**ILLINOIS STATE UNIVERSITY**

**FACULTY OF ARTS AND SCIENCES**

**DEPARTMENT OF APPLIED STATISTIC**

**MAT 450**

**COMPARISON OF FUEL ECONOMY FOR TOYOTA AND HONDA VEHICLES.**

**BY**

**AGYEMANG, ERIC**

**eagyem2@ilstu.edu**

**MAY 1, 2020**

**CHAPTER ONE**

**1.1 Background**

The description of ‘‘fuel economy’’ of vehicles driven in the United States (US) is often measured in miles per gallon (mpg)*.* This represents the number of miles a vehicle is expected to cover on a single gallon of fuel. When using a measure expressed as distance per fuel unit, a higher number means more efficient, while a lower number means less efficient and when expressed as units of fuel per fixed distance (L/100 km etc.), a lower number means more efficient, while a higher number means less efficient. The fuel is gasoline and can be purchased at a specific price per gallon from petroleum stations across the country.

A vehicle’s fuel economy may be viewed in two distinct ways. One observed on the expected mileage the vehicle can travel on a highway setting and the other representing the expected mileage the vehicle can travel within a city environment. The later is typically less than the forma due to the fact that it is attributed to the increase in stops and the zones of low speeding limits encountered in cities. Highway driving can be expected to have less frequent stops and areas of road in which a driver can travel upon at much higher speeds. This form of drive essentially optimizes the energy obtained from fuel better than driving in a city environment.

For this study, two popular Japanese automobile makes Toyota and Honda are taken into consideration. Toyota is a multinational automobile manufacturer headquartered in Toyota Aichi, Japan. It is the largest automobile manufacturer in Japan, and the second largest in the world behind Volkswagen. As at December, 2019, Toyota was the tenth-largest company in the world by revenue (Statista*. Retrieved December, 2019*). Honda on the other hand is a public multinational conglomerate corporation and the first Japanese automobile manufacturer to release a dedicated luxury brand, Acura, in 1986 (*OICA,2011*). In 2013, Honda invested about 5.7% (US$6.8 billion) of its revenues in research and development. It became the first Japanese automaker to be a net exporter from the United States, exporting 108,705 Honda and Acura models (Ross, Jeffrey N. 29 January, 2014).

**1.2 Research Questions**

The study was guided by the following questions.

* Which of the two companies Honda and Toyota manufacture the vehicle with the better, overall city and highway fuel economy?
* Which companies manufacture the vehicle with the better city and highway fuel economy based upon individual class?

**1.3 Significance of The Study**

Most people all over the world and therefore the US usually think of fuel economy of vehicles which influence their decision in purchasing and using vehicles since not all vehicles possess the same fuel economy values. The purpose of this study is to compare two popular Japanese multinational automobile makes; Toyota and Honda in an effort to show potential consumers which of the two makes produce better-performing vehicle in the field of fuel economy. By estimating the total cost expected for both options, this study’s conclusion can be used as a means for consumers wanting to make the smartest and economically as possible choice by presenting them with the cheapest available option.

**CHAPTER TWO**

**METHODOLOGY**

**2.1 Data Description**

Vehicles of various classes belonging to Toyota and Honda were observed for the analysis. Data from the years 2014 to 2018 was gathered from TrueCar.com and United States Environmental Protection Agency (EPA) database through the Kaggle.com. A total population of 49662 Honda vehicles and 49127 Toyota vehicles where obtained for the study. Each vehicle occupies an individual row and only the year, make, model, class, city fuel economy, and highway fuel economy for each vehicle were considered for this study. The data was imported onto R version 3.6.3 and SAS version 9.4 where analysis begun.

**2.1.1 Kernel Density Estimation**

Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample. It is used in this study to show the shape of the estimated population of City mpg and Highway mpg of the two vehicle makes. It is used to overcome the discreteness of the other methods like histogram by centering a smooth kernel function at each data point then summing to get a density estimate. The basic kernel estimator can be expressed as

where K is the kernel and h is the bandwidth (h=4 in our case). The kernel is a symmetric, usually positive function that integrates to one. Other common kernel functions are uniform, triangle, Epanechnikov, biweight, Gaussian(normal) and cosine. The bandwidth, h, is the smoothing parameter. Hence, large bandwidth result in very smooth estimates while small values result in wiggly estimates.

**2.2 Sampling Techniques.**

While an original population of possible observations is designated at the start of this study, it should be considered unrealistic in a real-world setting to measure each member of the population. As a result, a subsection of this population was chosen and observed as a sample. The goal of taking this sample is to obtain information that is closely representative of the population’s true values, while also keeping the use of time and resources to a minimum for parsimony

While several strategies can be implemented to take samples, some can be proven more efficient than others based on the nature of the population being studied. This study utilizes Simple Random Sampling (SRS) and Stratified Sampling methods for the analysis and using the survey designs we compute all estimates

**2.2.1 Sample Size**

In order to get a good representation of the population of Honda and Toyota vehicles, an overall proportion was decided upon before determining sample sizes for each vehicle class. The resulting number is designated as n. For comparative purpose, the value calculated from this proportion was kept the same for both the simple random sample (SRS) and a stratified sample. A margin of error of 0.10 was set up for the sample size and using the R package (*samplingbook)* and the function (*sample.size.mean),* we set up the margin of error to 0.10*,* population standard deviation, total population and the level of 0.95, and so we are able to compute the sample sizes required for the analysis. For Honda City mpg, a sample size selected was 15300 while Honda Highway mpg, a sample size of 15806 was chosen. For Toyota City mpg, the sample size selected was 10149 while Toyota Highway mpg had a sample size of 12154.

**2.2.2** **Simple Random Sampling (SRS) Method**

A Simple Random Sampling Replacement (SRS)is used as our basic method for our probability sampling where every possible subset of *n* distinct units in the population has the same probability of being selected as the sample. There are possible ways and the probability of selecting an individual sample is p(S)= where N is the total population and n sample size.

**2.2.3** **Stratified Sampling Method**

Suppose the population is partitioned into disjoint sets of sampling units called strata. If a sample is selected within each stratum, then this sampling procedure is known as stratified sampling. Each potential observation within a stratum can only belong to that strata. This method ensures a good representation of the population.

**2.2.3.1 Designation of the Strata**

The strata created for this project were established based upon the class in which each vehicle belongs to. Vehicles from both Honda and Toyota were determined to fit into one of eleven common classes. These classes were identified as:

Stratum A: Compact Cars Stratum F: Small sport Utility Cars

Stratum B: Large Cars Stratum G: Large Sport Utility cars

Stratum C: Midsize Station Wagons Stratum H: Standard Pickup Truck 4WD

Stratum D: Small Station Wagon Stratum I: Standard Utility cars 4WD

Stratum E: Small Pickup Truck 2WD Stratum J: Two Seaters

Stratum K: Van Cars

While the stratum Large Cars is present in Honda but not present in Toyota, Midsize Cars on the other side is present in Toyota but are present in Honda. Each class was designated as its own strata, containing only vehicles which were of that specific class. There were eleven classes in each case.

**2.2.3.2 Neyman Allocation to Strata**

This project assumes equal amount of cost of measuring a vehicle for a sample for every measurement made and so a special case of the optimal allocation (the Neyman allocation) was employed. The Neyman allocation is a type of optimal allocation which hold the sample chosen from a stratum, *,* proportional to the stratum sizemultiplied by the population standard deviation () in that same stratum.

**\*n** and was calculated for all stratum for all the four cases Honda City mpg and Highway mpg, Toyota City mpg and Highway mpg. The sample quantities chosen from each stratum were selected such that the sum of all 11 samples equaled ***n*** (the total sample size). Using the R package *strAlloc,* we are able to compute are compute all h stratum.

**2.3 Confidence Interval for Sample Means**

Another measure which can gauge the effectiveness of the SRS and the Stratified Random Sample is the computation of a confidence interval for the sample means which are the average fuel economy ascertained. Since the overall population mean can be estimated with the calculation of fort the simple random sample case and for the Stratified Random Sampled case, the 100(1*−α*)% confidence interval determined by using:

**\*SE( and \*SE()** respectively where and so  **=1.96** from the normal distribution table.

**2.4** **Comparison of The SRS and Stratified Random Sample Estimates on the Fuel Economies.**

For each of the four cases considered, the calculated estimates are compared to the true values corresponding to what each estimate is attempting to represent. The two-specific variance (variance for SRS and Stratified Random Sample) can be used in determining the efficiency of the stratified random sample taken. The stratified random sampling can be compared to a simple random sample of the same sample size by: ***Ratio Of Relative Gain From Stratification***=

The resulting ratio can then be multiplied by the common n chosen for both samples in order to determine what sample size would be needed with a stratified sample in order to obtain the same precision with an SRS of size n. Ideally, this ratio should be below 1 as this would show that taking a stratified random sample would require less of a sample size than a simple random sample to obtain a desired result.

Another approach is the design effect which provides a measure of the precision gained or lost by the use of a more complicated design instead of an SRS. Although it is a useful concept, it is not a way to avoid calculating variances. We need an estimate of the variance from the complex design to find the design effect. It is given as:

Different quantities in the same survey may have different design effects. In a stratified sampling, unless all of the stratum means are equal, the design effect for a stratified sample will usually be less than one. stratification generally gives more precision per observation unit than an SRS.

**CHAPTER THREE**

**3.1 RESULTS AND DISCUSSION**

**3.1.1 Kernel Density Estimation**

The results from the [[1]](#footnote-1)Kernel density estimate provided in the analysis of this paper shows a bandwidth 4 for all four cases. The plots obtained for Honda and Toyota City mpg and Highway mpg’s are normally distributed across vehicles at that bandwidth. Hence the data is smoothed and is underfitted with the bandwidth of 4.

**3.1.2 Comparison of estimates**

From the SAS and R output, the questions posed earlier in this paper can be formally addressed. The first question, which ask in a broad sense, “Which of the two companies Honda and Toyota manufacture the vehicle with the better, overall city and highway fuel economy?”, can be answered with the results obtained from the simple random sample. From the simple random sample conducted, it was found that the estimated total fuel economy for Honda city driving was 2868441mpg with a standard error of 5487.9 and an average fuel economy of 17.795mpg with a confidence interval of (17.7279, 17.86136). On the other hand, Toyota vehicles produced over the same period with the simple random sample shows an estimated total fuel economy of 3972278mpg for city driving with a standard error of 6191.1 and an average city driving fuel economy of 16.704mpg with a confidence interval of (16.65305, 16.7551). Based on these results, it can be concluded that Honda produced a better, overall city fuel economy, in regard to city driving, during the period of 2014-2018.

Next, using the same simple random sample, it was found that the estimated total fuel economy for Honda highway was 3711570mpg with a standard error of 5487.9 and an average city fuel economy across all vehicles of 23.787mpg with a confidence interval of (23.71821, 23.85486). For Toyota vehicles produced over the same period, the simple random sample showed that the estimated total fuel economy for highway driving was 5345873mpg with a standard error of 6956.4 and an average city fuel economy across all vehicles of 22.48mpg with a confidence interval of (22.42293, 22.5376). Based on these results, it can be concluded that Honda had produced the more fuel-efficient vehicles, in regard to highway driving, during the period of 2014-2018. Since Honda wins in both categories, the company can be considered the better option over Toyota if one were to compare the two makes in only a general sense.

The second question, “Which companies manufacture the vehicle with the better city and highway fuel economy based upon individual class?” can be considered more specific than the first question. This can be answered using the stratification method used in our analysis. It was found that the estimated total fuel economy for Honda city driving was 3573468mpg with a standard error of 7063.7 and an average city fuel economy across all vehicles of 17.732mpg with a confidence interval of (17.66293, 17.80033). The design effect was 1.4582. For Toyota vehicles produced over the same period, the stratified sampling method showed the estimated total fuel economy for city driving was 4052515mpg with a standard error of 6104.5 and an average city fuel economy across all vehicles of 16.647mpg with a confidence interval of (16.59789, 16.69619). The design effect was 1.2124. Based on these results, it can be concluded that Honda had produced the more fuel-efficient vehicles, in regard to city driving using stratified sampling technique. Once again, even after vehicles were broken down by class and analyzed separated, the combined results gathered from each of the 11 strata showed that Honda had produced the more fuel-efficient vehicles, in regard to city driving, during the period.

A similar outcome can be seen for highway fuel economy. From the stratified random sample It was found that the estimated total fuel economy for Honda highway driving was 4125383mpg with a standard error of 6083.6 and an average highway fuel economy across all vehicles of 23.741mpg with a confidence interval of (23.67229, 23.80953). The design effect was 1.4755. For Toyota vehicles produced over the same period, the stratified sampling method showed the estimated total fuel economy for highway driving was 5345873mpg with a standard error of 6956.4 and an average city fuel economy across all vehicles of 22.48mpg with a confidence interval of (22.42293, 22.5376). The design effect was 1.3089. Again, Honda had produced the more fuel-efficient vehicles, in regard to highway driving for the stratified sampling technique, during the period of 2014-2018.

**3.1.2** **Gain from Stratification**

The results produced from the analysis of this project shows that the gain from stratification for Honda vehicles City mpg and Highway mpg are 0.1008905and 0.1021133, respectively. These imply that if the value of the estimated variance from the stratified sampling and that from SRS were the population variances, we will expect that we would need observations and observations respectively with the stratified sample to obtain the same precision as from an SRS of 15300 observations.

Also, the results produced shows that the gain from stratification for Toyota vehicles City mpg and Highway mpg are 0.97222 and 0.04136722, respectively. These imply that if the value of the estimated variance from the stratified sampling and that from SRS were the population variances, we would expect that we would need only observations andobservations respectively with the stratified sample to obtain the same precision as from an SRS 14898 observations.

**3.2 CONCLUSION AND RECOMMENDATION**

The results from all techniques used in this project express the same results. Therefore, using the results presented in this project, it can be concluded that buying a Honda vehicle would be the better option over Toyota in terms of fuel economy cost efficiency. Hence, this paper recommends buyers to buy Honda vehicles to minimize cost.

**Reference**

1. Sampling: Design and Analysis: Edition 2 Sharon L. Lohr December 9, 2009
2. <https://vita.had.co.nz/papers/density-estimation.Pdf>
3. <https://www.kaggle.com/epa/fuel-economy>
4. <https://www.kaggle.com/jpayne/852k-used-car-listings>
5. <https://en.wikipedia.org/wiki/Fuel_economy_in_automobiles>

**Appendix A**

1. **SAS OUTPUT**

**Table 1 DESCRIPTIVE STATISTICS FOR THE MAKES (HONDA AND TOYOTA)**

**The MEANS Procedure**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Make | N Obs | Variable | Mean | Variance | Std Dev | Lower 95% CL for Mean | Upper 95% CL for Mean |
| Honda | 49662 | **Price**  **Year**  **Mileage**  **City mpg**  **Highway mpg** | 19804.31  2014.88  33723.94  17.7946317  23.7865370 | 35755327.98  0.9020238  426405922  57.5607911  60.3502160 | 26.8323234  0.0042618  92.6615394  0.0340448  0.0348600 | 19751.71  2014.87  33542.32  17.7279034  23.7182110 | 19856.90  2014.89  33905.55  17.8613600  23.8548630 |
| Toyota | 49127 | **Price**  **Year**  **Mileage**  **City mpg**  **Highway mpg** | 22794.60  2015.21  32665.02  16.7040731  22.4802654 | 82722914.85  0.9928105  369091505  33.2982497  42.0393807 | 41.0348445  0.0044954  86.6776050  0.0260346  0.0292528 | 22714.17  2015.20  32495.13  16.6530450  22.4229295 | 22875.03  2015.22  32834.91  16.7551012  22.5376014 |

**Table 2 Descriptive Statistics of HONDA**

**The MEANS Procedure**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | N | Mean | Std Dev | Minimum | Maximum |
| Price  Year  Mileage  City mpg  High mpg | 49662  49662  49662  49662  49662 | 19804.31  2014.88  33723.94  17.7946317  23.7865370 | 5979.58  0.9497493  20649.60  7.5868828  7.7685401 | 4950  2014  5  8  10 | 99999  2018  389781  132  110 |

**Table 3 Descriptive Statistics of HONDA on City\_Mpg Based on Class**

**The MEANS Procedure**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Analysis Variable : City\_mpg | | | | | | | | |
| Class | **N Obs** | **Median** | **Mean** | **Variance** | **Std Dev** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| Compact Car | **4961** | 17 | 18.50 | 70.9351308 | 8.4222996 | 18.2660810 | 18.7349269 |
| Large Cars | **7022** | 17 | 18.525 | 80.7607016 | 8.9866958 | 18.3154047 | 18.7358627 |
| Midsize Station Wagons | **80** | 15 | 16.275 | 14.0753165 | 3.7517085 | 15.4400982 | 17.1099018 |
| Small Station Wagon | **3070** | 15 | 16.842 | 49.4866013 | 7.0346714 | 16.5930799 | 17.0909592 |
| Small Pickup Truck 2WD | **7106** | 15 | 16.902 | 45.3145433 | 6.7316078 | 16.7455135 | 17.0585957 |
| Small Sport Utility Cars | **1528** | 15 | 15.776 | 16.3895873 | 4.0484055 | 15.5736834 | 15.9799815 |
| Large Sport U cars | **10988** | 16 | 17.855 | 55.8592519 | 7.4739047 | 17.7148082 | 17.9943290 |
| Standard Pickup Truck 4WD | **11566** | 16 | 17.993 | 51.4480294 | 7.1727282 | 17.8624363 | 18.1239030 |
| Standard U cars 4WD | **560** | 17 | 18.759 | 109.0634264 | 10.4433436 | 17.8920965 | 19.6257607 |
| Two Seaters | **1385** | 17 | 18.088 | 45.3477265 | 6.7340721 | 17.7331254 | 18.4430479 |
| Van Cars | **1396** | 16 | 17.7485 | 60.9525069 | 7.8072087 | 17.3386671 | 18.1584675 |

**Table 4 Descriptive Statistics of HONDA on High\_mpg** **Based on Class**

**The MEANS Procedure**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Analysis Variable : High\_mpg | | | | | | | |
| Class | **N Obs** | **Median** | **Mean** | **Variance** | **Std Dev** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| Compact Car | **4961** | 23 | 24.6736545 | 67.8670659 | 8.2381470 | 24.4443572 | 24.9029518 |
| Large Cars | **7022** | 23 | 24.5944175 | 74.3063525 | 8.6201133 | 24.3927642 | 24.7960709 |
| Midsize Station Wagons | **80** | 20 | 21.7125000 | 27.6757911 | 5.2607786 | 20.5417711 | 22.8832289 |
| Small Station Wagon | **3070** | 21 | 22.7462541 | 51.7423695 | 7.1932169 | 22.4917039 | 23.0008042 |
| Small Pickup Truck 2WD | **7106** | 21 | 22.7303687 | 50.4472043 | 7.1026195 | 22.5651999 | 22.8955375 |
| Small Sport Utility Cars | **1528** | 20 | 21.3560209 | 31.4527339 | 5.6082737 | 21.0745977 | 21.6374442 |
| Large Sport U cars | **10988** | 23 | 23.8704041 | 60.1064399 | 7.7528343 | 23.7254277 | 24.0153804 |
| Standard Pickup Truck 4WD | **11566** | 23 | 24.0180702 | 58.0286401 | 7.6176532 | 23.8792275 | 24.1569129 |
| Standard U cars 4WD | **560** | 24 | 24.9660714 | 91.4818522 | 9.5646146 | 24.1721768 | 25.7599661 |
| Two Seaters | **1385** | 23 | 24.1018051 | 54.1088479 | 7.3558717 | 23.7140680 | 24.4895421 |
| Van Cars | **1396** | 23 | 23.6489971 | 59.5154183 | 7.7146237 | 23.2439579 | 24.0540364 |

**Table 5 Descriptive Statistics For TOYOTA**

**The MEANS Procedure**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | N | Mean | Std Dev | Minimum | Maximum |
| Price  Year  Mileage  City\_mpg  High\_mpg | 49127  49127  49127  49127  49127 | 22794.60  2015.21  32665.02  16.7040731  22.4802654 | 9095.21  0.9963988  19211.75  5.7704636  6.4837783 | 5814.00  2014.00  5.00  8.00  10.00 | 183710.00  2018.00  475636.00  128.00  110.00 |

**Table 6 Descriptive Statistics for TOYOTA on City\_mpg Based on Class**

**The MEANS Procedure**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Analysis Variable: City\_mpg | | | | | | |
| Class | **N Obs** | **Mean** | **Variance** | **Std Dev** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| Compact Car | **7981** | 16.4136073 | 26.7791582 | 5.1748583 | 16.3000582 | 16.5271564 |
| Large Sport Utility cars | **2139** | 15.6568490 | 13.9130623 | 3.7300218 | 15.4986878 | 15.8150102 |
| Midsize Car | **6578** | 17.5185467 | 39.3796925 | 6.2753241 | 17.3668705 | 17.6702228 |
| Midsize Station Wagons | **908** | 16.2147577 | 19.5977092 | 4.4269300 | 15.9264291 | 16.5030863 |
| Small Pickup Truck 2WD | **3469** | 16.5269530 | 17.6553431 | 4.2018262 | 16.3870795 | 16.6668266 |
| Small Standard Wagons | **6769** | 16.9392820 | 42.9997111 | 6.5574165 | 16.7830406 | 17.0955235 |
| Small Sport Utility Cars | **5233** | 16.7729792 | 42.3956994 | 6.5111980 | 16.5965241 | 16.9494342 |
| Standard Pickup Truck 4WD | **7106** | 16.0963974 | 31.6575606 | 5.6265052 | 15.9655551 | 16.2272397 |
| Standard Utility cars 4WD | **3109** | 16.3965905 | 29.8983281 | 5.4679364 | 16.2043122 | 16.5888689 |
| Two Seaters | **5054** | 17.5182034 | 36.7618890 | 6.0631583 | 17.3510045 | 17.6854023 |
| Van Cars | **781** | 16.0204866 | 29.7124003 | 5.4509082 | 15.6376037 | 16.4033694 |

**Table 7 Descriptive Statistics For TOYOTA on High\_mpg Based on Class**

**The MEANS Procedure**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Analysis Variable : High\_mpg | | | | | | |
| Class | **N Obs** | **Mean** | **Variance** | **Std Dev** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| Compact Car | **7981** | 22.3124922 | 38.4584768 | 6.2014899 | 22.1764162 | 22.4485681 |
| Large Sport Utility cars | **2139** | 21.3749416 | 27.6993905 | 5.2630210 | 21.1517778 | 21.5981054 |
| Midsize Car | **6578** | 22.9702037 | 43.9696153 | 6.6309589 | 22.8099318 | 23.1304756 |
| Midsize Station Wagons | **908** | 21.8755507 | 33.4828415 | 5.7864360 | 21.4986766 | 22.2524247 |
| Small Pickup Truck 2WD | **3469** | 22.1640242 | 30.8401587 | 5.5533916 | 21.9791587 | 22.3488897 |
| Small Standard Wagons | **6769** | 23.0717979 | 50.6349152 | 7.1158215 | 22.9022515 | 23.2413443 |
| Small Sport Utility Cars | **5233** | 22.6287025 | 50.4410491 | 7.1021862 | 22.4362315 | 22.8211734 |
| Standard Pickup Truck 4WD | **7106** | 21.8335210 | 40.3831183 | 6.3547713 | 21.6857431 | 21.9812988 |
| Standard Utility cars 4WD | **3109** | 22.3116758 | 40.3201371 | 6.3498139 | 22.0883865 | 22.5349651 |
| Two Seaters | **5054** | 23.1228730 | 41.6969515 | 6.4573177 | 22.9448047 | 23.3009413 |
| Van Cars | **781** | 21.4788732 | 37.8832069 | 6.1549335 | 21.0465383 | 21.9112082 |

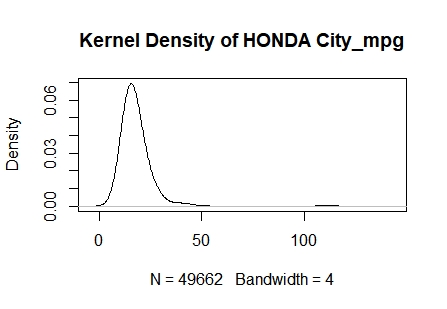
**R OUTPUT**

**#FIGURE 1**: R-CODE FOR KERNEL DENSITY ESTIMATION

> #The Kernel Density Estimation of Honda City Mpg

> den\_H1 <- density(HONDA$City\_mpg, bw = 4, na.rm = TRUE)

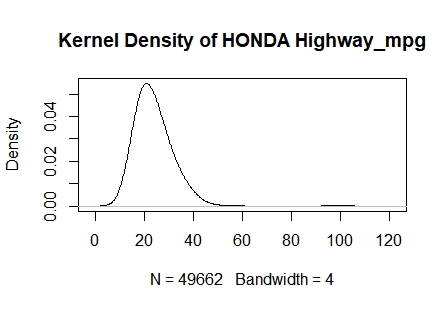
> plot1=plot(den\_H1, main="Kernel Density of HONDA City\_mpg", ylab = "Density")



**# FIGURE 2**: **The Kernel Density Estimation of Honda Highway Mpg**

> den\_H2 <- density(HONDA$High\_mpg, bw = 4, na.rm = TRUE)

> plot1=plot(den\_H2, main="Kernel Density of HONDA Highway\_mpg", ylab = "Density")

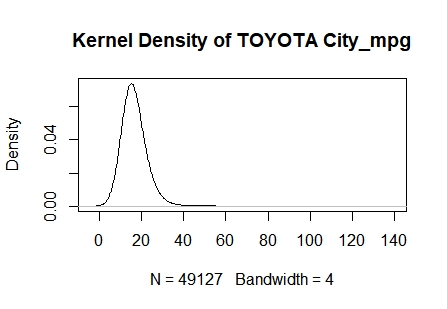


> **# FIGURE 3**: **The Kernel Density Estimation of Toyota City Mpg**

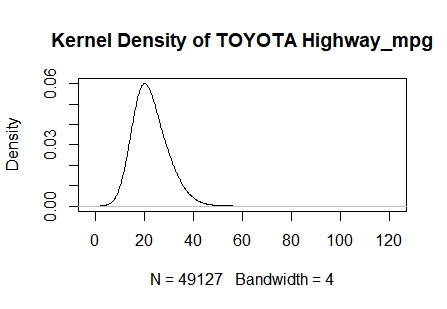
> den\_T1 <- density(TOYOTA$City\_mpg, bw = 4, na.rm = TRUE)

> plot1=plot(den\_T1, main="Kernel Density of TOYOTA City\_mpg", ylab = "Density")

>



|  |
| --- |
| # **FIGURE 4:** **The Kernel Density Estimation of Toyota Highway Mpg**  > den\_T2 <- density(TOYOTA$High\_mpg, bw = 4, na.rm = TRUE)  > plot1=plot(den\_T2, main="Kernel Density of TOYOTA Highway\_mpg", ylab = "Density") |



**Appendix B**

**i) R-OUTPUT FROM THE ANALYSIS OF THIS PROJECT**

**> ####STRATIFIED RANDOM SAMPLE FOR HONDA VEHICLE MAKE ##########**

>

> library(stratification)

> library(PracTools)

> **#Sample Size Estimation#**

> var(HONDA$City\_mpg)

[1] 57.56079

> library(samplingbook)

> sample.size.mean(e= 0.10 , S=sqrt(57.56079), N= 49662, level= 0.95)

sample.size.mean object: Sample size for mean estimate

With finite population correction: N=49662, precision e=0.1 and standard deviation S=7.5869

Sample size needed: 15300

> #Neyman Allocation,Stratified Sampling#

> Stratum=c("A", "B", "C", "D", "E", "F", "G", "H", "I", "J","K")

> Nh = c(4961, 7022, 80, 3070, 7106, 1528, 10988,11566,560,1385,1396)

> Sh = c(8.422, 8.9867, 3.752, 7.035, 6.732, 4.049, 7.474, 7.173, 10.443, 6.734, 7.807)

> sample=strAlloc(n.tot = 15300, Nh = Nh, Sh = Sh, alloc = "neyman")

> sample

allocation = neyman

Nh = 4961, 7022, 80, 3070, 7106, 1528, 10988, 11566, 560, 1385, 1396

Sh = 8.4220, 8.9867, 3.7520, 7.0350, 6.7320, 4.0490, 7.4740, 7.1730, 10.4430, 6.7340, 7.8070

nh = 1718.57901, 2595.64986, 12.34633, 888.35697, 1967.67945, 254.48147, 3377.97774, 3412.47170, 240.54611, 383.62590, 448.28544

nh/n = 0.1123254255, 0.1696503178, 0.0008069496, 0.0580625474, 0.1286064999, 0.0166327760, 0.2207828589, 0.2230373658, 0.0157219682, 0.0250735885, 0.0292997022

anticipated SE of estimated mean = 0.05016355

>

**> #Survey Design Stratified FOR HONDA CITY MPG#**

> HONDA$P <- NA

>

> #input Stratum weights#

> HONDA$P[HONDA$Class=="Compact Car"] <- 4961/1719

> HONDA$P[HONDA$Class=="Large Cars"] <- 7022/2596

> HONDA$P[HONDA$Class=="Large Sport Ucars"] <- 80/12

> HONDA$P[HONDA$Class=="Midsize Station Wagons"] <- 3070/888

> HONDA$P[HONDA$Class=="Small Pickup Truck 2WD"] <- 7106/1968

> HONDA$P[HONDA$Class=="Small sport Utility Cars"] <- 1528/254

> HONDA$P[HONDA$Class=="Small Station Wagon"] <- 10988/3378

> HONDA$P[HONDA$Class=="Standard Pickup Truck 4WD"]<- 11566/3412

> HONDA$P[HONDA$Class=="Standard Ucars 4WD"] <- 560/241

> HONDA$P[HONDA$Class=="Two Seaters"] <- 1385/384

> HONDA$P[HONDA$Class=="Van Cars"] <- 1396/448

**> ####STRATIFIED RANDOM SAMPLING DESIGN FOR HONDA CITY MPG######**

> stat\_stra=svydesign(id=~1, strata=~Class, weights=~P, data=HONDA)

>

> **#Mean Estimate of Str#**

> ybar\_str=svymean(~HONDA$City\_mpg, design = stat\_stra)

> ybar\_str

mean SE

HONDA$City\_mpg 17.732 0.0351

>

> #Variance of the estimated mean

> var\_ybar\_str=svyvar(svymean(~HONDA$City\_mpg, design = stat\_stra))

> var\_ybar\_str

[1] 0.00123201

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| #Confidence Interval of the Estimate mean#  > confint(ybar\_str)  2.5 % 97.5 %  HONDA$City\_mpg 17.66293 17.80033  >  > #Estimated total of Str  > t\_cup\_str=svytotal(~HONDA$City\_mpg, design = stat\_stra)  > t\_cup\_str  total SE  HONDA$City\_mpg 3573468 7063.7  #Variance of the estimated total  > var\_t\_cup\_str=svyvar(svytotal(~HONDA$City\_mpg, design = stat\_stra))  > var\_t\_cup\_str  [1] 3038523   |  | | --- | | > #Design Effect#  > svymean(~HONDA$City\_mpg, design = stat\_stra, deff = TRUE)  mean SE DEff  HONDA$City\_mpg 17.73163 0.03505 1.4582 | |  | | |  | | --- | |  | |   > **#########SIMPLE RANDOM SAMPLE FOR HONDA CITY MPG##############**  >  > **####estimate the statistics quantities for SRS FOR HONDA CITY MPG**  > N=49662  > n=15300  > HONDA$fpc\_srs=N  > HONDA$wgt\_stra=N/n  > View(HONDA)  > srs\_design=svydesign(ids=~1, weights =~wgt\_srs, data=HONDA)  > #Estimates SRS#  > #Mean Estimate of SRS#  > ybar\_srs=svymean(~HONDA$City\_mpg,srs\_design)  > ybar\_srs  mean SE  HONDA$City\_mpg 17.795 0.034   |  | | --- | | > #Confidence interval of the mean eastimate  > confint(svymean(~HONDA$City\_mpg,srs\_design),df=degf(srs\_design))  2.5 % 97.5 %  HONDA$City\_mpg 17.7279 17.86136 | |  | | |  | | --- | | > | | | > #Total Estimate of SRS#  > t\_cup\_srs=svytotal(~City\_mpg,srs\_design)  > t\_cup\_srs  total SE  City\_mpg 2868441 5487.9   |  | | --- | | > #Variance of the eastimated total for SRS for Honda City mpg  > var\_t\_cup\_srs=5487.9^2  > var\_t\_cup\_srs  [1] 30117046 | |  | | |  | | --- | |  | | |   **####STRATIFIED RANDOM SAMPLE HONDA VEHICLE MAKE HIGHWAY MPG #######**  >  > library(stratification)  > library(PracTools)  > **#Sample Size Estimation#**  > var(HONDA$High\_mpg)  [1] 60.35022  > sample.size.mean(e= 0.10, S=sqrt(60.35022), N= 49662, level= 0.95)  sample.size.mean object: Sample size for mean estimate  With finite population correction: N=49662, precision e=0.1 and standard deviation S=7.7685  Sample size needed: 15806  > #Neyman Allocation,Stratified Sampling#  > Stratum=c("A", "B", "C", "D", "E", "F", "G", "H", "I", "J","K")  > Nh =c(4961, 7022, 80, 3070, 7106, 1528, 10988,11566,560,1385,1396)  > Sh = c(8.2381470, 8.6201133, 5.2607786, 7.1932169, 7.1026195, 5.6082737, 7.7528343, 7.6176532, 9.5646146, 7.3558717, 7.7146237)  > sample=strAlloc(n.tot = 15806, Nh = Nh, Sh = Sh, alloc = "neyman")  > sample  allocation = neyman  Nh = 4961, 7022, 80, 3070, 7106, 1528, 10988, 11566, 560, 1385, 1396  Sh = 8.238147, 8.620113, 5.260779, 7.193217, 7.102620, 5.608274, 7.752834, 7.617653, 9.564615, 7.355872, 7.714624  nh = 1688.61274, 2500.95051, 17.38887, 912.41587, 2085.33124, 354.06570, 3519.73892, 3640.28749, 221.30274, 420.93517,  444.97075  nh/n = 0.106833654, 0.158227920, 0.001100143,  0.057725918, 0.131932889, 0.022400715, 0.222683723, 0.230310483, 0.014001186, 0.026631353, 0.028152015  anticipated SE of estimated mean = 0.05050903  **#Survey Design Stratified#**  > HONDA$P <- NA  >  > #input Stratum weights#  > HONDA$P[HONDA$Class=="Compact Car"] <- 4961/1689  > HONDA$P[HONDA$Class=="Large Cars"] <- 7022/2501  > HONDA$P[HONDA$Class=="Large Sport Ucars"] <- 80/17  > HONDA$P[HONDA$Class=="Midsize Station Wagons"] <- 3070/913  > HONDA$P[HONDA$Class=="Small Pickup Truck 2WD"] <-7106/2085  > HONDA$P[HONDA$Class=="Small sport Utility Cars"]<-1528/354  > HONDA$P[HONDA$Class=="Small Station Wagon"] <- 10988/3520  > HONDA$P[HONDA$Class=="Standard Pickup Truck 4WD"]<-11566/3640  > HONDA$P[HONDA$Class=="Standard Ucars 4WD"] <- 560/221  > HONDA$P[HONDA$Class=="Two Seaters"] <- 1385/421  > HONDA$P[HONDA$Class=="Van Cars"] <- 1396/445   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | > **###STRATIFIED RANDOM SAMPLING DESIG FOR HONDA HIGHWAY MPG###**  > stat\_stra=svydesign(id=~1, strata=~Class, weights=~P, data=HONDA)  >  > **#Mean Estimate of Str#**  > ybar\_str=svymean(~HONDA$High\_mpg, design = stat\_stra)  > ybar\_str  mean SE  HONDA$High\_mpg 23.741 0.035   |  | | --- | | > #Variance of the estimated mean  > var\_ybar\_str=0.035^2  > var\_ybar\_str  [1] 0.001225 | |  | | > #Confidence Interval of the Estimate mean#  > confint(ybar\_str)  2.5 % 97.5 %  HONDA$High\_mpg 23.67229 23.80953  >  >  > #Estimated total of Stratified sampling  > t\_cup\_str=svytotal(~HONDA$High\_mpg, design = stat\_stra)  > t\_cup\_str  total SE  HONDA$High\_mpg 4125383 1738.6  >  > #Variance of the estimated total  > var\_t\_cup\_str=1738.6^2  > var\_t\_cup\_str  [1] 3021234   |  | | --- | | > #Design Effect#  > svymean(~HONDA$High\_mpg, design = stat\_stra, deff = TRUE)  mean SE DEff  HONDA$High\_mpg 23.741972 0.035418 1.4755 | |  | | |  | | --- | |  | | | |  | | |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | |  | | --- | | **ii)** | | > **########SIMPLE RANDOM SAMPLE FOR HONDA HIGHWAY MPG########**  >  > **####estimate the statistics quantities for SRS ########**  > N=49662  > n=15806  > HONDA$fpc\_srs=N  > HONDA$wgt\_stra=N/n  > srs\_design=svydesign(ids=~1,weights=~wgt\_stra, data=HONDA)  >  > #Estimates SRS#  > #Mean Estimate of SRS#  > ybar\_srs=svymean(~HONDA$High\_mpg,srs\_design)  > ybar\_srs  mean SE  HONDA$High\_mpg 23.787 0.0349  >  > #Confidence interval of the mean eastimate  > confint(svymean(~HONDA$High\_mpg,srs\_design),df=degf(srs\_design))  2.5 % 97.5 %  HONDA$High\_mpg 23.71821 23.85486  >  > #Total Estimate of SRS#  > t\_cup\_srs=svytotal(~High\_mpg,srs\_design)  > t\_cup\_srs  total SE  High\_mpg 3711570 5439.4   |  | | --- | | > #Variance of the eastimate  > var\_t\_cup\_srs=5439.4^2  > var\_t\_cup\_srs  [1] 29587072 | |  | |  |   **iii)** | | |  | | --- | |  | |   **###SRATIFIED RANDOM SAMPLE FOR TOYOTA VEHICLE MAKE CITY MPG ###**  > library(stratification)  > library(PracTools)  >  > **#Sample Size Estimation#**  > var(TOYOTA$City\_mpg)  [1] 33.29825  > library(samplingbook)  > sample.size.mean(e= 0.10 , S=sqrt(33.29825), N= 49127, level= 0.95)  sample.size.mean object: Sample size for mean estimate  With finite population correction: N=49127, precision e=0.1 and standard deviation S=5.7705  Sample size needed: 10149  >  > #Neyman Allocation,Stratified Sampling#  > Stratum=c("A", "B", "C", "D", "E", "F", "G", "H", "I", "J","K")  > Nh = c(7981, 2139, 6578, 908, 3469, 6769, 5233, 7106,3109, 5054,781)  > Sh = c(5.1748583, 3.7300218, 6.2753241, 4.4269300, 4.2018262, 6.5574165, 6.5111980, 5.6265052, 5.4679364, 6.0631583, 5.4509082)  > sample=strAlloc(n.tot = 10149, Nh = Nh, Sh = Sh, alloc = "neyman")  > sample  allocation = neyman  Nh = 7981, 2139, 6578, 908, 3469, 6769, 5233, 7106, 3109, 5054, 781  Sh = 5.174858, 3.730022, 6.275324, 4.426930, 4.201826, 6.557417, 6.511198, 5.626505, 5.467936, 6.063158, 5.450908  nh = 1499.6950, 289.7139, 1498.9157, 145.9606, 529.2850, 1611.7752, 1237.2539, 1451.8144, 617.2930, 1112.7083, 154.5849  nh/n = 0.14776777, 0.02854605, 0.14769098, 0.01438177, 0.05215144, 0.15881123, 0.12190894, 0.14305000, 0.06082304, 0.10963724, 0.01523154  anticipated SE of estimated mean = 0.05017404 | | | |  |   **> #Survey Design Stratified#**  > TOYOTA$P <- NA  > #input Stratum weights#  > TOYOTA$P[TOYOTA$Class == "Compact Car"] <- 7981/1500  > TOYOTA$P[TOYOTA$Class == "Large Sport Utility cars"] <- 2139/290  > TOYOTA$P[TOYOTA$Class == "Midsize Car"] <- 6578/1499  > TOYOTA$P[TOYOTA$Class == "Midsize Station Wagons"] <- 908/146  > TOYOTA$P[TOYOTA$Class == "Small Pickup Truck 2WD"] <- 3469/529  > TOYOTA$P[TOYOTA$Class == "Small Standard Wagons"] <- 6769/1612  > TOYOTA$P[TOYOTA$Class == "Small Sport Utility Cars"] <- 5233/1237  > TOYOTA$P[TOYOTA$Class == "Standard Pickup Truck 4WD"] <- 7106/1452  > TOYOTA$P[TOYOTA$Class == "Standard Utility cars 4WD"] <- 3109/617  > TOYOTA$P[TOYOTA$Class == "Two Seaters"] <- 5054/1112  > TOYOTA$P[TOYOTA$Class == "Van Cars"] <- 781/154  >  > **########STRATIFIED RANDOM SAMPLING DESIGN FOR TOYOTA CITY MPG######**  > stat\_stra=svydesign(id=~1, strata=~Class, weights=~P, data=TOYOTA)  > **#Mean Estimate of Str#**  > ybar\_str=svymean(~TOYOTA$City\_mpg, design = stat\_stra)  > ybar\_str  mean SE  TOYOTA$City\_mpg 16.647 0.0251   |  | | --- | | > ##Variance of the estimated mean  > var\_ybar\_str=0.0251^2  > var\_ybar\_str  [1] 0.00063001 | |  |  |  | | --- | | > #Confidence interval of the mean eastimate  > confint(svymean(~TOYOTA$City\_mpg,stat\_stra),df=degf(stat\_stra))  2.5 % 97.5 %  TOYOTA$City\_mpg 16.59789 16.69619 | |  | | |  |  |  |  |  | | --- | --- | --- | --- | --- | | |  | | --- | | > #Estimated total of Str  > t\_cup\_str=svytotal(~TOYOTA$City\_mpg, design = stat\_stra)  > t\_cup\_str  total SE  TOYOTA$City\_mpg 4052515 6104.5 | |  | | |  | | --- | | > #Design Effect#  > svymean(~TOYOTA$City\_mpg, design = stat\_stra, deff = TRUE)  mean SE DEff  TOYOTA$City\_mpg 16.647041 0.025076 1.2124 | | | |  |  | | --- | | > #Variance of the estimated total  > var\_t\_cup\_str=6104.5^2  > var\_t\_cup\_str  [1] 37264920 | |  | | |  | | --- | |  |   **iv)** | |  |   **> #####################SIMPLE RANDOM SAMPLING FOR TOYOTA CITY MPG####**  >  > **####estimate the statistics quantities for SRS ########**  > N=49127  > n=10149  > TOYOTA$fpc\_srs=N  > TOYOTA$wgt\_stra=N/n  >  > srs\_design=svydesign(ids=~1, weights =~wgt\_stra, data=TOYOTA)  >  > #Estimates SRS#  > #Mean Estimate of SRS#  > ybar\_srs=svymean(~TOYOTA$City\_mpg,srs\_design)  > ybar\_srs  mean SE  TOYOTA$City\_mpg 16.704 0.026  >  > #Confidence interval of the mean eastimate  > confint(svymean(~TOYOTA$City\_mpg,srs\_design),df=degf(srs\_design))  2.5 % 97.5 %  TOYOTA$City\_mpg 16.65305 16.7551   |  | | --- | | > ####Variance of the estimated mean  > var\_ybar\_srs=0.026^2  > var\_ybar\_srs  [1] 0.000676 |   >  > #Total Estimate of SRS#  > t\_cup\_srs=svytotal(~City\_mpg,srs\_design)  > t\_cup\_srs  total SE  City\_mpg 3972278 6191.1  >  > #Variance of the eastimated total for SRS for TOYOTA City mpg  > var\_t\_cup\_srs=6191.1^2  > var\_t\_cup\_srs  [1] 38329719  **v)**  **###STRATIFIED RANDOM SAMPLING FOR TOYOTA VEHICLE MAKE HIGHWAY MPG ##**   |  | | --- | | > library(stratification)  > library(PracTools)  >  > **#Sample Size Estimation#**  > var(TOYOTA$High\_mpg)  [1] 42.03938  > library(samplingbook)  > sample.size.mean(e= 0.10 , S=sqrt(42.03938), N= 49127, level= 0.95)  sample.size.mean object: Sample size for mean estimate  With finite population correction: N=49127, precision e=0.1 and standard deviation S=6.4838  Sample size needed: 12154  > **#Neyman Allocation,Stratified Sampling#**  > Stratum=c("A", "B", "C", "D", "E", "F", "G", "H", "I", "J","K")  > Nh = c(7981, 2139, 6578, 908, 3469, 6769, 5233, 7106,3109, 5054,781)  > Sh = c(6.2014899, 5.2630210, 6.6309589, 5.7864360, 5.5533916, 7.1158215, 7.1021862, 6.3547713, 6.3498139, 6.4573177, 6.1549335)  > sample=strAlloc(n.tot = 12154, Nh = Nh, Sh = Sh, alloc = "neyman")  > sample    allocation = neyman  Nh = 7981, 2139, 6578, 908, 3469, 6769, 5233, 7106, 3109, 5054, 781  Sh = 6.201490, 5.263021, 6.630959, 5.786436, 5.553392, 7.115822, 7.102186, 6.354771, 6.349814, 6.457318, 6.154934  nh = 1900.2602, 432.2207, 1674.6726, 201.7236, 739.6433, 1849.3081, 1426.9294, 1733.7435, 757.9515, 1252.9886, 184.5585  nh/n = 0.15634854, 0.03556202, 0.13778777, 0.01659730, 0.06085595, 0.15215634, 0.11740410, 0.14264797, 0.06236231, 0.10309269, 0.01518500  anticipated SE of estimated mean = 0.05065602 | |  | | |  | | --- | | > | | | **####STRATIFIED RANDOM SAMPLING DESIGN FOR TOYOTA HIGHWAY MPG####**  > stat\_stra=svydesign(id=~1, strata=~Class, weights=~P, data=TOYOTA)  >  > **#Mean Estimate of Str#**  > ybar\_str=svymean(~TOYOTA$High\_mpg, design = stat\_stra)  >  > ybar\_str  mean SE  TOYOTA$High\_mpg 22.418 0.0288 | |  | | |  | | --- | |  | | | > #Variance of the estimated mean  > var\_ybar\_str=0.0288^2  > var\_ybar\_str  [1] 0.00082944 | |  | | |  | | --- | |  | | | > #Confidence Interval of the Estimate mean#  > confint(ybar\_str)  2.5 % 97.5 %  TOYOTA$High\_mpg 22.36108 22.47402   |  |  | | --- | --- | | |  | | --- | | > | | | #Estimated total of Stratified sampling  > t\_cup\_str=svytotal(~TOYOTA$High\_mpg, design = stat\_stra)  > t\_cup\_str  total SE  TOYOTA$High\_mpg 5457274 1414.86 | |  | | |  | | --- | | > | | | #Variance of the estimated total  > var\_t\_cup\_str=7014^2  > var\_t\_cup\_str  [1] 2001822 | |  | | |  |  |  |  |  | | --- | --- | --- | --- | --- | | |  | | --- | | > #Design Effect#  > svymean(~TOYOTA$High\_mpg, design = stat\_stra, deff = TRUE)  mean SE DEff  TOYOTA$High\_mpg 22.418333 0.028889 1.3089 | |  | | |  | | --- | |  | | | |   **vi)**   |  | | --- | | **##############SIMPLE RANDOM SAMPLE FOR TOYOTA HIGHWAY MPG########**  > N=49127  > n=10149  > TOYOTA$fpc\_srs=N  > TOYOTA$wgt\_stra=N/n  >  > srs\_design=svydesign(ids=~1, weights =~wgt\_stra, data=TOYOTA)  >  > **#Estimates SRS#**  **> #Mean Estimate of SRS#**  > ybar\_srs=svymean(~TOYOTA$High\_mpg,srs\_design)  > ybar\_srs  mean SE  TOYOTA$High\_mpg 22.48 0.0293  >  > #Confidence interval of the mean eastimate  > confint(svymean(~TOYOTA$High\_mpg,srs\_design),df=degf(srs\_design))  2.5 % 97.5 %  TOYOTA$High\_mpg 22.42293 22.5376  >  > var\_ybar\_srs=0.0293^2  > var\_ybar\_srs  [1] 0.00085849 | |  | |  | | #Total Estimate of SRS#  > t\_cup\_srs=svytotal(~High\_mpg,srs\_design)  > t\_cup\_srs  total SE  High\_mpg 5345873 6956.4 |   > #Variance of the eastimate  > var\_t\_cup\_srs=6956.4^2  > var\_t\_cup\_srs  [1] 48391501  **vii)**   |  | | --- | | > **####Gain from Stratification#####**  > # The Relative Gain From Stratification For Honda City Mpg  > Gain\_from\_str=var\_t\_cup\_str/var\_t\_cup\_srs  > Gain\_from\_str=3038523/30117046  > Gain\_from\_str  [1] 0.1008905  >  > # The Relative Gain From Stratification For Honda Highway Mpg  > Gain\_from\_str=var\_t\_cup\_str/var\_t\_cup\_srs  > Gain\_from\_str=3021234/29587072  > Gain\_from\_str  [1] 0.1021133  >  > # The Relative Gain From Stratification For TOYOTA City Mpg  > Gain\_from\_str=var\_t\_cup\_str/var\_t\_cup\_srs  > Gain\_from\_str=37264920/38329719  > Gain\_from\_str  [1] 0.97222  >  > # The Relative Gain From Stratification For TOYOTA City Mpg  > Gain\_from\_str=var\_t\_cup\_str/var\_t\_cup\_srs  > Gain\_from\_str=2001822/48391501  > Gain\_from\_str  [1] 0.04136722 | |  | | |  | | --- | | > | | | |  | | |  | | |
|  |
| |  | | --- | |  | |

1. Refer Kernel Density Estimation from Appendix A, figure 1 to 4 [↑](#footnote-ref-1)