PMA Data Analysis

library(readr)  
PMA\_DATA <- read\_csv("PMA2018\_KER7\_HHQFQ\_2Dec2019.csv")

library(knitr)  
library(dplyr)  
library(ggplot2)  
library(reshape2)  
library(tidyr)  
library(stringr)  
library(broom)

## Some basic descriptives.

Finding out the number of women within the study criteria

kable(PMA\_DATA%>%   
 select(county,age,gender, current\_method) %>%   
 subset(gender==2) %>%   
 filter(between (age,15,49)) %>%   
 group\_by(county) %>%   
 summarise(total=n()) %>%   
 mutate(rate=(total/sum(total))\*100))

| county | total | rate |
| --- | --- | --- |
| 1 | 581 | 10.058864 |
| 2 | 542 | 9.383657 |
| 3 | 484 | 8.379501 |
| 4 | 587 | 10.162742 |
| 5 | 553 | 9.574100 |
| 6 | 589 | 10.197368 |
| 7 | 547 | 9.470222 |
| 8 | 490 | 8.483379 |
| 9 | 412 | 7.132964 |
| 10 | 517 | 8.950831 |
| 11 | 474 | 8.206371 |

## Most preferred family planning method

FP\_preferred<-PMA\_DATA %>%   
 select(age,gender, current\_method) %>%   
 mutate(age=as.numeric(age)) %>%   
 subset(gender==2) %>%   
 filter(between (age,15,49)) %>%   
 group\_by(current\_method) %>%   
 summarize(total=n()) %>%   
 drop\_na()  
FP\_preferred$current\_method[FP\_preferred$current\_method== "-99"]<-"No Response"

## Awareness to FP

Tabulate the data in long format test for independence of the variables

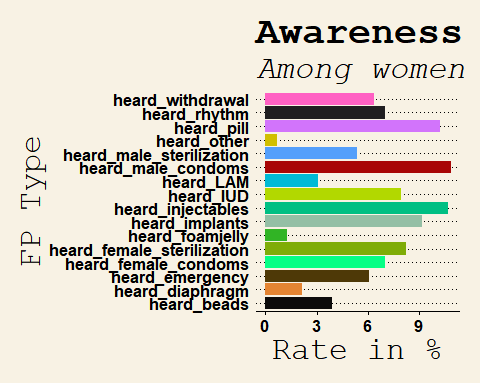
FP\_access<-PMA\_DATA %>%   
 select(age,gender, current\_method,heard\_female\_sterilization:heard\_other) %>%   
 mutate(age=as.numeric(age)) %>%   
 subset(gender==2) %>%   
 filter(between (age,15,49)) %>%   
 select(heard\_female\_sterilization:heard\_other) %>%   
 drop\_na()  
  
fpheard<-colnames(FP\_access)  
fpheard\_values<-tibble(colSums(FP\_access))  
accessfp<-data.frame(fpheard)  
d<-data.frame(accessfp,fpheard\_values)  
FP\_access2<-d %>%   
 mutate(Number\_of\_individuals=colSums(FP\_access),  
 Rate=round((Number\_of\_individuals/sum(Number\_of\_individuals))\*100,2),  
 Awareness\_Rate=paste0(Rate, "%"),  
 Not\_Aware=nrow(FP\_access)-Number\_of\_individuals) %>%  
 select(-colSums.FP\_access.)

Test for independence

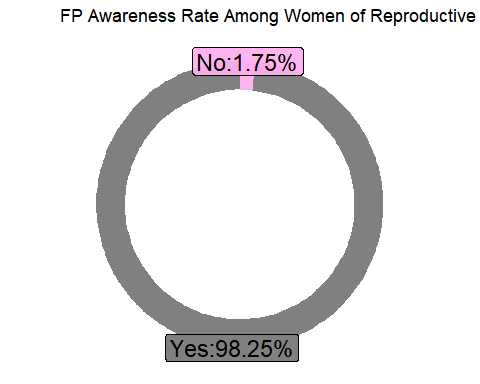
access\_long<-pivot\_longer(FP\_access, values\_to = "Heard\_of", names\_to = "FP\_Type", cols = (1:16))  
  
access\_prop<-access\_long [!(access\_long$Heard\_of== -99),]  
  
f<-table(access\_prop$FP\_Type,access\_prop$Heard\_of)  
  
chisq.test(f)

##   
## Pearson's Chi-squared test  
##   
## data: f  
## X-squared = 27217, df = 15, p-value < 2.2e-16

ggplot(FP\_access2) +  
 aes(x = fpheard, fill = fpheard, weight = Rate) +  
 geom\_bar() +  
 scale\_fill\_manual(values = c(heard\_beads = "#0C0B0B",   
 heard\_diaphragm = "#E48432", heard\_emergency = "#4E3907", heard\_female\_condoms = "#05FF85",  
 heard\_female\_sterilization = "#7FAC07",   
 heard\_foamjelly = "#31B425", heard\_implants = "#94BFA5", heard\_injectables = "#00BF83", heard\_IUD = "#B2D800",   
 heard\_LAM = "#00BAD5", heard\_male\_condoms = "#A80606", heard\_male\_sterilization = "#549FFB",   
 heard\_other = "#D1BF00", heard\_pill = "#D274FB", heard\_rhythm = "#201E20", heard\_withdrawal = "#FF61C3")) +  
 labs(x = "FP Type", y = "Rate in %", title = "Awareness of FP Methods",   
 subtitle = "Among women of ages 15-49") +  
 coord\_flip() +  
 ggthemes::theme\_wsj() +  
 theme(legend.position = "none", plot.title = element\_text(face = "bold"), plot.subtitle = element\_text(face = "italic"), axis.title = element\_text())



Awareness<-FP\_access %>%   
 mutate(Aware=as.numeric(if\_else(heard\_female\_sterilization|heard\_male\_sterilization|heard\_implants|heard\_IUD|  
 heard\_injectables|heard\_pill|heard\_emergency|heard\_male\_condoms|heard\_female\_condoms|  
 heard\_diaphragm|heard\_foamjelly|heard\_beads|heard\_LAM|heard\_rhythm|heard\_withdrawal  
 == "1", "1", "0")),  
 IsAware=if\_else(Aware=="1", "Yes","No")) %>%   
 group\_by(IsAware) %>%   
 summarise(Count=n()) %>%   
 mutate(Rate=round((Count/sum(Count))\*100,2),  
 Awareness\_Rate=paste0(Rate, "%"),  
 cumulative\_ratio=cumsum(Rate),  
 ration\_min=c(0, head(cumulative\_ratio,n=-1)),  
 labelPosition=(cumulative\_ratio+ration\_min)/2,  
 label=paste0(IsAware, ":", Awareness\_Rate))  
  
ggplot(Awareness,aes(ymax=cumulative\_ratio, ymin=ration\_min, xmax=4, xmin=3, fill=IsAware))+  
 geom\_rect()+  
 scale\_fill\_manual(values = c(No = "#FFAFEEEE",   
 Yes = "#808080"))+  
 geom\_label(x=4.0,aes(y=labelPosition, label=label),size=6)+  
 coord\_polar(theta="y")+  
 xlim(c(-1,4))+  
 labs(title="FP Awareness Rate Among Women of Reproductive Age(15-49)")+  
 theme\_void()+  
 theme(legend.position = "none")



## Logistic Regression

### Preping the data

FP\_Data<-PMA\_DATA %>%   
 mutate(age=as.numeric(age)) %>%   
 filter(between (age,15,49)) %>%   
 mutate(Aware=as.numeric(if\_else(heard\_female\_sterilization|heard\_male\_sterilization|heard\_implants|heard\_IUD|  
 heard\_injectables|heard\_pill|heard\_emergency|heard\_male\_condoms|heard\_female\_condoms|  
 heard\_diaphragm|heard\_foamjelly|heard\_beads|heard\_LAM|heard\_rhythm|heard\_withdrawal  
 == "1", "1", "0")),  
 IsAware=if\_else(Aware=="1", "Yes","No")) %>%   
 select(age,gender,current\_method,religion, method\_influences\_pro,ever\_birth, pregnant,   
 marital\_status, school, IsAware) %>%   
 subset(gender==2)  
  
FP\_Data$current\_method[FP\_Data$current\_method== "-99"]<-""  
FP\_Data$method\_influences\_pro[FP\_Data$method\_influences\_pro== "-77"]<-""  
FP\_Data$pregnant[FP\_Data$pregnant== "-88"]<-""  
FP\_Data$religion[FP\_Data$religion== "-77"]<-""  
FP\_Data$religion[FP\_Data$religion== "-88"]<-""  
FP\_Data$religion[FP\_Data$religion== "-99"]<-""  
  
FP\_Data\_Clean<-FP\_Data %>%   
 mutate(fp\_use=as.numeric(if\_else(current\_method=="NA","0", "1")),  
 been\_influenced=as.numeric(if\_else(method\_influences\_pro=="NA","0", "1")),  
 IsReligious=as.numeric(if\_else(religion=="NA","0","1")))  
  
FP\_Data\_Clean$fp\_use[is.na(FP\_Data\_Clean$fp\_use)]<-0  
FP\_Data\_Clean$been\_influenced[is.na(FP\_Data\_Clean$been\_influenced)]<-0  
FP\_Data\_Clean$IsReligious[is.na(FP\_Data\_Clean$IsReligious)]<-0  
FP\_Data\_Clean$pregnant[is.na(FP\_Data\_Clean$pregnant)]<-0  
FP\_Data\_Clean$ever\_birth[is.na(FP\_Data\_Clean$ever\_birth)]<-0  
  
Final\_FP<-FP\_Data\_Clean %>%   
 select(fp\_use,age,IsReligious,marital\_status,pregnant,been\_influenced,ever\_birth,school, IsAware) %>%   
 mutate(fp\_use=as.factor(fp\_use),  
 pregnant=as.numeric(pregnant))

### Logistic

model1<-glm(fp\_use~., family="binomial", data=Final\_FP)  
summary(model1)

##   
## Call:  
## glm(formula = fp\_use ~ ., family = "binomial", data = Final\_FP)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.461 0.000 0.000 0.356 1.029   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.839e+01 4.310e+02 -0.043 0.9660   
## age 1.273e-02 1.060e-02 1.201 0.2299   
## IsReligious -3.029e-01 1.924e-01 -1.574 0.1154   
## marital\_status -1.341e-01 5.538e-02 -2.422 0.0154 \*   
## pregnant -2.241e+01 1.201e+03 -0.019 0.9851   
## been\_influenced 2.996e+01 6.100e+02 0.049 0.9608   
## ever\_birth 1.390e+00 2.308e-01 6.022 1.72e-09 \*\*\*  
## school 3.111e-03 5.486e-02 0.057 0.9548   
## IsAwareYes -1.047e+01 4.328e+02 -0.024 0.9807   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 7792.4 on 5681 degrees of freedom  
## Residual deviance: 1564.8 on 5673 degrees of freedom  
## (94 observations deleted due to missingness)  
## AIC: 1582.8  
##   
## Number of Fisher Scoring iterations: 20

model2<-glm(fp\_use~school+ever\_birth+marital\_status, family="binomial", data=Final\_FP)  
summary(model2)

##   
## Call:  
## glm(formula = fp\_use ~ school + ever\_birth + marital\_status,   
## family = "binomial", data = Final\_FP)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.8236 -1.0841 -0.3486 1.0993 2.3796   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.06190 0.13939 -14.79 <2e-16 \*\*\*  
## school 0.31645 0.02436 12.99 <2e-16 \*\*\*  
## ever\_birth 2.13708 0.10639 20.09 <2e-16 \*\*\*  
## marital\_status -0.20502 0.02260 -9.07 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 7841.0 on 5721 degrees of freedom  
## Residual deviance: 6376.5 on 5718 degrees of freedom  
## (54 observations deleted due to missingness)  
## AIC: 6384.5  
##   
## Number of Fisher Scoring iterations: 4

model3<-glm(fp\_use~school+ever\_birth+marital\_status+IsReligious+IsAware, family="binomial", data=Final\_FP)  
summary(model3)

##   
## Call:  
## glm(formula = fp\_use ~ school + ever\_birth + marital\_status +   
## IsReligious + IsAware, family = "binomial", data = Final\_FP)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.8210 -1.0100 -0.3681 1.0671 2.6441   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.82316 1.01836 -5.718 1.08e-08 \*\*\*  
## school 0.29540 0.02458 12.017 < 2e-16 \*\*\*  
## ever\_birth 2.23500 0.11135 20.072 < 2e-16 \*\*\*  
## marital\_status -0.17203 0.02409 -7.142 9.19e-13 \*\*\*  
## IsReligious -0.37699 0.08936 -4.219 2.46e-05 \*\*\*  
## IsAwareYes 3.72985 1.01170 3.687 0.000227 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 7841.0 on 5721 degrees of freedom  
## Residual deviance: 6307.8 on 5716 degrees of freedom  
## (54 observations deleted due to missingness)  
## AIC: 6319.8  
##   
## Number of Fisher Scoring iterations: 7

tidy(model3)

## # A tibble: 6 × 5  
## term estimate std.error statistic p.value  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 (Intercept) -5.82 1.02 -5.72 1.08e- 8  
## 2 school 0.295 0.0246 12.0 2.90e-33  
## 3 ever\_birth 2.23 0.111 20.1 1.31e-89  
## 4 marital\_status -0.172 0.0241 -7.14 9.19e-13  
## 5 IsReligious -0.377 0.0894 -4.22 2.46e- 5  
## 6 IsAwareYes 3.73 1.01 3.69 2.27e- 4

library(sjPlot)  
library(sjmisc)  
library(sjlabelled)  
tab\_model(model3)

fp use

Predictors

Odds Ratios

CI

p

(Intercept)

0.00

0.00 – 0.01

<0.001

school

1.34

1.28 – 1.41

<0.001

ever birth

9.35

7.53 – 11.65

<0.001

marital status

0.84

0.80 – 0.88

<0.001

IsReligious

0.69

0.58 – 0.82

<0.001

IsAware [Yes]

41.67

9.10 – 738.95

<0.001

Observations

5722

R2 Tjur

0.237

FP<-Final\_FP %>%   
 select(fp\_use,IsAware) %>%   
 mutate(Aware=as.numeric(if\_else(IsAware=="Yes","1","0")),  
 fp=as.numeric(if\_else(fp\_use=="1","1","0"))) %>%   
 drop\_na()  
t.test(x=FP$Aware, y=FP$fp, alternative = "two.sided", conf.level = 0.95)

##   
## Welch Two Sample t-test  
##   
## data: FP$Aware and FP$fp  
## t = 80.443, df = 6515.7, p-value < 2.2e-16  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.5323172 0.5589096  
## sample estimates:  
## mean of x mean of y   
## 0.9825236 0.4369102

tidy(t.test(x=FP$Aware, y=FP$fp, alternative = "two.sided", conf.level = 0.95))

## # A tibble: 1 × 10  
## estimate estimate1 estimate2 statistic p.value parameter conf.low conf.high  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 0.546 0.983 0.437 80.4 0 6516. 0.532 0.559  
## # … with 2 more variables: method <chr>, alternative <chr>