# Introduction to R

# Code: Chi Squared Testing

# Author: J. Priestley, Ph.D.

# Import data
PSU <- read.csv("pennstate3.csv")

str(PSU)

summary(PSU)

# Chi-squared test
install.packages("prettyR")
library(prettyR)

# here we are creating a cross tab with the respective percentages
xtab(Sex~grtpers, data=PSU)

# here we are eliminating the blank values associated with great personality
PSU1 <- PSU[-which(PSU$grtpers==""),]
PSU1$grtpers<-factor(PSU1$grtpers)

xtab (Sex~grtpers, data=PSU1)

# note that when you generate a 2x2 table, you automatically generate the odds ratio and the relative risk:
# odds ratio here is ((%Female|No)/(%Female|Yes)) /((%Male|No)/(%Male|Yes))
# relative risk here is (%Yes|Male)/(%Yes|Female)
# here we are generating a simple chisquare test

```r
chisq.test(PSU1$Sex, PSU1$grtpers, correct=FALSE)
```

```r
chi <- chisq.test(PSU1$Sex, PSU1$grtpers, correct=FALSE)
```

```r
names(chi)
```

```r
chi$expected
d
```

```r
chi$observed
d
```

```r
chi$residuals
d
```

# note that you can also create a table

```r
tbla<-table(PSU1$Sex, PSU1$grtpers)
```

```r
?table
```

```r
chisq.test(tbla)
```

# you can create the proportions in the table based on column:

```r
prop.table(tbla, 2)
```

# you can create the proportions in the table based on row:

```r
prop.table(tbla, 1)
```

# this is an easy way to create the 100% stacked bar chart:

```r
barplot(prop.table(tbla, 2) )
```

```r
colors<-c("Yellow", "Blue")
```

```r
barplot(prop.table(tbla,2), col=colors, main = "Figure i: Bar Chart of Gender and Dating",
        xlab = "Gender")
```