ALLEGHENY COLLEGE
Department of Computer Science

1815

lab Assignment 8

Submitted To:
CMPSC 390
Prof. Oliver
Bonham-Carter

Submitted By :
Daniel Ocampo
Fall Semester

Contents

1   Objective

To explore statistical tools which are relevant for the evaluation of economic data. At this point in the class, there has already been much exposure to using many different types of statistical tools to handle different formats of data. It is therefore expected that the student will able to research code development and will be able to resolve data formatting issues during while working in the analysis of data. In particular, these skills comprise the ability to research R-statistics software packages for the application to the particular contexts for which they were designed and to extract knowledge from the produced visualizations and extracted interpretation of results. Furthermore, the student will be able to explain credible results and conclusions which are to be entirely supported by the methods and data.

2   Reading assignment

In the article by Heathcote et al. [1], female labor supply is discussed for its impact on the United States economy between the years of 1967 to 2002. The authors produce a model which was designed to help study and measure the below factors which are associated to women in the workforce.

Our data originates from [2] and concerns the lifestyles, living conditions and economic contri- butions made by women on the world's stage between the years of 1960 to 2017. In this lab, you are to use this data to gather direct or indirect proof to either support or refute the claims made by the four themed questions (listed above) of the study by Heathcote et al. For each of its four themes, determine the necessary variables from the data set in the spread- sheet which could be used to support or refute the claims of the article. Remember, the exact questions from the article may not be directly answerable by simply running tests over the in- cluded data, and so you will have to develop a strategy in which you will use the correlations of several variables, or multi-linear models to argue your point. For this step, your selected variables will amass proof that could be used to describe scenarios or conditions which will support or refute the conclusions of the authors.

2.1   The Article's Themed Questions

1. The decline in marriage rates,

2. The narrowing gender wage gap,

3. The preference (or cultural) shift towards market work, and

4. The change in womens bargaining power within the household.

3   R code that was Used

```r
library(gapminder)
library(dplyr)
library(tidyverse)
library(ggplot2)
library(psych)

library(ctv)
twoCountry <- read.csv(file = "/home/o/ocampod/fall2017/cs390-ochampoo/Data
View(twoCountry)
twoCountry[1,]

###### Question 1 ######The decline in marriage rates,

WifeBeater <- filter(
  twoCountry ,
  Indicator.Name ==   "Women who believe a husband is justified in beating h
  )

View(WifeBeater)

WifeBeater <- filter(
  WifeBeater, Country.Name == "Peru" | Country.Name == "Senegal")### Countr

write.csv(WifeBeater, file = "/home/o/ocampod/fall2017/cs390-ochampoo/Data-

WifeBeater <- read.csv(file = "/home/o/ocampod/fall2017/cs390-ochampoo/Data


qplot(Peru, unmarried, data = WifeBeater, alpha = I(1/4)) + geom_smooth(met
qplot(Senegal, unmarried, data = WifeBeater, alpha = I(1/4)) + geom_smooth
View(WifeBeater)
### we can see that that less women are less likely to let their husband be

LiteracyGender <- filter(#### this going to say if women are getting more e
  twoCountry ,
```

```
  Indicator.Name ==  "Literacy␣rate,␣adult␣female␣(%␣of␣females␣ages␣15␣an
)

View(LiteracyGender)


LiteracyGender <- filter(#### this going to say if women are getting more
  LiteracyGender,
  Country.Name ==  "Spain" | Country.Name ==  "Turkey"
)

write.csv(LiteracyGender, file = "/home/o/ocampod/fall2017/cs390-ochampoo/


LiteracyGender <- read.csv(file = "/home/o/ocampod/fall2017/cs390-ochampoo,

qplot(Spain, unmarried, data = LiteracyGender, alpha = I(1/4)) + geom_smoot
qplot(Turkey, unmarried, data = LiteracyGender, alpha = I(1/4)) + geom_smoo
View(LiteracyGender)



#################################### End Of Question 1 ################

############# Question 2 ############# The narrowing gender wage gap,
schooling <- filter(
  twoCountry ,
  Indicator.Name ==  "Expected␣years␣of␣schooling,␣female"

  )

schooling <- filter(
  schooling ,
  Country.Name ==  "Finland" | Country.Name ==  "Greece"
  )

write.csv(schooling, file = "/home/o/ocampod/fall2017/cs390-ochampoo/Data-

schooling <- read.csv(file = "/home/o/ocampod/fall2017/cs390-ochampoo/Data-
Finland <- select(schooling,Finland,Q1)
cor(Finland,method = "pearson")
Greece <- select(schooling,Greece,Q1)
```

```
cor(Finland)
pairs.panels(Finland)
pairs.panels(Greece)
describe(schooling)
corPlot(Greece)
View(schooling)

##################### end of  Question 2  ############

### Question 3 ###### preference (or cultural) shift towards market work

laborforceGender <- filter(
  twoCountry ,
  Indicator.Name ==  "Labor force, female")


laborforceGender <- filter(
  laborforceGender ,
  Country.Name  ==  "Mexico" | Country.Name  ==  "United States"
)

write.csv(laborforceGender, file = "/home/o/ocampod/fall2017/cs390-ochampod



laborforceGender <- read.csv(file = "/home/o/ocampod/fall2017/cs390-ochampo

test <- select(laborforceGender,MEX,Year)
View(test)
t.test(test)

laborforceGender[1,]

qplot(MEX, Year, data = laborforceGender, alpha = I(1/4)) + geom_smooth(met



laborforceGender %>% ggplot(aes(x = MEX, y = Year)) +
  geom_point(alpha = I(1/4)) + geom_smooth()
```

```
View(laborforceGender)
```

```
##### End of question 3 ##########
#### Question 4 ### The change in womens bargaining power within the house

homeowners <- filter(
  twoCountry ,
  Indicator.Name ==  "Women who own a house alone (\% of women age 15-49)"
)
View(homeowners)
#### was not able to obtain enough data for this one but most of the data
### A house

decisionMaker <- filter(
  twoCountry ,
 Indicator.Name ==  "Decision maker about a woman's own health care: mainly
)
View(decisionMaker)

write.csv(decisionMaker, file = "/home/o/ocampod/fall2017/cs390-ochampoo/Da

decisionMaker <- filter(
  decisionMaker ,
  Country.Name ==  "Peru" | Country.Name ==  "Bangladesh"
```

```
)

write.csv(decisionMaker, file = "/home/o/ocampod/fall2017/cs390-ochampoo/D

decisionMakerGraph <- read.csv(file = "/home/o/ocampod/fall2017/cs390-ocham

summary(decisionMakerGraph)


decisionMakerGraph \%>\% ggplot(aes(x = Peru, y = Year)) +
  geom_point(alpha = I(1/4)) + geom_smooth()

decisionMakerGraph \%>\% ggplot(aes(x = Bangladesh, y = Year)) +
  geom_point(alpha = I(1/4)) + geom_smooth()



View(decisionMakerGraph)


major<- filter(
   twoCountry,
  Indicator.Name ==  "Decision maker about major household purchases: main
  )

major<- filter(
  major,
  Country.Name ==  "Peru"
)

write.csv(major, file = "/home/o/ocampod/fall2017/cs390-ochampoo/Data-Analy

major <- read.csv(file = "/home/o/ocampod/fall2017/cs390-ochampoo/Data-Anal
ggplot(data = major) +
  geom_point(mapping = aes(x = Year, y = PER))
t.test(major)
View(major)


major \%>\% ggplot(aes(x = PER, y = Year)) +
  geom_point(alpha = I(1/4)) + geom_smooth()
```

```
###########End of question 4####

### Intresting Data

womenHiv <- filter(
  twoCountry ,
  Indicator.Name ==   "Women's␣share␣of␣population␣ages␣15+␣living␣with␣HIV␣
)


employ <- filter(
  twoCountry ,
  Indicator.Name ==   "Employment␣to␣population␣ratio,␣15+,␣female␣(\%)␣(moc
"  | Indicator.Name ==   "Employment␣to␣population␣ratio,␣15+,␣male␣(\%)␣(moc
"
)


married <- filter(
  twoCountry ,
  Indicator.Name ==   "Women␣who␣were␣first␣married␣by␣age␣18␣(\%␣of␣women␣a
)


SavedForEducation <- filter(
  twoCountry ,
  Indicator.Name ==   "Saved␣for␣education␣or␣school␣fees,␣female␣(\%␣age␣15
"Literacy␣rate,␣adult␣male␣(\%␣of␣males␣ages␣15␣and␣above)
")

View(SavedForEducation)

laborforceGender <- filter(
  twoCountry ,
  Indicator.Name ==   "Labor␣force,␣female")

View(laborforceGender)

femaleWage <- filter(
```

```
  twoCountry ,
  Indicator.Name ==   "Wage␣and␣salaried␣workers,␣female␣(\%␣of␣female␣emplo
"Wage␣and␣salaried␣workers,␣male␣(\%␣of␣male␣employment)"
)

Childemploy <- filter(
  twoCountry ,
  Indicator.Name ==   "Children␣in␣employment,␣female␣(\%␣of␣female␣childrer
"Children␣in␣employment,␣male␣(\%␣of␣male␣children␣ages␣7-14)
"
)


business <- filter(
  twoCountry ,
  Indicator.Name ==   "Cost␣of␣business␣start-up␣procedures,␣female␣(\%␣of␣(
" | Indicator.Name ==   "Cost␣of␣business␣start-up␣procedures,␣male␣(\%␣of␣(
)



View(twoCountry)
```

4 Personal Idea

The way we answered this question was by filtering the data From The world Bank
Gender statistics. After filtering I was able to pick up data by manually looking
at the data. I was considering using the like method but , the dataset was small
enough to look at.

1. The decline in marriage rates,

   The filter that was used to support or reject this theory was "Women who
   believe a husband is justified in beating his wife (any of five reasons) (%)". I
   chose the both Peru and Senegal, the reason I chose both of these country was
   because they had the most of data. Even then I had to manipulate the data
   so that it can fit. So for example I would to take the average of two data point
   so that I can get the middle missing data. In **figure 1** and In **figure 2** we see
   that women are less likely to think it is okay to beat women.

   The Next data set that we chose was "Literacy rate, adult female (% of females
   ages 15 and above)." The reason I chose this data set was because of the simple

fact that more educated would have a higher chance to have a job, and thus be more independent. In **Figure 3** We have a dataset

2. The narrowing gender wage gap

   For this dataset I look for country with most amount of data, which it happened to be Finland and Greece. These two countries had most data on the topic of, "Expected years of schooling, female". well the data shows that the more educated women get the more she will make.

3. The preference (or cultural) shift towards market work

   This one was women in the labor force, I feel that it more and more common to see women working instead of being a stay at home mom. there is more and more women joing the work force both in the US and Mexico.

4. The change in womens bargaining power within the household.

   This one is a little difficult to answer because of the simple fact that there was nearly no data. But with that data I did find I was able to conclude that women did not have may not have much of a bargaining power.

   I came to the conclusion based on my data of "Decision maker about a woman's own health care: mainly wife (% of women age 15-49)" For peru, there was a little spike in women who had less of say about their own health. While also Bangladesh there were little spikes but still increasing.

   The next one which was also difficult to answer because of lack of data was "Decision maker about major household purchases: mainly husband (% of women age 15-49)" The only country that enough data was Peru, and we can see that it had it ups and downs when it came making decision. ultimately based the correlation graph women were having a say.

5   Statement 1: The decline in marriage rates

Figure 1: The decline in marriage rates in peru:"Decision maker about major household purchases: mainly husband (% of women age 15-49)"
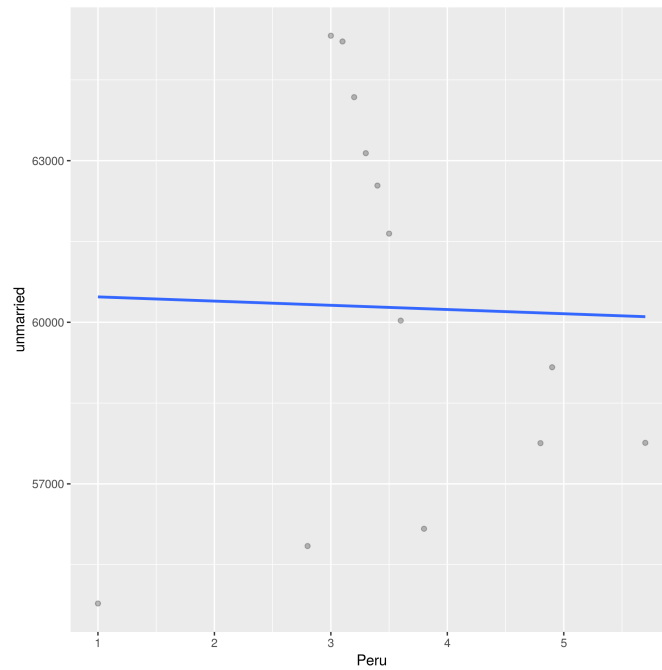
Figure 2: The decline in marriage rates in Senegal:"Decision maker about major household purchases: mainly husband (% of women age 15-49)"
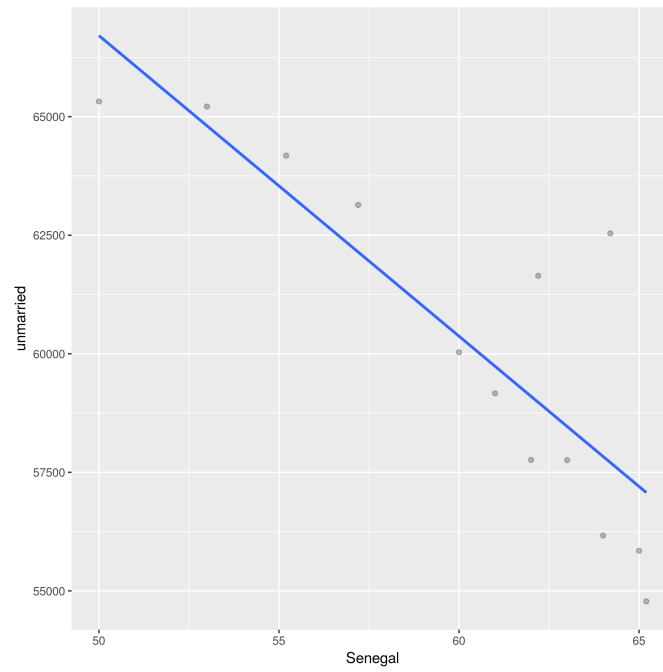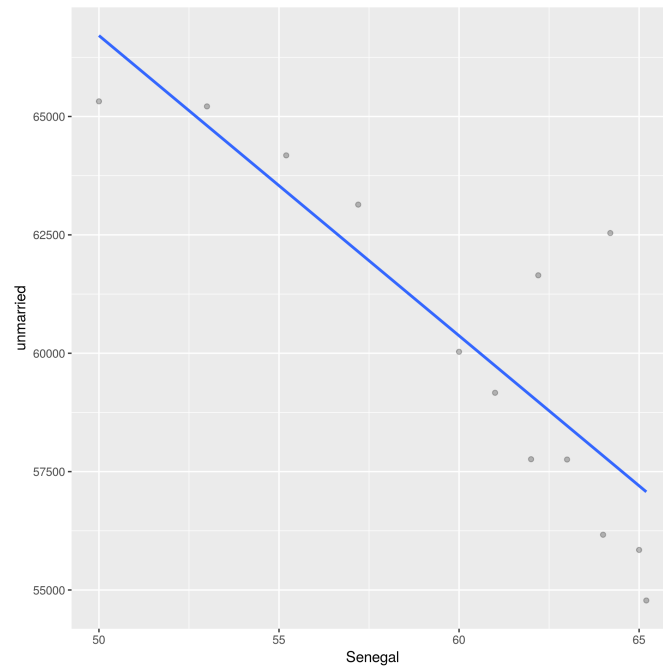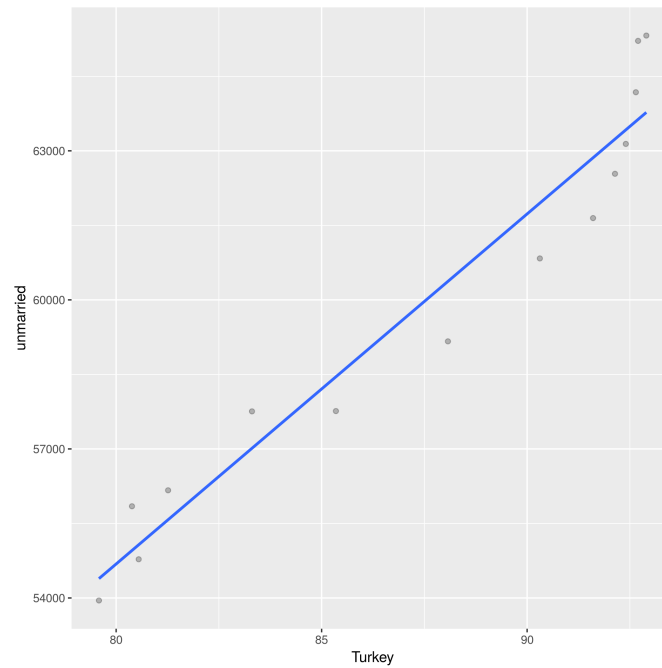
Figure 3: The decline in marriage rates in Senegal

Figure 4: The decline in marriage rates in Turkey

6   Statement 2:The narrowing gender wage gap

Figure 5: Increase of women going to school for Finland, and earning more as the years progress
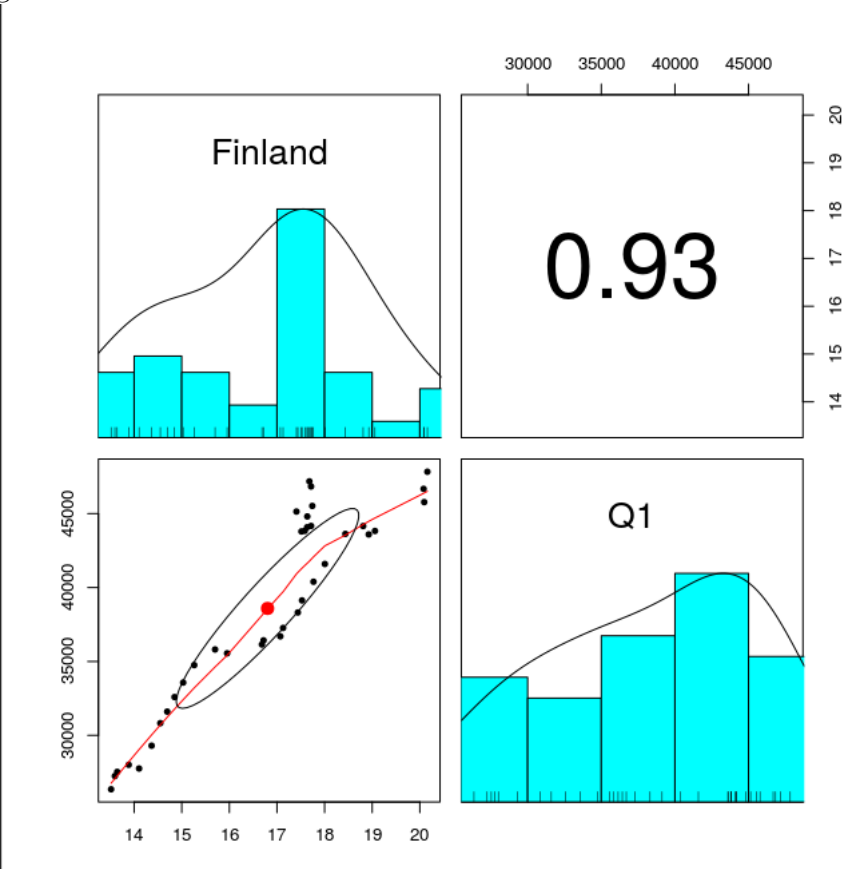
Figure 6: Increase in women going to school with an increase in pay Greece
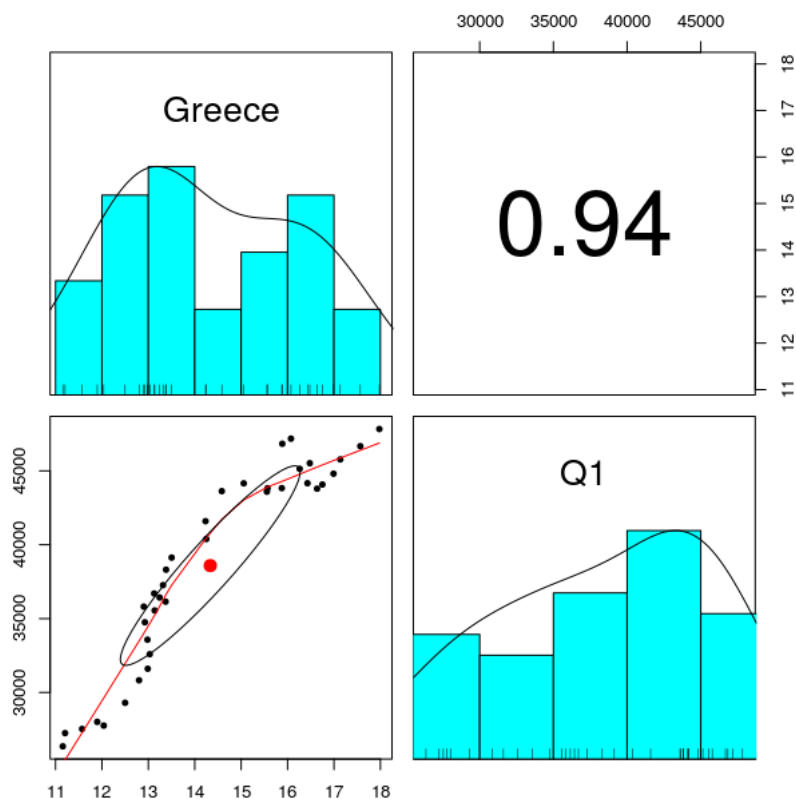
Figure 7: T test example

7 Statement 3: The preference (or cultural) shift towards market work

Figure 8: Increase of women in the Labor force for Mexico
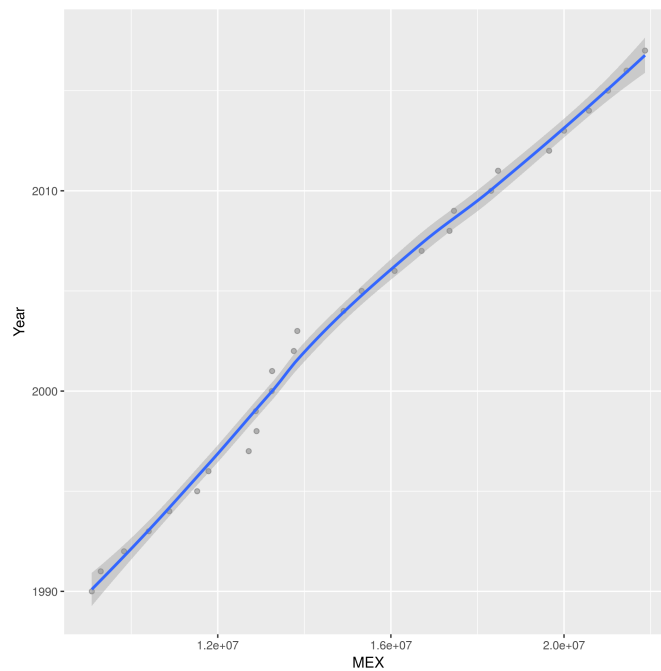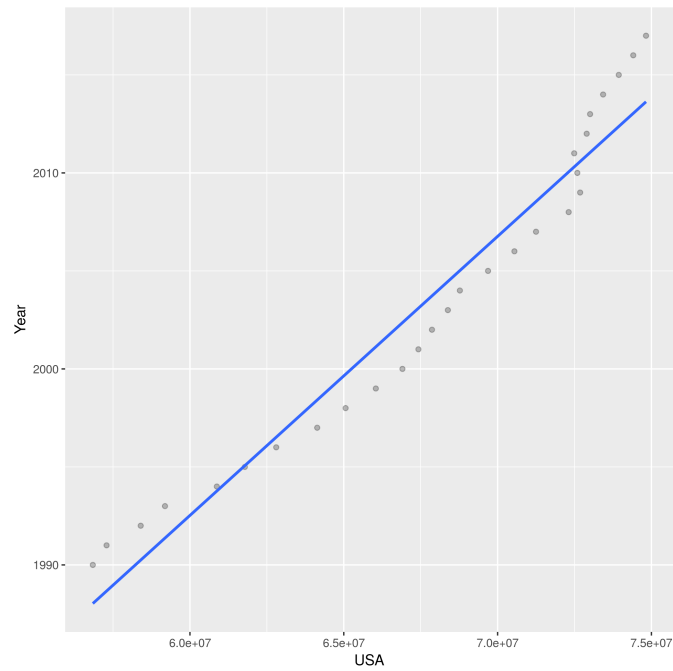
Figure 9: Increase in women in the labor force for Mexico
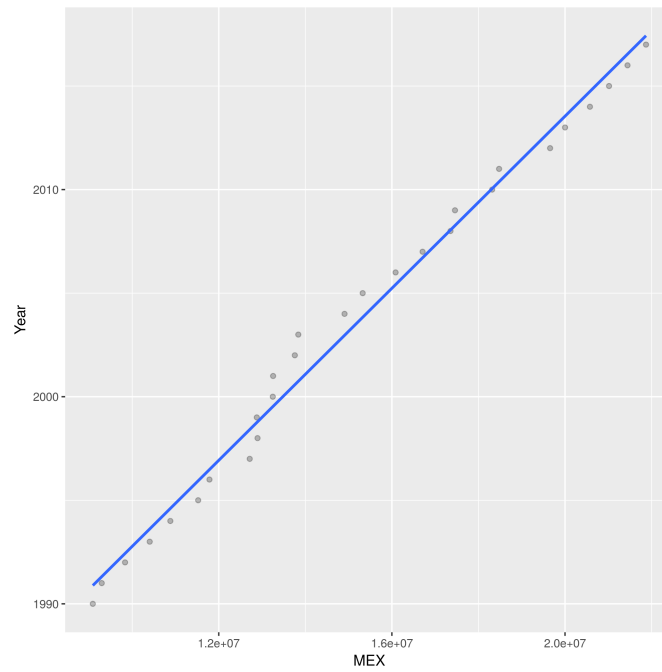
Figure 10: Increase in women in the labor force for USA

8 Statement 4: The change in womens bargaining power within the household.

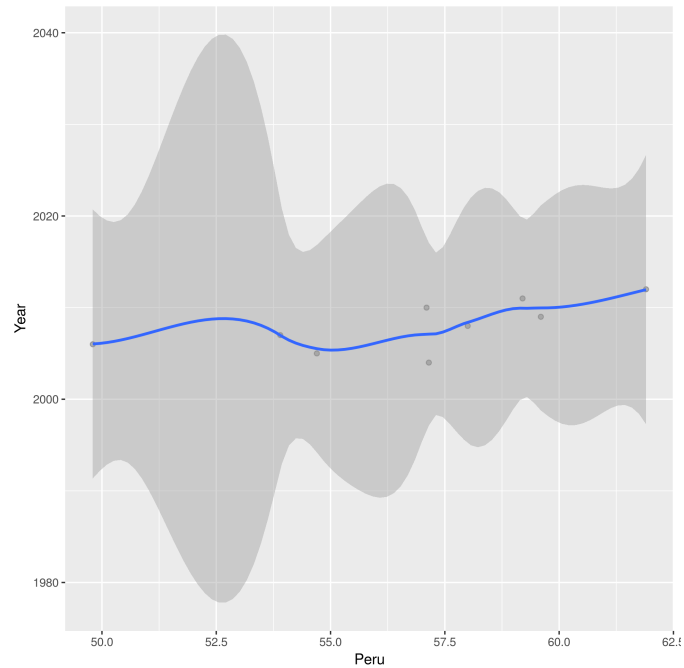Figure 11: Women in decision making about their health Peru

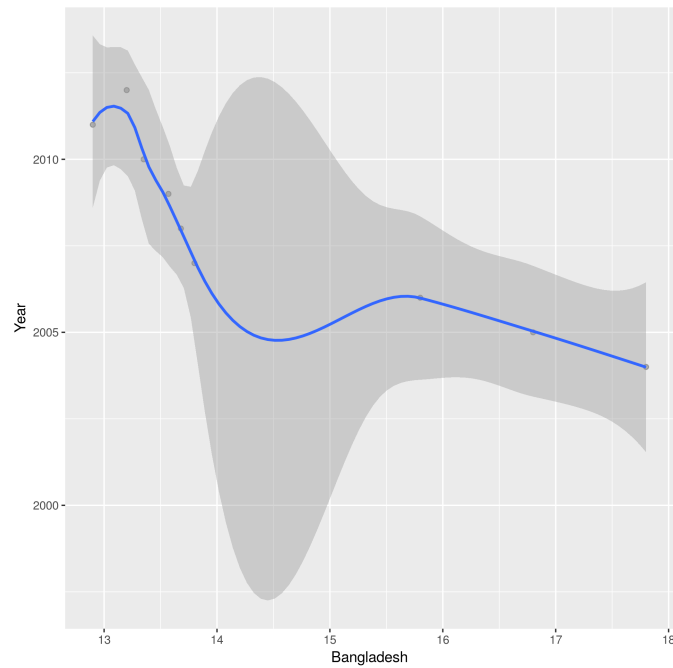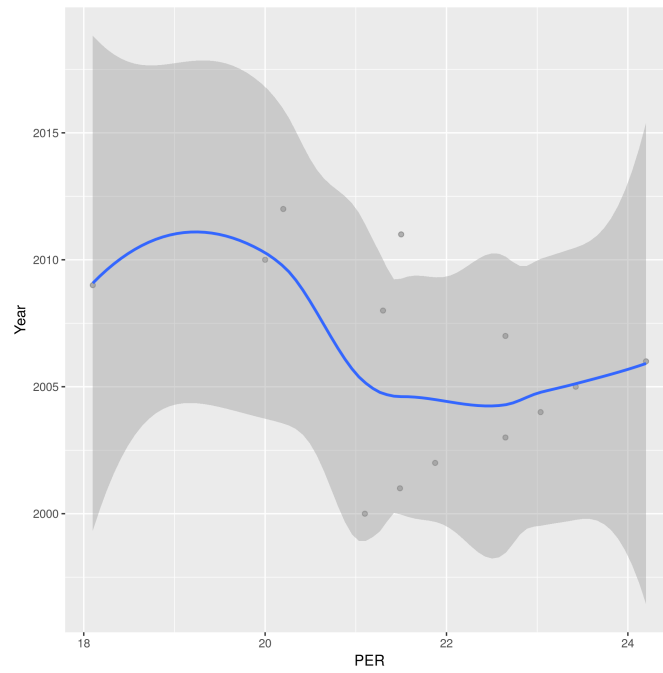Figure 12: Women in decision Making in about their health Bangladesh

Figure 13: Decision making about major household purchase

9 Psychology professor and Econ Professor

The first thing that I noted that was the biggest difference between Economics and psychology is that the data for economics is concrete, opposed to the psychology data. I mean the Data analyst or data collector has to figure out how much weight an answer would hold when asked a person for an answer. Question have to be manipulated for human error in the sense of asking question that would kind infer to the answer. They both have to make judgment in the amount of data they can collect. Usually the more data the better, but have difficulty collecting the data. In the sense of finding the right sample size to collect their data.