-wheel and body-style

Table 1 to 3 collection demonstrates data points distributions in three aspects:body style,drive wheels and aspiration. In dataset, there are more standard (165) than turbo (36) for aspiration type cars,which constitutes 82.1% of dataset.In addition, there are more front wheel drive (118) than four wheel drive (8) and rear wheel drive (75)for drive wheels type cars,which constitutes 58.7\% of dataset. Finally, there are more hatchback (68) and sedan (94) than other body styles cars,which constitute 33.8% and 46.8% of dataset.

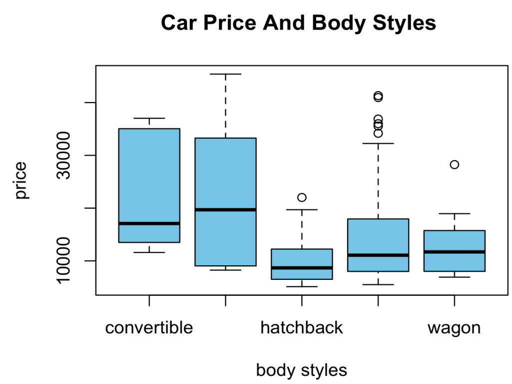
Following on summary of dataset, analyst is able to find some significant data on price variable. Below is the full table of summary on price variable.

	Price
Min.	5118
1st Qu.	7775
Median	13207
Mean	10295
3rd Qu.	16500
Max	45400
Sd.	7947.066

Table4. Price summary

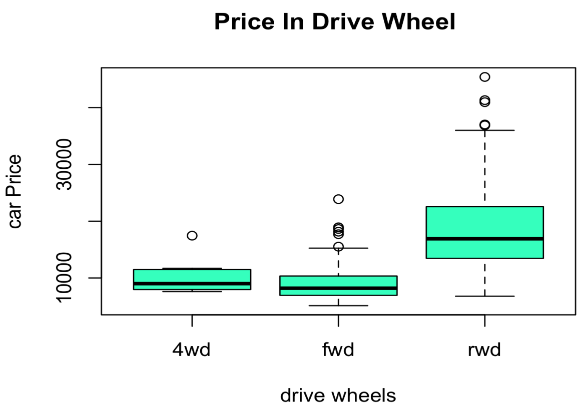
(Plot 1).Car Price Distribution

Based on the Plot 1, we are able to view the price distribution from the dataset.Blue line indicates the average car price ($13207) from the dataset, while purple line indicates the median car price ($10295) from the dataset.The distribution is right skewed with IQR=8725 (Q1=7775 and Q3=16500).



(Plot 2). Car Price and Body-Style Boxplot

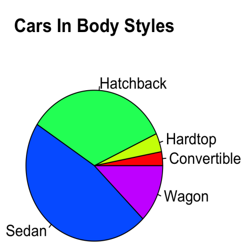
From Plot 2, analyst uses boxplot to demonstrate the price distribution in terms of body styles. There are five total body styles. The box plot shows the convertible and hardtop cars have relatively higher mean car price while they are also in very wide price range. In addition, sedan cars also has really wide price range and many outliers. Overall, this graph suggests convertible (mean=17084.5)and hardtop (19687.5)are relatively expensive than other cars, while the data summary suggests sedan (count=94) and hatchback (count=68) are the top two most used cars. However, analysis needs to have further analysis to testify if above estimations are true.



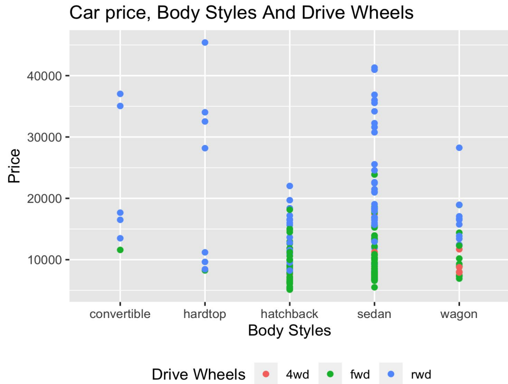
(Plot3).Car Price and Drive-Wheel Boxplot

The Plot 3 displays the price distribution in terms of drive wheels. There are three types of drive wheels: four wheel drive, front wheel drive and rear wheel drive. From the boxplot above, analyst finds the rear wheel drive cars have relatively higher price (mean=16900) than the other two types of drive wheels.From the data summary, analyst find there are more front wheel drive cars(count=118) than four wheel drive(count=8) and rear wheel drive(count=75) cars. Therefore, due to price differences between rear wheel drive and other two types of drive wheels,analyst combines the four wheel drive and front wheel drive cars to compare with rear wheel drive for further statistical analysis.

From Plot 4 on the left, analyst is able to find how car distributed in body styles in the datasets. Out of 201 cars, there are 6 convertible cars, 8 Hardtop cars,68 Hatchback cars, 94 Sedan cars, and 25 Wagon cars.



(Plot4).Cars in body-style



(Plot5).Car Price with body-style and drive-wheel

From above Plot 5, analyst demonstrates the car price distribution in terms of body styles and drive wheels. It shows the rear wheel drive cars have relatively higher price than other types of drive wheels. Most of expensive cars,regardless of which body styles, tend to have rear wheel drives than other two types of drive wheels. Convertible, hardtop and sedan cars are three car types that all have enormous fluctuation in price ranges,while hatchback and sedan cars are the top two most used cars.

Statistical Analysis

There are 5 tests for this study on price distribution in terms of car characteristics.

Test 1:One Sample T-Test

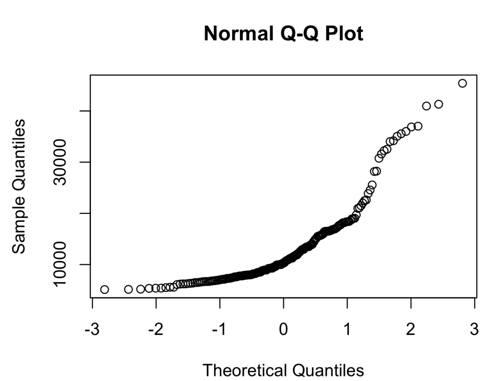
Question of Interest:
What is the average car price for all car population in 1985?

Statistic Test: One sample T-test approach.

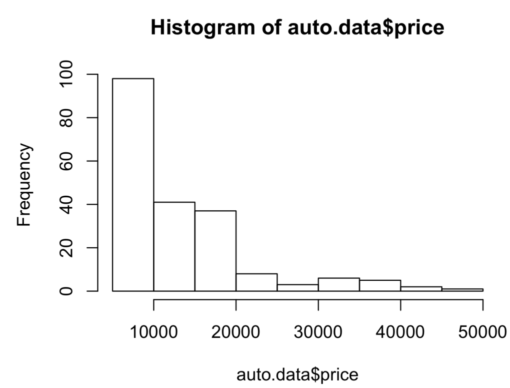
- **Explanation:** This question of interest focuses on the average car price in whole car population in 1985. From the previous research on average car price in the 1980s (Pearson), the estimated true mean car price is known, which is $7000, while the population variance of car price is unknown. Therefore, the best choice of this problem is one sample t-test.

Condition and Requirement

- Is the sample representative of the population?
  - The sample cannot surely confirm to be the representative of the population.Though it is from reliable source UCI Machine Learning Repository that complied by Jeffrey C. Schlimmer, dataset cannot be fully representative to the population due to limited resources.
- One quantitative variable of interest?
  - The car price from the dataset is quantitative variable.
- We want to make inferences about the population mean using the sample mean?
  - Yes.
- The population variance is unknown so we estimate it using sample data?
  - Yes.
- The sample comes from a single population?
  - Yes.
- The population data must be normally distributed?
  - To check this condition look at a QQ plot of the sample data.
  - Sample size is 201 which is larger than 30.



(Plot 6). Total sample Q-Q plot



(Plot 7). Total sample histogram

In Q-Q plot, we were expected to see that the data fell along the line with most points concentrated in the center of the line and fewer points towards the ends.The Q-Q plot shows there are more points concentrated in the middle and less towards the edges, but this is far from a straight line. And histogram also suggest the price distribution is right skewed.The normality assumption is not met and we should be concerned about the reliability of our p-value and confidence intervals.

Parameter

The population parameter we want to make inferences to is average car price μ in 1985.

Hypotheses

Null Hypothesis: H_0 : $\mu = 7000$

- The true mean car price in whole car population in 1985 is $7000.

Alternative Hypothesis*:* H_A : $\mu \neq 7000$

- The true mean car price in whole car population in 1985 is different than $7000.

Sample Statistic: The sample statistic is the sample mean car price $\bar{x}$.

Test Statistic:

From research, we find estimate true mean for car price is $7000 in 1980s,we will use it to compare with our sample.When we do not know the population variance for car price and we have to estimate it with the sample mean, the reference distribution of our test statistic shifts to a t-distribution. The shape of a t-distribution depends on the sample size.As n approaches infinity the t-distribution approaches the normal distribution.

Distribution of the test statistic

$$t_{n-1} = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$$

P Value: Two-sided based on alternative hypothesis.

```
## [1] 1.553377e-22
```

Confidence Interval:

$$P(\bar{x} - t_{(\frac{\alpha}{2},n-1)}\frac{s}{\sqrt{n}} < \mu < \bar{x} - t_{(\frac{\alpha}{2},n-1)}\frac{s}{\sqrt{n}}) = (1-\alpha)$$

- Lower Bound:

```
## [1] 12101.8
```

- Upper Bound:

```
## [1] 14312.46
```

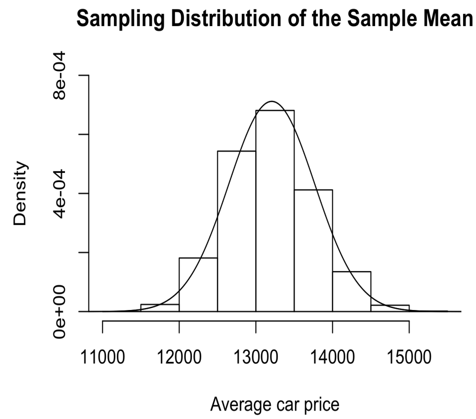
Comparison with R built-in T test

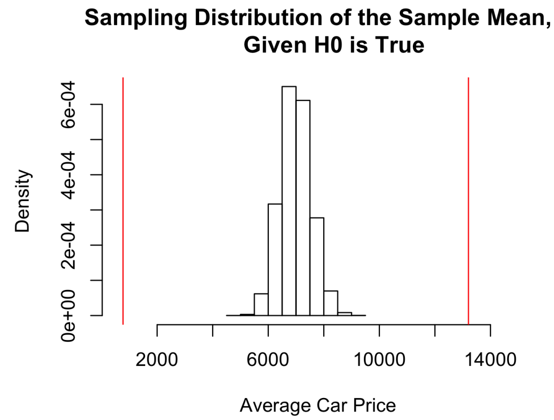
```
##
##  One Sample t-test
##
## data:  auto.data$price
## t = 11.073, df = 200, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 7000
## 95 percent confidence interval:
##  12101.80 14312.46
## sample estimates:
## mean of x
##  13207.13
```

Following comparison with R built-in T test, we find our result is fairly close to R built-in T test result.

Bootstrap Approach:

Due to the right skewed distribution, we need to simulate much more times. So we set 1000 times of simulation for sampling distribution of the sample mean. The graph Plot 8 shows that after simulation, sampling distribution is really close to normal distribution.





(Plot 8).Sample Distribution of the Sample Mean (Plot 9). Sampling Distribution of the Sample Mean,H0 True

To show the sampling distribution to be true under the null hypothesis, then the true mean car price is $7000. We need to shift distribution to satisfy the null hypothesis. Plot 9 shows the sampling distribution of the sample mean, giving H_0 is true. To satisfy the null hypothesis is true, sample mean is $7000, while two red lines indicate two extreme values. Upper end red line shows the average mean from the original dataset.

- **Bootstrap P value**

```
## [1] 0
```

Comparing to traditional one sample t test result, our p values of both approach is really close.

- **Bootstrap confidence interval**

```
## [1] 12096.17 14318.08
```

Interpretation:

There is strong evidence (p-value=1.553377e-22) to suggest that the true mean car price in 1985 is different than $7000. We reject the null hypothesis that the true mean car price in 1980s is $7000 at the level $\alpha = 0.05$. With 95% confidence, the true mean car price in 1985 is between

\$12101.8 and \$14312.46 which suggests that the true mean car price in 1985 is greater than \$7000.The interval of bootstrap approach is wider which agrees with our p-value being a bit more conservative. In this case,the traditional t-tools were making our results a bit more significant and our confidence intervals narrower than what we found using the empirical method.

Test 2:One Sample T-Test for Proportion

Question of Interest:
Is the rear wheel drive cars evenly distributed with the other two types wheel drive cars for all car population in 1985?
Statistic test: One Sample One Proportion T-Test approach.

- Explanation: This question of interest focuses on the proportion of rear wheel drive cars in whole car population in 1985. We have a categorical variable with three categories (front wheel, four wheel , rear wheel) but we can group the front wheel drive and four wheel drive cars together. So that instead of front wheel drive,four wheel drive and rear wheel drive, we have rear wheel drive and non-rear wheel drive cars.As we interested in the proportion of rear wheel drive cars, so the best choice of this test is one sample one proportion t test.

Condition and Requirement

- Exact Binomial Test?
  - No requirements
- Normal approximation?
  - $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$
  - Yes. $(75)\frac{75}{201} = 27.98 > 10$ and $(201-75)\frac{201-75}{201} = 78.99 > 10$

Condition and requirement satisfied for both exact binomial test and normal approximation.

Parameter

The population parameter we want to make inferences to the population proportion of rear wheel drive cars in 1985, p_r .

Hypotheses

Null Hypothesis: $H_0 : p_r = 0.333\overline{3}$

- The true proportion of rear wheel drive cars in 1985 is 33.33%.

Alternative Hypothesis: $H_A : p_r > 0.333\overline{3}$

- The true proportion of rear wheel drive cars in 1985 is greater than 33.33%.

Sample Statistic The sample statistic is $\hat{p} = \frac{75}{201} = 0.3731$

Test Statistic

- Exact test: there is no test statistic, find the probability directly.
- Normal approximation (using p_0 to find the test statistic)

$$z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.3731-0.3333}{\sqrt{\frac{0.3333(1-0.3333)}{201}}} = 1.1970$$

Distribution of the test statistic

$$z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0, 1)$$

P value

- Exact Test:

```
## data:  75 out of 201
## number of successes = 75, number of trials = 201, p-value = 0.1313
## alternative hypothesis: true probability of success is greater than
0.3333333
## 95 percent confidence interval:
```

```
##  0.3162255 1.0000000
## sample estimates:
## probability of success
##               0.3731343
```

- One Sided Upper Normal Approximation

```
## [1] 0.1156533
```

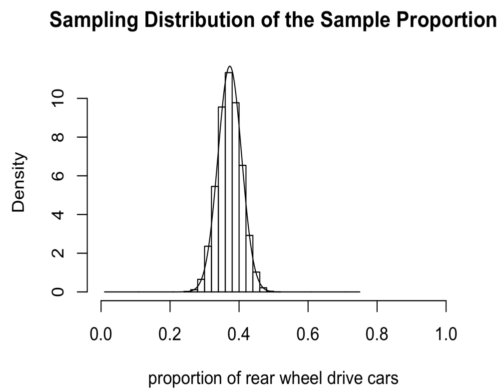
Confidence interval

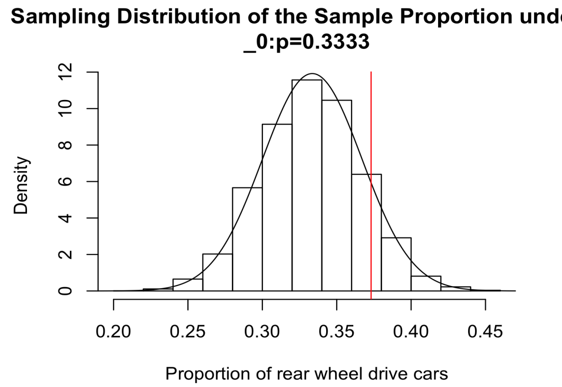
One-Sided Upper Confidence Interval - Exact: from binom.test - Normal Approximation: using the Wald statistic

```
## exact binomial test

## [1] 0.3162255 1.0000000
## attr(,"conf.level")
## [1] 0.95
## attr(,"method")
## [1] "Score"

## normal approx

## [1] 0.3171555 1.0000000
```

Bootstrap Approach

We set 1000 times of simulation for sampling distribution of the sample proportion. The graph below Plot 10 shows that after simulation, sampling distribution is really close to normal distribution.To show the sampling distribution of sampling proportion to be true under the null hypothesis,then the $H_0 : p = 0.3333$, we need to shift distribution to satisfy the null hypothesis. The graph Plot 11 shows it.





(Plot10).Sampling Distribution of Sample Proportion (Plot11).Sampling Distribution of Sample Porp.under H0 true

Bootstrap Confidence Interval

```
##        5%      100%
## 0.3184080 0.5124378
```

Bootstrap P Value

```
## [1] 0.1319
```

Interpretation :

There is very weak evidence (p-value=0.1313) that the true proportion of rear wheel drive cars is greater than .33333. We cannot reject the null hypothesis that the true proportion of rear wheel drive cars is equal to .3333 at significant level $\alpha = 0.05$. The true proportion of rear wheel drive cars in the population of rear wheel drive, front wheel drive and four wheel drive cars is between 0.3162255 and 1. Our bootstrapped p-value is closer to the exact binomial and it is more conservative than the normal approximation.

Test 3:Two Sample T-Test for Difference In Means

Question of Interest:

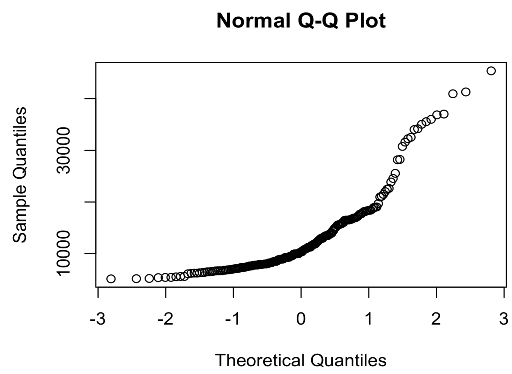
Is there a difference in price between turbo cars and standard cars?

Statistic test: Two sample T-test for difference in means approach.

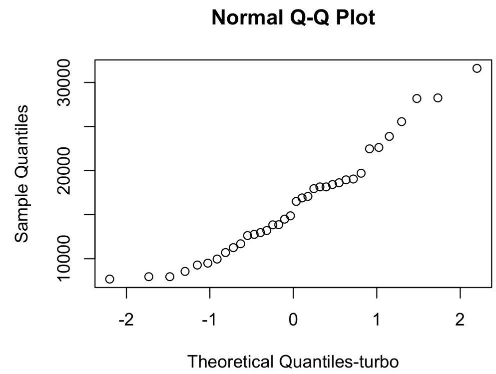
- Explanation: This question of interest focuses on the comparison of turo cars and standard cars as two populations in terms of price.The true mean car price as well as the true population of each group are unknown. Therefore, the best choice of this problem is one sample t-test.

Condition and Requirement

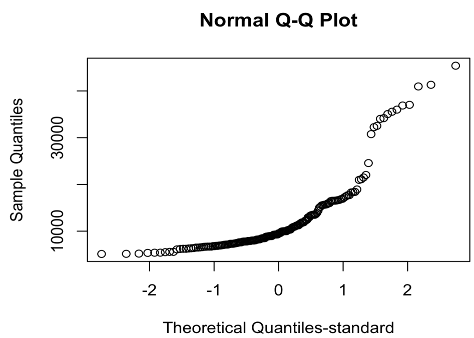
- ❖ Is the sample representative of the population?
  - ➢ The sample cannot surely confirm to be the representative of the population.
- ❖ Question of interest has to do with the difference of means between two populations?
  - ➢ Yes, turbo cars and standard cars populations and the difference in the mean price for each population.
- ❖ 2 independent samples from 2 populations?
  - ➢ Yes, turbo cars and standard cars can be considered as two different populations.
- ❖ The population data must be normally distributed?
  - ➢ To check this condition look at a QQ plot of the sample data.
  - ➢ Sample sizes are 36 for turbo cars and 165 for standard cars, which are larger than 30.



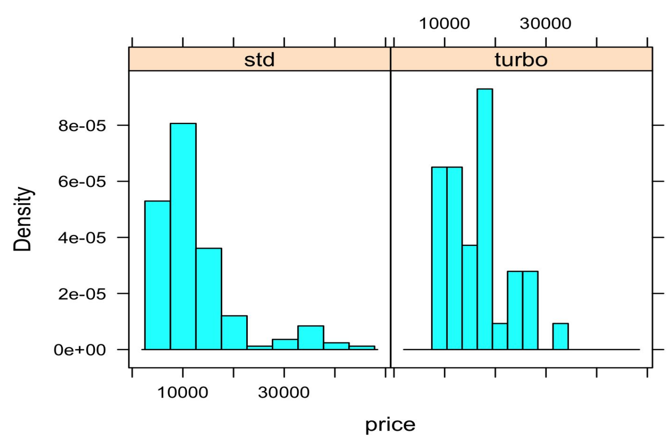
(Plot12).Total Sample Q-Q Plot



(Plot13).Turbo Aspiration Sample Q-Q Plot



(Plot 14).Standard Aspiration Sample Q-Q Plot



(Plot 15).Standardand Turbo Aspiration Sample Distribution

From Q-Q plots and histograms above, we find turbo samples' Q-Q plot is somewhat close to normal distribution while standard sample's Q-Q plot is not normally distributed but right skewed.Both histograms confirmed that both samples are not normally distributed. The normality assumption is not met and we should be concerned about the reliability of our p-value and confidence intervals.

Parameter:

We are interested in the true population mean difference in price between turbo cars and standard cars $\mu_t - \mu_s$

Hypotheses:

$H_0 : \mu_t - \mu_s = 0$

- The true population mean price for turbo cars is equal to the true population mean price for standard cars.

$H_A : \mu_t - \mu_s \neq 0$

- The true population mean price for turbo cars is different than the true population mean price for standard cars.

Sample Statistic

$$\overline{x_t} - \overline{x_s}$$

Test Statistic

$$t_{min(n_t-1, n_s-1)} = \frac{(\overline{x_t} - \overline{x_s}) - (\mu_t - \mu_s)}{\sqrt{\frac{s_t^2}{n_t} + \frac{s_s^2}{n_s}}}$$

Distribution of the test statistic

$$t_{min(n_t-1, n_s-1)} = \frac{(\overline{x_t} - \overline{x_s}) - (\mu_t - \mu_s)}{\sqrt{\frac{s_t^2}{n_t} + \frac{s_s^2}{n_s}}} \sim t_{min(n_t-1, n_s-1)}$$

P-Value for Two sided

```
## [1] 0.004128058
```

Confidence Interval

```
## [1] 1257.027

## [1] 5328.797
```

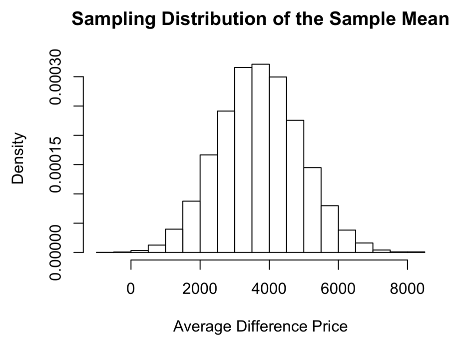
Comparison with Built In T-Test:

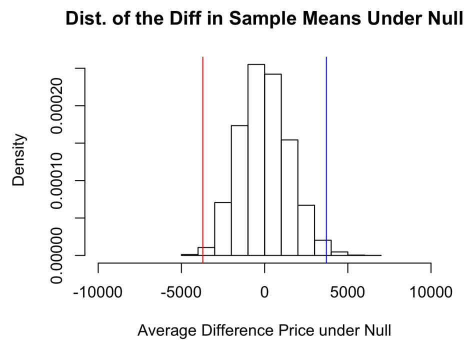
```
##  Welch Two Sample t-test
## data:  auto.data$price[auto.data$aspiration == "turbo"] and
auto.data$price[auto.data$aspiration == "std"]
## t = 3.0693, df = 64.602, p-value = 0.003135
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1296.625 6128.623
## sample estimates:
## mean of x mean of y
##  16254.81  12542.18
```

Comparing to the results of build in T-test, our traditional test is really close to them.

Bootstrap approach:

Due to the skewed distributions for both samples, we need to simulate much more times. So we set 1000 times of simulation for sampling distribution of the sample mean. The graph Plot 16 below shows that after simulation, sampling distribution is really close to normal distribution.





(Plot 16). Sampling Distribution of Difference in Sample Mean (Plot17).Dist. of the Diff in Sample Mean Under H0 True

Graph Plot 17 shows the distribution of the difference in sample mean, giving H_0 is true. To satisfy the null hypothesis is true, the difference in sample mean should be zero. Red line and blue line indicate two extreme values.

Bootstrap P-Value

```
## [1] 0.0096
```

Bootstrap Confidence Interval

```
##     2.5%    97.5%
## 1451.430 6088.783
```

Interpretation:

From the computation, analyst finds that there is strong evidence (p-value=0.004128058) to suggest that the true mean price for turbo cars is different than standard cars. We reject the null hypothesis that there is no difference between true population mean price for turbo cars at the level $\alpha = 0.05$. With 95% confidence, the true difference between the mean price between turbo cars and standard cars is between 1257.027 and 5328.797. The null hypothesized difference between the mean price is zero and zero is not in the 95% confidence interval which is consistent with our rejection of the null hypothesis and the values of the confidence interval suggest that on average turbo cars have higher prices than standard cars.Our p-value from traditional test is a bit more conservative compared to bootstrap approach. In this case the traditional t-tools were making our results a bit more significant and our confidence intervals narrower than what we found using the empirical method.

Test 4: Two Sample T-Test for Difference In Proportions

Question of Interest:
How turbo aspiration type distributed in luxury class compared to economy class in 1985?
Statistic test: Two Sample T-test for Difference In Proportions approach.

- Explanation: This question of interest focuses on the comparison of luxury and economy class as two populations in terms of turbo aspiration proportion in each population.The true difference in proportion of turbo cars in luxury class cars population and economy class cars population is what we seek. Therefore, the best choice of this problem is two sample t-test for difference in proportions approach.

Condition and Requirement

- ❖ Is the sample representative of the population?
  - ➢ The sample cannot surely confirm to be the representative of the population.
- ❖ Question of interest has to do with the difference of proportions between two populations?
  - ➢ Yes, Luxury and economy cars as two populations and the true difference in the proportion of turbo type aspiration for each population.
- ❖ Categorical response variable with 2 categories?
  - ➢ Yes.
- ❖ 2 independent samples from 2 populations?
  - ➢ Yes, luxury cars and economy cars can be considered as two different populations.
- ❖ $np \geq 10$ and $n(1-p) \geq 10$ for both populations?
  - ➢ Yes.

Parameter

We are interested in the difference between the true population proportion of turbo type from the luxury cars and true population proportion of turbo type from the economy cars, $p_l - p_e$.

Hypotheses

$H_0 : p_l - p_e = 0$

- There is no difference between the true population proportion of turbo type from the luxury cars and true population proportion of turbo type from the economy cars

$H_A : p_l - p_e \neq 0$

- There is a difference between the true population proportion of turbo type from the luxury cars and true population proportion of turbo type from the economy cars.

Sample Statistic

$$\hat{p}_l - \hat{p}_e$$

Test Statistic

$$z = \frac{(\hat{p}_l - \hat{p}_e) - (p_l - p_e)}{\sqrt{\frac{\hat{p}_l(1-\hat{p}_l)}{n_l} + \frac{\hat{p}_e(1-\hat{p}_e)}{n_e}}}$$

Distribution of the test statistic

$$z = \frac{(\hat{p}_l - \hat{p}_e) - (p_l - p_e)}{\sqrt{\frac{\hat{p}_l(1-\hat{p}_l)}{n_l} + \frac{\hat{p}_e(1-\hat{p}_e)}{n_e}}} \sim N(0, 1)$$

P value

```
## [1] 0.00181061
```

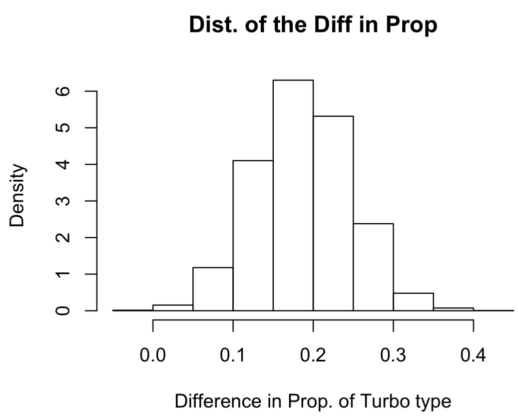
Confidence interval

```
## [1] 0.06953763
```

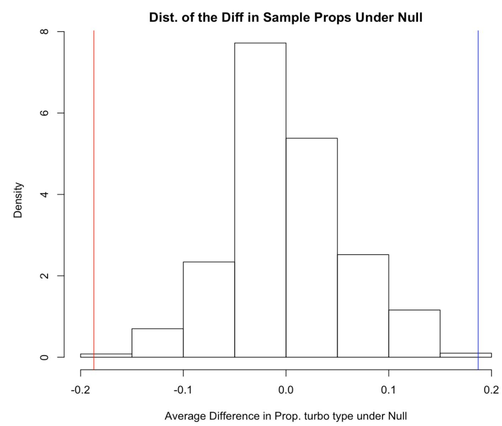
```
## [1] 0.3045845
```

Bootstrap Approach

Graph Plot 18 shows that after simulation, sampling distribution of difference in proportion is really close to normal distribution.But if the null hypothesis is true, then the true difference in proportions between the two groups is 0. We need to create a randomization distribution to satisfy the null hypothesis. The Plot 19 shows the distribution of difference in sample proportion given null hypothesis is true.



(Plot 18). Dist. of the Diff in Prop of sample



(Plot 19).Dist. of the Diff in Sample Props under H0 True

Bootstrap P-Value

```
## [1] 0.001
```

Bootstrap Confidence Interval

```
##       2.5%      97.5%
## 0.07224941 0.30527772
```

Interpretation:

Using randomization methods, there is strong evidence (p-value = 0.00181061) to suggest that there is a difference in the true proportion of turbo type between luxury cars and economy cars. We reject the null hypothesis that the true proportion of turbo type from luxury class is equal to the true proportion of turbo cars from economy class at the level $\alpha = 0.05$. Using confidence intervals created by the bootstrap method, we can say with 95% confidence that the true population proportion difference is between 0.07224941 to 0.30527772. The null hypothesized difference of 0 is not in the confidence interval which agrees with our rejection of the null hypothesis.

Test 5: Chi-Square Goodness of Fit Test

Question of Interest:

From the body-styles convertible, hardtop, hatchback, sedan, and wagon,are certain body style more preferred than others in 1985?

	Count	Percentage of Sample
Convertible	6	2.985075%
Hardtop	8	3.980100%
Hatchback	68	33.830846%
Sedan	94	46.766169%
Wagon	25	12.437811%

(Table5).Body-Style Summary

In Table 5, analyst is able to find each body style count and percentage of the dataset.

Parameter of interest

We are interested in the true proportion of each body style of cars, $p_c, p_{har}, p_{hat}, p_s, p_e$.

Statistic test: Chi-square Goodness of Fit Test

- Explantation: The body style is single categorical variable with 5 categories.The only right choice for this problem is Chi-square Goodness of Fit Test.

Condition and requirement

- ❖ Single categorical variable with more than 2 categories?
  - ➢ Yes
- ❖ The expected count of each count is at least 5?
  - ➢ Yes

Hypotheses

$H_0 : p_c = p_{har} = p_{hat} = p_s = p_e = 0.2$

- The proportion of each body style is the same and is equal to 0.2.

H_A : some $p_i \neq 0.2$

- At least one of the proportions is not equal to 0.2.

Under the null hypothesis, the expected count $np_i = 201 \times 0.2 = 40.1$

Sample Statistics

In this example we have 5 sample statistics.

$$\hat{p}_c, \hat{p}_{har}, \hat{p}_{hat}, \hat{p}_s, \hat{p}_e$$

Test Statistic and Distribution

$$\chi^2 = \sum_{i=1}^{k} \left(\frac{(O_i - E)^2}{E}\right) \sim \chi^2_{k-1}$$

Chi-square Value

```
## [1] 152.2406
```

P value

```
## [1] 6.738178e-32
```

Interpretation:

The data provides strong evidence (P=6.738178e-32) that some of the proportions of bodystyle is different than 0.2. We reject the null hypothesis that the proportions of the body styles are all equal to 0.2 at the level.

Discussion

Summary of Findings:

Based on the statistical analysis, at the $\alpha = 0.05$ and with a 95% confidence, this study

- Displays that rear wheel drive cars had relatively higher price than other drive wheels cars in 1985.
- Displays that convertible and hardtop had relatively higher price than other body style cars in 1985.
- Suggests that the average car price was greater than estimate mean price $7000 in 1980s.
- Shows the proportion of rear wheel drive cars was equal to the other two drive wheel cars.
- Indicates price of turbo aspiration car was significantly higher than standard aspiration cars.
- Presents turbo aspiration car took a greater percentage in luxury class population than economy class population.
- Displays some proportion of body styles of cars was different.

Implication of Findings:

These findings implicate that in 1985, the average car price had increased from previous years that it was greater than estimate mean price in the 1980s. The proportion of rear wheel drive, non-rear wheel drive are not significant different in the whole car proportion. For aspiration types, findings shows turbo aspiration cars had significantly higher price than standard aspiration cars. In addition, more luxury cars were more likely to have turbo aspiration type compared to economy cars. Finally, the availability of each body style in the market is also different. It was possible that market and customers had preferences on car body styles,which affected market demand.

Extensions, Limitations and Further Questions

This study focuses on the automobile's price and characteristics. The purpose of study is to provide statistical data as suggestions for car customer in 1985. Analyst can also find the trend of car features in 1985 to compare to what car features are most popular nowadays. From the findings on exploratory analysis, rear wheel cars were more likely to have higher prices, and convertible and hardtop had relatively higher price than other body style cars in 1985. In statistical analysis , this study finds that aspiration type was one of the factors affected automobile price.Therefore, these findings can be used as suggestions to car customers in 1985. Compared to car market, some of these findings still stands. Sports cars, which are turbo and rear wheel drive mostly, seem more likely to cost more than other non-turbo and non-rear wheel drive cars. But the comparison from 1985's car market and today's car market needs to be further studied to confirm above assumption.

These findings on automobile price, and its characteristics in 1985 are highly based on dataset from Jeffrey C. Schlimmer on limited sources.This study may have limitations due to confounding factors, potential biases environment, and limited sources so that this study could be biased.

These findings cannot fully support conclusions on causation but can be discussed with several further steps:

- Have a larger scope on the sample and more sources on our sampling selection to ensure the coverage of representative to population.

- Create a well-developed sampling strategy designed for this topic to limit sampling errors.
- Rule out potential confounding factors for the study.

By having these modifications, This study can be further extended on price comparisons with car's drive wheels, body styles and other undiscussed characteristics like fuel types, horsepowers, mpgs, and etc.

References

Aust, F., & Barth, M. (2018). papaja: Create APA manuscripts with R Markdown. Retrieved from https://github.com/crsh/papaja

Pearson, S. (n.d.). 1980's Collector Cars including Prices. Retrieved from http://www.thepeoplehistory.com/80scars.html.

Schlimmer , J. C. (n.d.). Retrieved from http://archive.ics.uci.edu/ml/datasets/Automobile.

Appendix

```
# Install devtools package if necessary
#if(!"devtools" %in% rownames(installed.packages()))
install.packages("devtools")

# Install the stable development verions from GitHub
#devtools::install_github("crsh/papaja")

# Install the latest development snapshot from GitHub
#devtools::install_github("crsh/papaja@devel")
library("papaja")
knitr::opts_chunk$set(echo = FALSE)# Seed for random number generation

set.seed(0)
knitr::opts_chunk$set(cache.extra = knitr::rand_seed)

library(ggplot2)
library(dplyr)          # data manipulation
library(mosaic)
library(Sleuth3)
library(dplyr)
auto.data <- read.csv( "~/Documents/IE6200/Final/imports-85.data.csv",header
= TRUE, stringsAsFactors = FALSE)
auto.data
summary(auto.data)
price<-select(auto.data,price)
summary(price)
#Get the mean, median, minimum, maximum and quartiles
lapply(price, sd)

table1to3<-table(auto.data$body.style,auto.data$drive.wheels,auto.data$aspira
tion)
addmargins(table1to3)
#Car Price Distribution
Plot1 <- ggplot(auto.data, aes(price)) + geom_density()+ggtitle("Car Price
Distribution")+geom_vline(xintercept = mean(auto.data$price),
                color = "skyblue", size=1)+geom_vline(xintercept =
median(auto.data$price),
                color = "#cc99ff",
size=1)+geom_text(aes(x=median(auto.data$price), label="\nMedian",
y=0.00001), colour="#cc99ff", angle=60) +
  geom_text(aes(x=mean(auto.data$price), label="\nMean", y=0.00001),
colour="skyblue", angle=60 )
Plot1
```

```
# Bodystyle and price
Plot2<-boxplot(auto.data$price~auto.data$body.style, ylab = 'price', xlab = 
'body styles', main = 'Car Price And Body Styles ', col = 'skyblue')
Plot2
# Aspiration And Price
Plot3<-boxplot( auto.data$price~auto.data$drive.wheels, ylab = 'car Price', 
xlab = 'drive wheels', main = 'Price In Drive Wheel ', col = "#33ffc9")
Plot3
#count body style
bs_count<-c(6,8,68,94,25)
Plot4<-pie(bs_count, main="Cars In Body Styles", 
col=rainbow(length(bs_count)),
   labels=c("Convertible","Hardtop","Hatchback","Sedan","Wagon"))
Plot4
auto.data$carclass<- ifelse(auto.data$price >7000,"luxury","economy")
# Car price, Body-styles and Drive wheels
ggplot(auto.data, aes(x = body.style, y = price,color = drive.wheels)) + 
geom_point() +ggtitle("Car price, Body Styles And Drive Wheels")+labs(x = 
"Body Styles", y = "Price", color = " Drive Wheels")+ theme(legend.position = 
"bottom")
#Combine to do comparison rwd+fwd+4wd -one sample proportion
auto.data.rwd<-filter(auto.data, auto.data$drive.wheels == "rwd")
auto.data.fwd<-filter(auto.data, auto.data$drive.wheels == "fwd")
auto.data.4wd<-filter(auto.data, auto.data$drive.wheels == "4wd")
auto.data.fwd<-rbind(auto.data.4wd,auto.data.fwd)
library(dplyr)
set.seed(0)
qqnorm(auto.data$price)
hist(auto.data$price)
# sample mean
x_bar <- mean(auto.data$price)
# null hypothesized population mean
mu_0 <- 7000
# sample st. dev
s <- sd(auto.data$price)
# sample size
n <- length(auto.data$price)
# t-test test statistic
t <- (x_bar - mu_0)/(s/sqrt(n))
# two-sided p-value
two_sided_t_pval <- pt(q = t, df = n-1, lower.tail = FALSE)*2
two_sided_t_pval
# lower bound
test1_lb<-x_bar+(qt(0.025, n-1)*(s/sqrt(n)))
test1_lb
# upper bound
test1_ub<-x_bar+(qt(0.975, n-1)*(s/sqrt(n)))
```

```
test1_ub
t.test(auto.data$price,
alternative = "two.sided",
mu =7000)
set.seed(0)
num_sims <- 10000
# A vector to store my results
results <- rep(NA, num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
results[i] <- mean(sample(x = auto.data$price,
size = n,
replace = TRUE))
}
# Finally plot the results
hist(results, freq = FALSE, main='Sampling Distribution of the Sample Mean',
xlab = 'Average car price', ylab = 'Density', ylim=c(0,0.0008))
# estimate a normal curve over it - this looks pretty good!
lines(x = seq(10000, 30000, .1), dnorm(seq(10000, 30000, .1), mean = x_bar,
sd =  7947.066/sqrt(201)))
price<-select(auto.data,price)

set.seed(0)
# Shift the sample so that the null hypothesis is true
time_given_H0_true <-auto.data$price-mean(auto.data$price)+mu_0


# This data is pretty skewed so even though n is large, I'm going to do a lot
of simulations
num_sims <- 10000
# A vector to store my results
results_given_H0_true <- rep(NA, num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
results_given_H0_true[i] <- mean(sample(x = time_given_H0_true,
size = n,
replace = TRUE))}
# Finally plot the results
hist(results_given_H0_true, freq = FALSE, main='Sampling Distribution of the
Sample Mean,
     Given H0 is True', xlab = 'Average Car Price', ylab = 'Density',
xlim=c(600,15000))
# add line to show values more extreme on upper end
abline(v=x_bar, col = "red")
# add line to show values more extreme on lower end
low_end_extreme <-
mean(results_given_H0_true)+(mean(results_given_H0_true)-x_bar)
```

```
abline(v=low_end_extreme, col="red")
# counts of values more extreme than the test statistic in our original
sample, given H0 is true
# two sided given the alternate hypothesis
count_of_more_extreme_lower_tail <- sum(results_given_H0_true <=
low_end_extreme)
count_of_more_extreme_upper_tail <- sum(results_given_H0_true >= x_bar)
bootstrap_pvalue <- (count_of_more_extreme_lower_tail +
count_of_more_extreme_upper_tail)/num_sims
bootstrap_pvalue
set.seed(0)
bootstrap_SE_X_bar <- sd(results)
# an estimate is to use the formula statistic +/- 2*SE
c(x_bar - 2*bootstrap_SE_X_bar, x_bar + 2*bootstrap_SE_X_bar)
binom.test(x=75, n = 201, p=(1/3), alternative="greater")
pnorm(1.1970, lower.tail = FALSE)
cat("exact binomial test")
binom.test(x=75, n = 201, p=(1/3), alternative="greater")$conf.int
cat("normal approx")
c(0.3731 - (1.64)*sqrt(((0.3731)*(1-0.3731))/201), 1)
set.seed(0)
rwd <- rep(c(1, 0), c(75, 201-75))
table(rwd)
num_sims <- 10000
# A vector to store my results
results <- rep(NA,num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
results[i] <- mean(sample(x=rwd,
size = 201,
replace = TRUE))
}
# Finally plot the results
hist(results,freq=FALSE,main='Sampling Distribution of the Sample
Proportion',xlab = 'proportion of rear wheel drive cars',ylab =
'Density',xlim=c(0,1))
# estimate a normal curve over it - this looks pretty good!
lines(x=seq(.01,.75,.001),dnorm(seq(.01,.75,.001),mean=mean(results),sd=sd(re
sults)))

cat("Bootstrap Confidence Interval")
c(quantile(results, c(.05, 1)))
cat("Exact Binomial test")
binom.test(x=75, n = 201, p=(.5), alternative="greater")$conf.int
cat("normal approx")
c(0.3731 - (1.64)*sqrt(((0.3731)*(1-0.3731))/201), 1)
# Under the assumption that the null hypothesis is true, we have 50% peanuts
```

```
rwd <- rep(c(1, 0), c(67, 201-67))
num_sims <- 10000
# A vector to store my results
results <- rep(NA, num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
results[i] <- mean(sample(x = rwd,
size = 201,
replace = TRUE))
}
# Finally plot the results
hist(results, freq = FALSE, main='Sampling Distribution of the Sample
Proportion under H
_0:p=0.3333', xlab = 'Proportion of rear wheel drive cars', ylab = 'Density')
# estimate a normal curve over it - this looks pretty good!
lines(x = seq(.1, .75, .001), dnorm(seq(.1, .75, .001), mean = mean(results),
sd = sd
(results)))
abline(v=0.3731, col="red")
count_of_more_extreme_upper_tail <- sum(results >= .3731)
bootstrap_pvalue <- count_of_more_extreme_upper_tail/num_sims
cat("Bootstrap p-value")
bootstrap_pvalue
cat("Exact Binomial p-value")
binom.test(x=75, n = 201, p=(1/3), alternative="greater")$p.value
cat("Normal Approximation p-value")
pnorm(1/3, lower.tail = FALSE)
turbo<-filter(auto.data,auto.data$aspiration=="turbo")
std<-filter(auto.data,auto.data$aspiration=="std")
qqnorm(auto.data$price)
qqnorm(auto.data$price[auto.data$aspiration=="turbo"],xlab = "Theoretical
Quantiles-turbo")
qqnorm(auto.data$price[auto.data$aspiration=="std"],xlab = "Theoretical
Quantiles-standard")
histogram(~ price | aspiration, data=auto.data)
# Sample Mean
X_bar_t <- mean(auto.data$price[auto.data$aspiration=="turbo"])
X_bar_s <- mean(auto.data$price[auto.data$aspiration=="std"])
X_bar_t
X_bar_s
# Sample Variance
s_t <- sd(auto.data$price[auto.data$aspiration=="turbo"])**2
s_s <- sd(auto.data$price[auto.data$aspiration=="std"])**2
s_t
s_s
# Sample Size
n_t <- length(auto.data$price[auto.data$aspiration=="turbo"])
n_s <- length(auto.data$price[auto.data$aspiration=="std"])
```

```
n_t
n_s
#Null Hypothises
mu <- 0
# T Test
t <- (X_bar_t - X_bar_s - mu)/sqrt((s_t/n_t) + (s_s/n_s))
t
# p-value for two sided upper
two_sided_diff <- pt(q=t, df = min(n_t, n_s)-1, lower.tail = FALSE) * 2
two_sided_diff
Alpha <- 0.05
Confidence_Interval <- 0.95
# Lower Bound
L_bound <- (X_bar_t - X_bar_s) + (qt(0.025, min(n_t, n_s)-1)* sqrt((s_t/n_t)
+(s_s/n_s)))
L_bound
# Upper Bound
U_bound <- (X_bar_t- X_bar_s) + (qt(0.975, min(n_t, n_s)-1)* sqrt((s_t/n_s)
+(s_s/n_s)))
U_bound
t.test(auto.data$price[auto.data$aspiration=="turbo"],
auto.data$price[auto.data$aspiration=="std"])

set.seed(0)
num_sims <- 10000
# A vector to store my results
results <- rep(NA, num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
mean_turbo <- mean(sample(x = auto.data$price[auto.data$aspiration=="turbo"],
size = 36,
replace = TRUE))
mean_std <- mean(sample(x = auto.data$price[auto.data$aspiration=="std"],
size = 165,
replace = TRUE))
results[i] <- mean_turbo - mean_std
}
# Finally plot the results
hist(results, freq = FALSE, main='Sampling Distribution of the Sample Mean',
xlab = 'Average Difference Price', ylab = 'Density')
c(quantile(results, c(.025, .975)))
# compare to our t-methods
t.test(auto.data$price[auto.data$aspiration=="turbo"],
auto.data$price[auto.data$aspiration=="std"])$conf.int
library(data.table)
a<-copy(auto.data)
transform(a,aspiration=sample(aspiration))
set.seed(0)
```

```
num_sims <- 10000
# A vector to store my results
results_given_H0_true <- rep(NA, num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
# idea here is if there is no relationshipm we should be able to shuffle the
groups
shuffled_groups <- transform(a,aspiration=sample(aspiration))
mean_t <- mean(shuffled_groups$price[shuffled_groups$aspiration=="turbo"])
mean_s <- mean(shuffled_groups$price[shuffled_groups$aspiration=="std"])
results_given_H0_true[i] <- mean_t - mean_s
}
# Finally plot the results
hist(results_given_H0_true, freq = FALSE,
main='Dist. of the Diff in Sample Means Under Null',
xlab = 'Average Difference Price under Null',
ylab = 'Density',xlim = c(-10000,10000))
diff_in_sample_means <- mean(auto.data$price[auto.data$aspiration=="turbo"])
- mean(auto.data$price[auto.data$aspiration=="std"])
abline(v=diff_in_sample_means, col = "blue")
abline(v=-(diff_in_sample_means), col = "red")
count_of_more_extreme_lower_tail <- sum(results_given_H0_true <=
-diff_in_sample_means)
count_of_more_extreme_upper_tail <- sum(results_given_H0_true >=
abs(diff_in_sample_means))
bootstrap_pvalue <- (count_of_more_extreme_lower_tail +
count_of_more_extreme_upper_tail)/num_sims
cat("Bootstrap p-value")
bootstrap_pvalue
t.test(auto.data$price[auto.data$aspiration=="turbo"],
auto.data$price[auto.data$aspiration=="std"])$p.value
auto.data$carclass<- ifelse(auto.data$price >13207,"luxury","economy")
#Comparison between luxury and standard -two sample proportion
#car class classfication
carclass.luxury<-filter(auto.data, auto.data$price > 13207)
carclass.economy<-filter(auto.data, auto.data$price <= 13207)
luxury.standard<-filter(carclass.luxury, carclass.luxury$aspiration=="turbo")
economy.standard<-filter(carclass.economy,
carclass.economy$aspiration=="turbo")

luxuryturbonumber<-count(carclass.luxury$aspiration=="turbo")
economytrbonumber<-count(carclass.economy$aspiration=="turbo")
# sample props
lt<-22
et<-14
p_hat_l <- 22/74
p_hat_e <- 14/127
p_hat_l
```

```
p_hat_e

# null hypothesized population prop difference between the two groups
p_0 <- 0
# sample size
n_l <- 74
n_e <- 127
# sample variances
den_p_l <- (p_hat_l*(1-p_hat_l))/n_l
den_p_e <- (p_hat_e*(1-p_hat_e))/n_e
# z-test test statistic
z <- (p_hat_l - p_hat_e - p_0)/sqrt(den_p_l + den_p_e)
# two sided p-value
two_sided_diff_prop_pval <- pnorm(q = z, lower.tail = FALSE)*2
two_sided_diff_prop_pval
#lower b
(p_hat_l - p_hat_e)+(qnorm(0.025)*sqrt(den_p_l + den_p_e))
# upper bound
(p_hat_l - p_hat_e)+(qnorm(0.975)*sqrt(den_p_l + den_p_e))
set.seed(2)
# Make the data
L <- rep(c(1, 0), c(lt, n_l - lt))
E<- rep(c(1,0), c(et, n_e - et))
num_sims <- 10000
# A vector to store my results
results <- rep(NA, num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
prop_L <- mean(sample(L,
size = n_l,
replace = TRUE))
prop_E <- mean(sample(x = E,
size = n_e,
replace = TRUE))
results[i] <- prop_L - prop_E
}
# Finally plot the results
hist(results, freq = FALSE, main='Dist. of the Diff in Prop', xlab =
'Difference in Prop. of Turbo type', ylab = 'Density')
cat("Bootstrap")
c(quantile(results, c(.025, .975)))
cat("Normal Approximation")
c((p_hat_l - p_hat_e)+(qnorm(0.025)*sqrt(den_p_l + den_p_e)), (p_hat_l -
p_hat_e)+(qnorm
(0.975)*sqrt(den_p_l + den_p_e)))
# Make the data
df_combined <- data.frame("Turbocar" = c(L, E),"group" = rep(c("luxury",
"economy"), c(n_l, n_e)))
```

```
# Sanity checks
summary(df_combined$group)
mean(df_combined$Turbocar[df_combined$group=="luxury"]) == p_hat_l

mean(df_combined$Turbocar[df_combined$group=="economy"]) == p_hat_e

num_sims <- 1000
# A vector to store my results
results_given_H0_true <- rep(NA, num_sims)
# A loop for completing the simulation
for(i in 1:num_sims){
# idea here is if there is no relationshipm we should be able to shuffle the
groups
shuffled_groups <- transform(df_combined, group=sample(group))
prop_l <- mean(shuffled_groups$Turbocar[shuffled_groups$group=="luxury"
])
prop_e <- mean(shuffled_groups$Turbocar[shuffled_groups$group=="economy"])
results_given_H0_true[i] <- prop_l - prop_e
}
# Finally plot the results
hist(results_given_H0_true, freq = FALSE,
main='Dist. of the Diff in Sample Sample Props Under Null',
xlab = 'Average Difference in Prop. turbo type under Null',
ylab = 'Density')
diff_in_sample_props <- p_hat_l - p_hat_e
abline(v=diff_in_sample_props, col = "blue")
abline(v=-diff_in_sample_props, col = "red")
# counts of values more extreme than the test statistic in our original
sample, given H0 is true
# two sided given the alternate hypothesis
count_of_more_extreme_lower_tail <- sum(results_given_H0_true <=
-diff_in_sample_props)
count_of_more_extreme_upper_tail <- sum(results_given_H0_true >=
diff_in_sample_props)
bootstrap_pvalue <- (count_of_more_extreme_lower_tail +
count_of_more_extreme_upper_tail)/num_sims
cat("Bootstrap p-value")
bootstrap_pvalue
cat("Normal Approx p-value")
two_sided_diff_prop_pval
table(auto.data$body.style)

prop.table(table(auto.data$body.style))
sum(((table(auto.data$body.style) - 40.1)^2)/40.1)

pchisq(152.2406, df = 5-1, lower.tail = FALSE)
```

Raw Dataset

```
auto.data

### Raw Dataset

 ##               make. fuel.type aspiration num.of.doors  body.style
drive.wheels

## 1  alfa-romero       gas         std         two convertible         rwd

## 2  alfa-romero    gas        std       two convertible         rwd

## 3  alfa-romero       gas         std         two   hatchback         rwd

## 4              audi        gas         std         four        sedan
fwd

## 5              audi        gas         std         four        sedan
4wd

## 6              audi        gas         std         two         sedan
fwd

## 7              audi        gas         std         four        sedan
fwd

## 8              audi        gas         std         four        wagon
fwd

## 9              audi        gas      turbo          four        sedan
fwd

## 10             bmw         gas         std         two         sedan
rwd

## 11             bmw         gas        std          four        sedan
rwd

## 12             bmw         gas         std         two         sedan
rwd

## 13             bmw         gas         std         four        sedan
rwd

## 14             bmw         gas         std           four        sedan
      rwd

## 15             bmw         gas         std         four        sedan
rwd
```

```
## 16         bmw       gas       std       two       sedan
rwd
## 17         bmw       gas       std       four      sedan
rwd
## 18   chevrolet       gas       std       two   hatchback
fwd
## 19   chevrolet       gas       std       two   hatchback
fwd
## 20   chevrolet       gas       std       four      sedan
fwd
## 21       dodge       gas       std       two   hatchback       fwd
## 22       dodge       gas       std       two   hatchback       fwd
## 23       dodge       gas     turbo       two   hatchback       fwd
## 24       dodge       gas       std       four  hatchback       fwd
## 25       dodge       gas       std       four      sedan       fwd
## 26       dodge       gas       std       four      sedan       fwd
## 27       dodge       gas     turbo       two       sedan       fwd
## 28       dodge       gas       std       four      wagon       fwd
## 29       dodge       gas     turbo       two   hatchback       fwd
## 30       honda       gas       std         two   hatchback
fwd
## 31       honda       gas       std       two   hatchback       fwd
## 32       honda       gas       std       two   hatchback       fwd
## 33       honda       gas       std       two   hatchback       fwd
## 34       honda       gas       std       two   hatchback       fwd
## 35       honda       gas       std       four      sedan       fwd
## 36       honda       gas       std       four      wagon       fwd
## 37       honda       gas       std       two   hatchback       fwd
## 38       honda       gas       std       two   hatchback       fwd
## 39       honda       gas       std       four      sedan       fwd
```

```
## 40        honda        gas        std        four        sedan        fwd
## 41        honda        gas        std        four        sedan        fwd
## 42        honda        gas        std        two         sedan        fwd
## 43        isuzu        gas        std        four        sedan        rwd
## 44        isuzu        gas        std        two     hatchback        rwd
## 45        jaguar       gas        std        four        sedan        rwd
## 46        jaguar       gas        std          four        sedan
rwd
## 47        jaguar       gas        std        two         sedan        rwd
## 48        mazda        gas        std        two     hatchback        fwd
## 49        mazda        gas        std        two     hatchback        fwd
## 50        mazda        gas        std        two     hatchback        fwd
## 51        mazda        gas        std        four        sedan        fwd
## 52        mazda        gas        std        four        sedan        fwd
## 53        mazda        gas        std        two     hatchback        rwd
## 54        mazda        gas        std        two     hatchback        rwd
## 55        mazda        gas        std        two     hatchback        rwd
## 56        mazda        gas        std        two     hatchback        rwd
## 57        mazda        gas        std        two     hatchback        fwd
## 58        mazda        gas        std        four        sedan        fwd
## 59        mazda        gas        std        two     hatchback        fwd
## 60        mazda        gas        std        four        sedan        fwd
## 61        mazda diesel      std          two          sedan       fwd
## 62        mazda        gas        std          four    hatchback
fwd
## 63        mazda        gas        std        four        sedan        rwd
## 64        mazda diesel      std          four         sedan       rwd
## 65  mercedes-benz   diesel      turbo      four        sedan        rwd
```

```
## 66  mercedes-benz    diesel      turbo       four        wagon       rwd
## 67  mercedes-benz    diesel      turbo       two   hardtop           rwd
## 68  mercedes-benz    diesel      turbo       four        sedan       rwd
## 69  mercedes-benz    gas         std         four        sedan       rwd
## 70  mercedes-benz    gas         std         two convertible         rwd
## 71  mercedes-benz    gas         std         four        sedan       rwd
## 72  mercedes-benz    gas         std         two   hardtop           rwd
## 73       mercury     gas      turbo          two   hatchback         rwd
## 74       mitsubishi        gas         std         two   hatchback
fwd
## 75       mitsubishi        gas        std          two   hatchback
fwd
## 76       mitsubishi        gas         std         two   hatchback
fwd
## 77       mitsubishi        gas      turbo          two   hatchback
fwd
## 78       mitsubishi        gas      turbo       two   hatchback
fwd
## 79       mitsubishi        gas         std         two   hatchback
fwd
## 80       mitsubishi        gas      turbo          two   hatchback
fwd
## 81       mitsubishi        gas      turbo          two   hatchback
fwd
## 82       mitsubishi        gas      turbo          two   hatchback
fwd
## 83       mitsubishi        gas         std         four        sedan
fwd
## 84       mitsubishi        gas         std         four        sedan
fwd
## 85       mitsubishi        gas      turbo          four        sedan
fwd
```

```
## 86       mitsubishi       gas        std        four       sedan
fwd
## 87       nissan      gas         std         two         sedan       fwd
## 88       nissan      diesel      std         two         sedan       fwd
## 89       nissan      gas         std         two         sedan       fwd
## 90       nissan      gas         std         four        sedan       fwd
## 91       nissan      gas        std          four        wagon       fwd
## 92       nissan      gas         std         two         sedan       fwd
## 93       nissan      gas         std         two   hatchback         fwd
## 94       nissan      gas         std           four        sedan
fwd
## 95       nissan      gas         std         four        wagon       fwd
## 96       nissan      gas         std         two   hardtop           fwd
## 97       nissan      gas         std         four   hatchback        fwd
## 98       nissan      gas         std         four        sedan       fwd
## 99       nissan      gas         std         four        sedan       fwd
## 100      nissan      gas         std         four        wagon       fwd
## 101      nissan      gas         std         four        sedan       fwd
## 102      nissan      gas         std         two   hatchback         rwd
## 103      nissan      gas       turbo         two   hatchback         rwd
## 104      nissan      gas         std         two   hatchback         rwd
## 105      peugot      gas         std         four        sedan       rwd
## 106      peugot      diesel      turbo       four        sedan       rwd
## 107      peugot      gas         std         four        wagon       rwd
## 108      peugot      diesel      turbo       four        wagon       rwd
## 109      peugot      gas         std         four        sedan       rwd
## 110      peugot      diesel      turbo       four        sedan       rwd
## 111      peugot      gas         std         four        wagon       rwd
```

```
## 112       peugot      diesel      turbo       four        wagon       rwd
## 113       peugot      gas         std           four        sedan
rwd
## 114       peugot      diesel      turbo       four        sedan       rwd
## 115       peugot      gas      turbo          four        sedan       rwd
## 116       plymouth    gas         std         two   hatchback         fwd
## 117       plymouth    gas      turbo          two   hatchback         fwd
## 118       plymouth    gas         std         four   hatchback        fwd
## 119       plymouth    gas         std         four        sedan       fwd
## 120       plymouth    gas         std         four        sedan       fwd
## 121       plymouth    gas         std         four        wagon       fwd
## 122       plymouth    gas      turbo          two   hatchback         rwd
## 123       porsche     gas         std         two   hatchback         rwd
## 124       porsche     gas         std         two   hardtop           rwd
## 125       porsche     gas         std         two   hardtop           rwd
## 126       porsche     gas        std          two convertible         rwd
## 127       renault     gas         std         four        wagon       fwd
## 128       renault     gas         std         two   hatchback         fwd
## 129             saab        gas         std             two   hatchback
      fwd
## 130             saab        gas         std         four        sedan
fwd
## 131             saab        gas         std         two   hatchback
fwd
## 132             saab        gas         std         four        sedan
fwd
## 133             saab        gas      turbo          two   hatchback
fwd
## 134             saab        gas      turbo          four        sedan
fwd
```

```
## 135      subaru      gas         std         two   hatchback         fwd
## 136      subaru      gas         std         two   hatchback         fwd
## 137      subaru      gas         std         two   hatchback         4wd
## 138      subaru      gas         std         four        sedan       fwd
## 139      subaru      gas         std         four        sedan       fwd
## 140      subaru      gas         std         four        sedan       fwd
## 141      subaru      gas         std         four        sedan       4wd
## 142      subaru      gas       turbo         four        sedan       4wd
## 143      subaru      gas         std         four        wagon       fwd
## 144      subaru      gas         std         four        wagon       fwd
## 145      subaru      gas         std           four        wagon
4wd
## 146      subaru      gas       turbo         four        wagon       4wd
## 147      toyota      gas         std         two   hatchback         fwd
## 148      toyota      gas         std         two   hatchback         fwd
## 149      toyota      gas         std         four   hatchback        fwd
## 150      toyota      gas         std         four        wagon       fwd
## 151      toyota      gas         std         four        wagon       4wd
## 152      toyota      gas         std         four        wagon       4wd
## 153      toyota      gas         std         four        sedan       fwd
## 154      toyota      gas         std         four   hatchback        fwd
## 155      toyota      diesel      std         four        sedan       fwd
## 156      toyota      diesel      std         four   hatchback        fwd
## 157      toyota      gas         std         four        sedan       fwd
## 158      toyota      gas         std         four   hatchback        fwd
## 159      toyota      gas         std         four        sedan       fwd
## 160      toyota      gas         std         two         sedan       rwd
## 161      toyota      gas        std          two   hatchback         rwd
```

```
## 162       toyota      gas         std         two         sedan       rwd
## 163       toyota      gas         std         two   hatchback         rwd
## 164       toyota      gas         std            two     hardtop
rwd
## 165       toyota      gas         std         two   hardtop           rwd
## 166       toyota      gas         std         two   hatchback         rwd
## 167       toyota      gas         std         two   hardtop           rwd
## 168       toyota      gas         std         two   hatchback         rwd
## 169       toyota      gas         std         two convertible         rwd
## 170       toyota      gas         std         four        sedan       fwd
## 171       toyota      diesel      turbo       four        sedan       fwd
## 172       toyota      gas         std         four   hatchback        fwd
## 173       toyota      gas         std         four        sedan       fwd
## 174       toyota      gas         std         four   hatchback        fwd
## 175       toyota      gas         std         two   hatchback         rwd
## 176       toyota      gas         std         two   hatchback         rwd
## 177       toyota      gas        std          four        sedan       rwd
## 178       toyota      gas         std         four        wagon       rwd
## 179       volkswagen  diesel      std         two         sedan       fwd
## 180       volkswagen        gas         std            two       sedan
      fwd
## 181       volkswagen  diesel      std         four        sedan       fwd
## 182       volkswagen        gas         std         four        sedan
fwd
## 183       volkswagen        gas         std         four        sedan
fwd
## 184       volkswagen  diesel      turbo       four        sedan       fwd
## 185       volkswagen        gas         std         four        sedan
fwd
```

```
## 186      volkswagen      gas        std        two convertible
fwd
## 187      volkswagen      gas        std        two   hatchback
fwd
## 188      volkswagen      gas        std        four       sedan
fwd
## 189      volkswagen  diesel      turbo      four       sedan      fwd
## 190      volkswagen      gas        std        four       wagon
fwd
## 191      volvo      gas        std        four       sedan      rwd
## 192      volvo      gas        std        four       wagon      rwd
## 193      volvo      gas        std        four       sedan      rwd
## 194      volvo      gas        std        four       wagon      rwd
## 195      volvo      gas      turbo        four       sedan      rwd
## 196      volvo      gas      turbo      four       wagon         rwd
## 197      volvo      gas        std        four       sedan      rwd
## 198      volvo      gas      turbo        four       sedan      rwd
## 199      volvo      gas        std        four       sedan      rwd
## 200      volvo diesel      turbo       four       sedan      rwd
## 201      volvo      gas      turbo        four       sedan      rwd
##    engine.location wheel.base length width height curb.weight engine.type
## 1             front       88.6  168.8  64.1   48.8        2548      dohc
## 2             front       88.6  168.8  64.1   48.8        2548      dohc
## 3             front       94.5  171.2  65.5   52.4        2823      ohcv
## 4             front       99.8  176.6  66.2   54.3        2337      ohc
## 5             front       99.4  176.6  66.4   54.3        2824      ohc
## 6             front       99.8  177.3  66.3   53.1        2507      ohc
## 7             front       105.8  192.7  71.4   55.7        2844
ohc
```

```
## 8                 front       105.8  192.7  71.4   55.7        2954
ohc
## 9                 front       105.8  192.7  71.4   55.9        3086
ohc
## 10                front       101.2  176.8  64.8   54.3        2395
ohc
## 11                front       101.2  176.8  64.8   54.3        2395
ohc
## 12                front       101.2  176.8  64.8   54.3        2710
ohc
## 13                front       101.2  176.8  64.8   54.3        2765
ohc
## 14                 front      103.5  189.0  66.9   55.7        3055
ohc
## 15                front       103.5  189.0  66.9   55.7        3230
ohc
## 16                front       103.5  193.8  67.9   53.7        3380
ohc
## 17                front       110.0  197.0  70.9   56.3        3505
ohc
## 18                front       88.4  141.1  60.3   53.2        1488
l
## 19                front       94.5  155.9  63.6   52.0        1874      ohc
## 20                front       94.5  158.8  63.6   52.0        1909      ohc
## 21                front       93.7  157.3  63.8   50.8        1876      ohc
## 22                front       93.7  157.3  63.8   50.8        1876      ohc
## 23                front       93.7  157.3  63.8   50.8        2128      ohc
## 24                front       93.7  157.3  63.8   50.6        1967      ohc
## 25                front       93.7  157.3  63.8   50.6        1989      ohc
## 26                front       93.7  157.3  63.8   50.6        1989      ohc
## 27                front       93.7  157.3  63.8   50.6        2191      ohc
```

```
## 28                  front        103.3  174.6  64.6   59.8         2535
ohc
## 29                  front        95.9  173.2  66.3   50.2         2811      ohc
## 30                  front        86.6  144.6  63.9   50.8         1713      ohc
## 31                  front        86.6  144.6  63.9   50.8         1819      ohc
## 32                  front        93.7  150.0  64.0   52.6         1837      ohc
## 33                  front        93.7  150.0  64.0   52.6         1940      ohc
## 34                  front        93.7  150.0  64.0   52.6         1956      ohc
## 35                  front        96.5  163.4  64.0   54.5         2010      ohc
## 36                  front        96.5  157.1  63.9   58.3         2024      ohc
## 37                 front         96.5  167.5  65.2   53.3         2236      ohc
## 38                  front        96.5  167.5  65.2   53.3         2289      ohc
## 39                  front        96.5  175.4  65.2   54.1         2304      ohc
## 40                  front        96.5  175.4  62.5   54.1         2372      ohc
## 41                  front        96.5  175.4  65.2   54.1         2465      ohc
## 42                  front        96.5  169.1  66.0   51.0         2293      ohc
## 43                  front        94.3  170.7  61.8   53.5         2337      ohc
## 44                  front        96.0  172.6  65.2   51.4         2734      ohc
## 45                  front        113.0  199.6  69.6   52.8         4066     dohc
## 46                  front        113.0  199.6  69.6   52.8         4066     dohc
## 47                  front        102.0  191.7  70.6   47.8         3950     ohcv
## 48                  front        93.1  159.1  64.2   54.1         1890      ohc
## 49                  front        93.1  159.1  64.2   54.1         1900      ohc
## 50                  front        93.1  159.1  64.2   54.1         1905      ohc
## 51                  front        93.1  166.8  64.2   54.1         1945      ohc
## 52                  front        93.1  166.8  64.2   54.1         1950      ohc
## 53                  front        95.3  169.0  65.7   49.6         2380      rotor
## 54                  front        95.3  169.0  65.7   49.6         2380      rotor
```

```
## 55              front      95.3  169.0  65.7   49.6        2385     rotor
## 56              front      95.3  169.0  65.7   49.6        2500     rotor
## 57              front      98.8  177.8  66.5   53.7        2385     ohc
## 58              front      98.8  177.8  66.5   55.5        2410     ohc
## 59              front      98.8  177.8  66.5   53.7        2385     ohc
## 60              front      98.8  177.8  66.5   55.5        2410     ohc
## 61              front      98.8  177.8  66.5   55.5        2443     ohc
## 62              front      98.8  177.8  66.5   55.5        2425        ohc
## 63              front      104.9  175.0  66.1   54.4        2670
ohc
## 64              front      104.9  175.0  66.1   54.4        2700
ohc
## 65              front      110.0  190.9  70.3   56.5        3515
ohc
## 66               front     110.0  190.9  70.3   58.7        3750
ohc
## 67              front      106.7  187.5  70.3   54.9        3495
ohc
## 68              front      115.6  202.6  71.7   56.3        3770
ohc
## 69              front      115.6  202.6  71.7   56.5        3740     ohcv
## 70              front      96.6  180.3  70.5   50.8        3685     ohcv
## 71              front      120.9  208.1  71.7   56.7        3900     ohcv
## 72              front      112.0  199.2  72.0   55.4        3715     ohcv
## 73              front      102.7  178.4  68.0   54.8        2910
ohc
## 74              front      93.7  157.3  64.4   50.8        1918     ohc
## 75              front      93.7  157.3  64.4   50.8        1944     ohc
## 76              front      93.7  157.3  64.4   50.8        2004     ohc
## 77              front      93.0  157.3  63.8   50.8        2145     ohc
```

```
## 78          front     96.3  173.0  65.4   49.4        2370     ohc
## 79          front     96.3  173.0  65.4   49.4        2328     ohc
## 80          front     95.9  173.2  66.3   50.2        2833     ohc
## 81          front     95.9  173.2  66.3   50.2        2921     ohc
## 82          front     95.9  173.2  66.3   50.2        2926     ohc
## 83          front     96.3  172.4  65.4   51.6        2365     ohc
## 84          front     96.3  172.4  65.4   51.6        2405     ohc
## 85          front     96.3  172.4  65.4   51.6        2403     ohc
## 86          front     96.3  172.4  65.4   51.6        2403     ohc
## 87          front     94.5  165.3  63.8   54.5        1889     ohc
## 88          front     94.5  165.3  63.8   54.5        2017     ohc
## 89         front      94.5  165.3  63.8   54.5        1918     ohc
## 90          front     94.5  165.3  63.8   54.5        1938     ohc
## 91          front     94.5  170.2  63.8   53.5        2024     ohc
## 92          front     94.5  165.3  63.8   54.5        1951     ohc
## 93          front     94.5  165.6  63.8   53.3        2028     ohc
## 94          front     94.5  165.3  63.8   54.5        1971     ohc
## 95          front     94.5  170.2  63.8   53.5        2037     ohc
## 96          front     95.1  162.4  63.8   53.3        2008     ohc
## 97          front     97.2  173.4  65.2   54.7        2324     ohc
## 98          front     97.2  173.4  65.2   54.7        2302     ohc
## 99          front     100.4  181.7  66.5   55.1        3095    ohcv
## 100         front     100.4  184.6  66.5   56.1        3296    ohcv
## 101         front     100.4  184.6  66.5   55.1        3060    ohcv
## 102         front     91.3  170.7  67.9   49.7        3071     ohcv
## 103         front     91.3  170.7  67.9   49.7        3139     ohcv
## 104         front     99.2  178.5  67.9   49.7        3139     ohcv
```

```
## 105             front      107.9  186.7  68.4   56.7        3020
l
## 106             front      107.9  186.7  68.4   56.7        3197
l
## 107             front      114.2  198.9  68.4   58.7        3230
l
## 108             front      114.2  198.9  68.4   58.7        3430
l
## 109             front      107.9  186.7  68.4   56.7        3075
l
## 110             front      107.9  186.7  68.4   56.7        3252
l
## 111             front      114.2  198.9  68.4   56.7        3285
l
## 112             front      114.2  198.9  68.4   58.7        3485
l
## 113             front      107.9  186.7  68.4   56.7        3075
l
## 114             front      107.9  186.7  68.4   56.7        3252
l
## 115              front     108.0  186.7  68.3   56.0        3130
l
## 116             front      93.7  157.3  63.8   50.8        1918      ohc
## 117             front      93.7  157.3  63.8   50.8        2128      ohc
## 118             front      93.7  157.3  63.8   50.6        1967      ohc
## 119             front      93.7  167.3  63.8   50.8        1989      ohc
## 120             front      93.7  167.3  63.8   50.8        2191      ohc
## 121             front      103.3  174.6  64.6   59.8        2535
ohc
## 122             front      95.9  173.2  66.3   50.2        2818      ohc
## 123             front      94.5  168.9  68.3   50.2        2778      ohc
## 124             rear       89.5  168.9  65.0   51.6        2756      ohcf
```

```
## 125             rear       89.5  168.9  65.0   51.6         2756     ohcf
## 126             rear       89.5  168.9  65.0   51.6         2800     ohcf
## 127             front      96.1  181.5  66.5   55.2         2579     ohc
## 128             front      96.1  176.8  66.6   50.5         2460     ohc
## 129             front      99.1  186.6  66.5   56.1         2658     ohc
## 130             front      99.1  186.6  66.5   56.1         2695     ohc
## 131             front      99.1  186.6  66.5   56.1         2707     ohc
## 132             front      99.1  186.6  66.5   56.1         2758     ohc
## 133             front      99.1  186.6  66.5   56.1         2808     dohc
## 134             front      99.1  186.6  66.5   56.1         2847     dohc
## 135             front      93.7  156.9  63.4   53.7         2050     ohcf
## 136             front      93.7  157.9  63.6   53.7         2120     ohcf
## 137             front      93.3  157.3  63.8   55.7         2240     ohcf
## 138            front       97.2  172.0  65.4   52.5         2145     ohcf
## 139             front      97.2  172.0  65.4   52.5         2190     ohcf
## 140             front      97.2  172.0  65.4   52.5         2340     ohcf
## 141             front      97.0  172.0  65.4   54.3         2385     ohcf
## 142             front      97.0  172.0  65.4   54.3         2510     ohcf
## 143             front      97.0  173.5  65.4   53.0         2290     ohcf
## 144             front      97.0  173.5  65.4   53.0         2455     ohcf
## 145             front      96.9  173.6  65.4   54.9         2420     ohcf
## 146             front      96.9  173.6  65.4   54.9         2650     ohcf
## 147             front      95.7  158.7  63.6   54.5         1985     ohc
## 148             front      95.7  158.7  63.6   54.5         2040     ohc
## 149             front      95.7  158.7  63.6   54.5         2015     ohc
## 150             front      95.7  169.7  63.6   59.1         2280     ohc
## 151             front      95.7  169.7  63.6   59.1         2290     ohc
```

```
## 152            front       95.7  169.7  63.6   59.1         3110      ohc
## 153            front       95.7  166.3  64.4   53.0         2081      ohc
## 154            front       95.7  166.3  64.4   52.8         2109      ohc
## 155            front       95.7  166.3  64.4   53.0         2275      ohc
## 156            front       95.7  166.3  64.4   52.8         2275      ohc
## 157            front       95.7  166.3  64.4   53.0         2094      ohc
## 158            front       95.7  166.3  64.4   52.8         2122      ohc
## 159            front       95.7  166.3  64.4   52.8         2140      ohc
## 160            front       94.5  168.7  64.0   52.6         2169         ohc
## 161            front       94.5  168.7  64.0   52.6         2204      ohc
## 162            front       94.5  168.7  64.0   52.6         2265      dohc
## 163            front       94.5  168.7  64.0   52.6         2300      dohc
## 164             front      98.4  176.2  65.6   52.0         2540      ohc
## 165            front       98.4  176.2  65.6   52.0         2536      ohc
## 166            front       98.4  176.2  65.6   52.0         2551      ohc
## 167            front       98.4  176.2  65.6   52.0         2679      ohc
## 168            front       98.4  176.2  65.6   52.0         2714      ohc
## 169            front       98.4  176.2  65.6   53.0         2975      ohc
## 170            front       102.4  175.6  66.5   54.9         2326
ohc
## 171            front       102.4  175.6  66.5   54.9         2480
ohc
## 172            front       102.4  175.6  66.5   53.9         2414
ohc
## 173            front       102.4  175.6  66.5   54.9         2414
ohc
## 174            front       102.4  175.6  66.5   53.9         2458
ohc
## 175            front       102.9  183.5  67.7   52.0         2976     dohc
## 176            front       102.9  183.5  67.7   52.0         3016     dohc
```

```
## 177                  front         104.5  187.8  66.5   54.1          3131     dohc
## 178                  front         104.5  187.8  66.5   54.1          3151     dohc
## 179                  front         97.3  171.7  65.5   55.7          2261     ohc
## 180                  front         97.3  171.7  65.5   55.7          2209     ohc
## 181                  front         97.3  171.7  65.5   55.7          2264     ohc
## 182                  front         97.3  171.7  65.5   55.7          2212     ohc
## 183                  front         97.3  171.7  65.5   55.7          2275     ohc
## 184                  front         97.3  171.7  65.5   55.7          2319     ohc
## 185                  front         97.3  171.7  65.5   55.7          2300     ohc
## 186                  front         94.5  159.3  64.2   55.6          2254     ohc
## 187                 front          94.5  165.7  64.0   51.4          2221     ohc
## 188                  front         100.4  180.2  66.9   55.1          2661
ohc
## 189                  front         100.4  180.2  66.9   55.1          2579
ohc
## 190                  front         100.4  183.1  66.9   55.1          2563
ohc
## 191                  front         104.3  188.8  67.2   56.2          2912
ohc
## 192                  front         104.3  188.8  67.2   57.5          3034
ohc
## 193                  front         104.3  188.8  67.2   56.2          2935
ohc
## 194                  front         104.3  188.8  67.2   57.5          3042
ohc
## 195                  front         104.3  188.8  67.2   56.2          3045
ohc
## 196                  front         104.3  188.8  67.2   57.5          3157
ohc
## 197                  front         109.1  188.8  68.9   55.5          2952
ohc
```

```
## 198              front       109.1  188.8  68.8   55.5        3049
ohc
## 199              front       109.1  188.8  68.9   55.5        3012     ohcv
## 200              front       109.1  188.8  68.9   55.5        3217
ohc
## 201              front       109.1  188.8  68.9   55.5        3062
ohc
##    num.of.cylinders engine.size fuel.system bore stroke compression.ratio
## 1                four        130         mpfi 3.47   2.68              9.0
## 2                four        130         mpfi 3.47   2.68              9.0
## 3                six         152         mpfi 2.68   3.47              9.0
## 4                four        109         mpfi 3.19   3.40             10.0
## 5                five        136         mpfi 3.19   3.40              8.0
## 6                five        136         mpfi 3.19   3.40              8.5
## 7                five        136         mpfi 3.19   3.40              8.5
## 8                five        136         mpfi 3.19   3.40              8.5
## 9                five        131         mpfi 3.13   3.40              8.3
## 10               four        108         mpfi 3.50   2.80              8.8
## 11               four        108         mpfi 3.50   2.80              8.8
## 12               six         164         mpfi 3.31   3.19              9.0
## 13               six         164         mpfi 3.31   3.19              9.0
## 14                   six         164         mpfi 3.31   3.19
9.0
## 15               six         209         mpfi 3.62   3.39              8.0
## 16               six         209         mpfi 3.62   3.39              8.0
## 17               six         209         mpfi 3.62   3.39              8.0
## 18               three       61          2bbl 2.91   3.03              9.5
## 19               four        90          2bbl 3.03   3.11              9.6
## 20               four        90          2bbl 3.03   3.11              9.6
```

```
## 21            four        90          2bbl 2.97   3.23              9.4
## 22            four        90          2bbl 2.97   3.23              9.4
## 23            four        98          mpfi 3.03   3.39              7.6
## 24            four        90          2bbl 2.97   3.23              9.4
## 25            four        90          2bbl 2.97   3.23              9.4
## 26            four        90          2bbl 2.97   3.23              9.4
## 27            four        98          mpfi 3.03   3.39              7.6
## 28            four        122         2bbl 3.34   3.46              8.5
## 29            four        156         mfi 3.60   3.90               7.0
## 30            four        92     1bbl 2.91   3.41             9.6
## 31            four        92          1bbl 2.91   3.41              9.2
## 32            four        79          1bbl 2.91   3.07              10.1
## 33            four        92          1bbl 2.91   3.41               9.2
## 34            four        92          1bbl 2.91   3.41              9.2
## 35            four        92          1bbl 2.91   3.41              9.2
## 36            four        92          1bbl 2.92   3.41              9.2
## 37             four       110         1bbl 3.15   3.58              9.0
## 38            four        110         1bbl 3.15   3.58              9.0
## 39            four        110         1bbl 3.15   3.58              9.0
## 40            four          110        1bbl 3.15   3.58             9.0
## 41            four        110         mpfi 3.15   3.58              9.0
## 42            four        110         2bbl 3.15   3.58              9.1
## 43            four        111         2bbl 3.31   3.23              8.5
## 44            four        119         spfi 3.43   3.23              9.2
## 45            six         258         mpfi 3.63   4.17              8.1
## 46            six         258         mpfi 3.63   4.17              8.1
## 47            twelve      326         mpfi 3.54   2.76              11.5
```

```
## 48                  four         91           2bbl 3.03   3.15                  9.0
## 49                  four         91           2bbl 3.03   3.15                  9.0
## 50                  four         91           2bbl 3.03   3.15                  9.0
## 51                  four         91           2bbl 3.03   3.15                  9.0
## 52                  four         91           2bbl 3.08   3.15                  9.0
## 53                  two          70           4bbl 0.00   0.00                  9.4
## 54                  two          70           4bbl 0.00   0.00                  9.4
## 55                  two          70           4bbl 0.00   0.00                  9.4
## 56                  two          80           mpfi 0.00   0.00                  9.4
## 57                  four         122          2bbl 3.39   3.39                  8.6
## 58                  four         122          2bbl 3.39   3.39                  8.6
## 59                  four         122          2bbl 3.39   3.39                   8.6
## 60                  four         122          2bbl 3.39   3.39                  8.6
## 61                  four         122          idi 3.39   3.39                   22.7
## 62                  four         122          2bbl 3.39   3.39                  8.6
## 63                     four         140          mpfi 3.76   3.16
8.0
## 64                  four         134          idi 3.43   3.64                   22.0
## 65                  five         183          idi 3.58   3.64                   21.5
## 66                  five         183          idi 3.58   3.64                   21.5
## 67                  five         183          idi 3.58   3.64                   21.5
## 68                  five         183          idi 3.58   3.64                   21.5
## 69                  eight        234          mpfi 3.46   3.10                  8.3
## 70                  eight        234          mpfi 3.46   3.10                  8.3
## 71                  eight        308          mpfi 3.80   3.35                  8.0
## 72                  eight        304          mpfi 3.80   3.35                  8.0
## 73                  four         140          mpfi 3.78   3.12                  8.0
## 74                  four         92           2bbl 2.97   3.23                  9.4
```

```
## 75                  four        92          2bbl 2.97   3.23                9.4
## 76                  four        92          2bbl 2.97   3.23                9.4
## 77                  four        98          spdi 3.03   3.39                7.6
## 78                  four        110         spdi 3.17   3.46                7.5
## 79                  four        122          2bbl 3.35   3.46               8.5
## 80                  four        156         spdi 3.58   3.86                7.0
## 81                  four        156         spdi 3.59   3.86                7.0
## 82                  four        156         spdi 3.59   3.86                 7.0
## 83                  four        122         2bbl 3.35   3.46                8.5
## 84                  four        122         2bbl 3.35   3.46                8.5
## 85                  four        110         spdi 3.17   3.46                7.5
## 86                   four       110         spdi 3.17   3.46                7.5
## 87                  four        97          2bbl 3.15   3.29                9.4
## 88                  four        103         idi 2.99   3.47                 21.9
## 89                  four           97        2bbl 3.15   3.29               9.4
## 90                  four        97          2bbl 3.15   3.29                9.4
## 91                  four        97          2bbl 3.15   3.29                9.4
## 92                  four        97          2bbl 3.15   3.29                9.4
## 93                  four        97          2bbl 3.15   3.29                9.4
## 94                  four        97          2bbl 3.15   3.29                9.4
## 95                  four        97          2bbl 3.15   3.29                9.4
## 96                  four        97          2bbl 3.15   3.29                9.4
## 97                  four        120         2bbl 3.33   3.47                8.5
## 98                  four        120         2bbl 3.33   3.47                8.5
## 99                  six         181         mpfi 3.43   3.27                9.0
## 100                 six         181         mpfi 3.43   3.27                9.0
## 101                 six         181         mpfi 3.43   3.27                9.0
```

```
## 102              six          181          mpfi 3.43   3.27                 9.0
## 103              six          181          mpfi 3.43   3.27                 7.8
## 104              six          181          mpfi 3.43   3.27                 9.0
## 105              four         120          mpfi 3.46   3.19                 8.4
## 106              four         152          idi 3.70   3.52                  21.0
## 107              four         120          mpfi 3.46   3.19                 8.4
## 108              four         152          idi 3.70   3.52                 21.0
## 109              four         120          mpfi 3.46   2.19                 8.4
## 110              four         152          idi 3.70   3.52                  21.0
## 111              four         120          mpfi 3.46   2.19                 8.4
## 112                 four         152          idi 3.70   3.52
21.0
## 113              four         120          mpfi 3.46   3.19                 8.4
## 114              four         152          idi 3.70   3.52                  21.0
## 115              four         134          mpfi 3.61   3.21                 7.0
## 116              four         90           2bbl 2.97   3.23                 9.4
## 117              four         98           spdi 3.03   3.39                 7.6
## 118              four         90           2bbl 2.97   3.23                 9.4
## 119              four         90           2bbl 2.97   3.23                 9.4
## 120              four         98           2bbl 2.97   3.23                 9.4
## 121              four         122          2bbl 3.35   3.46                 8.5
## 122              four         156          spdi 3.59   3.86                 7.0
## 123              four         151          mpfi 3.94   3.11                 9.5
## 124              six          194          mpfi 3.74   2.90                 9.5
## 125              six          194          mpfi 3.74   2.90                 9.5
## 126              six          194          mpfi 3.74   2.90                 9.5
## 127              four         132          mpfi 3.46   3.90                 8.7
## 128              four         132           mpfi 3.46   3.90                8.7
```

```
## 129          four      121       mpfi 3.54   3.07          9.3
## 130          four      121       mpfi 3.54   3.07          9.3
## 131          four      121       mpfi 2.54   2.07           9.3
## 132          four      121       mpfi 3.54   3.07          9.3
## 133          four      121       mpfi 3.54   3.07          9.0
## 134          four      121       mpfi 3.54   3.07          9.0
## 135           four     97        2bbl 3.62   2.36          9.0
## 136          four      108       2bbl 3.62   2.64          8.7
## 137          four      108       2bbl 3.62   2.64          8.7
## 138          four        108      2bbl 3.62   2.64         9.5
## 139          four      108       2bbl 3.62   2.64          9.5
## 140          four      108       mpfi 3.62   2.64          9.0
## 141          four      108       2bbl 3.62   2.64          9.0
## 142          four      108       mpfi 3.62   2.64          7.7
## 143          four      108       2bbl 3.62   2.64          9.0
## 144          four      108       mpfi 3.62   2.64          9.0
## 145          four      108       2bbl 3.62   2.64          9.0
## 146          four      108       mpfi 3.62   2.64          7.7
## 147          four      92        2bbl 3.05   3.03          9.0
## 148           four        92     2bbl 3.05   3.03          9.0
## 149          four      92        2bbl 3.05   3.03          9.0
## 150          four      92        2bbl 3.05   3.03          9.0
## 151          four      92        2bbl 3.05   3.03          9.0
## 152          four      92        2bbl 3.05   3.03          9.0
## 153          four      98        2bbl 3.19   3.03          9.0
## 154          four      98        2bbl 3.19   3.03          9.0
## 155          four      110       idi 3.27   3.35           22.5
```

```
## 156             four        110         idi 3.27   3.35                  22.5
## 157             four        98          2bbl 3.19   3.03                 9.0
## 158             four        98          2bbl 3.19   3.03                 9.0
## 159             four        98          2bbl 3.19   3.03                 9.0
## 160             four        98          2bbl 3.19   3.03                  9.0
## 161             four        98          2bbl 3.19   3.03                 9.0
## 162             four        98          mpfi 3.24   3.08                 9.4
## 163             four        98          mpfi 3.24   3.08                 9.4
## 164                four           146         mpfi 3.62   3.50
9.3
## 165             four        146         mpfi 3.62   3.50                 9.3
## 166             four        146         mpfi 3.62   3.50                 9.3
## 167             four        146         mpfi 3.62   3.50                 9.3
## 168             four        146         mpfi 3.62   3.50                 9.3
## 169             four        146         mpfi 3.62   3.50                 9.3
## 170             four        122         mpfi 3.31   3.54                 8.7
## 171             four        110         idi 3.27   3.35                  22.5
## 172             four        122         mpfi 3.31   3.54                 8.7
## 173             four        122         mpfi 3.31   3.54                 8.7
## 174             four        122         mpfi 3.31   3.54                 8.7
## 175             six         171         mpfi 3.27   3.35                 9.3
## 176             six         171         mpfi 3.27   3.35                 9.3
## 177             six         171         mpfi 3.27   3.35                 9.2
## 178             six         161         mpfi 3.27   3.35                 9.2
## 179             four        97          idi 3.01   3.40                  23.0
## 180             four        109          mpfi 3.19   3.40                9.0
## 181             four        97          idi 3.01   3.40                  23.0
## 182             four        109         mpfi 3.19   3.40                 9.0
```

```
## 183          four      109       mpfi 3.19   3.40               9.0
## 184          four      97        idi 3.01   3.40               23.0
## 185          four      109       mpfi 3.19   3.40              10.0
## 186          four      109       mpfi 3.19   3.40              8.5
## 187           four     109       mpfi 3.19   3.40              8.5
## 188          five      136       mpfi 3.19   3.40              8.5
## 189          four      97        idi 3.01   3.40               23.0
## 190          four        109      mpfi 3.19   3.40             9.0
## 191          four      141       mpfi 3.78   3.15              9.5
## 192          four      141       mpfi 3.78   3.15              9.5
## 193          four      141       mpfi 3.78   3.15              9.5
## 194          four      141       mpfi 3.78   3.15              9.5
## 195          four      130       mpfi 3.62   3.15              7.5
## 196          four      130       mpfi 3.62   3.15              7.5
## 197          four      141       mpfi 3.78   3.15              9.5
## 198          four      141       mpfi 3.78   3.15              8.7
## 199          six       173       mpfi 3.58   2.87              8.8
## 200          six       145       idi 3.01   3.40               23.0
## 201          four      141       mpfi 3.78   3.15              9.5
##    horsepower peak.rom city.mpg highway.mpg price carclass
## 1        111   5000       21           27 13495   luxury
## 2        111   5000       21            27 16500   luxury
## 3        154   5000       19            26 16500   luxury
## 4        102   5500       24            30 13950   luxury
## 5        115   5500       18          22 17450   luxury
## 6        110   5500       19            25 15250   luxury
## 7        110   5500       19            25 17710   luxury
```

```
## 8        110   5500       19            25 18920   luxury
## 9        140   5500       17            20 23875   luxury
## 10       101     5800      23            29 16430   luxury
## 11       101     5800      23            29 16925   luxury
## 12       121     4250      21            28 20970   luxury
## 13       121     4250      21            28 21105   luxury
## 14       121     4250      20            25 24565   luxury
## 15       182     5400      16            22 30760   luxury
## 16       182     5400      16            22 41315   luxury
## 17       182     5400      15            20 36880   luxury
## 18       48     5100       47            53  5151  economy
## 19       70     5400       38            43  6295  economy
## 20       70     5400       38            43  6575  economy
## 21       68     5500       37            41  5572  economy
## 22       68     5500       31            38  6377  economy
## 23       102     5500      24            30  7957  economy
## 24       68     5500       31            38  6229  economy
## 25       68     5500       31            38  6692  economy
## 26       68     5500       31            38  7609  economy
## 27       102     5500      24            30  8558  economy
## 28       88     5000       24            30  8921  economy
## 29       145     5000      19            24 12964  economy
## 30       58     4800       49            54  6479  economy
## 31       76     6000       31            38  6855  economy
## 32       60     5500       38            42  5399  economy
## 33       76     6000       30            34  6529  economy
## 34          76     6000    30            34  7129  economy
```

```
## 35        76     6000       30          34  7295  economy
## 36        76     6000       30          34  7295  economy
## 37        86     5800       27          33  7895  economy
## 38             86     5800        27           33  9095  economy
## 39        86     5800       27          33  8845  economy
## 40        86     5800       27          33 10295  economy
## 41        101     5800      24          28 12945  economy
## 42        100     5500      25          31 10345  economy
## 43        78     4800       24          29  6785  economy
## 44        90     5000       24          29 11048  economy
## 45        176     4750      15          19 32250   luxury
## 46        176     4750       15         19 35550   luxury
## 47        262     5000      13          17 36000   luxury
## 48        68     5000       30          31  5195  economy
## 49        68     5000       31          38  6095  economy
## 50        68     5000       31          38  6795  economy
## 51        68     5000       31          38  6695  economy
## 52        68     5000       31          38  7395  economy
## 53        101     6000      17          23 10945  economy
## 54        101     6000      17          23 11845  economy
## 55        101     6000      17          23 13645   luxury
## 56        135     6000      16          23 15645   luxury
## 57        84     4800       26          32  8845  economy
## 58           84     4800    26          32  8495  economy
## 59        84     4800       26          32 10595  economy
## 60        84     4800       26          32 10245  economy
## 61        64     4650       36          42 10795  economy
```

```
## 62          84     4800    26           32 11245  economy
## 63       120     5000     19           27 18280   luxury
## 64       72     4200      31           39 18344   luxury
## 65       123     4350     22           25 25552   luxury
## 66       123     4350     22           25 28248   luxury
## 67       123     4350     22           25 28176   luxury
## 68       123     4350     22           25 31600   luxury
## 69       155     4750     16           18 34184   luxury
## 70       155     4750     16           18 35056   luxury
## 71       184     4500     14           16 40960   luxury
## 72       184     4500     14           16 45400   luxury
## 73       175     5000     19           24 16503   luxury
## 74       68     5500      37           41  5389  economy
## 75       68     5500      31           38  6189  economy
## 76       68     5500      31           38  6669  economy
## 77       102     5500     24           30  7689  economy
## 78       116     5500      23           30  9959  economy
## 79       88     5000      25           32  8499  economy
## 80       145     5000     19           24 12629  economy
## 81       145     5000     19           24 14869   luxury
## 82       145     5000      19           24 14489   luxury
## 83       88     5000      25           32  6989  economy
## 84       88     5000      25           32  8189  economy
## 85       116     5500     23           30  9279  economy
## 86       116     5500     23           30  9279  economy
## 87       69     5200      31           37  5499  economy
## 88       55     4800      45           50  7099  economy
```

```
## 89       69     5200       31           37  6649  economy
## 90       69     5200       31         37  6849  economy
## 91       69     5200       31           37  7349  economy
## 92       69     5200       31           37  7299  economy
## 93       69     5200       31           37  7799  economy
## 94       69     5200       31            37  7499  economy
## 95       69     5200       31           37  7999  economy
## 96       69     5200       31           37  8249  economy
## 97       97     5200       27           34  8949  economy
## 98       97     5200       27           34  9549  economy
## 99       152     5200      17           22 13499   luxury
## 100      152     5200      17           22 14399   luxury
## 101      152     5200      19           25 13499   luxury
## 102      160     5200      19           25 17199   luxury
## 103      200     5200      17           23 19699   luxury
## 104      160     5200      19           25 18399   luxury
## 105      97     5000       19           24 11900  economy
## 106         95     4150    28           33 13200  economy
## 107      97     5000       19           24 12440  economy
## 108      95     4150       25           25 13860   luxury
## 109      95     5000       19           24 15580   luxury
## 110         95     4150    28           33 16900   luxury
## 111      95     5000       19           24 16695   luxury
## 112      95     4150       25           25 17075   luxury
## 113      97     5000       19           24 16630   luxury
## 114      95     4150       28           33 17950   luxury
## 115      142     5600      18           24 18150   luxury
```

```
## 116      68     5500       37           41  5572  economy
## 117      102     5500      24           30  7957  economy
## 118      68     5500       31           38  6229  economy
## 119      68     5500       31           38  6692  economy
## 120      68     5500       31           38  7609  economy
## 121      88     5000       24           30  8921  economy
## 122      145     5000      19           24 12764  economy
## 123      143     5500      19           27 22018   luxury
## 124      207     5900      17           25 32528   luxury
## 125      207     5900      17           25 34028   luxury
## 126      207     5900       17           25 37028   luxury
## 127      100     4150      23           31  9295  economy
## 128      100     4150      23           31  9895  economy
## 129      110     5250      21           28 11850  economy
## 130      110     5250       21           28 12170  economy
## 131      110     5250      21           28 15040   luxury
## 132      110     5250      21           28 15510   luxury
## 133      160     5500      19           26 18150   luxury
## 134      160     5500      19           26 18620   luxury
## 135      69     4900       31           36  5118  economy
## 136      73     4400       26           31  7053  economy
## 137      73     4400       26           31  7603  economy
## 138      82     4800       32          37  7126  economy
## 139      82     4400       28           33  7775  economy
## 140      94     5200       26           32  9960  economy
## 141      82     4800       24           25  9233  economy
## 142      111     4800      24            29 11259  economy
```

```
## 143      82     4800       28           32  7463  economy
## 144      94     5200       25           31 10198  economy
## 145      82     4800       23           29  8013  economy
## 146      111     4800      23           23 11694  economy
## 147      62     4800       35           39  5348  economy
## 148      62     4800       31           38  6338  economy
## 149      62     4800       31           38  6488  economy
## 150      62     4800       31           37  6918  economy
## 151      62     4800       27           32  7898  economy
## 152      62     4800       27           32  8778  economy
## 153      70     4800       30           37  6938  economy
## 154      70     4800       30           37  7198  economy
## 155      56     4500       34           36  7898  economy
## 156      56     4500       38           47  7788  economy
## 157      70     4800       38           47  7738  economy
## 158         70     4800    28           34  8358  economy
## 159      70     4800       28           34  9258  economy
## 160      70     4800       29           34  8058  economy
## 161      70     4800       29           34  8238  economy
## 162        112     6600    26           29  9298  economy
## 163      112     6600      26           29  9538  economy
## 164      116     4800      24           30  8449  economy
## 165      116     4800      24           30  9639  economy
## 166      116     4800      24           30  9989  economy
## 167      116     4800      24           30 11199  economy
## 168      116     4800      24           30 11549  economy
## 169      116     4800      24           30 17669   luxury
```

```
## 170      92     4200       29           34  8948  economy
## 171      73     4500       30           33 10698  economy
## 172      92     4200       27           32  9988  economy
## 173      92     4200       27           32 10898  economy
## 174      92     4200       27           32 11248  economy
## 175      161     5200      20           24 16558   luxury
## 176      161     5200      19           24 15998   luxury
## 177      156     5200      20           24 15690   luxury
## 178      156     5200       19           24 15750   luxury
## 179      52     4800       37           46  7775  economy
## 180      85     5250       27           34  7975  economy
## 181      52     4800       37           46  7995  economy
## 182      85     5250        27           34  8195  economy
## 183      85     5250       27           34  8495  economy
## 184      68     4500       37           42  9495  economy
## 185      100     5500      26           32  9995  economy
## 186      90     5500       24           29 11595  economy
## 187      90     5500       24           29  9980  economy
## 188      110     5500      19           24 13295   luxury
## 189      68     4500       33           38 13845   luxury
## 190      88     5500       25          31 12290  economy
## 191      114     5400      23           28 12940  economy
## 192      114     5400      23           28 13415   luxury
## 193      114     5400      24           28 15985   luxury
## 194      114     5400      24            28 16515   luxury
## 195      162     5100      17           22 18420   luxury
## 196      162     5100      17           22 18950   luxury
```

```
## 197       114     5400      23           28 16845   luxury
## 198       160     5300      19           25 19045   luxury
## 199       134     5500      18           23 21485   luxury
## 200       106     4800      26           27 22470   luxury
## 201       114     5400      19           25 22625   luxury
```