Sex and Financial Risk Aversion

Victoria G. Walter

Table of Contents

# Introduction

**“How More Women on Wall Street Could Have Prevented the Financial Crisis”**

This 2016 Fortune Magazine title and concept are obviously very intriguing– and controversial– but is there any merit to the idea? Or is it just click bait?

The article, an except of Newton-Small’s new book “Broad Influence: How Women Are Changing the Way America Works,” discusses the Great Recession, it’s causes, and the female regulators who were sent in to fix the situation.

*“The heart of the problem was risk, something mostly male Wall Street seemed to take irresponsibly and something the female regulators were sent in to mitigate. It’s a cliched image: the straying reckless man and a woman at home holding things together. But there is some underlying truth in it. Neuroscience has shown links between risk taking and testosterone, which is 15 times as prevalent in men as in women. Many world leaders, from Bair to International Monetary Fund chief Christine Lagarde to British Labour deput leader Harriet Harman, who was then Prime Minister Gordon Brown’s No. 2, go Japanese prime minister Shinzo Abe, became convinced that if more women had been working in senior Wall Street positions, the global financial crisis probably wouldn’t have happened. And many saw the crisis as a wake-up call for Wall Street to diversify its ranks.”*

The article goes on to explain a number of studies supporting the idea that women are more risk averse and that a greater female presence on Wall Street changes the overall behavior of trading floors to make fewer financially risky decisions.

Intrigued, I was inspired to create this report.

**I am curious regarding three specific points:**

**1. Are women more risk averse investors?**

**2. Are women more likely to invest in risk averse assets and less likely to invest in riskier assets?**

**3. Do women invest more of their portfolios in risk averse assets and less in riskier assets?**

Thus, I analyze investors’ sex, self-reported level of risk aversion, what assets they own, the market value of those assets, and how much of their portfolios are invested in those asset.

In this report, you will find my methodology for this study, as well as several plots and Cochran-Mantel-Haenszel tests to answer my questions.

# Methodology

To do this, I select four common asset classes because of their distinct levels of risk.

## Explanation of Asset Classes and Risk

An important note: Stocks (equities) are generally more risky than bonds (debt instruments) because bondholders (those who own the company’s debt) are paid before shareholders (those who own equity, stocks) if the company goes bankrupt. Thus, bondholders are more likely to receive their money back on their investment than shareholders/stock owners, and make bonds, generally, a less risky asset.

Stocks - individual shares of publicly traded equities. Of the four asset classes, they generally the most volatile and they are the most difficult to diversify and adjust for risk.

Stock Funds - a fund made up of stocks, including domestic stock funds, growth funds, index funds, global stock funds, sector funds. Still risky, but not as risky as stocks, because investors can diversify with many different shares to adjust for risk in companies and sectors.

Combination Funds - funds that hold both stocks and bonds. Less risky, because bonds are generally less risky than stocks, so the combination fund is reasonably diversified and risk adjusted.

Treasury Bond Funds - Treasury bonds are bonds backed by the United States government and are safest asset to own (except cash) because the United States has never defaulted. A treasury bond fund is just a fund made up of these bonds, most likely with different rates to maturity, such as a 2-year bond, 10-year bond, and 30-year bond.

## Data

I use the most recent The Survey of Consumer Finances (2016), a survey to assess the economic and financial health of those in the United States, including families’ balance sheets, pensions, incomes, and demographic information. It is conducted by the Federal Reserve Bank of New York every four years or so, partially in person and partially over the phone. The data set is very useful because it has a large sample size – over 30,000 respondents – and it offers a great depth of information – over 6,000 variables.

However, it is important to remember that the survey is voluntary, and very long and invasive, so population skews very high towards those that are older, likely retired, and likely to have more time on their hands. Additionally, the survey is predominantly White and thus isn’t properly representative of the United States racial makeup in 2016 – also a flaw of it being a very long voluntary survey.

# Data Cleaning

First I import and clean my data – the 2016 Survey of Consumer Finances.

## Importing 2016 Survey of Consumer Finances Data from Stata file

scf\_data <- read.dta13('../data/scfp2016s/p16i6.dta')

## Select Variables

Once the data is imported, I select the specific columns I am Using for my analysis and rename them so they make sense. I chose the following variables:

**Variable**…………………………………………………………………………*Description*

**Sex**………………………………………………………………………………….*What is the sex of the individual who makes the financial decisions in the Respondent’s family?*

**Financial\_Risk\_Willingness**………………………………………..*How willing to take financial risks is the individual who makes the financial decisions in the Respondent’s family? On a Scale of -1 (Not at All Willing) to 10 (Very Willing)*

**Owns\_Stocks**…………………………………………………………………*Does the respondent own stocks?*

**Owns\_Stock\_Funds**……………………………………………………..*Does the respondent own stock funds?*

**Owns\_Combination\_Funds**…………………………………………*Does the respondent own combination funds?*

**Owns\_Treasury\_Bonds\_Funds**…………………………………..*Does the respondent own treasury bond funds?*

**TMV\_Stocks**………………………………………………………………….*Total Market Value of the Stocks the Respondent Owns in 2016 U.S. Dollars*

**TMV\_Stock\_Funds**………………………………………………………*Total Market Value of the Stock Funds the Respondent Owns in 2016 U.S. Dollars*

**TMV\_Combination\_Funds**………………………………………….*Total Market Value of the Combination Funds the Respondent Owns in 2016 U.S. Dollars*

**TMV\_Treasury\_Bond\_Funds**……………………………………..*Total Market Value of the Treasury Bond Funds Funds the Respondent Owns in 2016 U.S. Dollars*

scf\_data %<>%  
 select("Sex" = "X8000",  
 "Financial\_Risk\_Willingness" = "X7557",   
 "Owns\_Stocks" = "X3913",   
 "Owns\_Stock\_Funds" = "X3821",   
 "Owns\_Combination\_Funds" = "X3829",   
 "Owns\_Treasury\_Bond\_Funds" = "X3825",   
 "TMV\_Stocks" = "X3915",   
 "TMV\_Stock\_Funds" = "X3822",   
 "TMV\_Combination\_Funds" = "X3830",  
 "TMV\_Treasury\_Bond\_Funds" = "X3826",)

## Recode and Factor Existing Categorical Variables

Next, I recode the variables that are categorical so that they make sense, and make them factored

scf\_data$Sex = dplyr::recode\_factor(as.character(scf\_data$Sex), "1"="Female","5"="Male")  
  
scf\_data$Owns\_Stocks <- dplyr::recode(as.character(scf\_data$Owns\_Stocks), "1"="Owns", "5"="Does Not Own")

For the fund variables – stock funds, combination funds, and treasury bond funds – the coding is as follows: 1 = Owns \_\_\_\_\_ funds, 5 = Doesn’t own \_\_\_\_\_ funds, but owns other funds, 0 = Doesn’t own any funds

scf\_data$Owns\_Stock\_Funds <- dplyr::recode(as.character(scf\_data$Owns\_Stock\_Funds),"1"="Owns",   
 "5"="Does Not Own",  
 "0"="Does Not Own")  
  
scf\_data$Owns\_Combination\_Funds <- dplyr::recode(as.character(scf\_data$Owns\_Combination\_Funds), "1"="Owns",   
 "5"="Does Not Own",  
 "0"="Does Not Own")  
  
scf\_data$Owns\_Treasury\_Bond\_Funds <- dplyr::recode(as.character(scf\_data$Owns\_Treasury\_Bond\_Funds),   
 "1"="Owns",   
 "5"="Does Not Own",  
 "0"="Does Not Own")

## Create New Variables

Additionally, I am creating a few new variables to account for how much a Respondent owns in one asset relative to the other assets.

First, I create “TMV\_Investments,” or the Total Market Value of Stocks, Stock Funds, Combination Funds, and Treasury Bond Funds the Respondent owns added together.

Then, I calculate the percentage of the Respondent’s portfolio in each asset. This is done by dividing the Total Market Value of the Asset by the Total Market Value of all four assets combined and then multiplied by 100%.

**Variable**………………………………………………………………………..*Description*

**“Percent\_Stocks”**………………………………………………………..*Percentage of the Respondent’s Portfolio in Stocks*

**“Percent\_Stock\_Funds”**…………………………………………….*Percentage of the Respondent’s Portfolio in Stock Funds*

**“Percent\_Combination\_Funds”**………………………………..*Percentage of the Respondent’s Portfolio in Combination Funds*

**“Percent\_Treasury\_Bonds”**………………………………………*Percentage of the Respondent’s Portfolio in Treasury Bond Funds*

scf\_data %<>%  
 mutate(TMV\_Investments = TMV\_Treasury\_Bond\_Funds + TMV\_Stocks + TMV\_Stock\_Funds + TMV\_Combination\_Funds,  
 Percent\_Treasury\_Bond\_Funds = TMV\_Treasury\_Bond\_Funds / TMV\_Investments \* 100,  
 Percent\_Combination\_Funds = TMV\_Combination\_Funds / TMV\_Investments \* 100,  
 Percent\_Stock\_Funds = TMV\_Stock\_Funds / TMV\_Investments \* 100,  
 Percent\_Stocks = TMV\_Stocks / TMV\_Investments \* 100)

Next, I create the “Risk\_Aversion” variable, which takes the Financial Risk Willingness scale of -1 to 10 and groups it into a condensed categorical variable.

**Risk Aversion (New Variable)**………………………………………..*Financial Risk Willingness (Variable from SCF)*

**Risk Averse**…………………………………………………………………………………………….*-1, 1, 2, 3*

**Risk Neutral**……………………………………………………………………………………………….*4, 5, 6*

**Risk Tolerant**…………………………………………………………………………………………..*7, 8, 9, 10*

scf\_data$Risk\_Aversion <- cut(scf\_data$Financial\_Risk\_Willingness,  
 breaks = c(-Inf, 4, 7, Inf),  
 labels = c("Risk Averse", "Risk Neutral", "Risk Tolerant"))

Then, I order the Risk Aversion variable from risk averse to risk tolerant.

scf\_data$Risk\_Aversion <- ordered(scf\_data$Risk\_Aversion, levels = c("Risk Averse", "Risk Neutral", "Risk Tolerant"))

## Filtering Survey Data for Only Respondents that Own Investments

The final step in cleaning is filtering only for Respondents who’s Total Market Value of Investments (Stocks + Stock Funds +Combination Funds + Treasury Bond Funds) is Greater than 0. Therefore this analysis will only be run on Respondents that own at least 1 of the four assets.

scf\_data\_cleaned <- scf\_data %>%  
 filter(TMV\_Investments>0)

This significantly reduces the data set, from 31,240 observations to 9,156. From this point on in the analysis, I will be referring to these filtered Respondents as “Investors.”

# Are women more risk averse investors?

I am going to tackle this question in 2 ways:

-a. How do male and female investors describe their risk tolerance?: Creating a pivot table and visualizing through a barplot

-b. Are female investors more likely to report they are risk averse?: Creating a pivot table and analyzing using a Chi-squared test

## a. How do male and female investors describe their risk tolerance?

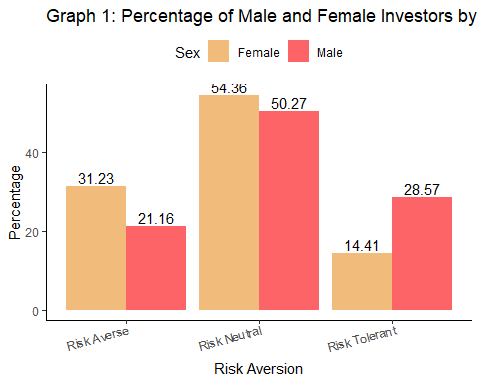
**Table 1:** I created a pivot table, “table\_1,” using the Sex and Risk Aversion variables, and then created a “Percentage” column. It makes more sense to use percentages instead of number of investors for this analysis, because there are so many more male investors than female investors.

table\_1 <- scf\_data\_cleaned %>%  
 select(Sex, Risk\_Aversion) %>%  
 group\_by(Sex, Risk\_Aversion) %>%  
 summarise(Number\_of\_Responses=n()) %>%  
 group\_by(Sex) %>%  
 mutate(Percentage=Number\_of\_Responses/sum(Number\_of\_Responses)\*100)

## `summarise()` regrouping output by 'Sex' (override with `.groups` argument)

**Graph 1:** I made a barplot with “table\_1” to visualize if there are any noticeable differences between how the sexes report their level of risk averison. I use the Wes Anderson color palette Grand Budapest 1 from one of my favorite movies.

ggplot(table\_1, aes(x = Risk\_Aversion, y = Percentage, fill = Sex))+   
 geom\_bar(stat='identity', position='dodge')+  
 labs(x = "Risk Aversion", y = "Percentage", title = "Graph 1: Percentage of Male and Female Investors by Risk Aversion")+   
 geom\_text(aes(label=round(Percentage, digits = 2)), position=position\_dodge(width=0.9), vjust=-0.25)+  
 theme\_classic()+  
 theme(legend.position = "top")+  
 theme(axis.text.x = element\_text(angle = 15, vjust = 1.2, hjust=1))+   
 scale\_fill\_manual(values = wes\_palette("GrandBudapest1", n = 2))



**Insights from Graph 1:** Male investors are risk tolerant at twice the level of female investors – 28.57% vs. 14.41%.

There is a much greater percentage of female investors who are risk averse than men – 31.23% vs. 21.16%.

A slightly greater percentage of female investors are risk neutral than male investors, but it is close – 54.36% vs. 50.27%.

## b. Are female investors more likely to report they are risk averse?

I make a table and then use a chi-squared test to test if there is a statistically significant relationship between sex and risk aversion.

**Table**

table\_5 <- table(scf\_data\_cleaned$Risk\_Aversion, scf\_data\_cleaned$Sex)   
  
table\_5

##   
## Female Male  
## Risk Averse 505 1595  
## Risk Neutral 879 3790  
## Risk Tolerant 233 2154

**Chi-Square Test**

chisq.test(table\_5)

##   
## Pearson's Chi-squared test  
##   
## data: table\_5  
## X-squared = 165.71, df = 2, p-value < 2.2e-16

*There is a statistically significant relationship between sex and risk aversion.*

# Are women more likely to invest in risk averse assets and less likely to invest in riskier assets?

I am going to assess this question in 3 ways:

-a. Do men and women own each asset at about the same rate?: Creating a pivot table and visualizing using a barplot

-b. Controlling for Risk Aversion, do men and women own each asset about the same rate?: Creating a pivot table and visualizing using a barplot

-c. Analyze Sex, Asset type, and Risk Aversion using a fancy test

## a. Do men and women own each asset at about the same rate?

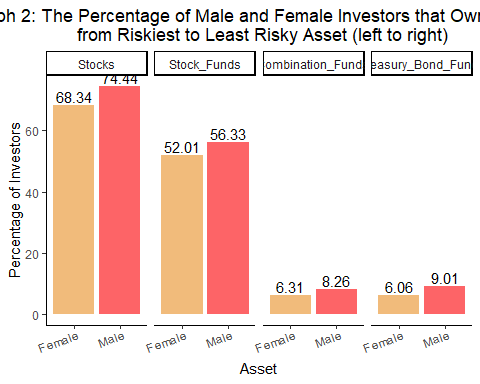
**Table 2:**

I start my analysis by creating a pivot table, “table\_2,” of the number of male and female investors that own and do not each asset, and then create a “percent” column. Then, I factor the Asset variable. Finally, I filter “table\_2” so that it only contains investors who own each asset, and create “table\_2\_filtered.”

table\_2 <- scf\_data\_cleaned %>%  
 select(Stocks=Owns\_Stocks, Stock\_Funds=Owns\_Stock\_Funds, Combination\_Funds=Owns\_Combination\_Funds, Treasury\_Bond\_Funds=Owns\_Treasury\_Bond\_Funds, Sex) %>%  
 pivot\_longer(-Sex, names\_to='Asset', values\_to='Response') %>%  
 count(Sex, Asset, Response) %>%  
 group\_by(Sex, Asset) %>%  
 mutate(percent = n / sum(n) \* 100)  
  
table\_2$Asset <- factor(table\_2$Asset, levels = c("Stocks", "Stock\_Funds", "Combination\_Funds", "Treasury\_Bond\_Funds", ordered = TRUE))  
  
table\_2\_filtered <- table\_2 %>%  
 filter(Response=="Owns")

**Graph 2:** I make a bar plot with “table\_2\_filtered” using ggplot2 to visualize the Percentage of Investors who own each asset, broken down by sex.

ggplot(table\_2\_filtered, aes(x=Sex, y=percent, fill=Sex))+  
 geom\_bar(stat = 'identity', position = 'dodge')+  
 facet\_grid(~Asset)+  
 labs(title = "Graph 2: The Percentage of Male and Female Investors that Own Each Asset, \n from Riskiest to Least Risky Asset (left to right)", y = "Percentage of Investors", x = "Asset") +  
 geom\_text(aes(label=round(percent, digits=2)), position=position\_dodge(width=0.9), vjust=-0.25)+  
 theme\_classic()+  
 theme(axis.text.x = element\_text(angle = 20, vjust = 1.2, hjust=1)) +  
 theme(legend.position = "non")+  
 theme(plot.title = element\_text(hjust = 0.4))+  
 scale\_fill\_manual(values = wes\_palette("GrandBudapest1", n = 3))



**Insights from Graph 2:** There is a greater share of men than women who own each asset, though the difference is not by much. Stocks are the most popularly owned asset, followed by Stock Funds, Treasury Bond Funds, and Combination Funds.

68.34% of women own stocks compared to 74.44% of men. This is a difference of 8.93%.

52.01% of women own stock funds versus 56.33% of men. This is a difference of 8.31%.

6.31% of women own combination funds compared to 8.26% men. This is a difference of 23.61%.

6.06% of women own treasury bond funds versus 9.01% men. This is a difference of 32.74%.

The differences in ownership for combination funds and treasury bond funds seem extreme, but it is important to note the small sample sizes since most of the investors in the study do not own those assets.

## b. Controlling for Risk Aversion, do men and women own each asset about the same rate?

**Table 3:** Next, I create the pivot table, “table\_3,” to observe what percentage of investors who own each asset identity as different levels of risk aversion. This is useful to see, considering that the each asset has a different level of risk. Do those that own risky assets have a higher percentage of people that are risk tolerant? Do those that own less risky assets have a higher percentage of people that are risk averse? “Table\_3” also breaks this information down by sex.

Additionally, I factor the Asset column and order the levels from most risky to least risky asset.

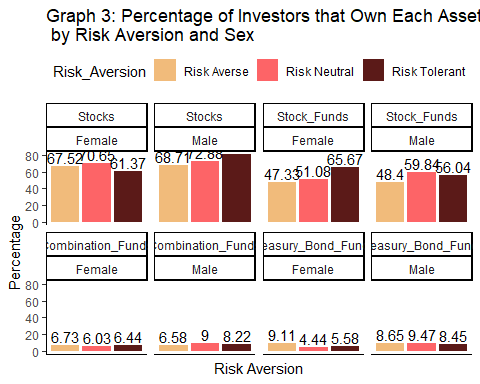
Finally, I filter the Owns column in “table\_3,” to create “table\_3\_filtered,” so that only investors that own that particular asset are in the data.

table\_3 <- scf\_data\_cleaned %>%  
 select(Stocks=Owns\_Stocks, Stock\_Funds=Owns\_Stock\_Funds, Combination\_Funds=Owns\_Combination\_Funds, Treasury\_Bond\_Funds=Owns\_Treasury\_Bond\_Funds, Sex, Risk\_Aversion) %>%  
 pivot\_longer(-c(Sex, Risk\_Aversion), names\_to= 'Asset', values\_to='Response') %>%  
 count(Sex, Asset, Risk\_Aversion, Response) %>%  
 group\_by(Sex, Asset, Risk\_Aversion) %>%  
 mutate(percent = n / sum(n) \* 100)  
  
table\_3$Asset <- factor(table\_3$Asset, levels = c("Stocks", "Stock\_Funds", "Combination\_Funds", "Treasury\_Bond\_Funds", ordered = TRUE))  
  
table\_3\_filtered <- table\_3 %>%  
 filter(Response=="Owns")  
  
table\_3

## # A tibble: 48 x 6  
## # Groups: Sex, Asset, Risk\_Aversion [24]  
## Sex Asset Risk\_Aversion Response n percent  
## <fct> <fct> <ord> <chr> <int> <dbl>  
## 1 Female Combination\_Funds Risk Averse Does Not Own 471 93.3   
## 2 Female Combination\_Funds Risk Averse Owns 34 6.73  
## 3 Female Combination\_Funds Risk Neutral Does Not Own 826 94.0   
## 4 Female Combination\_Funds Risk Neutral Owns 53 6.03  
## 5 Female Combination\_Funds Risk Tolerant Does Not Own 218 93.6   
## 6 Female Combination\_Funds Risk Tolerant Owns 15 6.44  
## 7 Female Stock\_Funds Risk Averse Does Not Own 266 52.7   
## 8 Female Stock\_Funds Risk Averse Owns 239 47.3   
## 9 Female Stock\_Funds Risk Neutral Does Not Own 430 48.9   
## 10 Female Stock\_Funds Risk Neutral Owns 449 51.1   
## # ... with 38 more rows

**Graph 3:** I made a barplot to visualize “table\_3\_filtered” and see if the investors in this data set invest according to their risk tolerance.

ggplot(table\_3\_filtered, aes(x = Risk\_Aversion, y = percent, fill = Risk\_Aversion))+  
 geom\_bar(stat='identity', position = 'dodge')+  
 facet\_wrap(~Asset + Sex, ncol=4)+  
 labs(x = "Risk Aversion", y = "Percentage", title = "Graph 3: Percentage of Investors that Own Each Asset, \n by Risk Aversion and Sex")+  
 geom\_text(aes(label=round(percent, digits = 2)), position=position\_dodge(width=0.9), vjust=-0.25)+  
 theme\_classic()+  
 theme(axis.text.x = element\_blank())+   
 theme(axis.ticks.x = element\_blank())+  
 theme(legend.position = "top")+  
 scale\_fill\_manual(values = wes\_palette("GrandBudapest1", n = 3))



**Insights from Graph 3:** There are three instances of investors investing in accordance with their risk aversion.

1. The share of men who are invested in stocks increases with increased risk tolerance: 68.71% of risk averse men own stocks, 72.88% of risk neutral men own stocks, and 81.42% of risk tolerant men own stocks. This makes sense because stocks are the riskiest of the four asssets.
2. The share of women who are invested in stock funds increases as women become risk tolerant: 47.33% of risk averse women own stock funds, 51.08% of risk neutral women own stock funds, and 65.67% of risk tolerant women own stock funds. This also makes sense because stock funds are the second riskiest of the four assets.
3. The share of risk averse women who own treasury bond funds is nearly twice that of the risk neutral and risk tolerant women who own treasury bond funds. This makes sense because treasury bond funds are the least risky of the four assets.

Aside from these instances, there seems to be very little variation in the percentage of each asset owned across the levels of risk aversion. Thus, risk aversion may not have an effect on which assets an investor chooses to invest in.

## c. Cochran-Mantel-Haenszel Tests

I use the Cochran-Mantel-Haenszel test to analyze if it is statistically significant that one sex is more likely than the other to invest – first in general, and then for each asset individually– than the other while controlling for Risk Aversion. The variables I will be using are Sex, Asset, Risk Aversion, and Owns – whether or not the investor owns that particular investment.

The Cochran-Mantel-Haenszel test is useful because it is used for categorical variables, can analyze up to three variables at a time (a chi-squared test can only analyze two at a time), and it controls for how there are six times as many male investors as female investors.

First, I will run this test on all of the asset classes together, and then each asset class individually. ### All Asset Classes **Create Table:** I start by creating a data table, “dt,” to organize the data which I will use throughout the process of checking for statistical significance. This includes the count and percentage for each sex, level of risk aversion, and asset owned.

For example: How many women are risk averse and owned stocks? What percentage of men are risk neutral and own combination funds?

dt <- scf\_data\_cleaned %>% select(Sex, Treasury\_Bond\_Funds=Owns\_Treasury\_Bond\_Funds, Stocks=Owns\_Stocks, Stock\_Funds=Owns\_Stock\_Funds, Combination\_Funds=Owns\_Combination\_Funds, Risk\_Aversion) %>% pivot\_longer(-c(Sex, Risk\_Aversion), names\_to='Asset', values\_to='Owns') %>%  
 group\_by(Sex, Asset, Risk\_Aversion, Owns) %>%  
 summarise(Number\_of\_Responses=n()) %>%  
 group\_by(Sex, Asset, Risk\_Aversion) %>%  
 mutate(Percentage=Number\_of\_Responses/sum(Number\_of\_Responses)\*100)

## `summarise()` regrouping output by 'Sex', 'Asset', 'Risk\_Aversion' (override with `.groups` argument)

Here, I combine the Asset and Risk Aversion variables into one column because the Cochran-Mantel-Haenszel test can only analyze three categorical variables at a time. Thus, "Asset\_\_Risk\_Aversion," collapses the two variables into one. Also, I condense the table by not selecting “Percentage.” Below is what the data looks like.

dt\_all <- xtabs(Number\_of\_Responses ~ Asset\_\_Risk\_Aversion +Sex + Owns, data = dt %>% mutate(Asset\_\_Risk\_Aversion = paste0(Asset,'\_', Risk\_Aversion)))  
  
ftable(dt\_all)

## Owns Does Not Own Owns  
## Asset\_\_Risk\_Aversion Sex   
## Combination\_Funds\_Risk Averse Female 471 34  
## Male 1490 105  
## Combination\_Funds\_Risk Neutral Female 826 53  
## Male 3449 341  
## Combination\_Funds\_Risk Tolerant Female 218 15  
## Male 1977 177  
## Stock\_Funds\_Risk Averse Female 266 239  
## Male 823 772  
## Stock\_Funds\_Risk Neutral Female 430 449  
## Male 1522 2268  
## Stock\_Funds\_Risk Tolerant Female 80 153  
## Male 947 1207  
## Stocks\_Risk Averse Female 164 341  
## Male 499 1096  
## Stocks\_Risk Neutral Female 258 621  
## Male 1028 2762  
## Stocks\_Risk Tolerant Female 90 143  
## Male 400 1754  
## Treasury\_Bond\_Funds\_Risk Averse Female 459 46  
## Male 1457 138  
## Treasury\_Bond\_Funds\_Risk Neutral Female 840 39  
## Male 3431 359  
## Treasury\_Bond\_Funds\_Risk Tolerant Female 220 13  
## Male 1972 182

**Cochran-Mantel-Haenszel Test**

mantelhaen.test(dt\_all)

##   
## Cochran-Mantel-Haenszel test  
##   
## data: dt\_all  
## Cochran-Mantel-Haenszel M^2 = 667.77, df = 11, p-value < 2.2e-16

**Findings:** It is statistically significant than men are more likely to invest than women.

### Stocks

Next, I analyze for stock ownership only.

**Create Table:** I filter the original data table, “dt” for only stocks to create “dt\_s.” Then, I condense the data because I no longer need the “Percentage” variable to create “dt\_stocks.” Below is what the data looks like.

dt\_s <- dt %>%  
 filter(Asset=="Stocks")  
  
dt\_stocks <- xtabs(Number\_of\_Responses ~ Risk\_Aversion+Sex+Owns, data = dt\_s)  
  
ftable(dt\_stocks)

## Owns Does Not Own Owns  
## Risk\_Aversion Sex   
## Risk Averse Female 164 341  
## Male 499 1096  
## Risk Neutral Female 258 621  
## Male 1028 2762  
## Risk Tolerant Female 90 143  
## Male 400 1754

**Cochran-Mantel-Haenszel Test**

mantelhaen.test(dt\_stocks)

##   
## Cochran-Mantel-Haenszel test  
##   
## data: dt\_stocks  
## Cochran-Mantel-Haenszel M^2 = 156.66, df = 2, p-value < 2.2e-16

**Findings:** It is a statistically significant finding that men are more likely to invest in stocks than women.

### Stock Funds

Next, I analyze for stock funds.

**Create Table:** I filter the original data table, “dt” for only stock funds to create “dt\_sf.” Then, I condense the data because I no longer need the “Percentage” variable to create “dt\_stock\_funds.” Below is what the data looks like.

dt\_sf <- dt %>%  
 filter(Asset=="Stock\_Funds")  
  
dt\_stock\_funds <- xtabs(Number\_of\_Responses ~ Risk\_Aversion+Sex+Owns, data = dt\_sf)  
  
ftable(dt\_stock\_funds)

## Owns Does Not Own Owns  
## Risk\_Aversion Sex   
## Risk Averse Female 266 239  
## Male 823 772  
## Risk Neutral Female 430 449  
## Male 1522 2268  
## Risk Tolerant Female 80 153  
## Male 947 1207

**Cochran-Mantel-Haenszel test**

mantelhaen.test(dt\_stock\_funds)

##   
## Cochran-Mantel-Haenszel test  
##   
## data: dt\_stock\_funds  
## Cochran-Mantel-Haenszel M^2 = 162.07, df = 2, p-value < 2.2e-16

**Findings:** It is a statistically significant finding that men are more likely to invest in stock funds than women.

### Combination Funds

**Create Table:** I filter the original data table, “dt” for only combination funds to create “dt\_cf.” Then, I condense the data because I no longer need the “Percentage” variable to create “dt\_combination\_funds.” Below is what the data looks like.

dt\_cf <- dt %>%  
 filter(Asset=="Combination\_Funds")  
  
dt\_combination\_funds <- xtabs(Number\_of\_Responses ~ Risk\_Aversion+Sex+Owns, data = dt\_cf)  
  
ftable(dt\_combination\_funds)

## Owns Does Not Own Owns  
## Risk\_Aversion Sex   
## Risk Averse Female 471 34  
## Male 1490 105  
## Risk Neutral Female 826 53  
## Male 3449 341  
## Risk Tolerant Female 218 15  
## Male 1977 177

**Cochran-Mantel-Haenszel test**

mantelhaen.test(dt\_combination\_funds)

##   
## Cochran-Mantel-Haenszel test  
##   
## data: dt\_combination\_funds  
## Cochran-Mantel-Haenszel M^2 = 164.81, df = 2, p-value < 2.2e-16

**Findings:** It is a statistically significant finding that men are more likely to invest in combination funds than women.

### Treasury Bond Funds

**Create Table:** I filter the original data table, “dt” for only treasury bond funds to create “dt\_tbf.” Then, I condense the data because I no longer need the “Percentage” variable to create “dt\_treasury\_bond\_funds.” Below is what the data looks like.

dt\_tbf <- dt %>%  
 filter(Asset=="Treasury\_Bond\_Funds")  
  
dt\_treasury\_bond\_funds <- xtabs(Number\_of\_Responses ~ Risk\_Aversion+Sex+Owns, data = dt\_tbf)  
  
ftable(dt\_treasury\_bond\_funds)

## Owns Does Not Own Owns  
## Risk\_Aversion Sex   
## Risk Averse Female 459 46  
## Male 1457 138  
## Risk Neutral Female 840 39  
## Male 3431 359  
## Risk Tolerant Female 220 13  
## Male 1972 182

**Cohran-Mantel-Haenszel Test**

mantelhaen.test(dt\_treasury\_bond\_funds)

##   
## Cochran-Mantel-Haenszel test  
##   
## data: dt\_treasury\_bond\_funds  
## Cochran-Mantel-Haenszel M^2 = 166.66, df = 2, p-value < 2.2e-16

**Findings:** It is a statistically significant finding that men are more likely to invest in treasury bond funds than women.

# Do women invest more of their portfolios in risk averse assets and less in riskier assets?

**Table 4:** This question is a little more straightforward.

I create a pivot table, “table\_4,” with the variables that represent the percent of each asset (“Percent\_Stocks,” “Percent\_Stock\_Funds,” “Percent\_Combination\_Funds,” “Percent\_Treasury\_Bond\_Funds”), sex, and risk aversion.

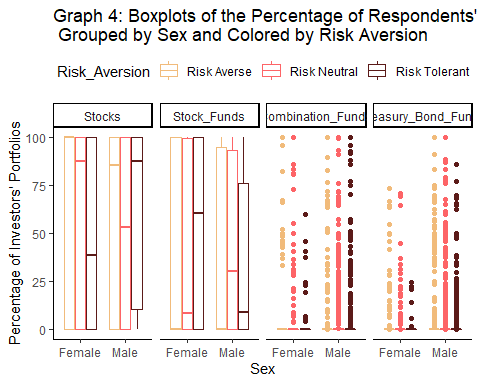
Then, I factor the Asset variable and order the levels from riskiest to least risk asset.

Finally, I filter “table\_4,” and create “table\_4\_filtered,” so that it only shows data for investors who own each asset, which effectively is if the percentage of their portfolio in that asset is greater than zero.

table\_4 <- scf\_data\_cleaned %>%  
 select(Stocks=Percent\_Stocks, Stock\_Funds=Percent\_Stock\_Funds, Combination\_Funds=Percent\_Combination\_Funds, Treasury\_Bond\_Funds=Percent\_Treasury\_Bond\_Funds, Sex, Risk\_Aversion) %>%  
 pivot\_longer(-c(Sex, Risk\_Aversion), names\_to='Asset', values\_to='Percentage\_of\_Portfolio') %>%  
 summarise(Sex, Asset, Risk\_Aversion, Percentage\_of\_Portfolio)  
  
table\_4$Asset <- factor(table\_4$Asset, levels = c("Stocks", "Stock\_Funds", "Combination\_Funds", "Treasury\_Bond\_Funds", ordered = TRUE))  
  
table\_4\_filtered <- table\_4 %>%  
 filter(Percentage\_of\_Portfolio > 0)

**Graph 4:** I create boxplots to visualize the distribution of portfolio makeup across the investors and grouped by sex and risk aversion.

ggplot(table\_4, aes(x = Sex, y = Percentage\_of\_Portfolio, col = Risk\_Aversion))+  
 geom\_boxplot()+   
 facet\_wrap(~Asset, ncol = 4)+   
 labs(x = "Sex", y = "Percentage of Investors' Portfolios", title = "Graph 4: Boxplots of the Percentage of Respondents' Portfolios in Each Asset, \n Grouped by Sex and Colored by Risk Aversion")+  
 theme\_classic()+  
 theme(legend.position = "top")+  
 scale\_color\_manual(values = wes\_palette("GrandBudapest1", n = 3))



**Insights from Graph 4:** *These results are pretty interesting!*

It seems that for most investors who own stocks, the majority of their portfolio is in stocks. For both men and women, the 50th and 75th percentiles are at 100% for all levels of risk aversion.

The results for stock funds are particularly interesting because we see distinct differences between male and female investors. The 50th and 75th percentiles are still at 100%, but for male investors, the 25th and 50th percentiles decrease as male investors become more comfortable with risk.

For Combination Funds and Treasury Bond Funds, the 25th, 50th, and 75th percentiles decrease as both male and female investors become more comfortable with risk. Combination and Treasury Bond Funds are the least risky. Thus, it seems that investors who are risk averse own a greater share of their portfolio in the least risky assets. Logically, this makes a lot of sense.

# Conclusion

There is a statistically significant relationship between sex and self-reported financial risk aversion and Graph 1 shows that there is a greater share of women who are risk averse and risk neutral than men, and thus a greater share of men who are risk tolerant in this study. However, when it comes to the actual portfolios of investors in this study, there wasn’t a big difference in how men and women invest. As shown in Graphs 2 and 3, there are a few examples of risk tolerant investors having a bigger portion of their portfolios in risky assets like stocks and stock funds, as well as risk averse investors having a bigger portion of their portfolios in risk averse assets like treasury bond funds. However, for the most part, self-reported risk aversion doesn’t seem to have a big impact on portfolio makeup.

Men are more likely to invest than women overall and for each asset class.

There is an important caveat to this analysis, however. Though stocks generally are the riskiest asset, there is a wide variety of risk levels for different types of stocks. For example, an investor could own a share of a Blue Chip stock that is not volatile and is known for having consistent and reliable growth – which would make it not very risky. Another investor could own a stock fund where the underlying stocks in the fund are very volatile – making it very risky. So in this example, the stock fund is actually riskier than the stock. Unfortunately, there is no way to control for this using this data because the Survey of Consumer Finances does not ask specific questions about the risk level of the assets they own.