**SAS CODES FOR DESCRIPTIVE STATISTICS OF THE DATA**

/\* Generated Code (IMPORT) \*/

/\* Source File: MAIN.xlsx \*/

/\* Source Path: /home/u45193287 \*/

/\* Code generated on: 4/21/20, 2:33 AM \*/

%web\_drop\_table(WORK.IMPORT1);

FILENAME REFFILE '/home/u45193287/MAIN.xlsx';

PROC IMPORT DATAFILE=REFFILE

DBMS=XLSX

OUT=WORK.IMPORT1;

GETNAMES=YES;

RUN;

PROC CONTENTS DATA=WORK.IMPORT1; RUN;

%web\_open\_table(WORK.IMPORT1);

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

TITLE"Descriptive Statistics Of Main Dataset on City\_Mpg Based on Make";

proc means Data=WORK.IMPORT1 MEAN VAR STD CLM;

CLASS Make;

var City\_Mpg;

Run;

/\* Generated Code (IMPORT) \*/

/\* Source File: HONDA.xlsx \*/

/\* Source Path: /home/u45193287 \*/

/\* Code generated on: 4/21/20, 12:51 AM \*/

%web\_drop\_table(WORK.IMPORT);

FILENAME REFFILE '/home/u45193287/HONDA.xlsx';

PROC IMPORT DATAFILE=REFFILE

DBMS=XLSX

OUT=WORK.IMPORT;

GETNAMES=YES;

RUN;

PROC CONTENTS DATA=WORK.IMPORT; RUN;

%web\_open\_table(WORK.IMPORT);

TITLE"Descriptive Statistics For HONDA on City\_Mpg Based on Class";

Proc means Data=WORK.IMPORT;

RUN;

proc means Data=WORK.IMPORT Median MEAN VAR STD CLM;

CLASS Class;

var City\_mpg;

run;

proc means Data=WORK.IMPORT Median MEAN VAR STD CLM;

CLASS Class;

var High\_mpg;

Run;

/\* Generated Code (IMPORT) \*/

/\* Source File: TOYOTA.xlsx \*/

/\* Source Path: /home/u45193287 \*/

/\* Code generated on: 4/23/20, 4:43 AM \*/

%web\_drop\_table(WORK.IMPORT);

FILENAME REFFILE '/home/u45193287/TOYOTA.xlsx';

PROC IMPORT DATAFILE=REFFILE

DBMS=XLSX

OUT=WORK.IMPORT;

GETNAMES=YES;

RUN;

PROC CONTENTS DATA=WORK.IMPORT; RUN;

%web\_open\_table(WORK.IMPORT);

TITLE"Descriptive Statistics For TOYOTA on City\_mpg Based on Class";

proc means Data=WORK.IMPORT MEAN VAR STD CLM;

CLASS Class;

var City\_mpg;

run;

**# R-CODES 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")

**#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")

#######################################################################

**#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")

**#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")

**R-CODES FOR SAMPLE SIZE DETERMINATION AND ESTIMATIONS**

**#NEYMAN ALLOCATION TO STRATA**

**library(Hmisc)**

**library(pastecs)**

**library(psych)**

**library(doBy)**

**library(stratification)**

**library(survey)**

**library(SDaA)**

**library(PracTools)**

**############# NEYMAN ALLOCATION FOR HONDA VEHICLES (CITY\_MPG)###################**

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 = 14898, Nh = Nh, Sh = Sh, alloc = "neyman")

sample

#As the sample size in each stratum

n\_h=c(1673,2527,12,865,1916,248,3289,3323,234,374,437)

N\_hS\_h=Nh\*Sh

**##SELECTING FROM EACH REATUM A SRS WITH SAMPLE SIZES CALCULATED USING NEYMAN ALLOCATION ABOVECITY FOR CITY\_MPG**

compact\_car1=subset(HONDA, select = City\_mpg, subset = Class=="Compact\_Car", drop = T)

Large\_Cars2=subset(HONDA, select = City\_mpg, subset = Class=="Large\_Cars", drop = T)

Midsize\_Station\_Wagons3=subset(HONDA, select = City\_mpg, subset = Class=="Midsize\_Station\_Wagons", drop = T)

Small\_Station\_Wagon4=subset(HONDA, select = City\_mpg, subset = Class=="Small \_Station\_Wagon", drop = T)

Small\_Pickup\_Truck\_2WD5=subset(HONDA, select = City\_mpg, subset = Class=="Small\_Pickup\_Truck\_2WD", drop = T)

Small\_sport\_Utility\_Cars6=subset(HONDA, select = City\_mpg, subset = Class=="Small\_sport\_Utility\_Cars", drop = T)

Large\_Sport\_Ucars7=subset(HONDA, select = City\_mpg, subset = Class=="Large\_Sport\_Ucars", drop = T)

Standard\_Pickup\_Truck\_4WD8=subset(HONDA, select = City\_mpg, subset = Class=="Standard\_Pickup\_Truck\_4WD", drop = T)

Standard\_Ucars\_4WD9=subset(HONDA, select = City\_mpg, subset = Class=="Standard\_Ucars\_4WD", drop = T)

Two\_Seaters10=subset(HONDA, select = City\_mpg, subset = Class=="Two\_Seaters", drop = T)

Van\_Cars11=subset(HONDA, select = City\_mpg, subset = Class=="Van\_Cars", drop = T)

**#FINDING THE MEANS OF THE SAMPLES SELECTED IN EACH STRATUM FOR CITY\_MPG**

compact\_car.sample=sample(compact\_car1,1676, F)

Large\_Cars.sample=sample(Large\_Cars2,2527, F)

Midsize\_Station\_Wagons.sample=sample(Midsize\_Station\_Wagons3,12, F)

Small\_Station\_Wagon.sample=sample(Small\_Station\_Wagon4,865, F)

Small\_Pickup\_Truck\_2WD.sample=sample(Small\_Pickup\_Truck\_2WD5,1916, F)

Small\_sport\_Utility\_Cars.sample=sample(Small\_sport\_Utility\_Cars6,248, F)

Large\_Sport\_Ucars7.sample=sample(Large\_Sport\_Ucars7,3289, F)

Standard\_Pickup\_Truck\_4WD.sample=sample(Standard\_Pickup\_Truck\_4WD8,3323, F)

Standard\_Ucars\_4WD.sample=sample(Standard\_Ucars\_4WD9,234, F)

Two\_Seaters.sample=sample(Two\_Seaters10,374, F)

Van\_Cars.sample=sample(Van\_Cars11,437, F)

**#FINDING Ybar\_h FOR EACH STRATUM FOR CITY\_MPG**

ybar\_1=mean(compact\_car.sample)

ybar\_2=mean(Large\_Cars.sample)

ybar\_3=mean(Midsize\_Station\_Wagons.sample)

ybar\_4=mean(Small\_Station\_Wagon.sample)

ybar\_5=mean(Small\_Pickup\_Truck\_2WD.sample)

ybar\_6=mean(Small\_sport\_Utility\_Cars.sample)

ybar\_7=mean(Large\_Sport\_Ucars7.sample)

ybar\_8=mean(Standard\_Pickup\_Truck\_4WD.sample)

ybar\_9=mean(Standard\_Ucars\_4WD.sample)

ybar\_10=mean(Two\_Seaters.sample)

ybar\_11=mean(Van\_Cars.sample)

**#FINDING VYbar\_h FOR City\_MPG**

VYh1=var(compact\_car.sample)

VYh2=var(Large\_Cars.sample)

VYh3=var(Midsize\_Station\_Wagons.sample)

VYh4=var(Small\_Station\_Wagon.sample)

VYh5=var(Small\_Pickup\_Truck\_2WD.sample)

VYh6=var(Small\_sport\_Utility\_Cars.sample)

VYh7=var(Large\_Sport\_Ucars7.sample)

VYh8=var(Standard\_Pickup\_Truck\_4WD.sample)

VYh9=var(Standard\_Ucars\_4WD.sample)

VYh10=var(Two\_Seaters.sample)

VYh11=var(Van\_Cars.sample)

Ybar\_h=c(ybar\_1, ybar\_2, ybar\_3, ybar\_4, ybar\_5, ybar\_6, ybar\_7, ybar\_8, ybar\_9, ybar\_10, ybar\_11)

vary\_h=c(VYh1, VYh2, VYh3, VYh4, VYh5, VYh6,VYh7, VYh8, VYh9, VYh10, VYh11)

**#THE ESTIMATED TOTALS FOR EACH STRATUM FOR CITY\_MPG IS GIVEN BY**

t\_cup\_h=Nh\*Ybar\_h

**#THE WEIGHT OF EACH STRATUM IS GIVEN BY**

whj=Nh/n\_h

**#THE ESTIMATED VARIANCE OF EACH STRATUM IS GIVEN BY**

Vcup\_t\_h=(1-n\_h/Nh)\*(Nh^2)\*(Sh^2)/n\_h

**##R OUTPUT FOR HONDA CITY\_MPG MEASURE OF FUEL ECONOMY**

**##THE REQUIRED DATA**

dat=data.frame(Stratum,Nh,Sh, N\_hS\_h, n\_h, whj, Ybar\_h , t\_cup\_h, Vcup\_t\_h)

dat

N=sum(Nh)

N

**# THE ESTIMATED AVERAGE FUEL ECONOMY FOR HONDA VEHICLE CITY\_MPG IS**

ybar\_str=sum(t\_cup\_h)/N

ybar\_str #As the estimated average fuel economy for HONDA Vehicles City\_MPG

**# THE CONFINDENCE INTERVAL FOR ybar\_str**

Varybar\_str=sum(Vcup\_t\_h)/(N^2)

Varybar\_str # As the Varinace of Y bar str

**# THE STANDARD ERROR OF** **ybar\_str**

SE\_ybar\_str=sqrt(Varybar\_str) # As the Standard Error of Y bar str

SE\_ybar\_str

**## CONFIDENCE INTERVAL FOR ybar\_str**

ybar\_str-1.96\*SE\_ybar\_str; ybar\_str+1.96\*SE\_ybar\_str

#(17.69507,17.89554) as the confidence interval for ybar\_str

**########### NEYMAN ALLOCATION FOR HONDA VEHICLES (HIGH\_MPG)################**

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 = 14898, Nh = Nh, Sh = Sh, alloc = "neyman")

sample

#As the sample size in each stratum

n\_h=c(1673,2527,12,865,1916,248,3289,3323,234,374,437)

N\_hS\_h=Nh\*Sh

**##SELECTING FROM EACH STRATUM A SRS WITH SAMPLE SIZES CALCULATED USING NEYMAN ALLOCATION ABOVE FOR HIGHWAY\_MPG**

compact\_car1=subset(HONDA, select = City\_mpg, subset = Class=="Compact\_Car", drop = T)

Large\_Cars2=subset(HONDA, select = City\_mpg, subset = Class=="Large\_Cars", drop = T)

Midsize\_Station\_Wagons3=subset(HONDA, select = City\_mpg, subset = Class=="Midsize\_Station\_Wagons", drop = T)

Small\_Station\_Wagon4=subset(HONDA, select = City\_mpg, subset = Class=="Small \_Station\_Wagon", drop = T)

Small\_Pickup\_Truck\_2WD5=subset(HONDA, select = City\_mpg, subset = Class=="Small\_Pickup\_Truck\_2WD", drop = T)

Small\_sport\_Utility\_Cars6=subset(HONDA, select = City\_mpg, subset = Class=="Small\_sport\_Utility\_Cars", drop = T)

Large\_Sport\_Ucars7=subset(HONDA, select = City\_mpg, subset = Class=="Large\_Sport\_Ucars", drop = T)

Standard\_Pickup\_Truck\_4WD8=subset(HONDA, select = City\_mpg, subset = Class=="Standard\_Pickup\_Truck\_4WD", drop = T)

Standard\_Ucars\_4WD9=subset(HONDA, select = City\_mpg, subset = Class=="Standard\_Ucars\_4WD", drop = T)

Two\_Seaters10=subset(HONDA, select = City\_mpg, subset = Class=="Two\_Seaters", drop = T)

Van\_Cars11=subset(HONDA, select = City\_mpg, subset = Class=="Van\_Cars", drop = T)

**#FINDING THE MEANS OF THE SAMPLES SELECTED IN EACH STRATUM**

compact\_car.sample=sample(compact\_car1,1676, F)

Large\_Cars.sample=sample(Large\_Cars2,2527, F)

Midsize\_Station\_Wagons.sample=sample(Midsize\_Station\_Wagons3,12, F)

Small\_Station\_Wagon.sample=sample(Small\_Station\_Wagon4,865, F)

Small\_Pickup\_Truck\_2WD.sample=sample(Small\_Pickup\_Truck\_2WD5,1916, F)

Small\_sport\_Utility\_Cars.sample=sample(Small\_sport\_Utility\_Cars6,248, F)

Large\_Sport\_Ucars7.sample=sample(Large\_Sport\_Ucars7,3289, F)

Standard\_Pickup\_Truck\_4WD.sample=sample(Standard\_Pickup\_Truck\_4WD8,3323, F)

Standard\_Ucars\_4WD.sample=sample(Standard\_Ucars\_4WD9,234, F)

Two\_Seaters.sample=sample(Two\_Seaters10,374, F)

Van\_Cars.sample=sample(Van\_Cars11,437, F)

**#FINDING Ybar\_h FOR EACH STRATUM**

ybar\_1=mean(compact\_car.sample)

ybar\_2=mean(Large\_Cars.sample)

ybar\_3=mean(Midsize\_Station\_Wagons.sample)

ybar\_4=mean(Small\_Station\_Wagon.sample)

ybar\_5=mean(Small\_Pickup\_Truck\_2WD.sample)

ybar\_6=mean(Small\_sport\_Utility\_Cars.sample)

ybar\_7=mean(Large\_Sport\_Ucars7.sample)

ybar\_8=mean(Standard\_Pickup\_Truck\_4WD.sample)

ybar\_9=mean(Standard\_Ucars\_4WD.sample)

ybar\_10=mean(Two\_Seaters.sample)

ybar\_11=mean(Van\_Cars.sample)

**#FINDING VYbar\_h FOR HIGH\_MPG**

VYh1=var(compact\_car.sample)

VYh2=var(Large\_Cars.sample)

VYh3=var(Midsize\_Station\_Wagons.sample)

VYh4=var(Small\_Station\_Wagon.sample)

VYh5=var(Small\_Pickup\_Truck\_2WD.sample)

VYh6=var(Small\_sport\_Utility\_Cars.sample)

VYh7=var(Large\_Sport\_Ucars7.sample)

VYh8=var(Standard\_Pickup\_Truck\_4WD.sample)

VYh9=var(Standard\_Ucars\_4WD.sample)

VYh10=var(Two\_Seaters.sample)

VYh11=var(Van\_Cars.sample)

Ybar\_h=c(ybar\_1, ybar\_2, ybar\_3, ybar\_4, ybar\_5, ybar\_6, ybar\_7, ybar\_8, ybar\_9, ybar\_10, ybar\_11)

vary\_h=c(VYh1, VYh2, VYh3, VYh4, VYh5, VYh6,VYh7, VYh8, VYh9, VYh10, VYh11)

**#THE ESTIMATED TOTALS FOR EACH STRATUM IS GIVEN BY**

t\_cup\_h=Nh\*Ybar\_h

**#THE WEIGHT OF EACH STRATUM IS GIVEN BY**

whj=Nh/n\_h

**#THE ESTIMATED VARIANCE OF EACH STRATUM IS GIVEN BY**

Vcup\_t\_h=(1-n\_h/Nh)\*(Nh^2)\*(Sh^2)/n\_h

####R OUTPUT FOR HONDA HIGHWAY\_MPG MEASURE OF FUEL ECONOMY

dat=data.frame(Stratum,Nh,Sh, N\_hS\_h, n\_h, whj, Ybar\_h , t\_cup\_h, Vcup\_t\_h)

dat

N=sum(Nh)

N

**# THE ESTIMATED AVERAGE FUEL ECONOMY FOR HONDA VEHICLE HIGHWAY\_MPG IS**

ybar\_str=sum(t\_cup\_h)/N

ybar\_str #As the estimated average fuel economy for HONDA Vehicles City\_MPG

**# THE CONFINDENCE INTERVAL FOR ybar\_str**

Varybar\_str=sum(Vcup\_t\_h)/(N^2)

Varybar\_str # As the Varinace of Y bar str

**# the Standard Error of Ybar\_str**

SE\_ybar\_str=sqrt(Varybar\_str) # As the Standard Error of Y bar str

SE\_ybar\_str

**##confidence interval for ybar\_str**

ybar\_str-1.96\*SE\_ybar\_str; ybar\_str+1.96\*SE\_ybar\_str

**########## NEYMAN ALLOCATION FOR TOYOTA VEHICLES (CITY\_MPG)###################**

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 = 14738, Nh = Nh, Sh = Sh, alloc = "neyman")

sample

**# The sample size in each stratum**

n\_h =c(2178,421, 2176,212,769, 2341,1797,2108,896, 1616, 224)

N\_hS\_h=Nh\*Sh

**##SELESELECTING FROM EACH STRATUM A SRS WITH SAMPLE SIZES CALCULATED USING NEYMAN ALLOCATION ABOVE**

Compact\_Car1=subset(TOYOTA, select = City\_mpg, subset = Class=="Compact\_Car", drop = T)

Large\_Sport\_Ucars2=subset(TOYOTA, select = City\_mpg, subset = Class=="Large\_Sport\_Ucars", drop = T)

Mid\_Size\_Car3=subset(TOYOTA, select = City\_mpg, subset = Class=="Mid\_Size\_Car", drop = T)

Midsize\_Station\_Wagons4=subset(TOYOTA, select = City\_mpg, subset = Class=="Midsize\_Station\_Wagons", drop = T)

Small\_Pickup\_Truck\_2WD5=subset(TOYOTA, select = City\_mpg, subset = Class=="Small\_Pickup\_Truck\_2WD", drop = T)

Small\_St\_Wags6=subset(TOYOTA, select = City\_mpg, subset = Class=="Small\_St\_Wags", drop = T)

Small\_sport\_Utility\_Cars7=subset(TOYOTA, select = City\_mpg, subset = Class=="Small\_sport\_Utility\_Cars", drop = T)

Standard\_Pickup\_Truck\_4WD8=subset(TOYOTA, select = City\_mpg, subset = Class=="Standard\_Pickup\_Truck\_4WD", drop = T)

Standard\_Ucars\_4WD9=subset(TOYOTA, select = City\_mpg, subset = Class=="Standard\_Ucars\_4WD", drop = T)

Two\_Seaters10=subset(TOYOTA, select = City\_mpg, subset = Class=="Two\_Seaters", drop = T)

Van\_Cars11=subset(TOYOTA, select = City\_mpg, subset = Class=="Van\_Cars", drop = T)

**#FINDING THE MEANS OF THE SAMPLES SELECTED IN EACH STRATUM**

compact\_car.sample=sample(Compact\_Car1, 2178, F)

Large\_Sport\_Ucars.sample=sample(Large\_Sport\_Ucars2,421, F)

Mid\_Size\_Car.sample=sample(Mid\_Size\_Car3, 2176, F)

Midsize\_Station\_Wagons.sample=sample(Midsize\_Station\_Wagons4, 212, F)

Small\_Pickup\_Truck\_2WD.sample=sample(Small\_Pickup\_Truck\_2WD5, 769, F)

Small\_St\_Wags.sample =sample(Small\_St\_Wags6, 2341, F)

Small\_sport\_Utility\_Cars.sample=sample(Small\_sport\_Utility\_Cars7,1797, F)

Standard\_Pickup\_Truck\_4WD.sample=sample(Standard\_Pickup\_Truck\_4WD8,2108, F)

Standard\_Ucars\_4WD.sample=sample(Standard\_Ucars\_4WD9,896, F)

Two\_Seaters.sample=sample(Two\_Seaters10,1616, F)

Van\_Cars.sample=sample(Van\_Cars11,224, F)

**#FINDING Ybar\_h for each stratum**

ybar\_1=mean(compact\_car.sample)

ybar\_2=mean(Large\_Sport\_Ucars.sample)

ybar\_3=mean(Mid\_Size\_Car.sample)

ybar\_4=mean(Midsize\_Station\_Wagons.sample)

ybar\_5=mean(Small\_Pickup\_Truck\_2WD.sample)

ybar\_6=mean(Small\_St\_Wags.sample)

ybar\_7=mean(Small\_sport\_Utility\_Cars.sample)

ybar\_8=mean(Standard\_Pickup\_Truck\_4WD.sample)

ybar\_9=mean(Standard\_Ucars\_4WD.sample)

ybar\_10=mean(Two\_Seaters.sample)

ybar\_11=mean(Van\_Cars.sample)

**#FINDING THE VARIANCE FOR EACH STRATUM CITY\_MPG**

VYh\_1=var(compact\_car.sample)

VYh\_2=var(Large\_Sport\_Ucars.sample)

VYh\_3=var(Mid\_Size\_Car.sample)

VYh\_4=var(Midsize\_Station\_Wagons.sample)

VYh\_5=var(Small\_Pickup\_Truck\_2WD.sample)

VYh\_6=var(Small\_St\_Wags.sample)

VYh\_7=var(Small\_sport\_Utility\_Cars.sample)

VYh\_8=var(Standard\_Pickup\_Truck\_4WD.sample)

VYh\_9=var(Standard\_Ucars\_4WD.sample)

VYh\_10=var(Two\_Seaters.sample)

VYh\_11=var(Van\_Cars.sample)

Ybar\_h=c(ybar\_1, ybar\_2, ybar\_3, ybar\_4, ybar\_5, ybar\_6, ybar\_7, ybar\_8, ybar\_9, ybar\_10, ybar\_11)

vary\_h=c(VYh\_1, VYh\_2, VYh\_3, VYh\_4, VYh\_5, VYh\_6 ,VYh\_7, VYh\_8, VYh\_9, VYh\_10, VYh\_11)

**#THE ESTIMATED TOTALS FOR EACH STRATUM CITY\_MPG IS GIVEN BY**

t\_cup\_h=Nh\*Ybar\_h

**#THE WEIGHT OF EACH STRATUM IS GIVEN BY**

whj=Nh/n\_h

**#THE ESTIMATED VARIANCE OF EACH STRATUM IS GIVEN BY**

Vcup\_t\_h=(1-n\_h/Nh)\*(Nh^2)\*(Sh^2)/n\_h

**##THE REQUIRED DATA FOR TOYOTA VEHICLE CITY\_MPG MEASURE OF FUEL ECONOMY**

dat=data.frame(Stratum,Nh,Sh, N\_hS\_h, n\_h, whj, Ybar\_h , t\_cup\_h, Vcup\_t\_h)

dat

N=sum(Nh)

N

**# THE ESTIMATED AVERAGE FUEL ECONOMY FOR TOYOTA VEHICLE CITY\_MPG IS**

ybar\_str=sum(t\_cup\_h)/N

ybar\_str

**# THE CONFINDENCE INTERVAL FOR ybar\_str**

Varybar\_str=sum(Vcup\_t\_h)/(N^2)

Varybar\_str # As the Varinace of ybar\_str

**# the Standard Error of Ybar\_str**

SE\_ybar\_str=sqrt(Varybar\_str)

SE\_ybar\_str

**##confidence interval for ybar\_str**

ybar\_str-1.96\*SE\_ybar\_str; ybar\_str+1.96\*SE\_ybar\_str

# as the confidence interval.

**############# Neyman allocation for TOYOTA Vehicles (HIGHWAY\_mpg)#############**

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 = 14738, Nh = Nh, Sh = Sh, alloc = "neyman")

sample

**# The selected sample size in each stratum**

n\_h =c(2304,524,2031,245, 897,2243,1730,2102, 919, 1519, 224)

N\_hS\_h=Nh\*Sh

**###SELESELECTING FROM EACH STRATUM A SRS WITH SAMPLE SIZES CALCULATED USING NEYMAN ALLOCATION ABOVE**

Compact\_Car1=subset(TOYOTA, select = High\_mpg, subset = Class=="Compact\_Car", drop = T)

Large\_Sport\_Ucars2=subset(TOYOTA, select = High\_mpg, subset = Class=="Large\_Sport\_Ucars", drop = T)

Mid\_Size\_Car3=subset(TOYOTA, select = High\_mpg, subset = Class=="Mid\_Size\_Car", drop = T)

Midsize\_Station\_Wagons4=subset(TOYOTA, select = High\_mpg, subset = Class=="Midsize\_Station\_Wagons", drop = T)

Small\_Pickup\_Truck\_2WD5=subset(TOYOTA, select = High\_mpg, subset = Class=="Small\_Pickup\_Truck\_2WD", drop = T)

Small\_St\_Wags6=subset(TOYOTA, select = High\_mpg, subset = Class=="Small\_St\_Wags", drop = T)

Small\_sport\_Utility\_Cars7=subset(TOYOTA, select = High\_mpg, subset = Class=="Small\_sport\_Utility\_Cars", drop = T)

Standard\_Pickup\_Truck\_4WD8=subset(TOYOTA, select = High\_mpg, subset = Class=="Standard\_Pickup\_Truck\_4WD", drop = T)

Standard\_Ucars\_4WD9=subset(TOYOTA, select = High\_mpg, subset = Class=="Standard\_Ucars\_4WD", drop = T)

Two\_Seaters10=subset(TOYOTA, select = High\_mpg, subset = Class=="Two\_Seaters", drop = T)

Van\_Cars11=subset(TOYOTA, select = High\_mpg, subset = Class=="Van\_Cars", drop = T)

**#FINDING THE MEANS OF THE SAMPLES SELECTED IN EACH STRATUM**

compact\_car.sample=sample(Compact\_Car1, 2304, F)

Large\_Sport\_Ucars.sample=sample(Large\_Sport\_Ucars2,524, F)

Mid\_Size\_Car.sample=sample(Mid\_Size\_Car3, 2031, F)

Midsize\_Station\_Wagons.sample=sample(Midsize\_Station\_Wagons4, 245, F)

Small\_Pickup\_Truck\_2WD.sample=sample(Small\_Pickup\_Truck\_2WD5, 897, F)

Small\_St\_Wags.sample =sample(Small\_St\_Wags6, 2243, F)

Small\_sport\_Utility\_Cars.sample=sample(Small\_sport\_Utility\_Cars7,1730, F)

Standard\_Pickup\_Truck\_4WD.sample=sample(Standard\_Pickup\_Truck\_4WD8,2102, F)

Standard\_Ucars\_4WD.sample=sample(Standard\_Ucars\_4WD9,919, F)

Two\_Seaters.sample=sample(Two\_Seaters10,1519, F)

Van\_Cars.sample=sample(Van\_Cars11,224, F)

**#FINDING Ybar\_h for each stratum**

ybar\_1=mean(compact\_car.sample)

ybar\_2=mean(Large\_Sport\_Ucars.sample)

ybar\_3=mean(Mid\_Size\_Car.sample)

ybar\_4=mean(Midsize\_Station\_Wagons.sample)

ybar\_5=mean(Small\_Pickup\_Truck\_2WD.sample)

ybar\_6=mean(Small\_St\_Wags.sample)

ybar\_7=mean(Small\_sport\_Utility\_Cars.sample)

ybar\_8=mean(Standard\_Pickup\_Truck\_4WD.sample)

ybar\_9=mean(Standard\_Ucars\_4WD.sample)

ybar\_10=mean(Two\_Seaters.sample)

ybar\_11=mean(Van\_Cars.sample)

**#FINDING THE VARIANCE FOR EACH STRATUM**

VYh\_1=var(compact\_car.sample)

VYh\_2=var(Large\_Sport\_Ucars.sample)

VYh\_3=var(Mid\_Size\_Car.sample)

VYh\_4=var(Midsize\_Station\_Wagons.sample)

VYh\_5=var(Small\_Pickup\_Truck\_2WD.sample)

VYh\_6=var(Small\_St\_Wags.sample)

VYh\_7=var(Small\_sport\_Utility\_Cars.sample)

VYh\_8=var(Standard\_Pickup\_Truck\_4WD.sample)

VYh\_9=var(Standard\_Ucars\_4WD.sample)

VYh\_10=var(Two\_Seaters.sample)

VYh\_11=var(Van\_Cars.sample)

Ybar\_h=c(ybar\_1, ybar\_2, ybar\_3, ybar\_4, ybar\_5, ybar\_6, ybar\_7, ybar\_8, ybar\_9, ybar\_10, ybar\_11)

vary\_h=c(VYh\_1, VYh\_2, VYh\_3, VYh\_4, VYh\_5, VYh\_6 ,VYh\_7, VYh\_8, VYh\_9, VYh\_10, VYh\_11)

**#THE ESTIMATED TOTALS FOR EACH STRATUM IS GIVEN BY**

t\_cup\_h=Nh\*Ybar\_h

**#THE WEIGHT OF EACH STRATUM IS GIVEN BY**

whj=Nh/n\_h

**#THE ESTIMATED VARIANCE OF EACH STRATUM IS GIVEN BY**

Vcup\_t\_h=(1-n\_h/Nh)\*(Nh^2)\*(Sh^2)/n\_h

**##THE REQUIRED DATA**

dat=data.frame(Stratum,Nh,Sh, N\_hS\_h, n\_h, whj, Ybar\_h , t\_cup\_h, Vcup\_t\_h)

dat

N=sum(Nh)

N

**# THE ESTIMATED AVERAGE FUEL ECONOMY FOR TOYOTA VEHICLE HIGHWAY\_MPG IS**

ybar\_str=sum(t\_cup\_h)/N

ybar\_str

**# THE CONFINDENCE INTERVAL FOR ybar\_str**

Varybar\_str=sum(Vcup\_t\_h)/(N^2)

Varybar\_str # As the Varinace of ybar\_str

**# the Standard Error of Ybar\_str**

SE\_ybar\_str=sqrt(Varybar\_str)

SE\_ybar\_str

**##confidence interval for ybar\_str**

ybar\_str-1.96\*SE\_ybar\_str; ybar\_str+1.96\*SE\_ybar\_str

#(22.34153, 22.51539) as the confidence interval.

**#USING SRS TO FIND THE FUEL ECONOMY OF HONDA AND TOYOTA VEHICLE MAKES**

library(survey)

library(SDaA)

**#HONDA VEHICLES**

**#######DRAW SAMPLES WITHOUT REPLACEMENT####**

**library(sampler)**

N=49662

n=14898 # As the sample size to be drawn

srs\_h=rsamp(df=HONDA, n=14898, rep=FALSE) ###replacement false

**#THE ESTIMATED AVERAGE FUEL ECONOMY FOR HONDA VEHICLES CITY\_MPG IS**

ybar\_srs1=sum(srs\_h[, "City\_mpg"])/n

ybar\_srs1

**# As the estimated average fuel economy for HONDA Vehicles City\_MPG**

Citysq\_diviation=(srs\_h$City\_mpg-ybar\_srs1)^2

**# the estimated variance of ybar\_srs1 for HONDA Vehicles City\_MPG is**

varybar\_srs1=(1-n/N)\*(sum(Citysq\_diviation)/n)

varybar\_srs1

**# Confidence interval of ybar\_srs1 for HONDA City\_mpg is**

ybar\_srs1-1.96\*sqrt(varybar\_srs1); ybar\_srs1+1.96\*sqrt(varybar\_srs1)

#(5.613796,29.9981) As the confindence interval

**#THE ESTIMATED AVERAGE FUEL ECONOMY FOR HONDA VEHICLES HIGHWAY\_MPG IS**

ybar\_srs2=sum(srs\_h[, "High\_mpg"])/n

ybar\_srs2

**# As the estimated average fuel economy for HONDA Vehicles Highway\_MPG**

Highsq\_diviation=(srs\_h$High\_mpg-ybar\_srs2)^2

**# the estimated variance of ybar\_srs2 for HONDA Vehicles Highway\_MPG is**

varybar\_srs2=(1-n/N)\*(sum(Highsq\_diviation)/n)

varybar\_srs2

**# Confidence interval of ybar\_srs forHONDA Highway\_mpg2 is**

ybar\_srs2-1.96\*sqrt(varybar\_srs2); ybar\_srs2+1.96\*sqrt(varybar\_srs2)

#(11.25233, 36.37809) As the confindence interval

**#SRS FOR TOYOTA VEHICLES**

**#######DRAW SAMPLES WITHOUT REPLACEMENT####**

library(sampler)

N=49127

n=14738 # As the sample size to be drawn

srs\_h=rsamp(df=TOYOTA, n=14738, rep=FALSE) ###replacement false

**#THE ESTIMATED AVERAGE FUEL ECONOMY FOR TOYOTA VEHICLES CITY\_MPG IS**

ybar\_srs1=sum(srs\_h[, "City\_mpg"])/n

ybar\_srs1

**# As the estimated average fuel economy for TOYOTA Vehicles City\_MPG**

Citysq\_diviation=(srs\_h$City\_mpg-ybar\_srs1)^2

**# the estimated variance of ybar\_srs1 for TOYOTA Vehicles City\_MPG is**

varybar\_srs1=(1-n/N)\*(sum(Citysq\_diviation)/n)

varybar\_srs1

**# Confidence interval of ybar\_srs1 for TOYOTA City\_mpg is**

ybar\_srs1-1.96\*sqrt(varybar\_srs1); ybar\_srs1+1.96\*sqrt(varybar\_srs1)

#(7.526573,25.92641) As the confindence interval

**#THE ESTIMATED AVERAGE FUEL ECONOMY FOR TOYOTA VEHICLES HIGHWAY\_MPG IS**

ybar\_srs2=sum(srs\_h[, "High\_mpg"])/n

ybar\_srs2

**# As the estimated average fuel economy for TOYOTA Vehicles Highway\_MPG**

Highsq\_diviation=(srs\_h$High\_mpg-ybar\_srs2)^2

**# the estimated variance of ybar\_srs2 for TOYOTA Vehicles Highway\_MPG is**

varybar\_srs2=(1-n/N)\*(sum(Highsq\_diviation)/n)

varybar\_srs2

**# Confidence interval of ybar\_srs for TOYOTA Highway\_mpg2 is**

ybar\_srs2-1.96\*sqrt(varybar\_srs2); ybar\_srs2+1.96\*sqrt(varybar\_srs2)

#(12.03635, 33.01711) As the confindence interval

dat1=data.frame(srs\_h, Citysq\_diviation, Highsq\_diviation)

dat1