Assessment 1: Statistical Analysis Presentation

Shaun Bell-Gibson

Scope

- Introduction
- Descriptive Statistics
- Inferential Statistics
- Discussion
- Conclusions and recommendations
- Appendix

Introduction

- Data from Health Survey for England, 2011
- 'The Health Survey for England (HSE) is an important annual survey looking at changes in the health and lifestyles of people all over the country.' (NHS, 2023)
- In England alcohol is the leading cause of ill-health and death among people aged 15-49 years, with an estimated 602,391 dependent drinkers only 18% of which are receiving treatment (Alcohol Change UK, nd).

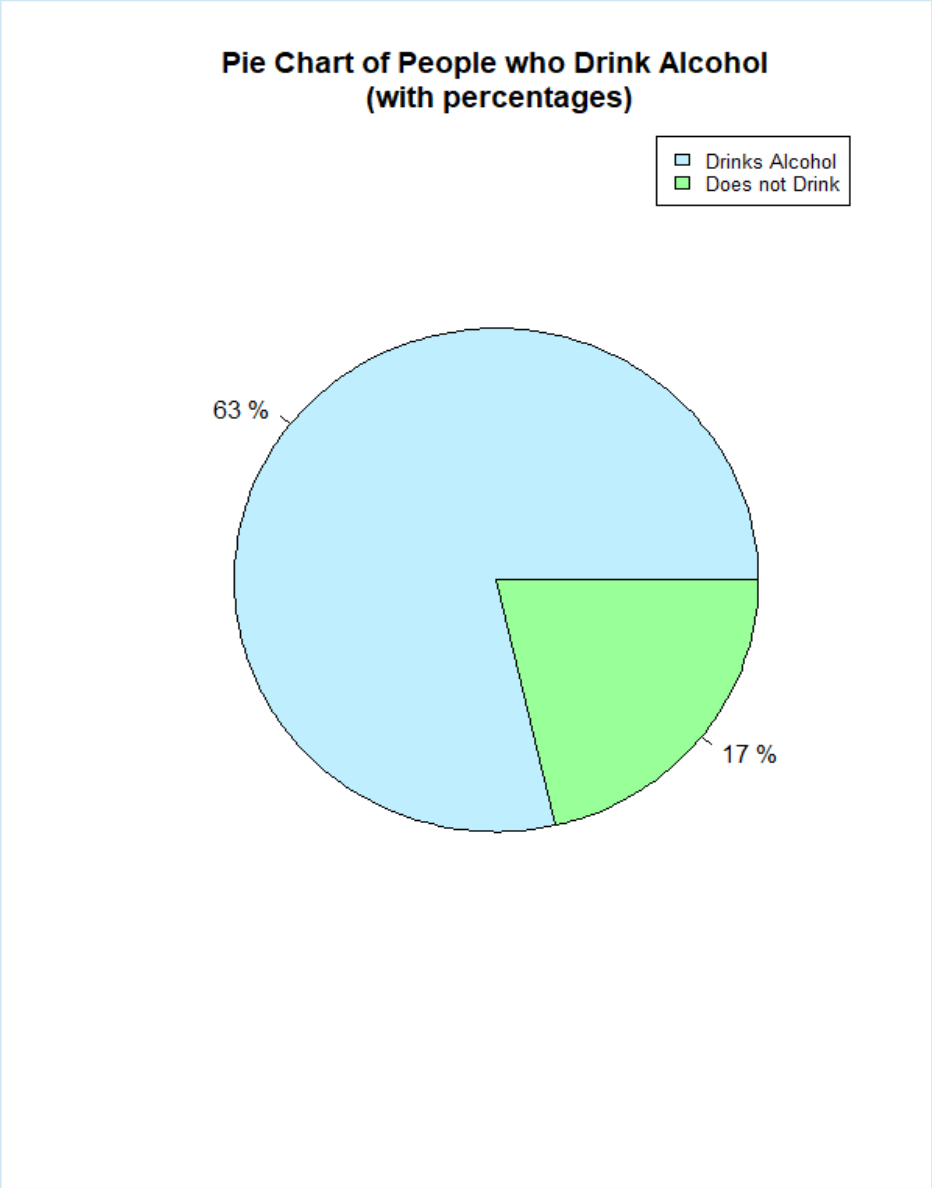
Sample Participants

- Sample size: 10,617
- Range of backgrounds

What is the percentage of people who drink alcohol?

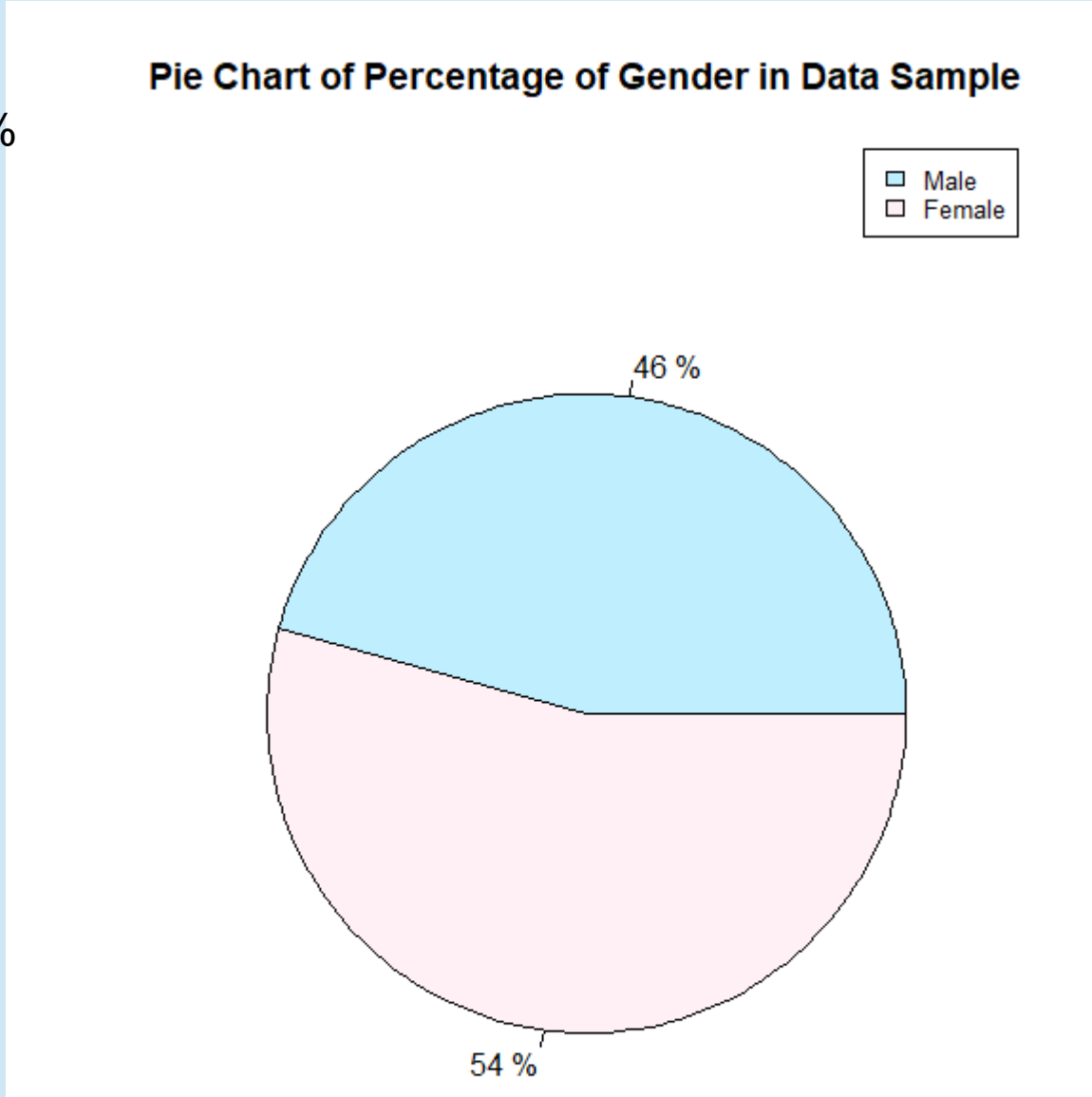
Drink Alcohol: 6,712
Does not drink Alcohol: 1,822

Drink Aware (2019):
48% consume alcohol >= once a week.
20% of adults are non-drinkers



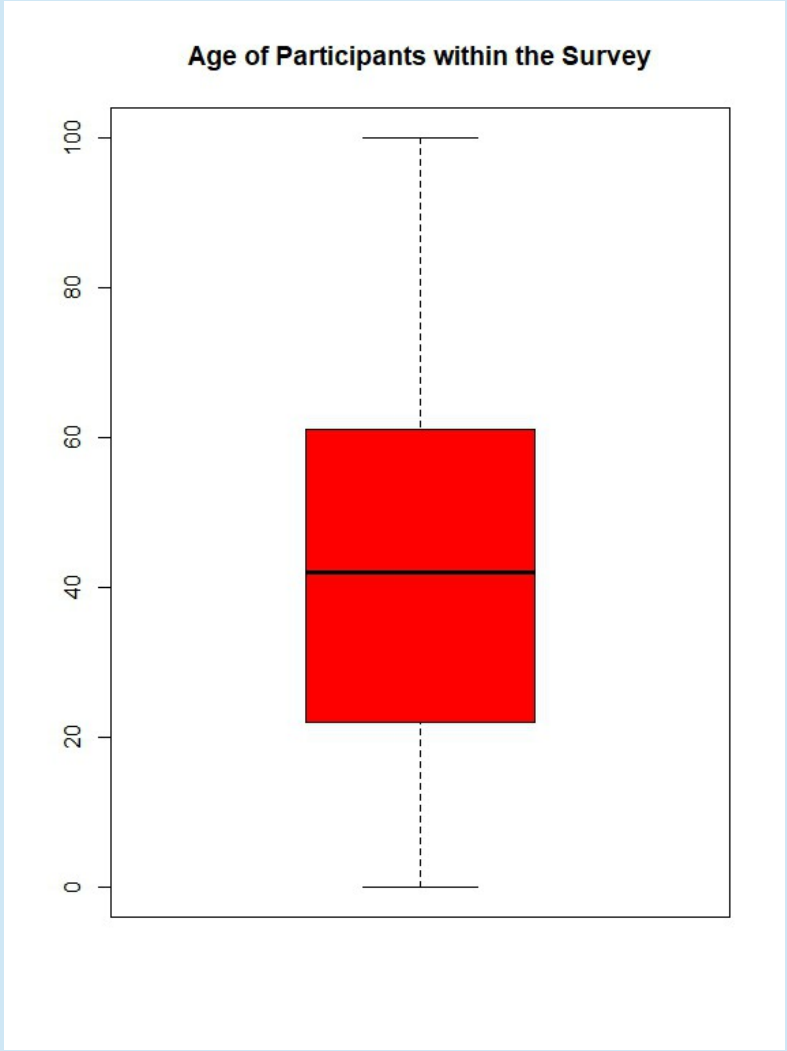
What is the percentage of women in the sample?

Gov.uk (2023) Women make up 51% of the population (30.4 million)



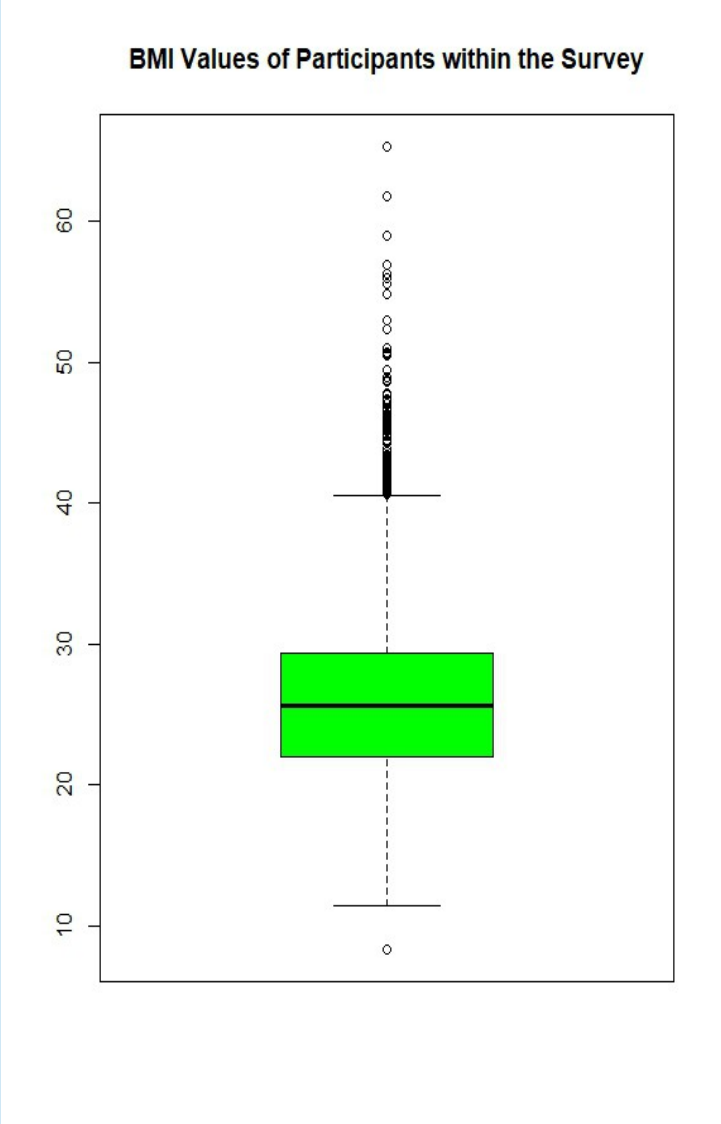
Age of Participants

	Age
Mean	41.56
Median	42
Mode	64 & 42
Minimum	0
Maximum	100
Range	100
Standard deviation	23.83203



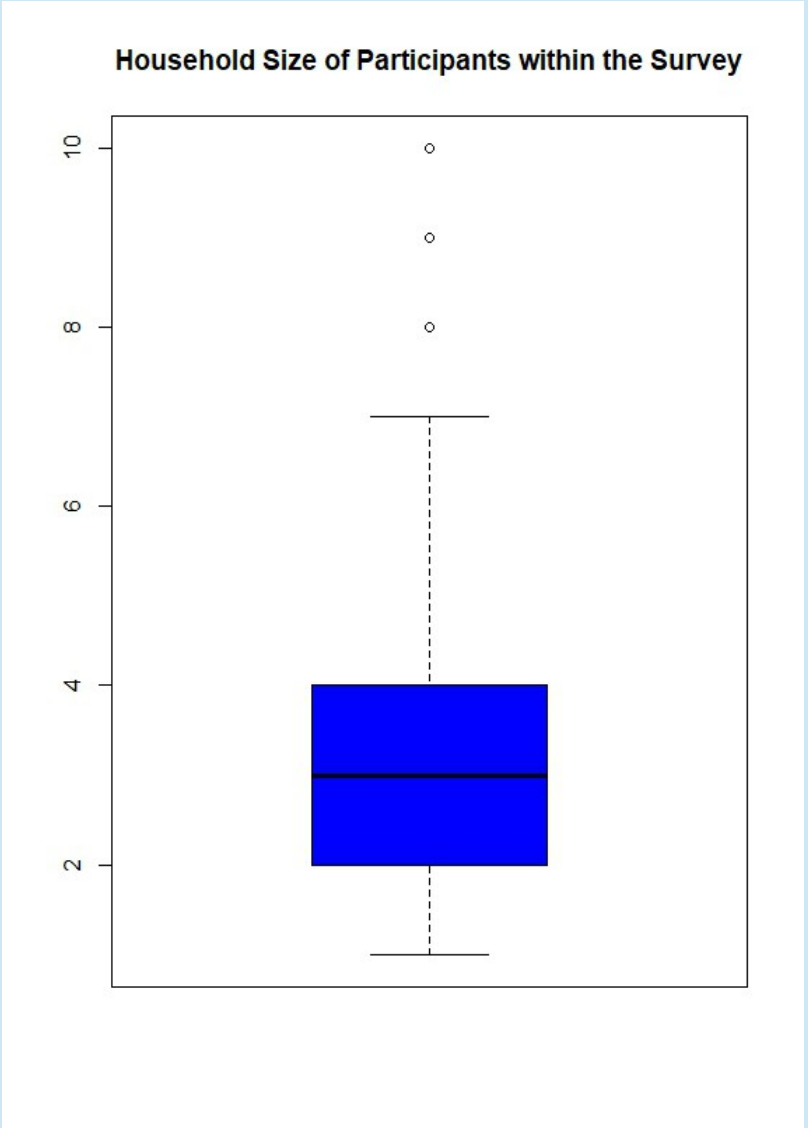
BMI of Participants

	BMI
Mean	25.92
Median	25.59
Mode	
Minimum	8.34
Maximum	65.28
Range	56.94
Standard deviation	



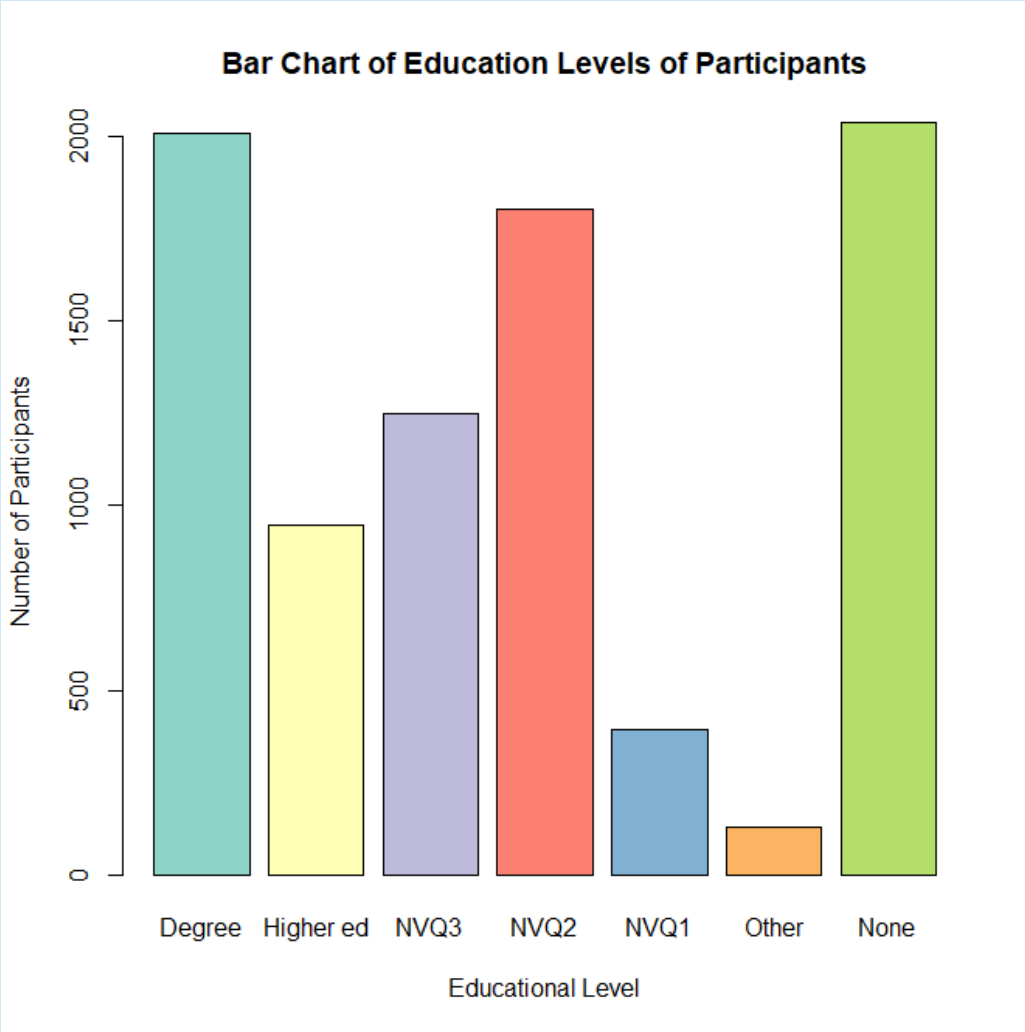
Household Size

	Household Size
Mean	2.851
Median	3
Mode	2
Minimum	1
Maximum	10
Range	9
Standard deviation	1.368528



What is the Highest Educational level?

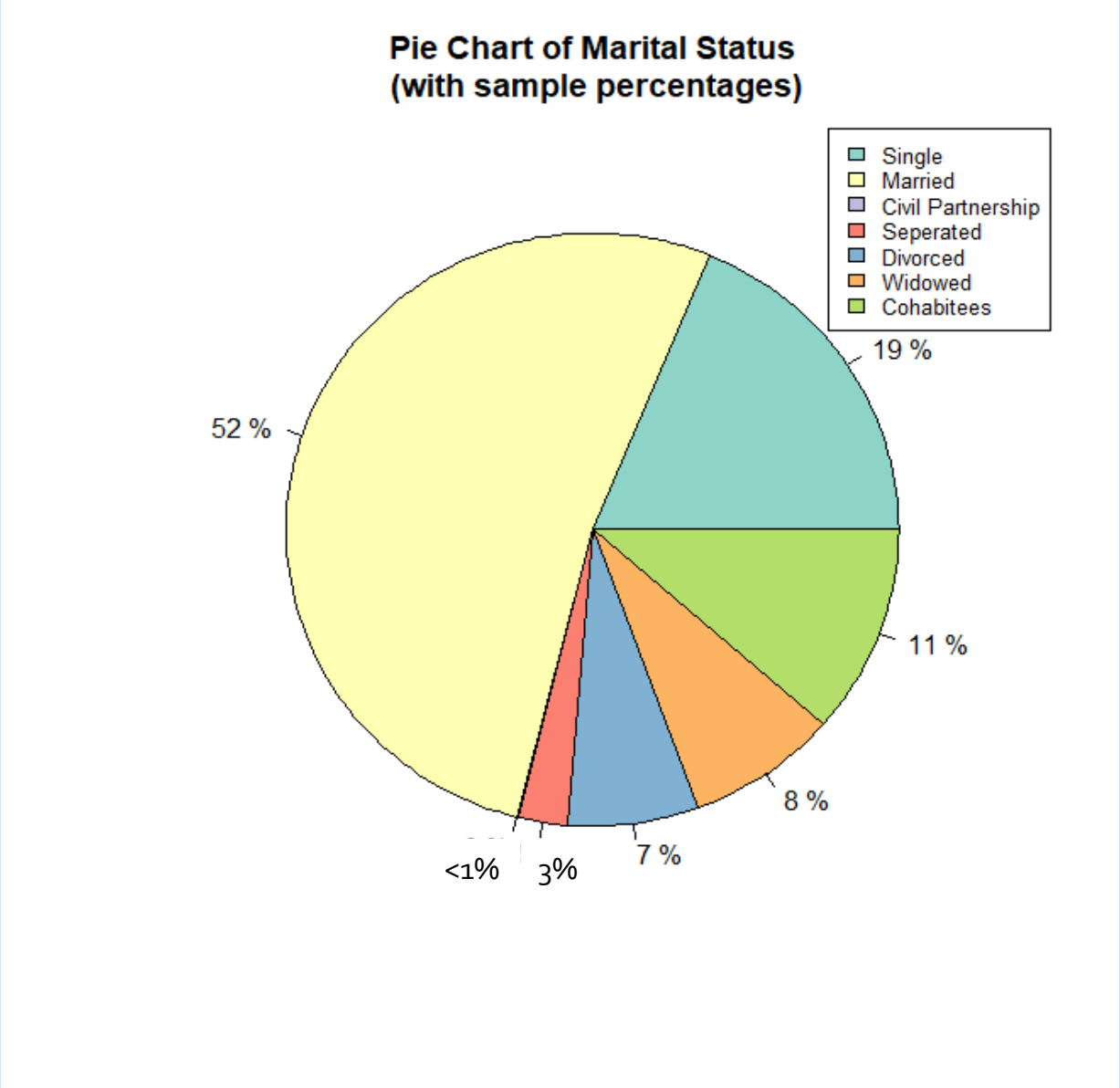
- Degree/equivalent: approx. 2000
- Rosoff et al. (2019)



What is percentage of divorced and separated people?

Separated/Divorced: 10%

Morgan (2019)



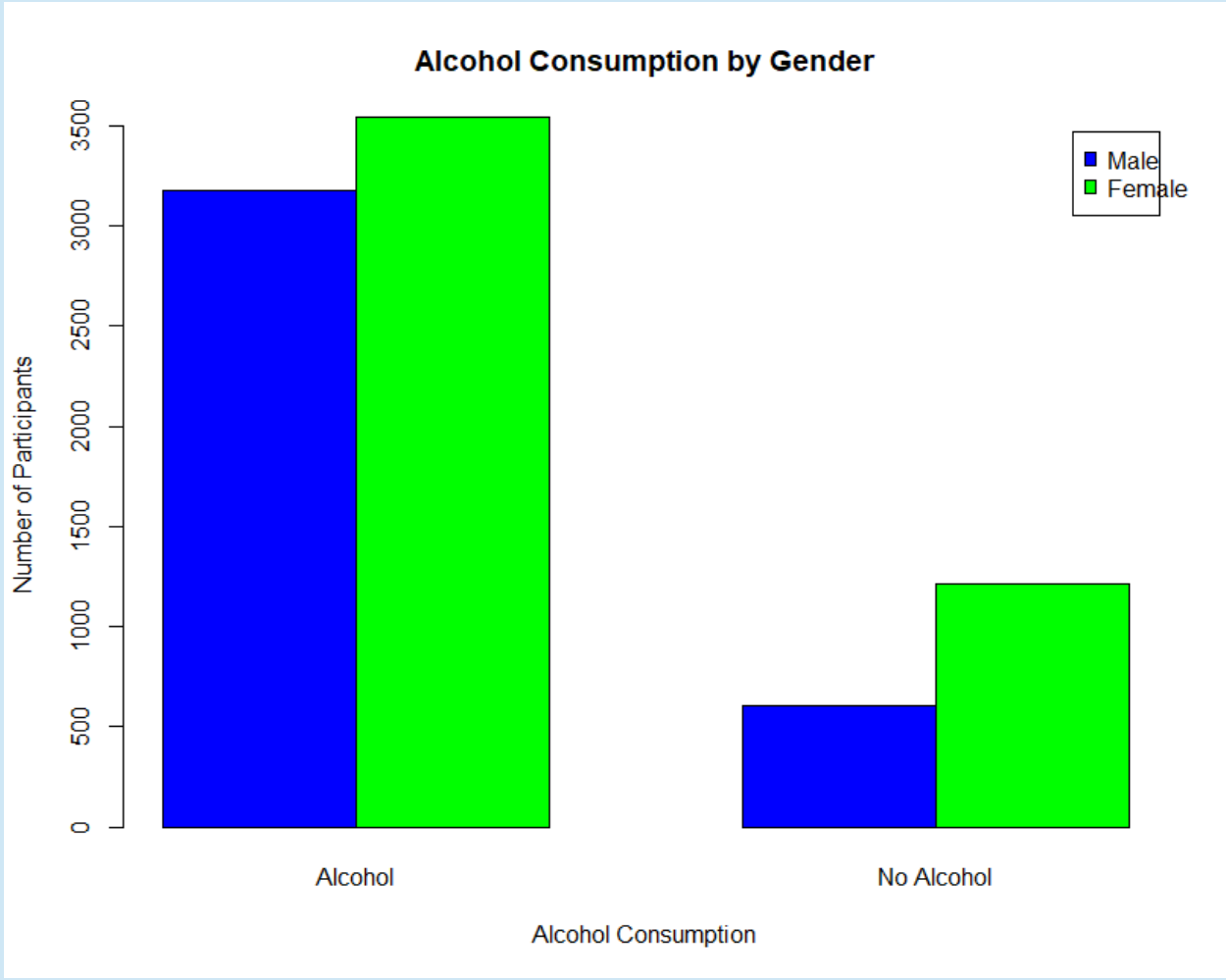
Which Gender Drinks the Most Alcohol?

Fisher's Exact Test for Count Data

p-value < 2.2e-16
alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:
1.614841 2.012964

sample estimates:
odds ratio
1.802317



Which Region Drinks the Most Alcohol?

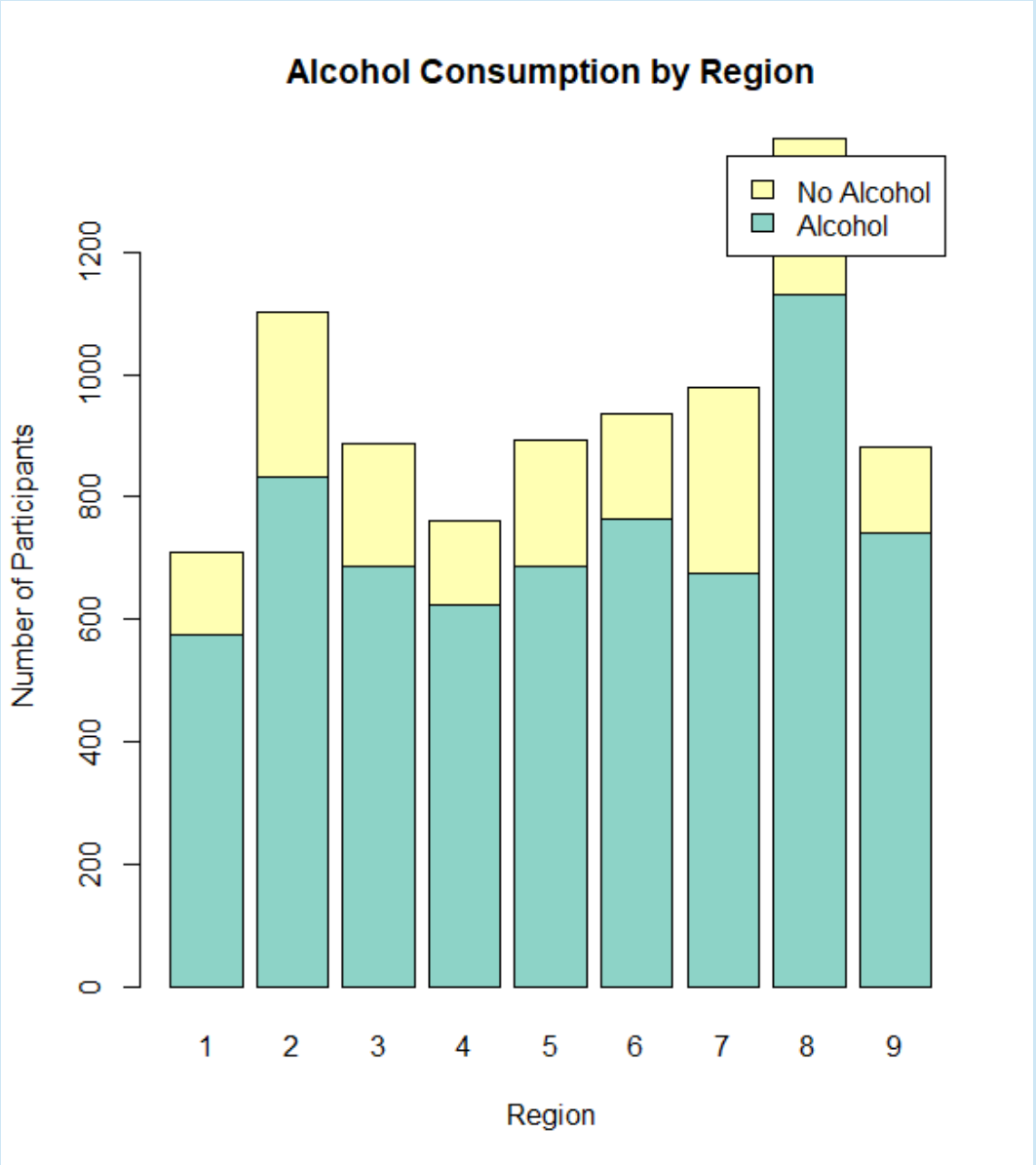
Key findings:

South East (Region 8)

South West (Region 9)

London (Region 7)

Research into effects of deprivation on alcohol consumption



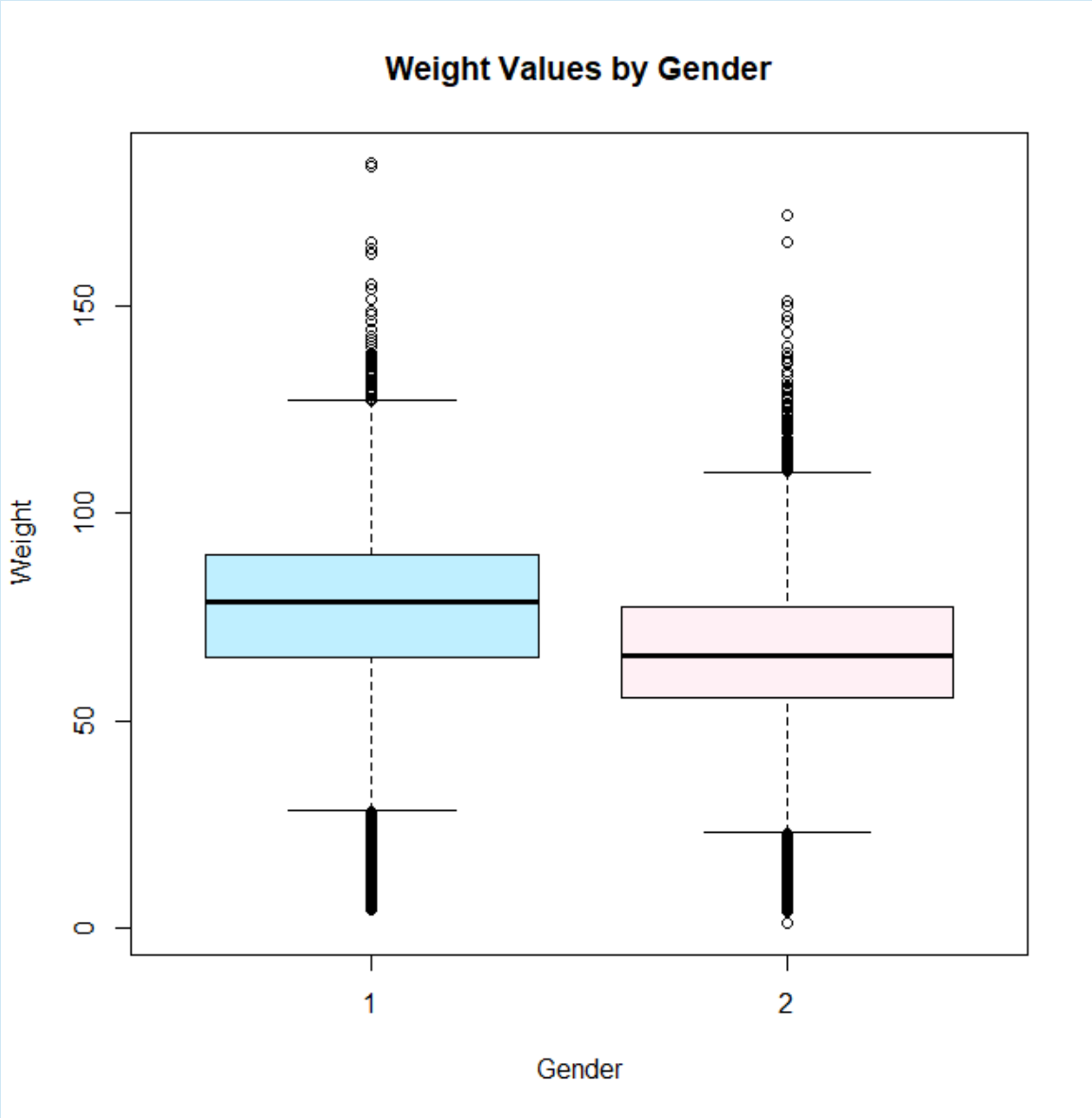
Difference between men and women in Weight.

Mean Values:

Men 74.27

Women 64.75

p-value < 2.2e-16



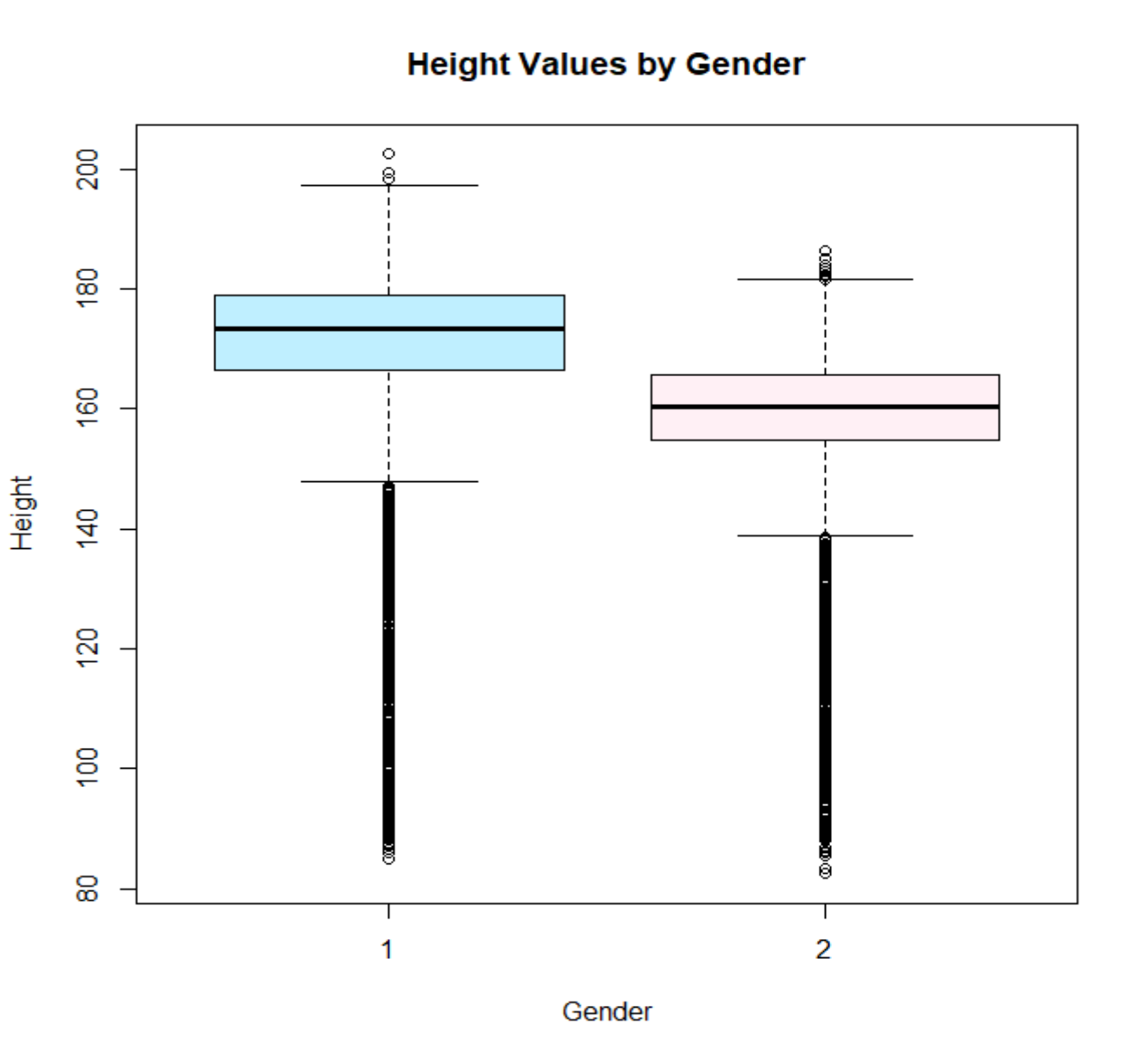
Difference between men and women in Height.

Mean Values:

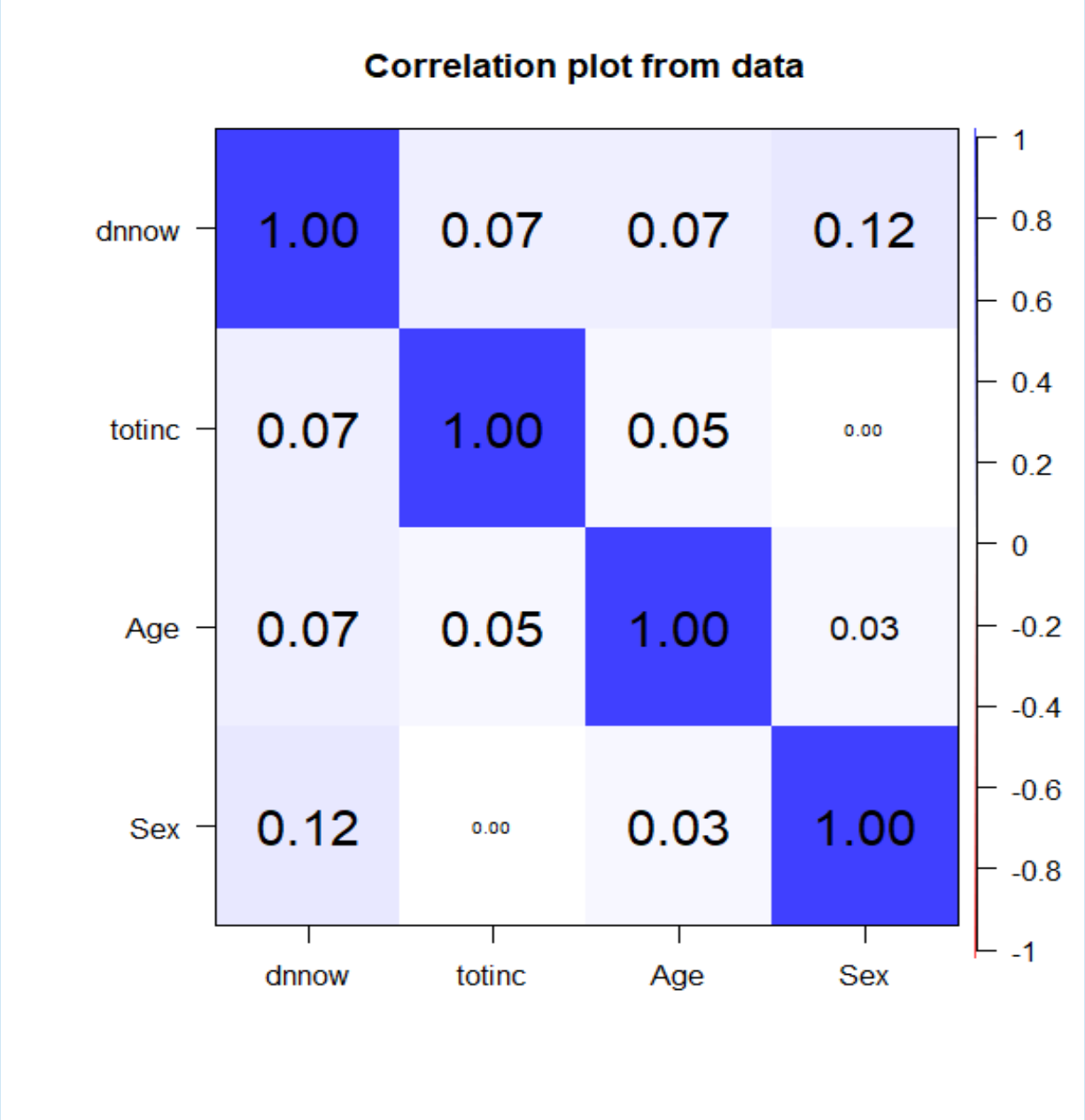
Men 167.39

Women 157.2

p-value < 2.2e-16



Correlation between multiple variables?



Conclusion

- 63% Participants consume alcohol
- Men more likely to consume alcohol
- South East: Most alcohol drinkers
- Educational level may effect drinking habits
- Weak positive correlations between drinking status, age, income and sex

Recommendations

- Future research into effect of education and marital status on levels of alcohol consumption
- Discover what factors cause men to consume more alcohol than women
- Effects of martial status on alcohol habits

Appendix

```r
#Get data
data("HSE_2011")

###2.Descriptive statistics###

#2.a  Participants = 10,617
total <- 10617

#b. What is the percentage of people who drink alcohol?
alcoholtable<- table(HSE_2011$dnnow)
AlcLable <- c("Drinks Alcohol", "Does not Drink")
colour<- c("#BFEFFF", "#99FF99")
pct <- round(100*alcoholtable/total)
pie(alcoholtable, col = colour, labels = paste(sep=" ", pct, "%"),
    main="Pie Chart of People who Drink Alcohol\n (with percentages)")
legend("topright", AlcLable, cex = 0.8, fill= colour)


#c. What is the percentage of women in the sample?
Sextable<- table(HSE_2011$Sex)
colour<- c("#BFEFFF", "#FFF0F5")
pct <- round(100*Sextable/total)
pie(Sextable, col = colour, labels = paste(sep=" ", pct, "%"),
    main="Pie Chart of Percentage of Gender in Data Sample")
legend("topright", AlcLable, cex = 0.8, fill= colour)


#d. What is the highest educational level?
install.packages("RColorBrewer")
library(RColorBrewer)
heltable<- table(HSE_2011$topqual3)
hellable<- c("Degree", "Higher ed", "NVQ3", "NVQ2", "NVQ1", "Other", "None")
barplot(heltable, xlab= "Educational Level", ylab = "Number of Participants", names.arg =hellable, col= brewer.pal(n = 7, name = "Set3"),
        main="Bar Chart of Education Levels of Participants")


#e. What is percentage of divorced and separated people?
Statustable<- table(HSE_2011$marstatc)
MarLabel <- c("Single", "Married", "Civil Partnership", "Seperated", "Divorced", "Widowed", "Cohabitees")
pct <- round(100*Statustable/sum(Statustable))
Mlbls <- paste(sep=" ", pct, "%")
pie(Statustable, labels = Mlbls,  col= brewer.pal(n = 7, name = "Set3"),
    main="Pie Chart of Marital Status\n (with sample percentages)")
legend("topright", AlcLable, cex = 0.8, fill= brewer.pal(n = 7, name = "Set3"))

##f.
#Summary
AgeSum <- summary(HSE_2011$Age)
HHSum <- summary(HSE_2011$HHSize)
BMISum <- summary(HSE_2011$bmival)

#Standard Deviation
sd(HSE_2011$Age)
sd(HSE_2011$HHSize)
sd(HSE_2011$bmival)


#Range
max(HSE_2011$Age)-min(HSE_2011$Age)
max(HSE_2011$HHSize)-min(HSE_2011$HHSize)
max(HSE_2011$bmival)-min(HSE_2011$bmival)

#Mode
find_mode <- function(x) {
  u <- unique(x)
  tab <- tabulate(match(x, u))
  u[tab == max(tab)]
}
find_mode(HSE_2011$Age)
find_mode(HSE_2011$HHSize)
find_mode(HSE_2011$bmival)
```

Appendix

```
#Box plot
boxplot(HSE_2011$Age, main="Age of Participants within the Survey", col= "red")
boxplot(HSE_2011$HHSize, main="Household Size of Participants within the Survey", col= "blue")
boxplot(HSE_2011$bmival, main="BMI Values of Participants within the Survey", col= "green")
############################################################

#3. Inferential Statistics.
#a. Run a significance test to find out which gender drinks more alcohol.
#Test for normal distribution
install.packages("nortest")
library(nortest)
install.packages("dplyr")
library(dplyr)
ad.test(HSE_2011$dnnow)
ad.test(HSE_2011$totinc)
ad.test(HSE_2011$Age)
ad.test(HSE_2011$Sex)


#Perform test
Gender_alcoholtable<- table(HSE_2011$Sex, HSE_2011$dnnow)
colnames(Gender_alcoholtable)[1] ="Alcohol"
colnames(Gender_alcoholtable)[2] ="No Alcohol"
rownames(Gender_alcoholtable)[1]= "Male"
rownames(Gender_alcoholtable)[2]= "Female"

fisher.test(Gender_alcoholtable)

#Visualization
Gender_alcoholtable_tr<- table(HSE_2011$dnnow, HSE_2011$Sex)
colnames(Gender_alcoholtable_tr)[1] ="Male"
colnames(Gender_alcoholtable_tr)[2] ="Female"
rownames(Gender_alcoholtable_tr)[1]= "Alcohol"
rownames(Gender_alcoholtable_tr)[2]= "No Alcohol"

barplot(Gender_alcoholtable_tr, xlab= "Gender", ylab = "Number of Participants", col= c("blue","green"),
        main="Alcohol Consumption by Gender", legend = rownames(Gender_alcoholtable))



#b. Run a significance test to find out which region drinks the most alcohol.
Region_Alcholdf<-table(HSE_2011[,c("gor1", "dnnow")])
colnames(Region_Alcholdf)[1]= "Alcohol"
colnames(Region_Alcholdf)[2]= "No Alcohol"
fisher.test(Region_Alcholdf)

#Visualisation
Region_Alchol_tr<-table(HSE_2011[,c("dnnow", "gor1")])
rownames(Region_Alchol_tr)[1]= "Alcohol"
rownames(Region_Alchol_tr)[2]= "No Alcohol"
barplot(Region_Alchol_tr, xlab= "Region", ylab = "Number of Participants", col= brewer.pal(n = 2, name = "Set3"),
  main="Alcohol Consumption by Region", legend = rownames(Region_Alchol_tr))

chisq.test(Region_Alcholdf)
fisher.test(Region_Alcholdf)
mosaicplot(Region_Alcholdf, shade = TRUE, type = "pearson", main = "")  # mosaic plot
corPlot(Region_Alchol_tr, cex = 1.2)
```

Appendix

```
131 #c. Investigate whether there is a statistical difference between men and women on the following variables:
132   #c i. Valid height.
133 gender_heightdf<-HSE_2011[,c("Sex","htval")]
134 gender_heightdf<-na.omit(gender_heightdf)
135 
136 boxplot(htval~Sex,data=gender_heightdf, col=colour, main="Height Values by Gender",
137         xlab="Gender", ylab="Height")
138 
139 wilcox.test(htval ~ Sex, data = gender_heightdf)
140 
141  #c ii. Valid weight.
142 gender_weightdf<-HSE_2011[,c("Sex","wtval")]
143 boxplot(wtval~Sex,data=gender_weightdf, col=colour, main="Weight Values by Gender",
144         xlab="Gender", ylab="Weight")
145 
146 wilcox.test(wtval ~ Sex, data = gender_weightdf)
147 
148 #d. What is the correlation between whether a person drinks nowadays, total household income, age at last birthday and gender?
149 qd<- HSE_2011[, c('dnnow','totinc','Age','Sex')]
150 library(psych)
151 corPlot(qd, cex = 1.2)
```

References

- Alcohol Change UK (n.d.) Alcohol Statistics, Alcohol Change UK. Available at: https://alcoholchange.org.uk/alcohol-facts/fact-sheets/alcohol-statistics (Accessed: April 6, 2023).
- Drink Aware (no date) Alcohol consumption UK, Drinkaware. Available at: https://www.drinkaware.co.uk/research/alcohol-facts-and-data/alcohol-consumption-uk#:~:text=In%202021%2C%2018%25%20of%20adults,compared%20to%2011%25%20of%20females. (Accessed: April 13, 2023).
- Gov.uk (2023) Male and female populations, GOV.UK Ethnicity facts and figures. Office for National Statistics. Available at: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/male-and-female-populations/latest (Accessed: April 13, 2023).
- Morgan, E. (2019) Population estimates by marital status and Living Arrangements, England and Wales: 2018, Population estimates by marital status and living arrangements, England and Wales - Office for National Statistics. Office for National Statistics. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationestimatesbymaritalstatusandlivingarrangements/2018 (Accessed: April 13, 2023).

References

- NHS (2023) Health Survey for England - Health, social care and lifestyles, NHS choices. NHS. Available at: https://digital.nhs.uk/data-and-information/areas-of-interest/public-health/health-survey-for-england---health-social-care-and-lifestyles#:~:text=The%20Health%20Survey%20for%20England,in%20the%20survey%20each%20year . (Accessed: April 6, 2023).
- NHS Digital (2022). Overweight and obesity in adults. NHS. Available at: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2021/overweight-and-obesity-in-adults . (Accessed: April 6, 2023).
- ONS (2022). Divorces in England and Wales: 2021, Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/divorce/bulletins/divorcesinenglandandwales/2021 (Accessed: April 13, 2023).
- Rosoff, D.B. *et al.* (2019) "Educational attainment impacts drinking behaviors and risk for alcohol dependence: Results from a two-sample Mendelian randomization study with ~780,000 participants," *Molecular Psychiatry*, 26(4), pp. 1119–1132. Available at: https://doi.org/10.1038/s41380-019-0535-9.
- Wilsnack, R.W. and Wilsnack, S.C. (1997) "Gender and Alcohol," INDIVIDUAL AND SOCIAL PERSPECTIVES. Available at: https://www.researchgate.net/profile/Michael-Frone/publication/259177012_Gender_Stress_Coping_and_Alcohol_Use/links/5af4ad19a6fdcc0c030af7ff/Gender-Stress-Coping-and-Alcohol-Use.pdf (Accessed: April 13, 2023).