Part III of Final Project

Morgan Raska

04 March 21

*Introduction into my Analysis*

I had selected data provided from the bureau of economic analysis. I downloaded data regarding personal consumption expenditures by state of government and private employees. After the first week there was a discussion with my classmates regarding the difference in inflation and prices among states. Because of this I narrowed my data sets down to regions, regional prices and income tend to be consistent with one another. I chose the plains region, so I omitted the columns of any state that were not in this region. States of this region include Nebraska, Iowa, Kansas, North Dakota, South Dakota, Missouri, and Minnesota. I continued to manipulate the data by binding two different data sets with information regarding PCE. One data set reflected consumption of services, another reflected consumption of durable and non-durable goods. This was done using the cbind function, this function could only be accomplished if the rows have identical names. I had to adjust the excel document because of varying label names. I did this by omitting data that was not essential to the project like titles regarding regions. I now have a data set with the measured consumption of services and goods. I then renamed my columns to make the data set easier to understand and visually appealing. Using the select function I ordered the data set.

I then cleaned the second data set that has the total consumption in millions, omitting the states that did not pertain to the project. After omitting the states, I organized the data in chronological order in terms of years 2015-2018 using the select function. I renamed the columns using the rename function to make the data set easier to understand and more visually appealing.

Since the data was large even after the clean I decided to narrow down my analysis to just finance, insurance, and health care. this makes it easier to dive into the data given without being overwhelmed with the findings. I was inspired by the data itself. I wanted to know, examining the correlation, and the linear regression models, how do these two variables affect one another. Not only that but I wanted to see how the plains viewed these two seemingly important services by examining the CPE.

*How I addressed the Consumer Price Expenditures*

I narrowed down the territory to just the plains. This would consist of the states Nebraska, North Dakota, South Dakota, Minnesota, Missouri, Kansas, and Iowa. This helped me not only make the data more manageable but if I were to use this data to analyze a different region I would have a possible baseline to compare it to. I used correlation analysis to examine the relationships between services and how much the plains spend yearly. I then used linear regression to create a predictive model of Finance and insurance against Health care that each state uses. This allowed me to see the relationship as a visual rather than just a intellectual representation. I produced a correlation matrix of every service, and goods to give me an idea of how this CPE is with other variables.

*What I found*

I first wanted to get my correlation findings. What variables are affected by Finance and Insurance? I found that Miscellaneous services that are not specifically listed (haircuts, delivery fees, etc.) have a minuscule relationship with finance and insurance. In fact it is so small that it can be considered unimportant. There is an almost identical relationship in terms of gas as well. Both variables have a - .01 correlation coefficient. I think the fun begins when we examine who has the strongest positive relationship with this variable. The top three variables with the highest coefficient is, Utilities, Auto, and Misc goods. Individuals in the plains who spend more money on finance and insurance tend to spend more money on utilities, automobiles, and misc goods. When compared to the total CPE of the plains it seems to have a correlation coefficient of .33. As the CPE increases there is only .33 percent chance that finance and insurance increase. When this is repeated with the health care variable it is found that there are no negative relationships. However, those who tend to spend more on health care have less than a .15 correlation coefficient with, utilities, automobiles, and misc goods. They have a strong positive relationship with transportation, clothing, and food. Coefficients for these variables are greater than .55. In terms of total CPE it also has a strong positive relationship, with a coefficient of .57.

*What I found Continued*

*Regression Model? Cor Plot?*

Out of sheer curiosity I wanted to construct a regression model of the two variables in comparison to total CPE. I wanted to see the magnitude of variance between the variables. I First created a generalized linear model. I found that the null deviance was 7.02 on 7 degrees of freedom and the residual deviance was 4.6 with a AIC of 26.35. I adjusted the family from gaussian to Gamma with a link of “inverse”. This is because the distribution was so randomized it was hard to interpret. This is a better model than the linear regression model I then constructed. However both are not ideal or a great fit in terms of data. I Then decided that the best use of information that was found would be constructing a cor plot. It is a great visual to get across what I am essentially interested in, relationships with the variables.

*So did I find out what I was looking for?*

I set out to discover more about the relationships of services and goods, as well as what the citizens in the plains region are more apt to consume. I wanted to do this because not only can consumers find this information valuable, but producers as well. When running a business or providing a service, it is important to know your customer. No matter how large or small the company is the customer is what drives you. If there is a way to better know how to serve a customer in order to increase sales, or overall experience a business should want to do it. I believe looking into consumer expenditures should be important to not only consumers, but those providing a service. It is easier to predict what a customer will do by using the correaltion analysis that I have provided. We can see visually and numerically what a customer who spends on one thing will likeley spend on something else.

*Limitations*

It is important to remember this is not a one size fits all analysis, there are limitations to what was stated here. The biggest one is lack of data. If I were to repeat this analysis with unlimited resources, I would compile data from over 21 years. The more data that is compiled the better the prediction of the model. It also highlights relationships. In 2008 there was the economic plummet in the housing market. How did this effect consumer spending, how did consumers act differently where did their relationships lie in terms of goods and services? In 2012 there was a myth that the world was going to end due to the lack of days following in the Myan calendar, how did consumers respond? In 2020 there was a pandemic that swept the nation, and while I had information from 2015-2018 there was not information relesased from 2020 and so I could not analyze these years. I essentially just wanted to discover possible patterns of consumers to see what makes them tick. Because of the lack of data, there was not much of a conclusion that could be drawn.

*Code, Results, Graphs*

I have compiled my code, results, and graphs below. I found that it is easier for me to read about information completely and then see the visuals separately.

install.packages("tidyr", repos="http://cran.us.r-project.org")

## Installing package into 'C:/Users/morga/Documents/R/win-library/4.0'  
## (as 'lib' is unspecified)

##   
## There is a binary version available but the source version is later:  
## binary source needs\_compilation  
## tidyr 1.1.2 1.1.3 TRUE  
##   
## Binaries will be installed  
## package 'tidyr' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'tidyr'

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:  
## \Users\morga\Documents\R\win-library\4.0\00LOCK\tidyr\libs\x64\tidyr.dll to C:  
## \Users\morga\Documents\R\win-library\4.0\tidyr\libs\x64\tidyr.dll: Permission  
## denied

## Warning: restored 'tidyr'

##   
## The downloaded binary packages are in  
## C:\Users\morga\AppData\Local\Temp\Rtmp27YVGl\downloaded\_packages

library("tidyr")

## Warning: package 'tidyr' was built under R version 4.0.4

install.packages("dplyr", repos="http://cran.us.r-project.org")

## Installing package into 'C:/Users/morga/Documents/R/win-library/4.0'  
## (as 'lib' is unspecified)

##   
## There is a binary version available but the source version is later:  
## binary source needs\_compilation  
## dplyr 1.0.4 1.0.5 TRUE  
##   
## Binaries will be installed  
## package 'dplyr' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'dplyr'

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:  
## \Users\morga\Documents\R\win-library\4.0\00LOCK\dplyr\libs\x64\dplyr.dll to C:  
## \Users\morga\Documents\R\win-library\4.0\dplyr\libs\x64\dplyr.dll: Permission  
## denied

## Warning: restored 'dplyr'

##   
## The downloaded binary packages are in  
## C:\Users\morga\AppData\Local\Temp\Rtmp27YVGl\downloaded\_packages

library(dplyr)

## Warning: package 'dplyr' was built under R version 4.0.4

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

install.packages("ggplot", repos="http://cran.us.r-project.org")

## Installing package into 'C:/Users/morga/Documents/R/win-library/4.0'  
## (as 'lib' is unspecified)

## Warning: package 'ggplot' is not available for this version of R  
##   
## A version of this package for your version of R might be available elsewhere,  
## see the ideas at  
## https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages

library(ggplot2)  
install.packages("ggfortify",repos="http://cran.us.r-project.org")

## Installing package into 'C:/Users/morga/Documents/R/win-library/4.0'  
## (as 'lib' is unspecified)

## package 'ggfortify' successfully unpacked and MD5 sums checked  
##   
## The downloaded binary packages are in  
## C:\Users\morga\AppData\Local\Temp\Rtmp27YVGl\downloaded\_packages

install.packages("factoextra",repos="http://cran.us.r-project.org")

## Installing package into 'C:/Users/morga/Documents/R/win-library/4.0'  
## (as 'lib' is unspecified)

## package 'factoextra' successfully unpacked and MD5 sums checked  
##   
## The downloaded binary packages are in  
## C:\Users\morga\AppData\Local\Temp\Rtmp27YVGl\downloaded\_packages

library(ggfortify)

## Warning: package 'ggfortify' was built under R version 4.0.4

library(factoextra)

## Warning: package 'factoextra' was built under R version 4.0.4

## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa

PCE\_1 <- read.csv("data/goodsCPE.csv")  
PCE\_2 <- read.csv("data/servicesPCE.csv")  
  
##Combine the two sets  
PCE\_final <- cbind(PCE\_1,PCE\_2)  
  
##Clean the final data set  
  
drop\_na(PCE\_final)

## X Motor.vehicles.and.parts  
## 1 United States1 -0.3  
## 2 Iowa -0.4  
## 3 Kansas 0.0  
## 4 Minnesota 1.3  
## 5 Missouri -0.1  
## 6 Nebraska -0.6  
## 7 North Dakota -0.1  
## 8 South Dakota 0.1  
## Furnishings.and.durable.household.equipment Recreational.goods.and.vehicles  
## 1 4.1 8.6  
## 2 0.7 5.3  
## 3 0.9 6.0  
## 4 3.8 8.6  
## 5 1.4 6.7  
## 6 5.9 6.2  
## 7 2.0 4.6  
## 8 4.5 3.7  
## Other.durable.goods Off.premises.food.and.beverages Clothing.and.footwear  
## 1 2.6 2.7 2.3  
## 2 1.1 0.0 0.1  
## 3 1.4 0.3 0.1  
## 4 5.0 1.3 0.2  
## 5 1.1 0.1 1.4  
## 6 1.9 0.3 0.3  
## 7 -0.7 0.1 -0.1  
## 8 -0.5 0.3 1.1  
## Gasoline.and.other.energy.goods Other.nondurable.goods Total.CPE  
## 1 -4.0 5.7 4.3  
## 2 -5.0 3.4 3.7  
## 3 -5.0 5.1 2.9  
## 4 -4.3 4.0 4.7  
## 5 -3.6 4.8 3.9  
## 6 -4.6 3.5 4.5  
## 7 -5.9 2.8 5.7  
## 8 -5.9 4.2 5.9  
## Housing.and.utilities Health.care Transportation.services Recreation.services  
## 1 4.5 3.6 3.3 4.0  
## 2 2.7 1.4 1.9 1.4  
## 3 4.7 0.2 1.0 2.1  
## 4 3.3 2.7 1.9 2.3  
## 5 5.3 2.7 1.4 2.9  
## 6 2.8 2.1 2.3 2.2  
## 7 3.2 2.8 5.1 1.6  
## 8 7.1 2.2 2.7 1.6  
## Food.services.and.accommodations Financial.services.and.insurance  
## 1 5.0 5.5  
## 2 3.0 3.5  
## 3 3.3 4.9  
## 4 2.4 5.2  
## 5 4.5 5.1  
## 6 3.9 3.6  
## 7 4.3 5.2  
## 8 4.0 5.5  
## Other.services  
## 1 0.1  
## 2 0.6  
## 3 -0.3  
## 4 1.1  
## 5 0.3  
## 6 1.0  
## 7 4.1  
## 8 4.7

dim(PCE\_final)

## [1] 8 17

summary(PCE\_final)

## X Motor.vehicles.and.parts  
## Length:8 Min. :-0.6000   
## Class :character 1st Qu.:-0.3250   
## Mode :character Median :-0.1000   
## Mean :-0.0125   
## 3rd Qu.: 0.0250   
## Max. : 1.3000   
## Furnishings.and.durable.household.equipment Recreational.goods.and.vehicles  
## Min. :0.700 Min. :3.700   
## 1st Qu.:1.275 1st Qu.:5.125   
## Median :2.900 Median :6.100   
## Mean :2.913 Mean :6.213   
## 3rd Qu.:4.200 3rd Qu.:7.175   
## Max. :5.900 Max. :8.600   
## Other.durable.goods Off.premises.food.and.beverages Clothing.and.footwear  
## Min. :-0.700 Min. :0.0000 Min. :-0.100   
## 1st Qu.: 0.700 1st Qu.:0.1000 1st Qu.: 0.100   
## Median : 1.250 Median :0.3000 Median : 0.250   
## Mean : 1.488 Mean :0.6375 Mean : 0.675   
## 3rd Qu.: 2.075 3rd Qu.:0.5500 3rd Qu.: 1.175   
## Max. : 5.000 Max. :2.7000 Max. : 2.300   
## Gasoline.and.other.energy.goods Other.nondurable.goods Total.CPE   
## Min. :-5.900 Min. :2.800 Min. :2.90   
## 1st Qu.:-5.225 1st Qu.:3.475 1st Qu.:3.85   
## Median :-4.800 Median :4.100 Median :4.40   
## Mean :-4.787 Mean :4.188 Mean :4.45   
## 3rd Qu.:-4.225 3rd Qu.:4.875 3rd Qu.:4.95   
## Max. :-3.600 Max. :5.700 Max. :5.90   
## Housing.and.utilities Health.care Transportation.services  
## Min. :2.70 Min. :0.200 Min. :1.000   
## 1st Qu.:3.10 1st Qu.:1.925 1st Qu.:1.775   
## Median :3.90 Median :2.450 Median :2.100   
## Mean :4.20 Mean :2.212 Mean :2.450   
## 3rd Qu.:4.85 3rd Qu.:2.725 3rd Qu.:2.850   
## Max. :7.10 Max. :3.600 Max. :5.100   
## Recreation.services Food.services.and.accommodations  
## Min. :1.400 Min. :2.400   
## 1st Qu.:1.600 1st Qu.:3.225   
## Median :2.150 Median :3.950   
## Mean :2.263 Mean :3.800   
## 3rd Qu.:2.450 3rd Qu.:4.350   
## Max. :4.000 Max. :5.000   
## Financial.services.and.insurance Other.services   
## Min. :3.500 Min. :-0.30   
## 1st Qu.:4.575 1st Qu.: 0.25   
## Median :5.150 Median : 0.80   
## Mean :4.812 Mean : 1.45   
## 3rd Qu.:5.275 3rd Qu.: 1.85   
## Max. :5.500 Max. : 4.70

PCE\_rename1<- rename(PCE\_final, PCE = X)  
PCE\_rename2 <- rename(PCE\_rename1, c(Auto = Motor.vehicles.and.parts, House = Furnishings.and.durable.household.equipment,   
 Recreation = Recreational.goods.and.vehicles,   
 Misc = Other.durable.goods, Food = Off.premises.food.and.beverages,  
 Clothing = Clothing.and.footwear, Gas = Gasoline.and.other.energy.goods,   
 Misc2 = Other.nondurable.goods, Utilities = Housing.and.utilities,  
 Transportation = Transportation.services, Activities = Recreation.services,   
 Dining = Food.services.and.accommodations, Finance.Insurance = Financial.services.and.insurance))  
PCE\_final\_rename <- select(PCE\_rename2, Total.CPE, Auto, House, Recreation, Misc, Misc2, Food, Clothing, Gas, Utilities, Transportation, Activities,  
 Dining, Finance.Insurance, Health.care, Other.services)  
PCE\_final\_rename

## Total.CPE Auto House Recreation Misc Misc2 Food Clothing Gas Utilities  
## 1 4.3 -0.3 4.1 8.6 2.6 5.7 2.7 2.3 -4.0 4.5  
## 2 3.7 -0.4 0.7 5.3 1.1 3.4 0.0 0.1 -5.0 2.7  
## 3 2.9 0.0 0.9 6.0 1.4 5.1 0.3 0.1 -5.0 4.7  
## 4 4.7 1.3 3.8 8.6 5.0 4.0 1.3 0.2 -4.3 3.3  
## 5 3.9 -0.1 1.4 6.7 1.1 4.8 0.1 1.4 -3.6 5.3  
## 6 4.5 -0.6 5.9 6.2 1.9 3.5 0.3 0.3 -4.6 2.8  
## 7 5.7 -0.1 2.0 4.6 -0.7 2.8 0.1 -0.1 -5.9 3.2  
## 8 5.9 0.1 4.5 3.7 -0.5 4.2 0.3 1.1 -5.9 7.1  
## Transportation Activities Dining Finance.Insurance Health.care Other.services  
## 1 3.3 4.0 5.0 5.5 3.6 0.1  
## 2 1.9 1.4 3.0 3.5 1.4 0.6  
## 3 1.0 2.1 3.3 4.9 0.2 -0.3  
## 4 1.9 2.3 2.4 5.2 2.7 1.1  
## 5 1.4 2.9 4.5 5.1 2.7 0.3  
## 6 2.3 2.2 3.9 3.6 2.1 1.0  
## 7 5.1 1.6 4.3 5.2 2.8 4.1  
## 8 2.7 1.6 4.0 5.5 2.2 4.7

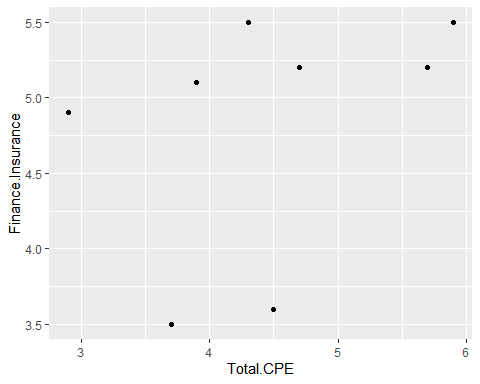
PCE\_total<- read.csv("data/millionstotalCPE.csv")  
PCE\_total\_rename <- rename(PCE\_total, c("2015" = X2015, "2016" =X2016, "2017"=X2017, "2018" =X2018))  
drop\_na(PCE\_total\_rename)

## [1] TOTAL\_STATE 2015 2016 2017 2018 X2015.1   
## [7] X2016.1 X2017.1 X2018.1 X2015.2 X2016.2 X2017.2   
## [13] X2018.2   
## <0 rows> (or 0-length row.names)

PCE\_total\_final <- PCE\_total\_rename %>% select(-(X2015.1:X2018.2))  
  
correlation\_PCE\_services <- cor(PCE\_final\_rename, use = "pairwise.complete.obs")  
correlation\_PCE\_services

## Total.CPE Auto House Recreation Misc  
## Total.CPE 1.00000000 0.17683819 0.51845621 -0.38116643 -0.33562081  
## Auto 0.17683819 1.00000000 0.03764125 0.37106380 0.57636597  
## House 0.51845621 0.03764125 1.00000000 0.17505351 0.25330764  
## Recreation -0.38116643 0.37106380 0.17505351 1.00000000 0.89333835  
## Misc -0.33562081 0.57636597 0.25330764 0.89333835 1.00000000  
## Misc2 -0.46974447 0.02784765 -0.01223472 0.52348183 0.30604618  
## Food -0.00384608 0.20667411 0.39280956 0.75869412 0.58179876  
## Clothing 0.03215056 -0.17701052 0.27433092 0.38086737 0.07771614  
## Gas -0.53966229 0.07461668 0.02491527 0.82715196 0.65992399  
## Utilities 0.22282046 0.07540637 0.07597870 -0.28587019 -0.36000464  
## Transportation 0.73455282 -0.15269654 0.21822273 -0.24598924 -0.40492823  
## Activities -0.25037688 -0.04470697 0.23924336 0.77679997 0.45604847  
## Dining 0.23108868 -0.58729740 0.18025814 -0.05846549 -0.46005853  
## Finance.Insurance 0.37333127 0.47767354 0.05270985 0.15721491 -0.01956400  
## Health.care 0.56415643 0.12019018 0.45455865 0.39275775 0.17255975  
## Other.services 0.91630703 0.09663848 0.26216192 -0.68938124 -0.59872364  
## Misc2 Food Clothing Gas Utilities  
## Total.CPE -0.46974447 -0.00384608 0.03215056 -0.53966229 0.22282046  
## Auto 0.02784765 0.20667411 -0.17701052 0.07461668 0.07540637  
## House -0.01223472 0.39280956 0.27433092 0.02491527 0.07597870  
## Recreation 0.52348183 0.75869412 0.38086737 0.82715196 -0.28587019  
## Misc 0.30604618 0.58179876 0.07771614 0.65992399 -0.36000464  
## Misc2 1.00000000 0.61064869 0.74589708 0.56693288 0.52004873  
## Food 0.61064869 1.00000000 0.65925988 0.44820186 0.02549168  
## Clothing 0.74589708 0.65925988 1.00000000 0.50237847 0.54058266  
## Gas 0.56693288 0.44820186 0.50237847 1.00000000 -0.12017351  
## Utilities 0.52004873 0.02549168 0.54058266 -0.12017351 1.00000000  
## Transportation -0.42550349 0.15949667 0.01839734 -0.48817116 -0.14811072  
## Activities 0.77963769 0.80871851 0.81814626 0.76895750 0.13653321  
## Dining 0.33288189 0.26765583 0.70217653 0.12265613 0.36404377  
## Finance.Insurance 0.47030933 0.43982189 0.48297386 -0.01943278 0.64534921  
## Health.care 0.05308396 0.57366770 0.58771354 0.30658118 0.02379913  
## Other.services -0.54374023 -0.29728669 -0.16821676 -0.79065142 0.32373938  
## Transportation Activities Dining Finance.Insurance  
## Total.CPE 0.73455282 -0.25037688 0.23108868 0.37333127  
## Auto -0.15269654 -0.04470697 -0.58729740 0.47767354  
## House 0.21822273 0.23924336 0.18025814 0.05270985  
## Recreation -0.24598924 0.77679997 -0.05846549 0.15721491  
## Misc -0.40492823 0.45604847 -0.46005853 -0.01956400  
## Misc2 -0.42550349 0.77963769 0.33288189 0.47030933  
## Food 0.15949667 0.80871851 0.26765583 0.43982189  
## Clothing 0.01839734 0.81814626 0.70217653 0.48297386  
## Gas -0.48817116 0.76895750 0.12265613 -0.01943278  
## Utilities -0.14811072 0.13653321 0.36404377 0.64534921  
## Transportation 1.00000000 -0.05397507 0.46555895 0.28166160  
## Activities -0.05397507 1.00000000 0.54731144 0.39070058  
## Dining 0.46555895 0.54731144 1.00000000 0.35442789  
## Finance.Insurance 0.28166160 0.39070058 0.35442789 1.00000000  
## Health.care 0.57218171 0.55103282 0.52925112 0.43480323  
## Other.services 0.65631436 -0.53018514 0.13798251 0.31820633  
## Health.care Other.services  
## Total.CPE 0.56415643 0.91630703  
## Auto 0.12019018 0.09663848  
## House 0.45455865 0.26216192  
## Recreation 0.39275775 -0.68938124  
## Misc 0.17255975 -0.59872364  
## Misc2 0.05308396 -0.54374023  
## Food 0.57366770 -0.29728669  
## Clothing 0.58771354 -0.16821676  
## Gas 0.30658118 -0.79065142  
## Utilities 0.02379913 0.32373938  
## Transportation 0.57218171 0.65631436  
## Activities 0.55103282 -0.53018514  
## Dining 0.52925112 0.13798251  
## Finance.Insurance 0.43480323 0.31820633  
## Health.care 1.00000000 0.23310091  
## Other.services 0.23310091 1.00000000

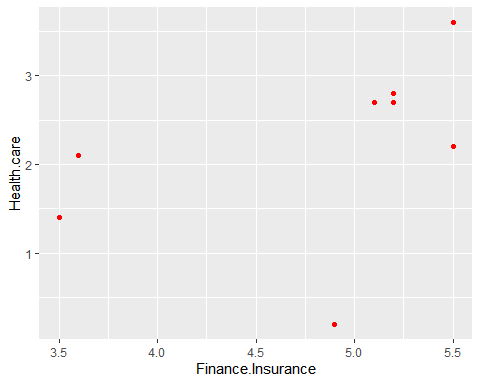
PCE\_plot <- ggplot(PCE\_final\_rename, aes(x=Total.CPE, y=Finance.Insurance)) + geom\_point()  
PCE\_plot



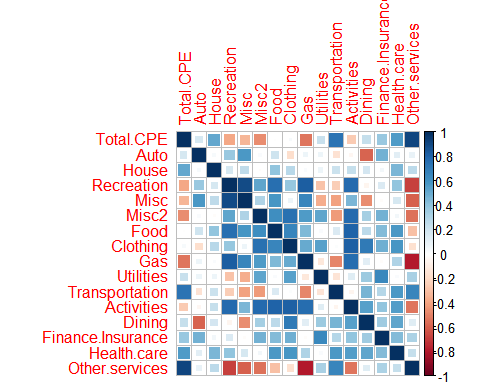
regression\_model<- glm(Total.CPE~Finance.Insurance+Health.care, family = Gamma(link = 'inverse'), data =PCE\_final\_rename)  
regression\_model

##   
## Call: glm(formula = Total.CPE ~ Finance.Insurance + Health.care, family = Gamma(link = "inverse"),   
## data = PCE\_final\_rename)  
##   
## Coefficients:  
## (Intercept) Finance.Insurance Health.care   
## 0.329275 -0.008795 -0.026517   
##   
## Degrees of Freedom: 7 Total (i.e. Null); 5 Residual  
## Null Deviance: 0.3671   
## Residual Deviance: 0.2426 AIC: 26.3

graph\_1 <- ggplot(regression\_model, aes(Finance.Insurance,Health.care), colour= 'red')  
graph\_1 + geom\_point(colour='red')



corrplot::corrplot(correlation\_PCE\_services, method = "square")



PCE\_lm\_Finance <- lm(Finance.Insurance~Health.care, data = PCE\_final\_rename)  
summary(PCE\_lm\_Finance)

##   
## Call:  
## lm(formula = Finance.Insurance ~ Health.care, data = PCE\_final\_rename)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.1744 -0.1675 0.2031 0.3398 0.7689   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.0634 0.6912 5.879 0.00107 \*\*  
## Health.care 0.3386 0.2863 1.183 0.28167   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7824 on 6 degrees of freedom  
## Multiple R-squared: 0.1891, Adjusted R-squared: 0.0539   
## F-statistic: 1.399 on 1 and 6 DF, p-value: 0.2817

PCE\_predict\_df <- data.frame(Finance.Insurance = predict(PCE\_lm\_Finance, PCE\_final\_rename), Health=PCE\_final\_rename$Health.care)  
ggplot(data = PCE\_predict\_df, aes(y = Finance.Insurance, x = PCE\_final\_rename$Health.care)) +  
 geom\_line(color='blue') +  
 geom\_point(color='red',data = PCE\_lm\_Finance, aes(y=Finance.Insurance, x= PCE\_final\_rename$Health.care)) +  
 labs(x = 'Health Care CPE by Percentage',  
 y = 'Finance and Insurance CPE by Percentage ',  
 title = "Consumer Price Expenditures 2015-2018") +  
 theme\_bw()+  
 theme(axis.text.x = element\_text(angle = 90, vjust = 0.5, hjust = 1))

