

```
---
title: "IS4250 Group18 Project Code"
author: "jy"
date: "April 5, 2016"
output: md_document
---

##### Import library

```{r, message=FALSE, echo=FALSE}
library(MASS)
library(effects)
library(vcd)
library(car)
library(mlogit)
library(nnet)
library(dplyr)
library(data.table)
library(ggplot2)
library(lattice)
```

##### Convert excel data file into .csv format. Read all data from csv file in R.

```{r, message=FALSE, echo=FALSE}
OriginalHealthData <- read.csv("C:/Users/ASUS/Documents/Health Data.csv")
healthData <- dplyr::select(OriginalHealthData, everything())
```

##### Rename 24 data attributes according to codebook provided.
##### (Detected some discrepancy and several missing attributes such as self-reported health status)

```{r, message=FALSE, echo=FALSE}
healthData <- dplyr::rename(healthData, Occupation = b1)
healthData <- dplyr::rename(healthData, WorkplaceScale = b2)
healthData <- dplyr::rename(healthData, Gender = q102)
healthData <- dplyr::rename(healthData, Age = q103)
healthData <- dplyr::rename(healthData, MaritalStatus = q104)
healthData <- dplyr::rename(healthData, EducationAttainment = q105)
healthData <- dplyr::rename(healthData, AccompanyChild = q106)
healthData <- dplyr::rename(healthData, Salary = q107)
healthData <- dplyr::rename(healthData, YearsAsMigrant = q108a1)
healthData <- dplyr::rename(healthData, NumberOfCitiesResidedIn = q108a2)
healthData <- dplyr::rename(healthData, DailyWorkingHours = q108b)
healthData <- dplyr::rename(healthData, WeeklyWorkingDays = q108c1)
healthData <- dplyr::rename(healthData, TypeOfResidence = q108c2)
healthData <- dplyr::rename(healthData, HundredCigarettes = q410)
healthData <- dplyr::rename(healthData, CurrentSmoker = q411)
healthData <- dplyr::rename(healthData, NumberOfCigarettesPerDay = q412)
healthData <- dplyr::rename(healthData, DrinkingFrequency = q431)
healthData <- dplyr::rename(healthData, AlcoholDrinksPerDay = q432)
healthData <- dplyr::rename(healthData, SixOrMoreDrinksFrequency = q433)
healthData <- dplyr::rename(healthData, DailySleepingHours = q444)
healthData <- dplyr::rename(healthData, SleepingQuality = q446)
healthData <- dplyr::rename(healthData, DailyBreakfast = q450)
healthData <- dplyr::rename(healthData, DailyFruitsVegetables = q451)
healthData <- dplyr::rename(healthData, DailyRegularMeals = q452)
```

##### Compute Mental Health Score:
```

```
* Read self-assessment scores of q201-q290, 90 questions from Symptom Checklist-90-Revised.
* Each question carries 1-5 marks, representing none, mild, moderate, severe, extreme.
* Sum up 90 scores
```

```
```{r, message=FALSE, echo=FALSE}
mentalChecklist <- dplyr::select(OriginalHealthData, contains("q2"))
mentalChecklist <- mentalChecklist %>% transmute(MentalHealth = q201 + q202 + q203 + q204
+q205 + q206 + q207 + q208 +q209 + q210 + q211 + q212 + q213 + q214 + q215 + q216 + q217 +
q218 + q219 + q220 + q221 + q222 + q223 + q224 + q225 + q226 + q227 + q228 + q229 + q230 +
q231 + q232 + q233 + q234 + q235 + q236 + q237 + q238 + q239 + q240 + q241 + q242 + q243 +
q244 + q245 + q246 + q247 + q248 + q249 + q250 + q251 + q252 + q253 + q254 + q255 + q256 +
q257 + q258 + q259 + q260 + q261 + q262 + q263 + q264 + q265 + q266 + q267 + q268 + q269 +
q270 + q271 + q272 + q273 + q274 + q275 + q276 + q277 + q278 + q279 + q280 + q281 + q282 +
q283 + q284 + q285 + q286 + q287 + q288 + q289 + q290)
healthData <- dplyr::bind_cols(healthData, mentalChecklist)
```
```

#### ##### Compute Drinking Behaviour Score

```
* 3 question
+ Frequency of drinking
+ How many drinks containing alcohol do you have on a typical day when you are drinking?
+ How often do you have six or more drinks on one occasion?
* Each question carries 1-5 marks, representing degree of alcohol consumption
* Sum up 3 scores
```

```
```{r, message=FALSE, echo=FALSE}
drinkingStatus <- dplyr::select(healthData, DrinkingFrequency, AlcoholDrinksPerDay,
SixOrMoreDrinksFrequency)
drinkingStatus <- drinkingStatus %>% transmute(AlcoholConsumption = DrinkingFrequency +
AlcoholDrinksPerDay + SixOrMoreDrinksFrequency - 3)
healthData <- dplyr::bind_cols(healthData, drinkingStatus)
```
```

#### ##### Further process data (mainly Categorization)

```
* Education Attainment
+ 4 catrgories: Elementary or lower, Junior high school, High school, College or above
* Gender
+ swap sequence into 2 categories: Male, Female
* Number of Accompanying Children
+ 3 categories: No children, Cohabiting children, Children residing elsewhere
* Salary
+ 4 categories: <1500RMB, 1500-2500RMB, 2500-3500RMB, >=3500RMB
* Number of Cities Resided in
+ 2 categories: 1-2, >=3
* Daily Working Hours
* Weekly Working Days
* BMI:
+ BMI = weight(kg) / (height(m)^2)
+ 3 categories: < 18.5; >= 18.5 and < 24; >=24
* Mental Health
+ 2 categories: if MentalHealth score <= 160, Normal; if MentalHealth score > 160, Abnormal
* Smoking
+ Current Smoker
+ 2 categories: Unhealthy, Healthy
+ if smoked in the past 30 days, Unhealthy; else Healthy
* Drinking
+ Hazardous Alcohol Consumption
+ 2 categories: Unhealthy, Healthy
+ if Alcohol Consumption Score >=4 for Male or Alcohol Consumption Score >=3 for Female,
Unhealthy; else Healthy
* Daily Sleeping Hours
```

```

+ 2 categories: Unhealthy, Healthy
+ if Daily Sleeping Hours 7-9 hours/night, Health; else Unhealthy
* Sleeping Quality
+ 2 categories: 1-2 Good/Fair; 3-4 Poor/Very Poor
* Daily Breakfasts
+ 2 categories: Unhealthy, Healthy
+ if almost Daily Breakfasts, Health; else Unhealthy
* Daily Fruits and Vegetables Consumption
+ 2 categories: Unhealthy, Healthy
+ if almost Daily Fruits and Vegetables Consumption, Health; else Unhealthy
* Daily Regular Meals
+ 2 categories: Unhealthy, Healthy
+ if almost Daily Regular Meals, Health; else Unhealthy

```{r, message=FALSE, echo=FALSE}
healthData <- dplyr::mutate(healthData, EducationAttainmentX = ifelse(EducationAttainment ==
1, 1, ifelse(EducationAttainment == 2, 2, ifelse(EducationAttainment == 3, 3, 4))))
healthData <- dplyr::mutate(healthData, BMI = weight/(height/100)/(height/100))
healthData <- dplyr::mutate(healthData, BMIX = ifelse(BMI < 18.5, 1, ifelse(BMI < 24, 2, 3)))
healthData <- dplyr::mutate(healthData, GenderX = ifelse(Gender == 1, 2, 1))
healthData <- dplyr::mutate(healthData, AgeX = ifelse(Age <= 32, 1, 2))
healthData <- dplyr::mutate(healthData, AccompanyChildX = ifelse(AccompanyChild == 1, 1,
ifelse(AccompanyChild == 4, 3, 2)))
healthData <- dplyr::mutate(healthData, SalaryX = ifelse(Salary == 1, 1, ifelse(Salary == 2,
2, ifelse(Salary == 3, 3, 4))))
healthData <- dplyr::mutate(healthData, NumberOfCitiesResidedInX =
ifelse(NumberOfCitiesResidedIn <= 2, 1, 2))
healthData <- dplyr::mutate(healthData, DailyWorkingHoursX = ifelse(DailyWorkingHours == 8,
1, ifelse(DailyWorkingHours < 8, 2, ifelse(DailyWorkingHours < 11, 3, 4))))
healthData <- dplyr::mutate(healthData, WeeklyWorkingDaysX = ifelse(WeeklyWorkingDays <= 4,
1, ifelse(WeeklyWorkingDays == 5, 2, ifelse(WeeklyWorkingDays == 6, 3, 4))))
healthData <- dplyr::mutate(healthData, MentalHealthX = ifelse(MentalHealth <= 160, 1, 2))
healthData <- dplyr::mutate(healthData, CurrentSmokerX = ifelse(CurrentSmoker == 1 |
is.na(CurrentSmoker), 0, 1))
healthData <- dplyr::mutate(healthData, HazardousDrinkX = ifelse(AlcoholConsumption >= 4, 1,
ifelse(GenderX == 2 & AlcoholConsumption == 3, 1, 0)))
healthData <- dplyr::mutate(healthData, DailySleepingHoursX = ifelse(DailySleepingHours < 7
| DailySleepingHours > 9, 1, 0))
healthData <- dplyr::mutate(healthData, SleepingQualityX = ifelse(SleepingQuality <= 2, 0, 1))
healthData <- dplyr::mutate(healthData, DailyBreakfastX = ifelse(DailyBreakfast == 1, 0, 1))
healthData <- dplyr::mutate(healthData, DailyFruitsVegetablesX =
ifelse(DailyFruitsVegetables == 1, 0, 1))
healthData <- dplyr::mutate(healthData, DailyRegularMealsX = ifelse(DailyRegularMeals == 1,
0, 1))
```

##### Compute Lifestyle Behaviour Score
* 6 factors: CurrentSmoker, HazardousDrink, DailySleepingHours, DailyBreakfast,
DailyFruitsVegetables, DailyRegularMeals
* 1 score for every Unhealthy Behavior
* Sum: 0, Healthy; 1-2 Relatively Healthy; 3-6 Unhealthy

```{r, message=FALSE, echo=FALSE}
lifestyleScore <- dplyr::select(healthData, CurrentSmokerX, HazardousDrinkX,
DailySleepingHoursX, DailyBreakfastX, DailyFruitsVegetablesX, DailyRegularMealsX)
lifestyleScore <- lifestyleScore %>% transmute(Score = CurrentSmokerX + HazardousDrinkX +
DailySleepingHoursX + DailyBreakfastX + DailyFruitsVegetablesX + DailyRegularMealsX)
healthData <- dplyr::bind_cols(healthData, lifestyleScore)

healthData <- dplyr::mutate(healthData, ScoreX = ifelse(Score == 0, 1, ifelse(Score == 1 |
Score == 2, 2, 3)))

```

```

healthDataTbl <- data.table(healthData)

...

##### Till here, end of Section 0 - Data Preparation

##### Section 1 - Chi-square Test for Socio-demographic characteristics of Rural-to-Urban
Migrants by Gender

```{r, message=FALSE, echo=FALSE}
Age <- table(healthDataTbl$AgeX, healthDataTbl$GenderX)
dimnames(Age) <- list(AgeX = c("<=32",">32"), GenderX = c("Male","Female"))

Age

chisq.test(Age,correct=F)

...

```{r, message=FALSE, echo=FALSE}
Occupation <- table(healthDataTbl$Occupation, healthDataTbl$GenderX)
dimnames(Occupation) <- list(Occupation =
c("Manufacturing","Construction","Hospitality","Domestic service", "Small
business","Recreation/leisure"), GenderX = c("Male","Female"))

Occupation

chisq.test(Occupation,correct=F)

...

```{r, message=FALSE, echo=FALSE}
WorkplaceScale <- table(healthDataTbl$WorkplaceScale, healthDataTbl$GenderX)
dimnames(WorkplaceScale) <- list(WorkplaceScale = c("Large","Moderate","Small"), GenderX =
c("Male","Female"))

WorkplaceScale

chisq.test(WorkplaceScale,correct=F)

...

```{r, message=FALSE, echo=FALSE}
EducationAttainment <- table(healthDataTbl$EducationAttainmentX, healthDataTbl$GenderX)
dimnames(EducationAttainment) <- list(EducationAttainmentX = c("Elementary or lower","Junior
high school","High school","College"), GenderX = c("Male","Female"))

EducationAttainment

chisq.test(EducationAttainment,correct=F)

...

```{r, message=FALSE, echo=FALSE}
MaritalStatus <- table(healthDataTbl$MaritalStatus, healthDataTbl$GenderX)
dimnames(MaritalStatus) <- list(MaritalStatus =
c("Married","Single","Cohabitating","Divorced or widowed"), GenderX = c("Male","Female"))

MaritalStatus

```

```
chisq.test(MaritalStatus,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
AccompanyChild <- table(healthDataTbl$AccompanyChildX, healthDataTbl$GenderX)
dimnames(AccompanyChild) <- list(AccompanyChildX = c("No children", "Cohabiting
children", "Children residing elsewhere"), GenderX = c("Male", "Female"))
```

```
AccompanyChild
```

```
chisq.test(AccompanyChild,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
Salary <- table(healthDataTbl$SalaryX, healthDataTbl$GenderX)
dimnames(Salary) <- list(SalaryX = c("<1500RMB", "1500-2500RMB", "2500-3500RMB", ">=3500RMB"),
GenderX = c("Male", "Female"))
```

```
Salary
```

```
chisq.test(Salary,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
NumberOfCitiesResidedIn <- table(healthDataTbl$NumberOfCitiesResidedInX,
healthDataTbl$GenderX)
dimnames(NumberOfCitiesResidedIn) <- list(NumberOfCitiesResidedInX = c("1-2", ">=3"),
GenderX = c("Male", "Female"))
```

```
NumberOfCitiesResidedIn
```

```
chisq.test(NumberOfCitiesResidedIn,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
DailyWorkingHours <- table(healthDataTbl$DailyWorkingHoursX, healthDataTbl$GenderX)
dimnames(DailyWorkingHours) <- list(DailyWorkingHoursX = c("=8", "<8", "8-11", ">=11"), GenderX
= c("Male", "Female"))
```

```
DailyWorkingHours
```

```
chisq.test(DailyWorkingHours,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
WeeklyWorkingDays <- table(healthDataTbl$WeeklyWorkingDaysX, healthDataTbl$GenderX)
dimnames(WeeklyWorkingDays) <- list(WeeklyWorkingDaysX = c("<=4", "5", "6", "7"), GenderX =
c("Male", "Female"))
```

```
WeeklyWorkingDays
```

```
chisq.test(WeeklyWorkingDays,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
TypeOfResidence <- table(healthDataTbl$TypeOfResidence, healthDataTbl$GenderX)
dimnames(TypeOfResidence) <- list(TypeOfResidence = c("Collective dormitory", "Renting with
others", "Renting as a family", "Renting alone", "Owning a living place"), GenderX =
c("Male", "Female"))
```

TypeOfResidence

```
chisq.test(TypeOfResidence,correct=F)
```

```
```
```

##### Section 2 - Chi-square Test for Physical and Mental Status and Lifestyle Behaviours of Rural-to-Urban Migrants by Gender

```
```{r, message=FALSE, echo=FALSE}
```

```
BMI <- table(healthDataTbl$BMIX, healthDataTbl$GenderX)
```

```
dimnames(BMI) <- list(BMIX = c("Under Weight", "Normal", "Overweight/Obesity"), GenderX =  
c("Male", "Female"))
```

BMI

```
chisq.test(BMI,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
MentalHealth <- table(healthDataTbl$MentalHealthX, healthDataTbl$GenderX)
```

```
dimnames(MentalHealth) <- list(MentalHealthX = c("Normal", "Abnormal"), GenderX =  
c("Male", "Female"))
```

MentalHealth

```
chisq.test(MentalHealth,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
Smoking <- table(healthDataTbl$CurrentSmokerX, healthDataTbl$GenderX)
```

```
dimnames(Smoking) <- list(CurrentSmokerX = c("Previous smoker/Never smoked", "Current  
smoker"), GenderX = c("Male", "Female"))
```

Smoking

```
chisq.test(Smoking,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
Drinking <- table(healthDataTbl$HazardousDrinkX, healthDataTbl$GenderX)
```

```
dimnames(Drinking) <- list(HazardousDrinkX = c("Nonhazardous alcohol use", "Hazardous  
alcohol use"), GenderX = c("Male", "Female"))
```

Drinking

```
chisq.test(Drinking,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
```

```
SleepingTime <- table(healthDataTbl$DailySleepingHoursX, healthDataTbl$GenderX)
```

```
dimnames(SleepingTime) <- list(DailySleepingHoursX = c("7-9 Hours/night", "<7 or >9  
Hours/night"), GenderX = c("Male", "Female"))
```

SleepingTime

```
chisq.test(SleepingTime,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
SleepingQuality <- table(healthDataTbl$SleepingQualityX, healthDataTbl$GenderX)
dimnames(SleepingQuality) <- list(SleepingQualityX = c("Nonhazardous alcohol use",
"Hazardous alcohol use"), GenderX = c("Male","Female"))
```

```
SleepingQuality
```

```
chisq.test(SleepingQuality,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
DailyBreakfast <- table(healthDataTbl$DailyBreakfastX, healthDataTbl$GenderX)
dimnames(DailyBreakfast) <- list(DailyBreakfastX = c("Almost Daily Breakfast", "Not Almost
Daily Breakfast"), GenderX = c("Male","Female"))
```

```
DailyBreakfast
```

```
chisq.test(DailyBreakfast,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
DailyFruitsVegetables <- table(healthDataTbl$DailyFruitsVegetablesX, healthDataTbl$GenderX)
dimnames(DailyFruitsVegetables) <- list(DailyFruitsVegetablesX = c("Almost Daily Fruits
Vegetables", "Not Almost Daily Fruits Vegetables"), GenderX = c("Male","Female"))
```

```
DailyFruitsVegetables
```

```
chisq.test(DailyFruitsVegetables,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
DailyRegularMeals <- table(healthDataTbl$DailyRegularMealsX, healthDataTbl$GenderX)
dimnames(DailyRegularMeals) <- list(DailyRegularMealsX = c("Almost Daily Regular Meals",
"Not Almost Daily Regular Meals "), GenderX = c("Male","Female"))
```

```
DailyRegularMeals
```

```
chisq.test(DailyRegularMeals,correct=F)
```

```
```
```

```
```{r, message=FALSE, echo=FALSE}
LifestyleScoreX <- table(healthDataTbl$ScoreX, healthDataTbl$GenderX)
dimnames(LifestyleScoreX) <- list(ScoreX = c("Healthy", "Relatively Healthy", "Unhealthy"),
GenderX = c("Male","Female"))
```

```
LifestyleScoreX
```

```
chisq.test(LifestyleScoreX,correct=F)
```

```
healthDataTbl$ScoreX <- factor(healthDataTbl$ScoreX,levels=c(1,2,3),
labels=c("Healthy", "Relatively Healthy", "Unhealthy"))
healthDataTbl$GenderX <- factor(healthDataTbl$GenderX,levels=c(1,2),
labels=c("Male","Female"))
```

```

densityplot(~healthDataTbl$Score|healthDataTbl$GenderX, main="Lifestyle Score - Gender",
xlab="Lifestyle Score")
qplot(GenderX, data=healthDataTbl, geom="bar", fill=ScoreX, xlab="", ylab="Lifestyle Score",
main="Lifestyle Score - Gender")

...

##### Section 3 - Association between lifestyle Score and socio-demographic characteristics
and phiscal and mental status (By Gender)

3.1 For Male rural-to-urban migrants

```{r, message=FALSE, echo=FALSE}

Male <- dplyr::filter(healthData, GenderX == 1)

AGE1 <- cut(Male$AgeX, c(0.5, 1.5,2.5), labels = c("<=32",">32"))
OCCUPATION1 <- cut(Male$Occupation, c(0.5, 1.5,2.5,3.5,4.5,5.5, 6.5), labels =
c("Manufacturing", "Construction", "Hospitality", "DomesticServices", "SmallBusiness",
"Recreation/Leisures"))
WORKPLACESCALE1 <- cut(Male$WorkplaceScale, c(0.5, 1.5,2.5,3.5), labels =
c("Large","Moderate","Small"))
EDUCATIONATTAINMENT1 <- cut(Male$EducationAttainmentX, c(0.5, 1.5,2.5,3.5,4.5), labels =
c("Elementary/Lower","JuniorHigh","HighSchool","College"))
MARITALSTATUS1 <- cut(Male$MaritalStatus, c(0.5, 1.5,2.5,3.5,4.5), labels = c("Married",
"Single","Cohabitating","Divorced/Widowed"))
ACCOMPANYCHILD1 <- cut(Male$AccompanyChildX, c(0.5, 1.5,2.5,3.5), labels =
c("NoChildren","Cohabitate","ResideElsewhere"))
SALARY1 <- cut(Male$SalaryX, c(0.5, 1.5,2.5,3.5,4.5), labels =
c("<1500RMB","1500-2500RMB","2500-3500RMB",">=3500RMB"))
NUMBEROFCITIESRESIDEDIN1 <- cut(Male$NumberOfCitiesResidedInX, c(0.5, 1.5,2.5), labels =
c("1-2", ">=3"))
DAILYWORKINGHOURS1 <- cut(Male$DailyWorkingHoursX, c(0.5, 1.5,2.5,3.5,4.5), labels =
c("=8","<8","8-11",">=11"))
WEEKLYWORKINGDAYS1 <- cut(Male$WeeklyWorkingDaysX, c(0.5, 1.5,2.5,3.5,4.5), labels =
c("<=4","5","6","7"))
RESIDENCE1 <- cut(Male$TypeOfResidence, c(0.5, 1.5,2.5,3.5,4.5,5.5), labels =
c("CollectiveDormitory", "Rent(with others)", "Rent(family)", "Rent(alone)",
"OwnLivingPlace"))
BMI1 <- cut(Male$BMIX, c(0.5, 1.5,2.5,3.5), labels = c("Underweight", "NormalWeight",
"Overweight"))
MENTALHEALTH1 <- cut(Male$MentalHealthX, c(0.5, 1.5,2.5), labels = c("Normal", "Abnormal"))

test1 <- glm(ScoreX~ AGE1 + OCCUPATION1 + WORKPLACESCALE1 + EDUCATIONATTAINMENT1 +
MARITALSTATUS1 + ACCOMPANYCHILD1 + SALARY1 + NUMBEROFCITIESRESIDEDIN1 + DAILYWORKINGHOURS1
+ WEEKLYWORKINGDAYS1 + RESIDENCE1 + BMI1 + MENTALHEALTH1, data = Male)
summary(test1)

cil <- confint(test1)

exp(cbind(COR = coef(test1), cil))

...

##### Male - significant risk factor COR visualization

```{r, message=FALSE, echo=FALSE}
COR = c(1.0, 1.0651334, 1.1669024, 0.9563699, 1.0786757, 1.3927085)
Occupation = c("1Manufacturing", "2Construction", "3Hospitality", "4DomesticServices",
"5SmallBusiness", "6Recreation/Leisures")
COROccupation = data.frame(COR,Occupation)

```



```

ggplot(COROccupation, aes(x = Occupation, y = COR, fill = Occupation)) + geom_bar(stat =
"identity", width = 0.8)
```

```{r, message=FALSE, echo=FALSE}
COR = c(1.0, 0.9919265, 1.1223432)
WorkplaceScale = c("1Large", "2Moderate", "3Small")
CORWorkplaceScale = data.frame(COR, WorkplaceScale)
ggplot(CORWorkplaceScale, aes(x = WorkplaceScale, y = COR, fill = WorkplaceScale)) +
geom_bar(stat = "identity")
```

```{r, message=FALSE, echo=FALSE}
COR = c(0.9747957, 1.0, 1.0039951, 1.1209372)
DailyWorkingHours = c("1.<8hr", "2.=8hr", "3.8-11hr", "4.>=11hr")
CORDailyWorkingHours = data.frame(COR, DailyWorkingHours)
ggplot(CORDailyWorkingHours, aes(x = DailyWorkingHours, y = COR, fill = DailyWorkingHours))
+ geom_bar(stat = "identity")
```

```{r, message=FALSE, echo=FALSE}
COR = c(1.0, 1.4310271)
MentalHealth = c("1Normal", "2Abnormal")
CORMentalHealth = data.frame(COR, MentalHealth)
ggplot(CORMentalHealth, aes(x = MentalHealth, y = COR, fill = MentalHealth)) + geom_bar(stat
= "identity")
```

```

### 3.2 For Female rural-to-urban migrants

```

```{r, message=FALSE, echo=FALSE}

Female <- dplyr::filter(healthData, GenderX == 2)

AGE2 <- cut(Female$AgeX, c(0.5, 1.5, 2.5), labels = c("<=32", ">32"))
OCCUPATION2 <- cut(Female$Occupation, c(0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5), labels =
c("Manufacturing", "Construction", "Hospitality", "DomesticServices", "SmallBusiness",
"Recreation/Leisures"))
WORKPLACESCALE2 <- cut(Female$WorkplaceScale, c(0.5, 1.5, 2.5, 3.5), labels =
c("Large", "Moderate", "Small"))
EDUCATIONATTAINMENT2 <- cut(Female$EducationAttainmentX, c(0.5, 1.5, 2.5, 3.5, 4.5), labels =
c("Elementary/Lower", "JuniorHigh", "HighSchool", "College"))
MARITALSTATUS2 <- cut(Female$MaritalStatus, c(0.5, 1.5, 2.5, 3.5, 4.5), labels = c("Married",
"Single", "Cohabiting", "Divorced/Widowed"))
ACCOMPANYCHIL2 <- cut(Female$AccompanyChildX, c(0.5, 1.5, 2.5, 3.5), labels =
c("NoChildren", "Cohabitate", "ResideElsewhere"))
SALARY2 <- cut(Female$SalaryX, c(0.5, 1.5, 2.5, 3.5, 4.5), labels =
c("<1500RMB", "1500-2500RMB", "2500-3500RMB", ">=3500RMB"))
NUMBEROFCITIESRESIDEDIN2 <- cut(Female$NumberOfCitiesResidedInX, c(0.5, 1.5, 2.5), labels =
c("1-2", ">=3"))
DAILYWORKINGHOURS2 <- cut(Female$DailyWorkingHoursX, c(0.5, 1.5, 2.5, 3.5, 4.5), labels =
c("=8", "<8", "8-11", ">=11"))
WEEKLYWORKINGDAYS2 <- cut(Female$WeeklyWorkingDaysX, c(0.5, 1.5, 2.5, 3.5, 4.5), labels =
c("<=4", "5", "6", "7"))
RESIDENCE2 <- cut(Female$TypeOfResidence, c(0.5, 1.5, 2.5, 3.5, 4.5, 5.5), labels =
c("CollectiveDormitory", "Rent(with others)", "Rent(family)", "Rent(alone)",
"OwnLivingPlace"))

```

```

BMI2 <- cut(Female$BMIX, c(0.5, 1.5,2.5,3.5), labels = c("Underweight", "NormalWeight",
"Overweight"))
MENTALHEALTH2 <- cut(Female$MentalHealthX, c(0.5, 1.5,2.5), labels = c("Normal", "Abnormal"))

test2 <- glm(ScoreX~ AGE2 + OCCUPATION2 + WORKPLACESCALE2 + EDUCATIONATTAINMENT2 +
MARITALSTATUS2 + ACCOMPANYCHILD2 + SALARY2 + NUMBEROFCITIESRESIDEDIN2 + DAILYWORKINGHOURS2
+ WEEKLYWORKINGDAYS2 + RESIDENCE2 + BMI2 + MENTALHEALTH2, data = Female)
summary(test2)

ci2 <- confint(test2)

exp(cbind(COR = coef(test2), ci2))

...

##### Female - significant risk factor COR visualization

```{r, message=FALSE, echo=FALSE}
COR = c(1.0, 0.9787387 , 1.1584592 , 1.1146430 , 1.1034555 , 1.6781331 )
Occupation = c("1Manufacturing", "2Construction", "3Hospitality", "4DomesticServices",
"5SmallBusiness", "6Recreation/Leisures")
COROccupation = data.frame(COR,Occupation)
ggplot(COROccupation, aes(x = Occupation, y = COR, fill = Occupation)) + geom_bar(stat =
"identity", width = 0.8)
```

```{r, message=FALSE, echo=FALSE}
COR = c(1.0, 1.2135993 , 1.1061900 , 1.1118359 )
MaritalStatus = c("1Married", "2Single","3Cohabiting","4Divorced/Widowed")
CORMaritalStatus = data.frame(COR,MaritalStatus)
ggplot(CORMaritalStatus, aes(x = MaritalStatus, y = COR, fill = MaritalStatus)) +
geom_bar(stat = "identity")
```

```{r, message=FALSE, echo=FALSE}
COR = c(1.0, 0.9193689 , 0.9299730 , 1.1547884 )
Salary = c("1.<1500RMB", "2.1500-2500RMB", "3.2500-3500RMB", "4.>=3500RMB")
CORSalary = data.frame(COR,Salary)
ggplot(CORSalary, aes(x = Salary, y = COR, fill = Salary)) + geom_bar(stat = "identity")
```

```{r, message=FALSE, echo=FALSE}
COR = c(1.0, 1.0890634 )
NumberOfCitiesResidedIn = c("1. 1-2 cities", "2. >=3 cities")
CORNumberOfCitiesResidedIn = data.frame(COR,NumberOfCitiesResidedIn)
ggplot(CORNumberOfCitiesResidedIn, aes(x = NumberOfCitiesResidedIn, y = COR, fill =
NumberOfCitiesResidedIn)) + geom_bar(stat = "identity")
```

```{r, message=FALSE, echo=FALSE}
COR = c(1.0464129 , 1.0, 1.0785105 , 1.0507422 )
DailyWorkingHours = c("1.<8hr", "2.=8hr", "3.8-11hr", "4.>=11hr")
CORDailyWorkingHours = data.frame(COR,DailyWorkingHours)
ggplot(CORDailyWorkingHours, aes(x = DailyWorkingHours, y = COR, fill = DailyWorkingHours))
+ geom_bar(stat = "identity")

```

```
```\n\n
```

```
```\n  {r, message=FALSE, echo=FALSE}\n  COR = c(1.0, 1.3779398 )\n  MentalHealth = c("1Normal", "2Abnormal")\n  CORMentalHealth = data.frame(COR,MentalHealth)\n  ggplot(CORMentalHealth, aes(x = MentalHealth, y = COR, fill = MentalHealth)) + geom_bar(stat\n= "identity")\n  ```\n\n
```