Chi-Squared Test

# Create One DataSet for Person

person\_data <- PerAccPlu.wide  
  
person\_data <- merge(person\_data, PerADMN.wide[, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_PerADMN"))  
  
person\_data <- merge(person\_data, PerDatPlu.wide[, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_PerDatPlu"))  
  
person\_data <- merge(person\_data, PerDatSin.wide[, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_PerDatSing"))  
  
  
person\_data <- merge(person\_data, PerGenPlu.wide[, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_PerGenPlu"))  
  
  
person\_data <- merge(person\_data, PerNomPlu.wide[, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_PerNomPlu"))  
  
person\_data <- merge(person\_data, PerNomSin.wide[, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_PerNomSing"))

## Create Binary column

# List of columns to check  
columns\_to\_check <- c("tot.form", "tot.form\_PerADMN", "tot.form\_PerDatPlu", "tot.form\_PerDatSing", "tot.form\_PerGenPlu", "tot.form\_PerNomPlu", "tot.form\_PerNomSing")  
  
# Create the new binary column  
person\_data$multiform <- ifelse(rowSums(person\_data[, columns\_to\_check] > 1) > 0, 1, 0)

# Create one DataDet for Year

year\_data <- YrAccPlu.wide  
  
year\_data <- merge(year\_data, YrAccSin.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrAccSin"))  
  
year\_data <- merge(year\_data, YrADMN.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrADMN"))  
  
year\_data <- merge(year\_data, YrGenPlu.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrGenPlu"))  
  
year\_data <- merge(year\_data, YrGenSin.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrGenSin"))  
  
year\_data <- merge(year\_data, YrINSPlu.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrINSPlu"))  
  
year\_data <- merge(year\_data, YrLocPlu.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrLocPlu"))  
  
year\_data <- merge(year\_data, YrLocSin.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrLocSin"))  
  
year\_data <- merge(year\_data, YrNomPlu.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrNomPlu"))  
  
year\_data <- merge(year\_data, YrNomSin.wide [, c("Nid", "tot.form")], by="Nid", suffixes=c("", "\_YrNomSin"))

## Create Binary column

columns\_to\_check\_year <- c("tot.form", "tot.form\_YrAccSin", "tot.form\_YrADMN", "tot.form\_YrGenPlu", "tot.form\_YrGenSin", "tot.form\_YrINSPlu", "tot.form\_YrLocPlu", "tot.form\_YrLocSin", "tot.form\_YrNomPlu", "tot.form\_YrNomSin")  
  
# Create the new binary column for year\_data  
year\_data$multiform <- ifelse(rowSums(year\_data[, columns\_to\_check\_year] > 1) > 0, 1, 0)

# Chi Squared Test for Person

# Contingency table for Area with person\_data  
table\_area\_person <- table(person\_data$multiform, person\_data$Area)  
print(table\_area\_person)

##   
## East Poland West  
## 0 21 10 15  
## 1 4 6 6

# Chi-squared test for Area with person\_data  
#chisq\_area\_person <- chisq.test(table(person\_data$multiform, person\_data$Area))  
#print(chisq\_area\_person)  
  
chisq\_area\_person <- chisq.test(table(person\_data$multiform, person\_data$Area), simulate.p.value = TRUE, B = 2000)  
print(chisq\_area\_person)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: table(person\_data$multiform, person\_data$Area)  
## X-squared = 2.4822, df = NA, p-value = 0.3108

# Contingency table for Length with person\_data  
table\_length\_person <- table(person\_data$multiform, person\_data$Length)  
print(table\_length\_person)

##   
## 1 2 3 4  
## 0 10 20 11 5  
## 1 0 4 8 4

# Chi-squared test for Length with person\_data  
chisq\_length\_person <- chisq.test(table(person\_data$multiform, person\_data$Length), simulate.p.value = TRUE, B = 2000)  
print(chisq\_length\_person)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: table(person\_data$multiform, person\_data$Length)  
## X-squared = 8.7944, df = NA, p-value = 0.03598

# Contingency table for Gender.age with person\_data  
table\_gender\_age\_person <- table(person\_data$multiform, person\_data$Gender.age)  
print(table\_gender\_age\_person)

##   
## female 75 plus female under 75 male  
## 0 18 14 14  
## 1 6 4 6

# Chi-squared test for Gender.age with person\_data  
chisq\_gender\_age\_person <- chisq.test(table(person\_data$multiform, person\_data$Gender.age), simulate.p.value = TRUE, B = 2000)  
print(chisq\_gender\_age\_person)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: table(person\_data$multiform, person\_data$Gender.age)  
## X-squared = 0.31262, df = NA, p-value = 0.8741

# Chi Squared Test for Year

# Contingency table for Area with year\_data  
table\_area\_year <- table(year\_data$multiform, year\_data$Area)  
print(table\_area\_year)

##   
## East Poland West  
## 0 17 10 10  
## 1 8 6 11

# Chi-squared test for Area with year\_data  
chisq\_area\_year <- chisq.test(table(year\_data$multiform, year\_data$Area))  
print(chisq\_area\_year)

##   
## Pearson's Chi-squared test  
##   
## data: table(year\_data$multiform, year\_data$Area)  
## X-squared = 2.0415, df = 2, p-value = 0.3603

# Contingency table for Length with year\_data  
table\_length\_year <- table(year\_data$multiform, year\_data$Length)  
print(table\_length\_year)

##   
## 1 2 3 4  
## 0 9 18 8 2  
## 1 1 6 11 7

# Chi-squared test for Length with year\_data  
chisq\_length\_year <- chisq.test(table(year\_data$multiform, year\_data$Length), simulate.p.value = TRUE, B = 2000)  
print(chisq\_length\_year)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: table(year\_data$multiform, year\_data$Length)  
## X-squared = 13.848, df = NA, p-value = 0.001999

# Contingency table for Gender.age with year\_data  
table\_gender\_age\_year <- table(year\_data$multiform, year\_data$Gender.age)  
print(table\_gender\_age\_year)

##   
## female 75 plus female under 75 male  
## 0 13 12 12  
## 1 11 6 8

# Chi-squared test for Gender.age with year\_data  
chisq\_gender\_age\_year <- chisq.test(table(year\_data$multiform, year\_data$Gender.age))  
print(chisq\_gender\_age\_year)

##   
## Pearson's Chi-squared test  
##   
## data: table(year\_data$multiform, year\_data$Gender.age)  
## X-squared = 0.66915, df = 2, p-value = 0.7156