Exam R Cheat Sheet

|  |  |  |
| --- | --- | --- |
| Chapter | Topic | Code |
| 2 | Data – load existing datasets | data(*dataset name*) |
| 2 | Frequency tables | table(*column name*) |
| 2 | Frequency tables as percentages | table(*column name*) / length(*column name*) \* 100 |
| 2 | Grouped frequency table | stem(*column name*, scale = *#*) |
| 2 | Ggplot2 | library(ggplot2) |
| 2 | Histograms | myplot = ggplot(*dataset*, aes(*column name*))  myplot +  geom\_histogram(binwidth = #) |
| 2 | Frequency Polygons | myplot = ggplot(*dataset*, aes(*column name*))  myplot +  geom\_freqpoly(binwidth = #) |
| 3 | Scatterplots | myplot = ggplot(*dataset*, aes(*X axis column, Y axis column*))  myplot +  geom\_point() +  geom\_smooth(method =”lm”) |
| 3 | Theme coding for plots | theme = theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank(), panel.background = element\_blank(), axis.line = element\_line(colour = "black"), legend.key = element\_rect(fill = "white")) |
| 3 | X and Y axis labels | xlab(“Text goes here”)  ylab(“Text goes here”) |
| 3 | Bar graph | myplot = ggplot(*dataset*, aes(*X axis column, Y axis column*))  myplot +  stat\_summary(fun.y = mean,  geom = "bar",  fill = "White",  color = "Black") +  stat\_summary(fun.data = mean\_cl\_normal,  geom = "errorbar",  position = position\_dodge(width = 0.90),  width = 0.2) |
| 4 | Combine data | mycolumn = c(*#, #, #*) |
| 4 | Mean | summary(*column name*)  mean(*column name*, na.rm = T) |
| 4 | Median | summary(*column name*)  median(*column name*, na.rm = T) |
| 4 | Mode | temp <- table(as.vector(*column*))  names(temp)[temp == max(temp)] |
| 4 | Range | summary(*column name*)  max(*column*, na.rm = T) – min(*column*, na.rm = T) |
| 4 | Unbiased variance | var(*column name*, na.rm = T) |
| 4 | Unbiased standard deviation | sd(*column name,* na.rm = T) |
| 4 | Biased variance | pop.var <- function(x) var(x) \* (length(x)-1) / length(x)  pop.sd <- function(x) sqrt(pop.var(x))  pop.var(*column name*) |
| 4 | Biased standard deviation | pop.var <- function(x) var(x) \* (length(x)-1) / length(x)  pop.sd <- function(x) sqrt(pop.var(x))  pop.sd(*column name*) |
| 4 | IQR | summary(*column name*)  IQR(*column name*, na.rm = T) |
| 6-7 | pnorm | pnorm(*Z score*, lower.tail = F) |
| 6-7 | qnorm | qnorm(*p value*, lower.tail = F) |
| 8 | Confidence interval | M + qnorm(*p value*, lower.tail = F)  M – qnorm(*p value,* lower.tail = F) |
| 8 | Power | ##enter numbers here  popmean = 80  popsd = 30  N = 5  alpha = .05  samplemean = 120  lower = F  ##auto calculate  popse = popsd / sqrt(N)  mneed = popmean + popse\*qnorm(alpha, lower.tail = lower)  z = (mneed - samplemean)/popse  pnorm(z, lower.tail = lower) |