seminarscript\_ver\_2.R

urizafrir

2022-07-16

library(dplyr)

##   
## 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

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ ggplot2 3.3.5 ✓ purrr 0.3.4  
## ✓ tibble 3.1.6 ✓ stringr 1.4.0  
## ✓ tidyr 1.1.4 ✓ forcats 0.5.1  
## ✓ readr 2.0.2

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

library(sandwich)  
  
# General Settings and Data Read ------------------------------------------  
  
options(scipen=999) # tell R not to use Scientific notation  
options(digits = 5) # controls how many digits are printed by default  
rm(list=ls()) # del all objects and functions  
df1 <- read.csv("H20181362Data.csv") # load data  
  
# Cleaning up variables, ordering them ---------------------------------------------------  
  
#subset MerutzeChaim, minn, MatzavMishp, DatiutYehudi, DatiutLoYehudi, TeudaGvoha, ShnotlimudKlali\_C, hachnasa\_lenefesh, Dat, mispYeladim, Gil  
df<-df1 %>% select(MerutzeChaim, Minn, MatzavMishp, DatiutYehudi, DatiutLoYehudi, TeudaGvoha, ShnotlimudKlali\_C, hachnasa\_lenefesh, Dat, MispYeladim, Gil, HachnasaAvoda)  
  
  
# clean merutzechaim values of 888888  
table(df$MerutzeChaim)

##   
## 1 2 3 4 888888   
## 2836 3782 667 126 37

df$MerutzeChaim[df$MerutzeChaim==888888] <- NA  
table(df$MerutzeChaim)

##   
## 1 2 3 4   
## 2836 3782 667 126

# recode MerutzeChaim - to go up from 1 to 4  
table(df$MerutzeChaim)

##   
## 1 2 3 4   
## 2836 3782 667 126

df$MerutzeChaim[df$MerutzeChaim==4] <- 5  
df$MerutzeChaim[df$MerutzeChaim==1] <- 4  
df$MerutzeChaim[df$MerutzeChaim==5] <- 1  
df$MerutzeChaim[df$MerutzeChaim==2] <- 6  
df$MerutzeChaim[df$MerutzeChaim==3] <- 2  
df$MerutzeChaim[df$MerutzeChaim==6] <- 3  
table(df$MerutzeChaim)

##   
## 1 2 3 4   
## 126 667 3782 2836

#create log\_MerutzeChaim  
df$log\_MerutzeChaim<-log(df$MerutzeChaim)  
table(df$log\_MerutzeChaim)

##   
## 0 0.693147180559945 1.09861228866811 1.38629436111989   
## 126 667 3782 2836

#create MerutzeChaim\_percentage  
df$MerutzeChaim\_percentage<-df$MerutzeChaim  
df$MerutzeChaim\_percentage[df$MerutzeChaim\_percentage==1] <- 0  
df$MerutzeChaim\_percentage[df$MerutzeChaim\_percentage==2] <- 33.3  
df$MerutzeChaim\_percentage[df$MerutzeChaim\_percentage==3] <- 66.6  
df$MerutzeChaim\_percentage[df$MerutzeChaim\_percentage==4] <- 100  
table(df$MerutzeChaim\_percentage)

##   
## 0 33.3 66.6 100   
## 126 667 3782 2836

#create MerutzeChaim\_zscore  
df$MerutzeChaim\_zscores<-(df$MerutzeChaim-mean(df$MerutzeChaim,na.rm=TRUE))/sd(df$MerutzeChaim, na.rm=TRUE)  
table(df$MerutzeChaim\_zscores)

##   
## -3.281242405078 -1.82851001580546 -0.37577762653292 1.07695476273962   
## 126 667 3782 2836

sd(df$MerutzeChaim, na.rm=TRUE)

## [1] 0.68836

#create MerutzeChaim\_dummy  
df$MerutzeChaim\_dummy<-ifelse(df$MerutzeChaim==3 | df$MerutzeChaim==4,1,0)  
table(df$MerutzeChaim)

##   
## 1 2 3 4   
## 126 667 3782 2836

table(df$MerutzeChaim\_dummy)

##   
## 0 1   
## 793 6618

# recode minn , rename minn to female  
table(df$Minn)

##   
## 1 2   
## 3682 3766

df$Minn[df$Minn==1] <- 0  
df$Minn[df$Minn==2] <- 1  
table(df$Minn)

##   
## 0 1   
## 3682 3766

names(df)[names(df) == 'Minn'] <- "Female"  
table(df$Female)

##   
## 0 1   
## 3682 3766

#recode matzav mishpachti  
df$Married<-ifelse(df$MatzavMishp==1,1,0)  
#check  
table(df$Married)

##   
## 0 1   
## 2695 4753

table(df$MatzavMishp)

##   
## 1 2 3 4 5   
## 4753 42 580 408 1665

#recode datiyehudi, leave 888888, 999999 as is  
table(df$DatiutYehudi)

##   
## 1 2 3 4 5 888888 999999   
## 563 640 744 1289 2579 23 1610

df$DatiutYehudi[df$DatiutYehudi==5] <- 6  
df$DatiutYehudi[df$DatiutYehudi==1] <- 5  
df$DatiutYehudi[df$DatiutYehudi==6] <- 1  
df$DatiutYehudi[df$DatiutYehudi==4] <- 6  
df$DatiutYehudi[df$DatiutYehudi==2] <- 4  
df$DatiutYehudi[df$DatiutYehudi==6] <- 2  
table(df$DatiutYehudi)

##   
## 1 2 3 4 5 888888 999999   
## 2579 1289 744 640 563 23 1610

#recode datiutloyehudi leave 888888, 999999 as is  
table(df$DatiutLoYehudi)

##   
## 1 2 3 4 888888 999999   
## 43 790 348 213 3 6051

df$DatiutLoYehudi[df$DatiutLoYehudi==4] <- 5  
df$DatiutLoYehudi[df$DatiutLoYehudi==1] <- 4  
df$DatiutLoYehudi[df$DatiutLoYehudi==5] <- 1  
df$DatiutLoYehudi[df$DatiutLoYehudi==2] <- 6  
df$DatiutLoYehudi[df$DatiutLoYehudi==3] <- 2  
df$DatiutLoYehudi[df$DatiutLoYehudi==6] <- 3  
table(df$DatiutLoYehudi)

##   
## 1 2 3 4 888888 999999   
## 213 348 790 43 3 6051

#clean TeudaGvoha  
table(df$TeudaGvoha)

##   
## 0 1 2 3 4 5 6 7 888888   
## 85 1182 1485 1195 1428 831 91 1138 13

df$TeudaGvoha[df$TeudaGvoha==888888] <- NA  
df$TeudaGvoha[df$TeudaGvoha==7] <- 0  
table(df$TeudaGvoha)

##   
## 0 1 2 3 4 5 6   
## 1223 1182 1485 1195 1428 831 91

#create teudagvoha\_factor, rename teudagvoha\_factor  
df$TeudaGvoha\_factor<-factor(df$TeudaGvoha)  
table(df$TeudaGvoha\_factor)

##   
## 0 1 2 3 4 5 6   
## 1223 1182 1485 1195 1428 831 91

#clean ShnotlimudKlali\_C  
table(df$ShnotlimudKlali\_C)

##   
## 1 2 3 4 5 6 999999   
## 93 356 375 1610 2126 2796 92

df$ShnotlimudKlali\_C[df$ShnotlimudKlali\_C==999999] <- NA  
table(df$ShnotlimudKlali\_C)

##   
## 1 2 3 4 5 6   
## 93 356 375 1610 2126 2796

#turn ShnotlimudKlali\_C to discrete years  
df$ShnotLimud\_Discrete<-df$ShnotlimudKlali\_C  
table(df$ShnotLimud\_Discrete)

##   
## 1 2 3 4 5 6   
## 93 356 375 1610 2126 2796

df$ShnotLimud\_Discrete[df$ShnotLimud\_Discrete==6] <- 18  
df$ShnotLimud\_Discrete[df$ShnotLimud\_Discrete==5] <- 15  
df$ShnotLimud\_Discrete[df$ShnotLimud\_Discrete==4] <- 12  
df$ShnotLimud\_Discrete[df$ShnotLimud\_Discrete==3] <- 9  
df$ShnotLimud\_Discrete[df$ShnotLimud\_Discrete==2] <- 6  
df$ShnotLimud\_Discrete[df$ShnotLimud\_Discrete==1] <- 3  
table(df$ShnotlimudKlali\_C)

##   
## 1 2 3 4 5 6   
## 93 356 375 1610 2126 2796

table(df$ShnotLimud\_Discrete)

##   
## 3 6 9 12 15 18   
## 93 356 375 1610 2126 2796

#add factor to ShnotlimudKlali\_C, ShnotLimud\_Discrete  
df$ShnotlimudKlali\_C\_factor <- factor(df$ShnotlimudKlali\_C)  
df$ShnotlimudKlali\_C\_discrete\_factor <- factor(df$ShnotLimud\_Discrete)  
  
  
table(df$ShnotlimudKlali\_C\_factor)

##   
## 1 2 3 4 5 6   
## 93 356 375 1610 2126 2796

table(df$ShnotlimudKlali\_C\_discrete\_factor)

##   
## 3 6 9 12 15 18   
## 93 356 375 1610 2126 2796

table(df$ShnotlimudKlali\_C)

##   
## 1 2 3 4 5 6   
## 93 356 375 1610 2126 2796

#create dummy variable - finished\_ba\_or\_more  
table(df$ShnotlimudKlali\_C)

##   
## 1 2 3 4 5 6   
## 93 356 375 1610 2126 2796

df$finished\_ba\_or\_more<-ifelse(df$ShnotlimudKlali\_C==6,1,0)  
table(df$finished\_ba\_or\_more)

##   
## 0 1   
## 4560 2796

#clean hachnasa\_lenefesh  
table(df$hachnasa\_lenefesh)

##   
## 1 2 3 888888   
## 1159 1732 3075 1482

df$hachnasa\_lenefesh[df$hachnasa\_lenefesh==888888] <- NA  
table(df$hachnasa\_lenefesh)

##   
## 1 2 3   
## 1159 1732 3075

#clean mispYeladim  
table(df$MispYeladim)

##   
## 0 1 2 3 4 5 6 7 888888   
## 1937 750 1510 1543 797 372 192 336 11

df$MispYeladim[df$MispYeladim==888888] <- NA  
table(df$MispYeladim)

##   
## 0 1 2 3 4 5 6 7   
## 1937 750 1510 1543 797 372 192 336

#mispYeladim squared  
df$MispYeladim\_sq<-df$MispYeladim^2  
table(df$MispYeladim\_sq)

##   
## 0 1 4 9 16 25 36 49   
## 1937 750 1510 1543 797 372 192 336

#HachnasaAvoda  
table(df$HachnasaAvoda)

##   
## 1 2 3 4 5 6 7 8 9 10 11   
## 323 193 238 289 445 586 844 701 474 471 98   
## 888888 999999   
## 436 2350

df$HachnasaAvoda[df$HachnasaAvoda==888888] <- NA  
df$HachnasaAvoda[df$HachnasaAvoda==999999] <- NA  
df$HachnasaAvoda[df$HachnasaAvoda==11] <- 0  
table(df$HachnasaAvoda)

##   
## 0 1 2 3 4 5 6 7 8 9 10   
## 98 323 193 238 289 445 586 844 701 474 471

#HachnasaAvoda squared  
df$HachnasaAvoda\_sq<-df$HachnasaAvoda^2  
table(df$HachnasaAvoda)

##   
## 0 1 2 3 4 5 6 7 8 9 10   
## 98 323 193 238 289 445 586 844 701 474 471

table(df$HachnasaAvoda\_sq)

##   
## 0 1 4 9 16 25 36 49 64 81 100   
## 98 323 193 238 289 445 586 844 701 474 471

#clean Dat  
table(df$Dat)

##   
## 1 2 3 4 5 6 888888   
## 5838 1041 233 106 17 199 14

df$Dat[df$Dat==888888] <-NA  
table(df$Dat)

##   
## 1 2 3 4 5 6   
## 5838 1041 233 106 17 199

#check all variables MerutzeChaim, female, MatzavMishp, DatiutYehudi, DatiutLoYehudi, TeudaGvoha, ShnotlimudKlali\_C, hachnasa\_lenefesh, Dat, mispYeladim  
table(df$Gil)

##   
## 1 2 3 4 5 6 7 8 9 10 11   
## 829 740 740 732 710 665 531 551 501 839 610

table(df$Female)

##   
## 0 1   
## 3682 3766

table(df$TeudaGvoha)

##   
## 0 1 2 3 4 5 6   
## 1223 1182 1485 1195 1428 831 91

table(df$MispYeladim)

##   
## 0 1 2 3 4 5 6 7   
## 1937 750 1510 1543 797 372 192 336

table(df$hachnasa\_lenefesh)

##   
## 1 2 3   
## 1159 1732 3075

table(df$Married)

##   
## 0 1   
## 2695 4753

table(df$MerutzeChaim)

##   
## 1 2 3 4   
## 126 667 3782 2836

table(df$MispYeladim)

##   
## 0 1 2 3 4 5 6 7   
## 1937 750 1510 1543 797 372 192 336

table(df$MatzavMishp)

##   
## 1 2 3 4 5   
## 4753 42 580 408 1665

table(df$DatiutYehudi)

##   
## 1 2 3 4 5 888888 999999   
## 2579 1289 744 640 563 23 1610

table(df$DatiutLoYehudi)

##   
## 1 2 3 4 888888 999999   
## 213 348 790 43 3 6051

table(df$ShnotlimudKlali\_C)

##   
## 1 2 3 4 5 6   
## 93 356 375 1610 2126 2796

#divide population to jews and arabs. create dummy variable - jewish=1, non jewish=0  
df$jewish<-ifelse(df$Dat==1,1,0)  
table(df$jewish)

##   
## 0 1   
## 1596 5838

#create age squared  
df$Gil\_sq<-df$Gil^2  
  
  
  
# Create database of Jews and non Jews and Clean ---------------------------------------------  
  
#since reilgiousness level is divided to jews and non jews, we need to subset our data accordingly.  
  
# subset all jews  
df\_jews<-subset(df, Dat==1)   
#clean df\_jews datiut yehudi  
table(df\_jews$DatiutYehudi)

##   
## 1 2 3 4 5 888888   
## 2579 1289 744 640 563 23

df\_jews$DatiutYehudi[df\_jews$DatiutYehudi==888888] <-NA  
table(df\_jews$DatiutYehudi)

##   
## 1 2 3 4 5   
## 2579 1289 744 640 563

# subset all non jews  
df\_non\_jews<-subset(df, Dat!=1)   
#clean df\_non\_jews datiut yehudi  
table(df\_non\_jews$DatiutLoYehudi)

##   
## 1 2 3 4 888888 999999   
## 213 348 790 43 3 199

df\_non\_jews$DatiutLoYehudi[df\_non\_jews$DatiutLoYehudi==888888] <-NA  
df\_non\_jews$DatiutLoYehudi[df\_non\_jews$DatiutLoYehudi==999999] <-NA  
table(df\_non\_jews$DatiutLoYehudi)

##   
## 1 2 3 4   
## 213 348 790 43

df\_arab\_muslim\_and\_christian<-subset(df, Dat==2 | Dat==3)   
  
  
  
  
#create dummy for religiousness level  
df\_jews$DatiutYehudi\_dummy<-ifelse(df\_jews$DatiutYehudi==3 | df\_jews$DatiutYehudi==4 | df\_jews$DatiutYehudi==5,1,0)  
table(df\_jews$DatiutYehudi)

##   
## 1 2 3 4 5   
## 2579 1289 744 640 563

table(df\_jews$DatiutYehudi\_dummy)

##   
## 0 1   
## 3868 1947

#create dummy for religiousness level  
table(df\_non\_jews$DatiutLoYehudi)

##   
## 1 2 3 4   
## 213 348 790 43

df\_non\_jews$DatiutLoYehudi\_dummy<-ifelse(df\_non\_jews$DatiutLoYehudi==3 | df\_non\_jews$DatiutLoYehudi==4,1,0)  
table(df\_non\_jews$DatiutLoYehudi\_dummy)

##   
## 0 1   
## 561 833

# Regressions 1 -----------------------------------------------------------  
#possible options:  
  
# df$ShnotLimud\_Discrete  
# df$ShnotlimudKlali\_C  
# table(df$ShnotlimudKlali\_C\_discrete)  
# table(df$ShnotlimudKlali\_C\_factor)  
# df$ShnotlimudKlali\_C\_discrete\_factor  
# df$MerutzeChaim  
# df$MerutzeChaim\_dummy  
# df$MerutzeChaim\_zscores  
# df$MerutzeChaim\_percentage  
# table(df1$TeudaGvoha)  
  
# basic first regression life satisfaction and schooling years  
model1 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete, data=df)   
summary(model1)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.446 -0.541 -0.265 0.912 1.602   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.66336 0.04825 -13.8 <0.0000000000000002 \*\*\*  
## ShnotLimud\_Discrete 0.04602 0.00321 14.3 <0.0000000000000002 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.982 on 7324 degrees of freedom  
## (122 observations deleted due to missingness)  
## Multiple R-squared: 0.0273, Adjusted R-squared: 0.0271   
## F-statistic: 205 on 1 and 7324 DF, p-value: <0.0000000000000002

lmtest::bptest(model1)

##   
## studentized Breusch-Pagan test  
##   
## data: model1  
## BP = 57.1, df = 1, p-value = 0.000000000000042

# basic first regression life satisfaction and and schooling years divided to categories  
model1 <-lm(MerutzeChaim\_zscores ~ ShnotlimudKlali\_C\_discrete\_factor, data=df)   
summary(model1)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotlimudKlali\_C\_discrete\_factor,   
## data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.427 -0.521 -0.346 0.932 1.580   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.5035 0.1028 -4.90 0.00000098756  
## ShnotlimudKlali\_C\_discrete\_factor6 0.0581 0.1152 0.50 0.61  
## ShnotlimudKlali\_C\_discrete\_factor9 0.0768 0.1147 0.67 0.50  
## ShnotlimudKlali\_C\_discrete\_factor12 0.4739 0.1057 4.49 0.00000738215  
## ShnotlimudKlali\_C\_discrete\_factor15 0.5341 0.1050 5.09 0.00000037146  
## ShnotlimudKlali\_C\_discrete\_factor18 0.6489 0.1045 6.21 0.00000000055  
##   
## (Intercept) \*\*\*  
## ShnotlimudKlali\_C\_discrete\_factor6   
## ShnotlimudKlali\_C\_discrete\_factor9   
## ShnotlimudKlali\_C\_discrete\_factor12 \*\*\*  
## ShnotlimudKlali\_C\_discrete\_factor15 \*\*\*  
## ShnotlimudKlali\_C\_discrete\_factor18 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.981 on 7320 degrees of freedom  
## (122 observations deleted due to missingness)  
## Multiple R-squared: 0.0307, Adjusted R-squared: 0.03   
## F-statistic: 46.3 on 5 and 7320 DF, p-value: <0.0000000000000002

lmtest::bptest(model1)

##   
## studentized Breusch-Pagan test  
##   
## data: model1  
## BP = 59.7, df = 5, p-value = 0.000000000014

# our result from the the basic regression - schooling years has a positive and significant effect on happiness  
  
  
# basic first regression life satisfaction and highest degree with z scores  
model1 <-lm(MerutzeChaim\_zscores ~ TeudaGvoha, data=df)   
summary(model1)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ TeudaGvoha, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.431 -0.467 -0.292 1.044 1.220   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.14256 0.02058 -6.93 0.0000000000046 \*\*\*  
## TeudaGvoha 0.05841 0.00698 8.37 < 0.0000000000000002 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.996 on 7403 degrees of freedom  
## (43 observations deleted due to missingness)  
## Multiple R-squared: 0.00938, Adjusted R-squared: 0.00925   
## F-statistic: 70.1 on 1 and 7403 DF, p-value: <0.0000000000000002

lmtest::bptest(model1)

##   
## studentized Breusch-Pagan test  
##   
## data: model1  
## BP = 73.2, df = 1, p-value <0.0000000000000002

# basic first regression life satisfaction and highest degree with z scores  
model1 <-lm(MerutzeChaim\_zscores ~ TeudaGvoha\_factor, data=df)   
summary(model1)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ TeudaGvoha\_factor, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.421 -0.481 -0.331 0.972 1.331   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.2541 0.0284 -8.94 < 0.0000000000000002 \*\*\*  
## TeudaGvoha\_factor1 0.2098 0.0405 5.18 0.00000022975 \*\*\*  
## TeudaGvoha\_factor2 0.3593 0.0383 9.37 < 0.0000000000000002 \*\*\*  
## TeudaGvoha\_factor3 0.2202 0.0404 5.45 0.00000005287 \*\*\*  
## TeudaGvoha\_factor4 0.3943 0.0387 10.18 < 0.0000000000000002 \*\*\*  
## TeudaGvoha\_factor5 0.2851 0.0447 6.38 0.00000000019 \*\*\*  
## TeudaGvoha\_factor6 0.4850 0.1078 4.50 0.00000688918 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.992 on 7398 degrees of freedom  
## (43 observations deleted due to missingness)  
## Multiple R-squared: 0.0179, Adjusted R-squared: 0.0171   
## F-statistic: 22.4 on 6 and 7398 DF, p-value: <0.0000000000000002

lmtest::bptest(model1)

##   
## studentized Breusch-Pagan test  
##   
## data: model1  
## BP = 106, df = 6, p-value <0.0000000000000002

#second regression - with supervising variables  
  
# second regression life satisfaction and everything else with schooling years  
model2 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+Female+Gil+Gil\_sq+Married+MispYeladim+HachnasaAvoda, data=df)  
summary(model2)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + Female +   
## Gil + Gil\_sq + Married + MispYeladim + HachnasaAvoda, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.746 -0.545 -0.211 0.862 1.933   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.33635 0.07693 -4.37 0.0000125894376095 \*\*\*  
## ShnotLimud\_Discrete 0.03818 0.00475 8.03 0.0000000000000012 \*\*\*  
## Female 0.06372 0.02888 2.21 0.027 \*   
## Gil -0.27402 0.02487 -11.02 < 0.0000000000000002 \*\*\*  
## Gil\_sq 0.01882 0.00211 8.90 < 0.0000000000000002 \*\*\*  
## Married 0.21309 0.03516 6.06 0.0000000014554750 \*\*\*  
## MispYeladim 0.07805 0.01040 7.50 0.0000000000000749 \*\*\*  
## HachnasaAvoda 0.04130 0.00599 6.90 0.0000000000059917 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.924 on 4633 degrees of freedom  
## (2807 observations deleted due to missingness)  
## Multiple R-squared: 0.0677, Adjusted R-squared: 0.0662   
## F-statistic: 48 on 7 and 4633 DF, p-value: <0.0000000000000002

lmtest::bptest(model2)

##   
## studentized Breusch-Pagan test  
##   
## data: model2  
## BP = 104, df = 7, p-value <0.0000000000000002

# second regression life satisfaction and everything else with schooling years  
model2 <-lm(MerutzeChaim\_zscores ~ ShnotlimudKlali\_C\_discrete\_factor+Female+Gil+Gil\_sq+Married+MispYeladim+HachnasaAvoda, data=df)  
summary(model2)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotlimudKlali\_C\_discrete\_factor +   
## Female + Gil + Gil\_sq + Married + MispYeladim + HachnasaAvoda,   
## data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.725 -0.553 -0.211 0.860 2.143   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) -0.09431 0.20943 -0.45  
## ShnotlimudKlali\_C\_discrete\_factor6 -0.09595 0.22180 -0.43  
## ShnotlimudKlali\_C\_discrete\_factor9 -0.13523 0.21439 -0.63  
## ShnotlimudKlali\_C\_discrete\_factor12 0.25651 0.20448 1.25  
## ShnotlimudKlali\_C\_discrete\_factor15 0.30013 0.20398 1.47  
## ShnotlimudKlali\_C\_discrete\_factor18 0.41392 0.20379 2.03  
## Female 0.06314 0.02888 2.19  
## Gil -0.26728 0.02499 -10.69  
## Gil\_sq 0.01839 0.00212 8.67  
## Married 0.21576 0.03513 6.14  
## MispYeladim 0.07840 0.01041 7.53  
## HachnasaAvoda 0.04152 0.00599 6.93  
## Pr(>|t|)   
## (Intercept) 0.653   
## ShnotlimudKlali\_C\_discrete\_factor6 0.665   
## ShnotlimudKlali\_C\_discrete\_factor9 0.528   
## ShnotlimudKlali\_C\_discrete\_factor12 0.210   
## ShnotlimudKlali\_C\_discrete\_factor15 0.141   
## ShnotlimudKlali\_C\_discrete\_factor18 0.042 \*   
## Female 0.029 \*   
## Gil < 0.0000000000000002 \*\*\*  
## Gil\_sq < 0.0000000000000002 \*\*\*  
## Married 0.000000000886217 \*\*\*  
## MispYeladim 0.000000000000059 \*\*\*  
## HachnasaAvoda 0.000000000004929 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.923 on 4629 degrees of freedom  
## (2807 observations deleted due to missingness)  
## Multiple R-squared: 0.0705, Adjusted R-squared: 0.0683   
## F-statistic: 31.9 on 11 and 4629 DF, p-value: <0.0000000000000002

lmtest::bptest(model2)

##   
## studentized Breusch-Pagan test  
##   
## data: model2  
## BP = 110, df = 11, p-value <0.0000000000000002

# second regression life satisfaction and everything else with highest degree  
model2 <-lm(MerutzeChaim\_zscores ~ TeudaGvoha\_factor+Female+Gil+Gil\_sq+Married+MispYeladim+HachnasaAvoda, data=df)  
summary(model2)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ TeudaGvoha\_factor + Female +   
## Gil + Gil\_sq + Married + MispYeladim + HachnasaAvoda, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.796 -0.541 -0.216 0.867 1.974   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.10849 0.06898 -1.57 0.11587   
## TeudaGvoha\_factor1 0.19777 0.05663 3.49 0.00048 \*\*\*  
## TeudaGvoha\_factor2 0.33447 0.05647 5.92 0.0000000033930381 \*\*\*  
## TeudaGvoha\_factor3 0.24267 0.05651 4.29 0.0000179046450678 \*\*\*  
## TeudaGvoha\_factor4 0.35037 0.05637 6.22 0.0000000005572892 \*\*\*  
## TeudaGvoha\_factor5 0.30964 0.06205 4.99 0.0000006258089630 \*\*\*  
## TeudaGvoha\_factor6 0.48638 0.12791 3.80 0.00015 \*\*\*  
## Female 0.05848 0.02950 1.98 0.04748 \*   
## Gil -0.25248 0.02517 -10.03 < 0.0000000000000002 \*\*\*  
## Gil\_sq 0.01719 0.00213 8.08 0.0000000000000008 \*\*\*  
## Married 0.23182 0.03532 6.56 0.0000000000583833 \*\*\*  
## MispYeladim 0.07947 0.01051 7.56 0.0000000000000473 \*\*\*  
## HachnasaAvoda 0.04365 0.00615 7.10 0.0000000000014153 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.926 on 4638 degrees of freedom  
## (2797 observations deleted due to missingness)  
## Multiple R-squared: 0.0646, Adjusted R-squared: 0.0622   
## F-statistic: 26.7 on 12 and 4638 DF, p-value: <0.0000000000000002

lmtest::bptest(model2)

##   
## studentized Breusch-Pagan test  
##   
## data: model2  
## BP = 113, df = 12, p-value <0.0000000000000002

#our result - gender has no effect on reported happiness. Age has a negative effect on happiness, but the effect becomes positive after a certain age. number of children has a positive effect on happiness. income has a positive effect on happiness. being married has a positive effect on happiness.  
  
#life satisfaction comparison for jews and non jews  
model6 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+Female+Gil+Married+MispYeladim+HachnasaAvoda+jewish, data=df)  
summary(model6)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + Female +   
## Gil + Married + MispYeladim + HachnasaAvoda + jewish, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.533 -0.562 -0.146 0.814 2.065   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.70220 0.07273 -9.65 < 0.0000000000000002 \*\*\*  
## ShnotLimud\_Discrete 0.02308 0.00479 4.82 0.0000015019 \*\*\*  
## Female 0.02567 0.02859 0.90 0.36933   
## Gil -0.06536 0.00583 -11.21 < 0.0000000000000002 \*\*\*  
## Married 0.20362 0.03463 5.88 0.0000000044 \*\*\*  
## MispYeladim 0.05657 0.01001 5.65 0.0000000167 \*\*\*  
## HachnasaAvoda 0.02140 0.00568 3.77 0.00017 \*\*\*  
## jewish 0.44129 0.03507 12.58 < 0.0000000000000002 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.916 on 4629 degrees of freedom  
## (2811 observations deleted due to missingness)  
## Multiple R-squared: 0.0834, Adjusted R-squared: 0.082   
## F-statistic: 60.1 on 7 and 4629 DF, p-value: <0.0000000000000002

lmtest::bptest(model6)

##   
## studentized Breusch-Pagan test  
##   
## data: model6  
## BP = 110, df = 7, p-value <0.0000000000000002

#life satisfaction comparison for jews and non jews with interactions of gender on years of study.  
model6 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+Female+Female\*ShnotLimud\_Discrete+Gil+Married+MispYeladim+HachnasaAvoda+jewish, data=df)  
summary(model6)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + Female +   
## Female \* ShnotLimud\_Discrete + Gil + Married + MispYeladim +   
## HachnasaAvoda + jewish, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.524 -0.560 -0.142 0.818 2.040   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.73472 0.08848 -8.30 < 0.0000000000000002 \*\*\*  
## ShnotLimud\_Discrete 0.02529 0.00589 4.29 0.0000180581 \*\*\*  
## Female 0.11667 0.14379 0.81 0.41719   
## Gil -0.06537 0.00583 -11.21 < 0.0000000000000002 \*\*\*  
## Married 0.20455 0.03467 5.90 0.0000000039 \*\*\*  
## MispYeladim 0.05650 0.01001 5.64 0.0000000175 \*\*\*  
## HachnasaAvoda 0.02149 0.00568 3.78 0.00016 \*\*\*  
## jewish 0.43982 0.03515 12.51 < 0.0000000000000002 \*\*\*  
## ShnotLimud\_Discrete:Female -0.00593 0.00919 -0.65 0.51846   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.916 on 4628 degrees of freedom  
## (2811 observations deleted due to missingness)  
## Multiple R-squared: 0.0834, Adjusted R-squared: 0.0819   
## F-statistic: 52.7 on 8 and 4628 DF, p-value: <0.0000000000000002

lmtest::bptest(model6)

##   
## studentized Breusch-Pagan test  
##   
## data: model6  
## BP = 110, df = 8, p-value <0.0000000000000002

#jews are happiers on average than non jews.  
  
#what is the effect of being jewish on the effect of education on happiness?  
model7 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+Female+Gil+Married+MispYeladim+HachnasaAvoda+jewish+jewish\*ShnotLimud\_Discrete, data=df) #   
summary(model7)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + Female +   
## Gil + Married + MispYeladim + HachnasaAvoda + jewish + jewish \*   
## ShnotLimud\_Discrete, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.527 -0.559 -0.167 0.817 2.186   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.03497 0.12104 -8.55 < 0.0000000000000002 \*\*\*  
## ShnotLimud\_Discrete 0.04664 0.00836 5.58 0.00000002553 \*\*\*  
## Female 0.02234 0.02857 0.78 0.43430   
## Gil -0.06591 0.00583 -11.31 < 0.0000000000000002 \*\*\*  
## Married 0.20597 0.03460 5.95 0.00000000283 \*\*\*  
## MispYeladim 0.06031 0.01005 6.00 0.00000000215 \*\*\*  
## HachnasaAvoda 0.02201 0.00567 3.88 0.00011 \*\*\*  
## jewish 0.92712 0.14561 6.37 0.00000000021 \*\*\*  
## ShnotLimud\_Discrete:jewish -0.03399 0.00989 -3.44 0.00059 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.915 on 4628 degrees of freedom  
## (2811 observations deleted due to missingness)  
## Multiple R-squared: 0.0857, Adjusted R-squared: 0.0841   
## F-statistic: 54.2 on 8 and 4628 DF, p-value: <0.0000000000000002

lmtest::bptest(model7)

##   
## studentized Breusch-Pagan test  
##   
## data: model7  
## BP = 114, df = 8, p-value <0.0000000000000002

#for jews, happiness is less depenedent on the level of education. for arabs happiness is more dependent on level of education  
  
#is there an effect of gender on happiness? seems like none. p value is wrong. after correcting still not significant  
model9 <-lm(MerutzeChaim\_zscores ~ Female, data=df) #   
summary(model9)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ Female, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.302 -0.397 -0.355 1.056 1.098   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.0209 0.0165 1.27 0.205   
## Female -0.0415 0.0232 -1.79 0.074 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1 on 7409 degrees of freedom  
## (37 observations deleted due to missingness)  
## Multiple R-squared: 0.00043, Adjusted R-squared: 0.000295   
## F-statistic: 3.19 on 1 and 7409 DF, p-value: 0.0741

lmtest::bptest(model9)

##   
## studentized Breusch-Pagan test  
##   
## data: model9  
## BP = 0.00274, df = 1, p-value = 0.96

coeftest(model9, vcov = vcovHC(model9, "HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.0209 0.0165 1.27 0.205   
## Female -0.0415 0.0232 -1.79 0.074 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#for jews - is there an effect of gender on happiness? seems like none. p value is wrong. after correcting still not significant  
model9\_jews <-lm(MerutzeChaim\_zscores ~ Female, data=df\_jews) #   
summary(model9\_jews)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ Female, data = df\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.396 -0.490 -0.442 0.962 1.010   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.1146 0.0179 6.42 0.00000000015 \*\*\*  
## Female -0.0480 0.0253 -1.90 0.058 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.964 on 5809 degrees of freedom  
## (27 observations deleted due to missingness)  
## Multiple R-squared: 0.000619, Adjusted R-squared: 0.000447   
## F-statistic: 3.6 on 1 and 5809 DF, p-value: 0.0579

lmtest::bptest(model9)

##   
## studentized Breusch-Pagan test  
##   
## data: model9  
## BP = 0.00274, df = 1, p-value = 0.96

coeftest(model9, vcov = vcovHC(model9, "HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.0209 0.0165 1.27 0.205   
## Female -0.0415 0.0232 -1.79 0.074 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#for non jews - is there an effect of gender on happiness? seems like none. p value is wrong. after correcting still not significant  
model9\_non\_jews <-lm(MerutzeChaim\_zscores ~ Female, data=df\_non\_jews) #   
summary(model9\_non\_jews)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ Female, data = df\_non\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.9540 -0.0485 -0.0367 -0.0367 1.4161   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.3391 0.0385 -8.81 <0.0000000000000002 \*\*\*  
## Female 0.0119 0.0530 0.22 0.82   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.06 on 1589 degrees of freedom  
## (5 observations deleted due to missingness)  
## Multiple R-squared: 3.16e-05, Adjusted R-squared: -0.000598   
## F-statistic: 0.0502 on 1 and 1589 DF, p-value: 0.823

lmtest::bptest(model9)

##   
## studentized Breusch-Pagan test  
##   
## data: model9  
## BP = 0.00274, df = 1, p-value = 0.96

coeftest(model9, vcov = vcovHC(model9, "HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.0209 0.0165 1.27 0.205   
## Female -0.0415 0.0232 -1.79 0.074 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# for jews effect of level of religiousness on happiness   
model10 <-lm(MerutzeChaim\_zscores ~ DatiutYehudi, data=df\_jews) #   
summary(model10)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ DatiutYehudi, data = df\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.743 -0.440 -0.308 0.880 1.145   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.20056 0.02364 -8.48 <0.0000000000000002 \*\*\*  
## DatiutYehudi 0.13251 0.00916 14.47 <0.0000000000000002 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.947 on 5787 degrees of freedom  
## (49 observations deleted due to missingness)  
## Multiple R-squared: 0.0349, Adjusted R-squared: 0.0347   
## F-statistic: 209 on 1 and 5787 DF, p-value: <0.0000000000000002

lmtest::bptest(model10)

##   
## studentized Breusch-Pagan test  
##   
## data: model10  
## BP = 15.4, df = 1, p-value = 0.000088

#for jews more religious is more happy for basic regression. now add additional supervising parameters  
  
model10\_add\_supervising\_variables <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+Female+Gil+Gil\_sq+Married+MispYeladim+HachnasaAvoda+DatiutYehudi, data=df\_jews) #   
summary(model10\_add\_supervising\_variables)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + Female +   
## Gil + Gil\_sq + Married + MispYeladim + HachnasaAvoda + DatiutYehudi,   
## data = df\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.455 -0.559 -0.206 0.829 1.772   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.20069 0.09667 -2.08 0.038 \*   
## ShnotLimud\_Discrete 0.01801 0.00548 3.29 0.001 \*\*   
## Female 0.04919 0.03013 1.63 0.103   
## Gil -0.22687 0.02695 -8.42 < 0.0000000000000002 \*\*\*  
## Gil\_sq 0.01477 0.00224 6.60 0.000000000048297 \*\*\*  
## Married 0.19274 0.03677 5.24 0.000000168207495 \*\*\*  
## MispYeladim 0.05471 0.01233 4.44 0.000009432446397 \*\*\*  
## HachnasaAvoda 0.04364 0.00632 6.91 0.000000000005790 \*\*\*  
## DatiutYehudi 0.09671 0.01281 7.55 0.000000000000054 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.87 on 3702 degrees of freedom  
## (2127 observations deleted due to missingness)  
## Multiple R-squared: 0.0862, Adjusted R-squared: 0.0842   
## F-statistic: 43.6 on 8 and 3702 DF, p-value: <0.0000000000000002

lmtest::bptest(model10\_add\_supervising\_variables)

##   
## studentized Breusch-Pagan test  
##   
## data: model10\_add\_supervising\_variables  
## BP = 73.8, df = 8, p-value = 0.00000000000085

# after adding supervising variables the effect of religiousness level on happniess is still strong in jews  
  
# for non jews effect of level of religiousness on happiness p value is wrong?  
model11 <-lm(MerutzeChaim\_zscores ~ DatiutLoYehudi, data=df\_non\_jews) #   
summary(model11)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ DatiutLoYehudi, data = df\_non\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.0196 -0.0652 -0.0652 0.0326 1.4853   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.4573 0.0941 -4.86 0.0000013 \*\*\*  
## DatiutLoYehudi 0.0489 0.0362 1.35 0.18   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.06 on 1390 degrees of freedom  
## (204 observations deleted due to missingness)  
## Multiple R-squared: 0.00131, Adjusted R-squared: 0.000594   
## F-statistic: 1.83 on 1 and 1390 DF, p-value: 0.177

lmtest::bptest(model11)

##   
## studentized Breusch-Pagan test  
##   
## data: model11  
## BP = 0.607, df = 1, p-value = 0.44

#regression of religion on happiness with factor p value is wrong? after correcting not much difference  
model12 <-lm(MerutzeChaim\_zscores ~ factor(Dat), data=df\_non\_jews) #  
summary(model12)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ factor(Dat), data = df\_non\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.514 -0.106 0.038 0.038 1.490   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.4135 0.0324 -12.77 < 0.0000000000000002 \*\*\*  
## factor(Dat)3 0.1437 0.0757 1.90 0.058 .   
## factor(Dat)4 0.6465 0.1069 6.05 0.0000000018 \*\*\*  
## factor(Dat)5 0.4650 0.2552 1.82 0.069 .   
## factor(Dat)6 0.0970 0.0813 1.19 0.233   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.04 on 1586 degrees of freedom  
## (5 observations deleted due to missingness)  
## Multiple R-squared: 0.0248, Adjusted R-squared: 0.0223   
## F-statistic: 10.1 on 4 and 1586 DF, p-value: 0.0000000479

lmtest::bptest(model12)

##   
## studentized Breusch-Pagan test  
##   
## data: model12  
## BP = 0.898, df = 4, p-value = 0.92

coeftest(model12, vcov = vcovHC(model12, "HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.4135 0.0325 -12.74 < 0.0000000000000002 \*\*\*  
## factor(Dat)3 0.1437 0.0746 1.93 0.054 .   
## factor(Dat)4 0.6465 0.1083 5.97 0.0000000029 \*\*\*  
## factor(Dat)5 0.4650 0.2924 1.59 0.112   
## factor(Dat)6 0.0970 0.0802 1.21 0.227   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#regression of religion on happiness with factor   
model12 <-lm(MerutzeChaim\_zscores ~ factor(Dat), data=df) #  
summary(model12)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ factor(Dat), data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.514 -0.466 -0.427 0.986 1.490   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.0907 0.0129 7.04 0.0000000000021 \*\*\*  
## factor(Dat)2 -0.5042 0.0331 -15.25 < 0.0000000000000002 \*\*\*  
## factor(Dat)3 -0.3605 0.0656 -5.49 0.0000000404909 \*\*\*  
## factor(Dat)4 0.1423 0.0967 1.47 0.14   
## factor(Dat)5 -0.0392 0.2385 -0.16 0.87   
## factor(Dat)6 -0.4072 0.0713 -5.71 0.0000000117586 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.982 on 7396 degrees of freedom  
## (46 observations deleted due to missingness)  
## Multiple R-squared: 0.0362, Adjusted R-squared: 0.0356   
## F-statistic: 55.6 on 5 and 7396 DF, p-value: <0.0000000000000002

lmtest::bptest(model12)

##   
## studentized Breusch-Pagan test  
##   
## data: model12  
## BP = 14.1, df = 5, p-value = 0.015

#same effect for non jews and for arabs and muslems.  
#for non jews there is no significant effect for religious level on happiness.  
  
# for jews - regression of education and happiness  
model13 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete, data=df\_jews)   
summary(model13)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete, data = df\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.491 -0.585 -0.356 0.867 1.440   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.47740 0.05915 -8.07 0.00000000000000085 \*\*\*  
## ShnotLimud\_Discrete 0.03817 0.00384 9.93 < 0.0000000000000002 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.951 on 5767 degrees of freedom  
## (69 observations deleted due to missingness)  
## Multiple R-squared: 0.0168, Adjusted R-squared: 0.0167   
## F-statistic: 98.7 on 1 and 5767 DF, p-value: <0.0000000000000002

lmtest::bptest(model13)

##   
## studentized Breusch-Pagan test  
##   
## data: model13  
## BP = 37.1, df = 1, p-value = 0.0000000011

# for non jews - regression of education and happiness  
model13 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete, data=df\_non\_jews)   
summary(model13)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete, data = df\_non\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.1245 -0.2191 -0.0221 0.2734 1.7261   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.74766 0.08714 -8.58 < 0.0000000000000002 \*\*\*  
## ShnotLimud\_Discrete 0.03283 0.00643 5.10 0.00000038 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.05 on 1547 degrees of freedom  
## (47 observations deleted due to missingness)  
## Multiple R-squared: 0.0166, Adjusted R-squared: 0.0159   
## F-statistic: 26 on 1 and 1547 DF, p-value: 0.000000375

lmtest::bptest(model13)

##   
## studentized Breusch-Pagan test  
##   
## data: model13  
## BP = 8.62, df = 1, p-value = 0.0033

#conclusion -for basic regression, for jews, the effect of education on happiness is stronger than for arabs. For both jews and non jews the effect is significant and positive.  
  
#adding all parameters  
#for each group jews and non jews, calculating the effect of education on happiness, with additional supervising variables.  
model15 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+ShnotLimud\_Discrete\*jewish+Female+Gil+Married+MispYeladim+MispYeladim\_sq+HachnasaAvoda, data=df)   
summary(model15)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + ShnotLimud\_Discrete \*   
## jewish + Female + Gil + Married + MispYeladim + MispYeladim\_sq +   
## HachnasaAvoda, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.522 -0.555 -0.169 0.820 2.206   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.04912 0.12161 -8.63 < 0.0000000000000002 \*\*\*  
## ShnotLimud\_Discrete 0.04723 0.00837 5.64 0.00000001798 \*\*\*  
## jewish 0.93910 0.14595 6.43 0.00000000014 \*\*\*  
## Female 0.02615 0.02875 0.91 0.36305   
## Gil -0.06348 0.00617 -10.29 < 0.0000000000000002 \*\*\*  
## Married 0.22184 0.03706 5.99 0.00000000231 \*\*\*  
## MispYeladim 0.03095 0.02654 1.17 0.24354   
## MispYeladim\_sq 0.00478 0.00400 1.20 0.23196   
## HachnasaAvoda 0.02296 0.00573 4.01 0.00006217568 \*\*\*  
## ShnotLimud\_Discrete:jewish -0.03477 0.00991 -3.51 0.00045 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.915 on 4627 degrees of freedom  
## (2811 observations deleted due to missingness)  
## Multiple R-squared: 0.086, Adjusted R-squared: 0.0842   
## F-statistic: 48.4 on 9 and 4627 DF, p-value: <0.0000000000000002

lmtest::bptest(model15)

##   
## studentized Breusch-Pagan test  
##   
## data: model15  
## BP = 117, df = 9, p-value <0.0000000000000002

#for extended model, for jews, the effect of education on happiness is less strong than for arabs. For both jews and non jews the effect is significant and positive.  
  
#effect of level or religiousness on return of education to happiness  
  
# for jews - simple regression of education and happiness with level of religiousness  
model17 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+DatiutYehudi\_dummy\*ShnotLimud\_Discrete, data=df\_jews)   
summary(model17)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + DatiutYehudi\_dummy \*   
## ShnotLimud\_Discrete, data = df\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.768 -0.457 -0.251 0.995 1.512   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) -0.53802 0.07540 -7.14  
## ShnotLimud\_Discrete 0.03442 0.00486 7.08  
## DatiutYehudi\_dummy 0.11135 0.11894 0.94  
## ShnotLimud\_Discrete:DatiutYehudi\_dummy 0.01631 0.00777 2.10  
## Pr(>|t|)   
## (Intercept) 0.0000000000011 \*\*\*  
## ShnotLimud\_Discrete 0.0000000000017 \*\*\*  
## DatiutYehudi\_dummy 0.349   
## ShnotLimud\_Discrete:DatiutYehudi\_dummy 0.036 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.936 on 5743 degrees of freedom  
## (91 observations deleted due to missingness)  
## Multiple R-squared: 0.0473, Adjusted R-squared: 0.0468   
## F-statistic: 95 on 3 and 5743 DF, p-value: <0.0000000000000002

lmtest::bptest(model17)

##   
## studentized Breusch-Pagan test  
##   
## data: model17  
## BP = 58, df = 3, p-value = 0.0000000000015

# for non jews - simple regression of education and happiness with level of religiousness  
model18 <-lm(MerutzeChaim\_zscores ~ ShnotLimud\_Discrete+DatiutLoYehudi\_dummy\*ShnotLimud\_Discrete, data=df\_non\_jews)   
summary(model18)

##   
## Call:  
## lm(formula = MerutzeChaim\_zscores ~ ShnotLimud\_Discrete + DatiutLoYehudi\_dummy \*   
## ShnotLimud\_Discrete, data = df\_non\_jews)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.256 -0.220 -0.048 0.334 1.787   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.80546 0.16739 -4.81 0.0000017  
## ShnotLimud\_Discrete 0.03187 0.01197 2.66 0.0079  
## DatiutLoYehudi\_dummy -0.00796 0.19900 -0.04 0.9681  
## ShnotLimud\_Discrete:DatiutLoYehudi\_dummy 0.01195 0.01467 0.81 0.4155  
##   
## (Intercept) \*\*\*  
## ShnotLimud\_Discrete \*\*   
## DatiutLoYehudi\_dummy   
## ShnotLimud\_Discrete:DatiutLoYehudi\_dummy   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.05 on 1346 degrees of freedom  
## (246 observations deleted due to missingness)  
## Multiple R-squared: 0.026, Adjusted R-squared: 0.0238   
## F-statistic: 12 on 3 and 1346 DF, p-value: 0.0000000984

lmtest::bptest(model18)

##   
## studentized Breusch-Pagan test  
##   
## data: model18  
## BP = 15.1, df = 3, p-value = 0.0018