TA

CK

23/11/2018

```
# Import ATUS Summary file
ATUS.SUM <- import('atussum_0317.csv', setclass = "tibble")
# Import ATUS CPS File
ATUS.CPS <- import('atuscps_0317.csv', setclass = "tibble")
# Create a diary date column
ATUS.SUM$TUDIARYDATE <- as.Date(paste0(substr(ATUS.SUM$TUCASEID, 1, 6), '01'), "%Y%m%d")
# Add months as a column as a derivative of the TUCASEID variable
ATUS.SUM$TUMONTH <- as.numeric(substr(ATUS.SUM$TUCASEID, 5, 6))
# Get birth year as a variable for purpose of classifying generation
ATUS.SUM$TUBIRTHYEAR <- ATUS.SUM$TUYEAR - ATUS.SUM$TEAGE
#Classify Generations
ATUS.SUM$TUGENERATION <- ifelse(ATUS.SUM$TUBIRTHYEAR < 1946, "Silent Generation",
                                ifelse(ATUS.SUM$TUBIRTHYEAR < 1965, "Baby Boomers",
                                       ifelse(ATUS.SUM$TUBIRTHYEAR < 1981, "Generation X", "Millenials"]
# Filter CPS down to just relevant variables. GEDSTFIPS is state, GEDIV is division and GEREG is region
ATUS.CPS.States <- select(ATUS.CPS, TUCASEID, GESTFIPS, GEDIV, GEREG)
ATUS.CPS.States <- distinct(ATUS.CPS.States)
# Create variables to be added to a data frame with relevant gender activities
Working <- ATUS.SUM[, grep("^t05", names(ATUS.SUM))]
House.Maintenance <- ATUS.SUM[, grep("^t0204", names(ATUS.SUM))]
Vehicle.Maintenance <- ATUS.SUM[, grep("^t0207", names(ATUS.SUM))]
Housework <- ATUS.SUM[, grep("^t0201", names(ATUS.SUM))]
Food.Preparation <- ATUS.SUM[, grep("^t0202", names(ATUS.SUM))] + ATUS.SUM$t070101 + ATUS.SUM$t070103
Childcare <- ATUS.SUM[, grep("^t03", names(ATUS.SUM))]
```

```
# Create a data frame from relevant variables within ATUS Summary
# Childnum there for future use
Gender.Roles.Data <- select(ATUS.SUM, TUCASEID, TESEX, TEAGE, TUFNWGTP, TUYEAR, TRCHILDNUM, TUMONTH, TU
# Join this dataset with the CPS states info by unique id number of particiapnt
# Add in relevant categories
Gender.Roles.Data$Working <- rowSums(Working)
Gender.Roles.Data$House.Maintenance <- rowSums(House.Maintenance)
Gender.Roles.Data$Vehicle.Maintenance <- rowSums(Vehicle.Maintenance)
Gender.Roles.Data$Housework <- rowSums(Housework)
Gender.Roles.Data$Food.Preparation <- rowSums(Food.Preparation)
Gender.Roles.Data$Childcare <- rowSums(select(Childcare, t030101:t030399))
# Join this dataset with the CPS states info by unique id number of particiapnt
Gender.Roles.Data <- inner_join(Gender.Roles.Data, ATUS.CPS.States, by = 'TUCASEID')
Gender.Roles.Data <- filter(Gender.Roles.Data)
Gender.Roles.Data$TESEX[Gender.Roles.Data$TESEX == 1] <- "M"
Gender.Roles.Data$TESEX[Gender.Roles.Data$TESEX == 2] <- "F"
```

```
weighted.valuesM <- c()
weighted.valuesF <- c()
for (variable in c("Working", "House.Maintenance", "Vehicle.Maintenance", "Housework", "Food.Preparatio
    Gender.Roles.Data[,paste0('weighted.', variable)] <- 0
    for (sex in c("M", "F")) {
      for (region in 1:4) {
```

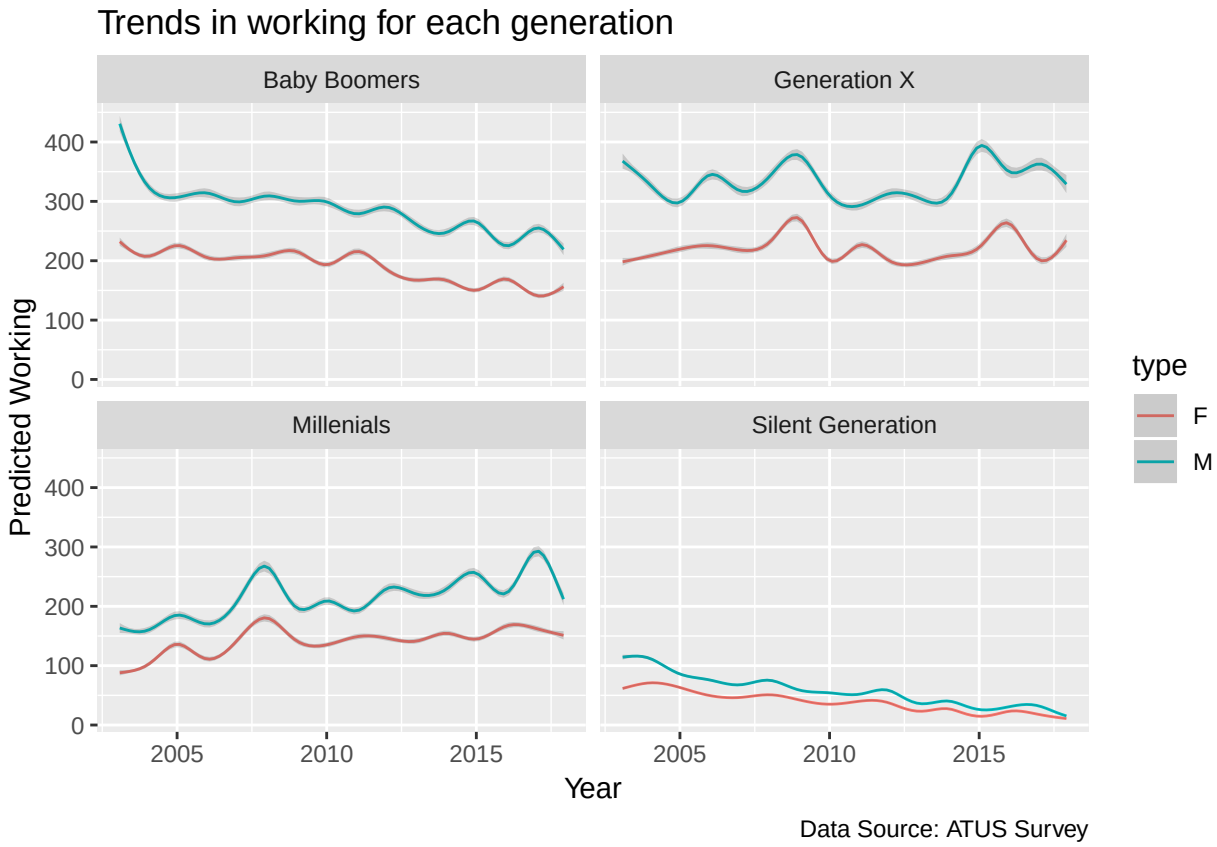
```
        for (year in unique(Gender.Roles.Data$TUYEAR)) {
            for (month in unique(Gender.Roles.Data$TUMONTH)) {
              for(generation in c("Silent Generation", "Baby Boomers", "Generation X", "Millenials")){
                Filtered.Data <- filter(Gender.Roles.Data, TESEX == sex, TUYEAR == year, TUMONTH == mon
                bottom.sum <- sum(Filtered.Data$TUFNWGTP)
                top.sum <- sum(Filtered.Data$TUFNWGTP * Filtered.Data[,variable])
                weighted.values <- top.sum / bottom.sum

                Gender.Roles.Data[Gender.Roles.Data$TESEX == sex & Gender.Roles.Data$TUYEAR == year & G
                if (Filtered.Data$TESEX[1] == "M") {
                  weighted.valuesM <- c(weighted.valuesM, weighted.values)
                } else {
                  weighted.valuesF <- c(weighted.valuesF, weighted.values)
                  }
                }
              }
            }
        }
    }
}
Difference <- abs(weighted.valuesM - weighted.valuesF)
```

```
Gender.Differences <- data.frame(Year = rep(rep(2003:2017, each = 48), 4), Month = rep(rep(1:12, each =
Gender.Differences <- data.frame(Gender.Differences, Difference[1:2880], Difference[2881:5760], Differe
Gender.Roles.Train <- filter(Gender.Roles.Data, TUMONTH %% 2 == 0)
Gender.Roles.Validate <- filter(Gender.Roles.Data, TUMONTH %% 2 != 0)
```

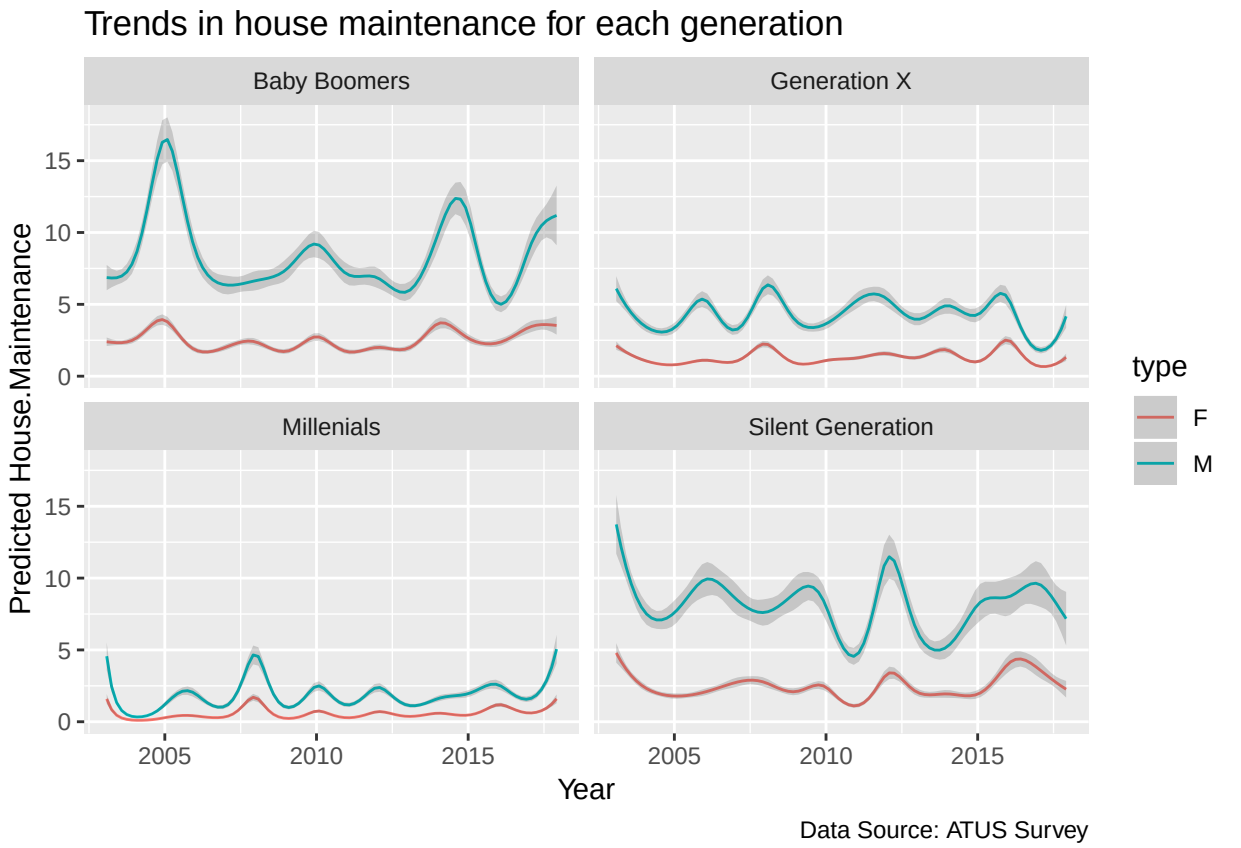
```
# Make a variable for all years.
Dates = c()
for (i in 4:17) {
    if (i < 10) {
        Dates[i-3] <- paste(paste0("200", i), "01", "01", sep = "-")
    } else {
        Dates[i-3] <- paste(paste0("20", i), "01", "01", sep = "-")
    }
}
# Commented out lines show how to introduce knots to the models
for (variable in c("Working", "House.Maintenance", "Vehicle.Maintenance", "Housework", "Food.Preparatio
    # Start building models for the data
    n <- nrow(Gender.Roles.Train)
    # model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE, knots = (c(as.Date("2006
    model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE)')), data = Gender.Roles.Tra
    model2 <- update(model1, . ~ . + TESEX)
    # model3 <- update(model1, . ~ -1 + TESEX + TESEX:ns(TUDIARYDATE, knots = (c(as.Date('2006-01-01'),
    model3 <- update(model1, . ~ -1 + TESEX + TESEX:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    model4 <- update(model3, .~. + GEREG + GEREG:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    model5 <- update(model3, .~. + TUGENERATION + TUGENERATION:ns(TUDIARYDATE, knots = c(as.Date(Dates)
    # model5 <- update(model3, .~. + GEREG + GEREG:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    # Use same method as in lab 5 to plot them
    # change the model number in the predict functions below for different models
    male.prediction <- predict.glm(model5, newdata = filter(Gender.Roles.Train, TESEX == 'M'), se = TRU
    female.prediction <- predict.glm(model5, newdata = filter(Gender.Roles.Train, TESEX == 'F'), se = T
    df <- data.frame(age = c(filter(Gender.Roles.Train, TESEX == 'M')$TEAGE, filter(Gender.Roles.Train,
    model3_plot <- ggplot(df, aes(x=date, y=prediction, colour=type)) +
```

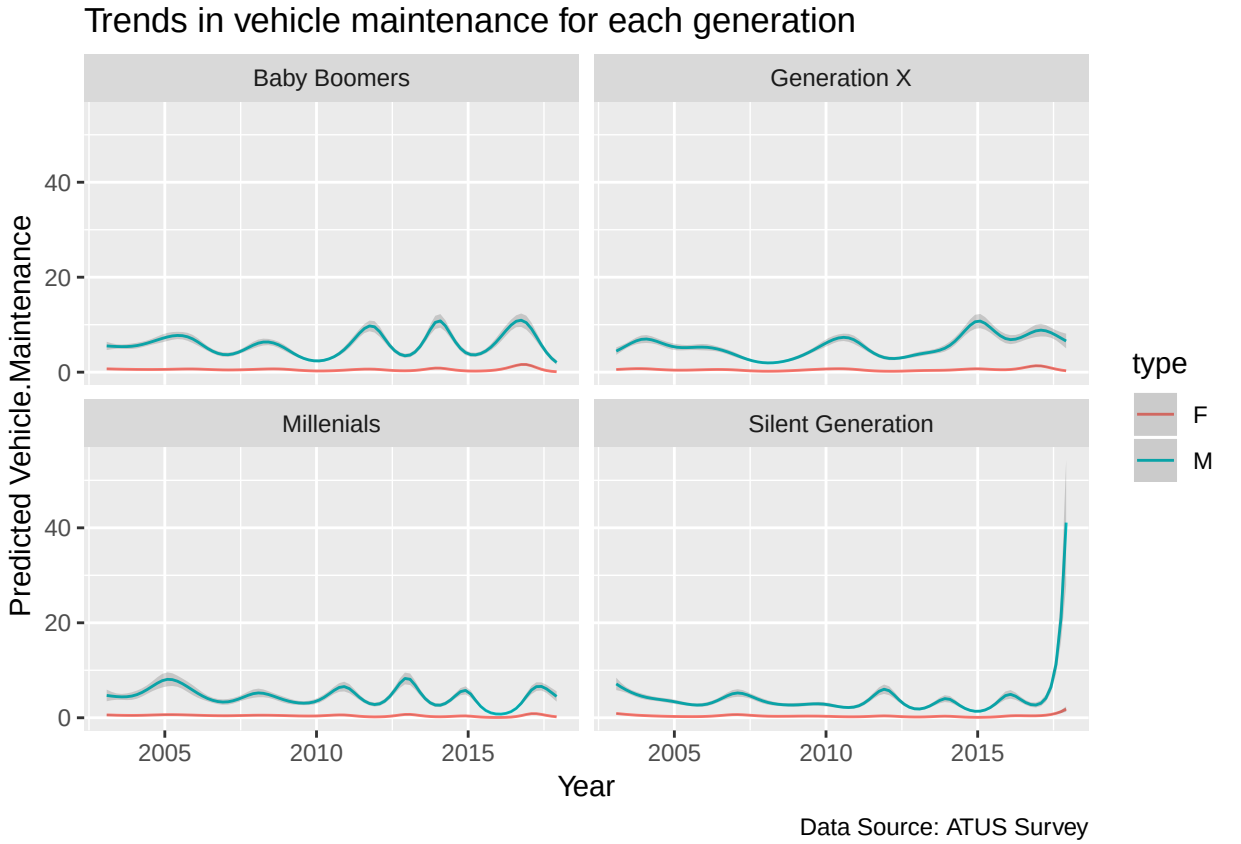
```
    geom_line() +
    geom_ribbon(aes(ymin = prediction - prediction_error, ymax = prediction + prediction_error), alpha =
    print(model3_plot)
}
```

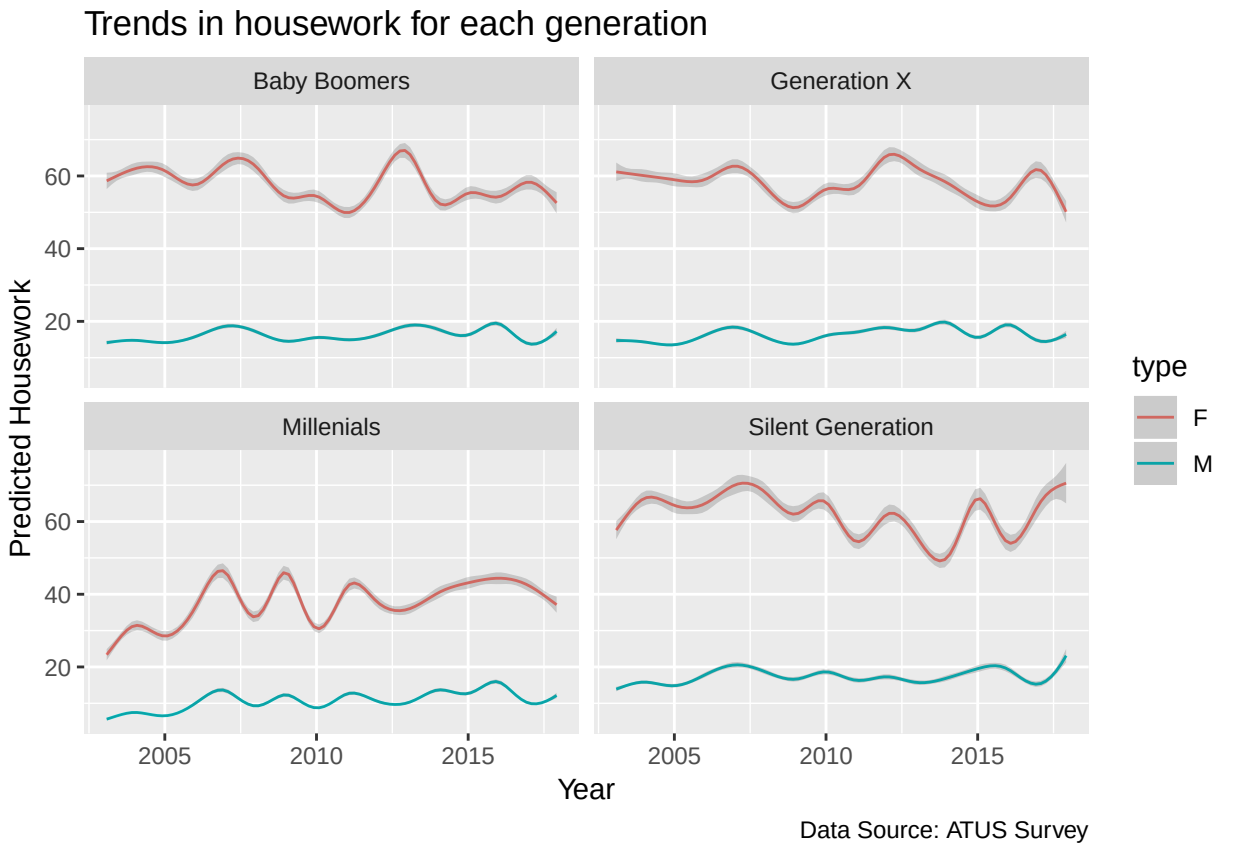
```
## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge
```

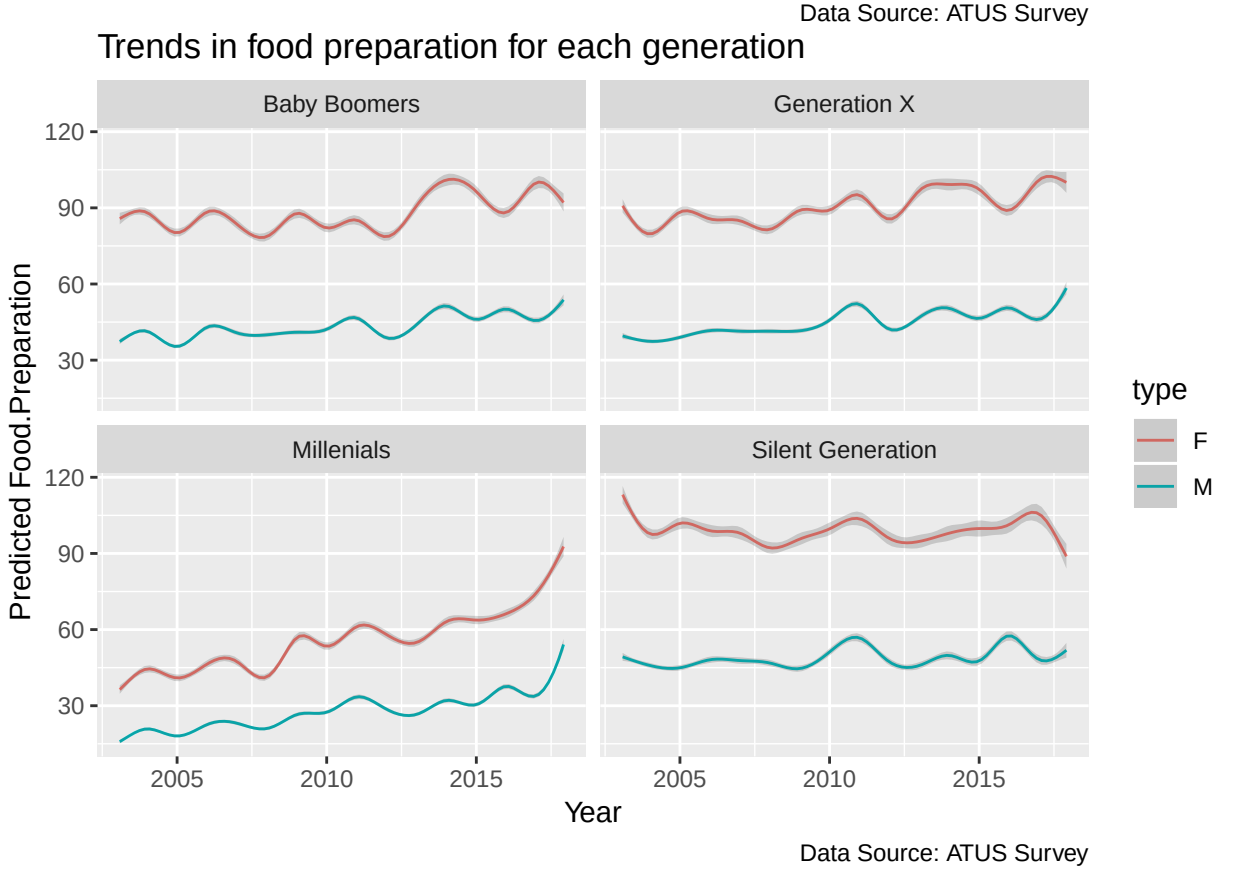


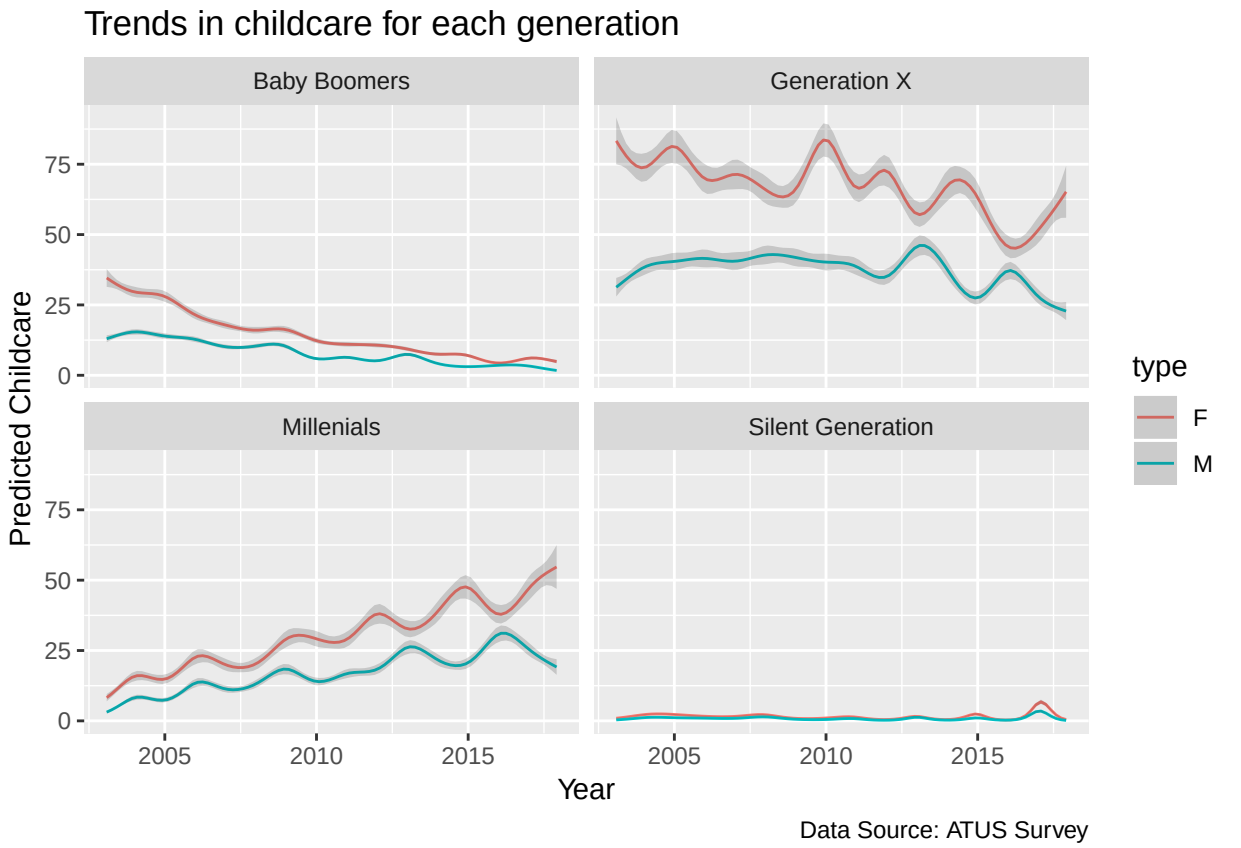

```
## Warning: glm.fit: algorithm did not converge
```


Trends in house maintenance for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted House.Maintenance
0
5
10
15
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in vehicle maintenance for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Vehicle.Maintenance
0
20
40
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in housework for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Housework
60
40
20
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in food preparation for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Food.Preparation
120
90
60
30
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey



Validate the data

```
# Make a variable for all years.
Dates <- c()
for (i in 4:17) {
    if (i < 10) {
        Dates[i-3] <- paste(paste0("200", i), "01", "01", sep = "-")
    } else {
        Dates[i-3] <- paste(paste0("20", i), "01", "01", sep = "-")
    }
}
# Commented out lines show how to introduce knots to the models
for (variable in c("Working", "House.Maintenance", "Vehicle.Maintenance", "Housework", "Food.Preparatio
    # Start building models for the data
    n <- nrow(Gender.Roles.Validate)
    # model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE, knots = (c(as.Date("2006
    model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE)')), data = Gender.Roles.Va
    model2 <- update(model1, . ~ . + TESEX)
    # model3 <- update(model1, . ~ -1 + TESEX + TESEX:ns(TUDIARYDATE, knots = (c(as.Date('2006-01-01'),
    model3 <- update(model1, . ~ -1 + TESEX + TESEX:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    model4 <- update(model3, .~. + GEREG + GEREG:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    model5 <- update(model3, .~. + TUGENERATION + TUGENERATION:ns(TUDIARYDATE, knots = c(as.Date(Dates)
    # model5 <- update(model3, .~. + GEREG + GEREG:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    # Use same method as in lab 5 to plot them
    # change the model number in the predict functions below for different models
    male.prediction <- predict.glm(model5, newdata = filter(Gender.Roles.Validate, TESEX == 'M'), se =
    female.prediction <- predict.glm(model5, newdata = filter(Gender.Roles.Validate, TESEX == 'F'), se
    df <- data.frame(age = c(filter(Gender.Roles.Validate, TESEX == 'M')$TEAGE, filter(Gender.Roles.Vali
```

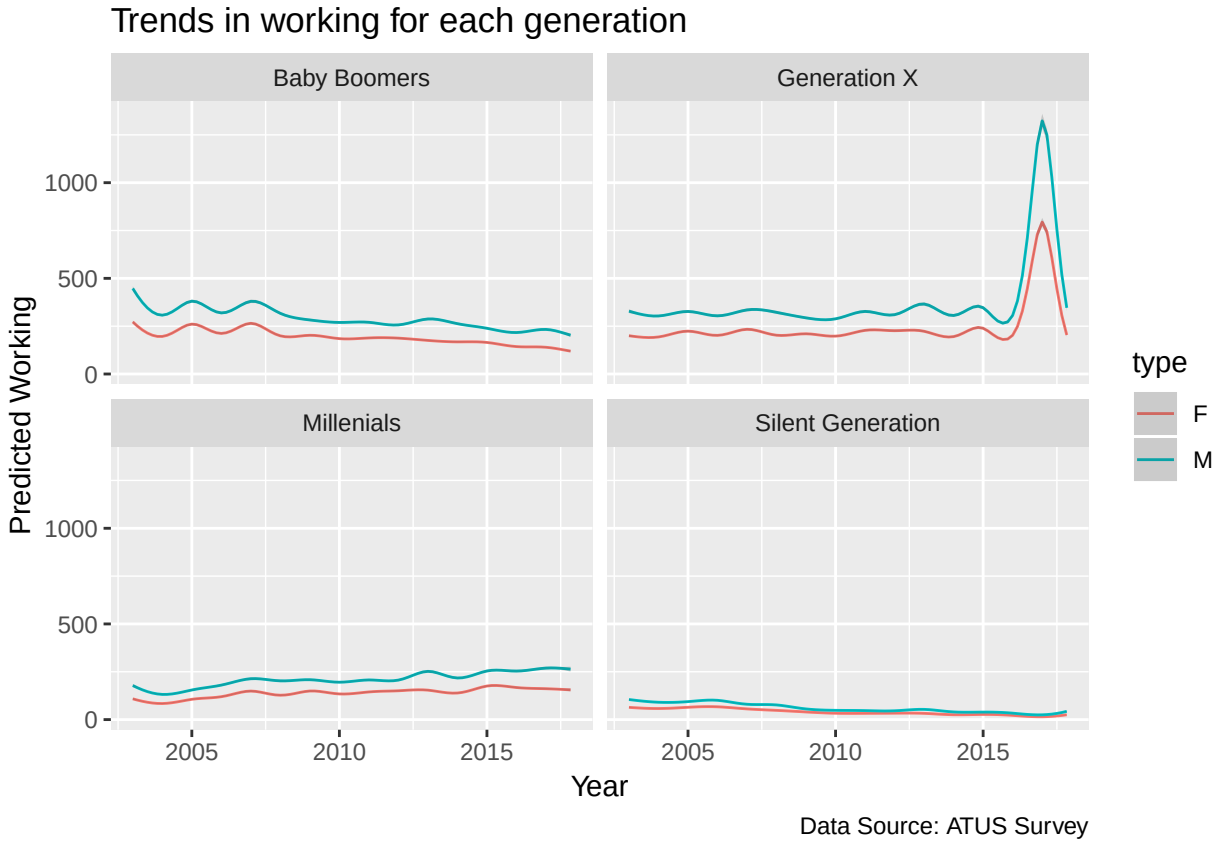
```
    model3_plot <- ggplot(df, aes(x=date, y=prediction, colour=type)) +
    geom_line() +
    geom_ribbon(aes(ymin = prediction - prediction_error, ymax = prediction + prediction_error), alpha =
    print(model3_plot)
}
```

```
## Warning: glm.fit: algorithm did not converge
```

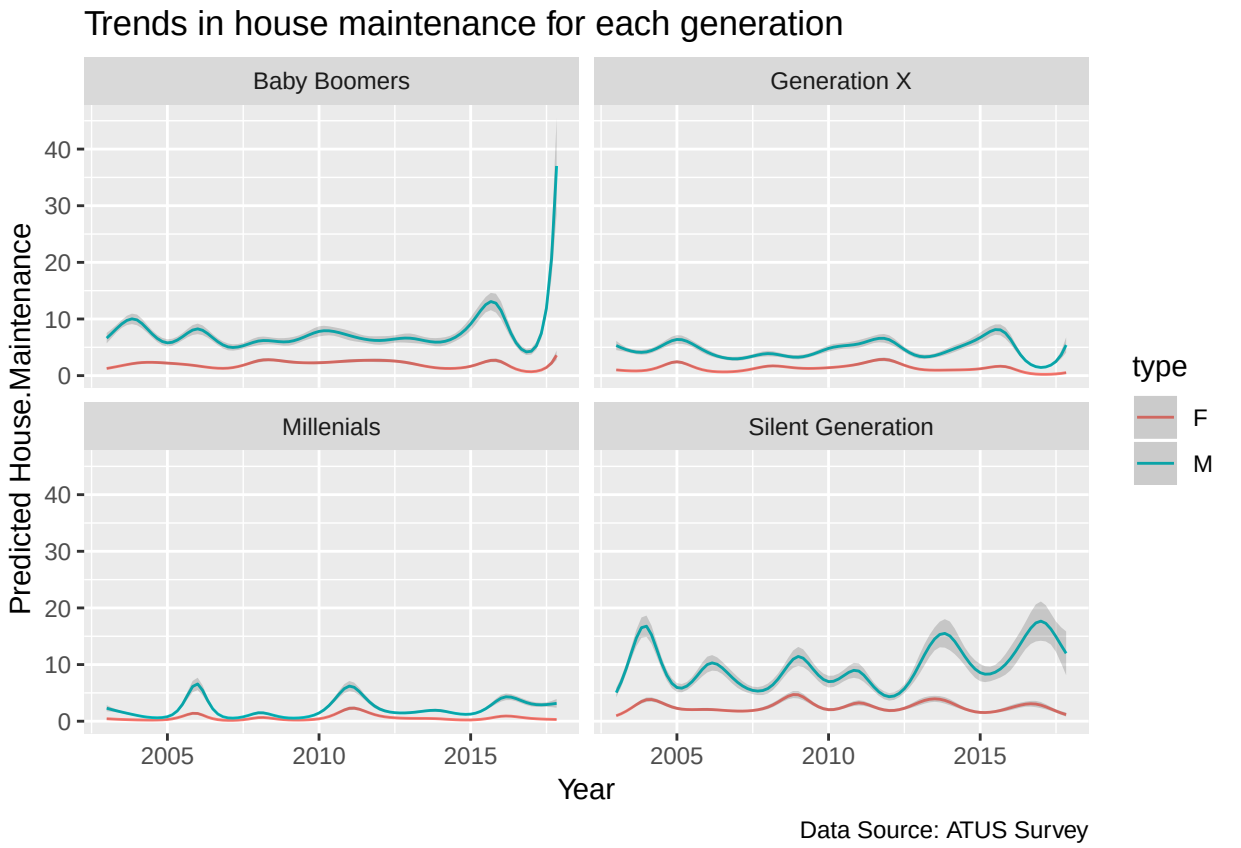
```
## Warning: glm.fit: algorithm did not converge
```

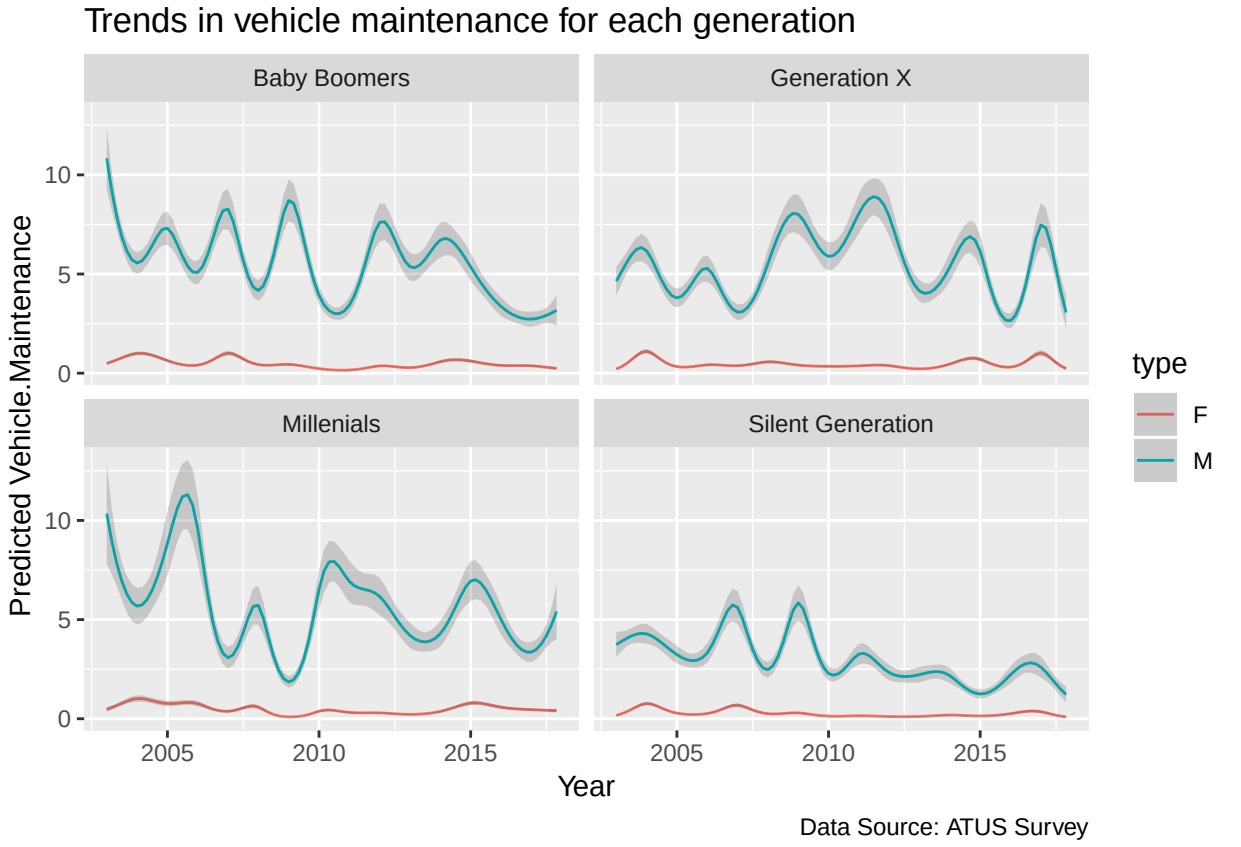
```
## Warning: glm.fit: algorithm did not converge
```

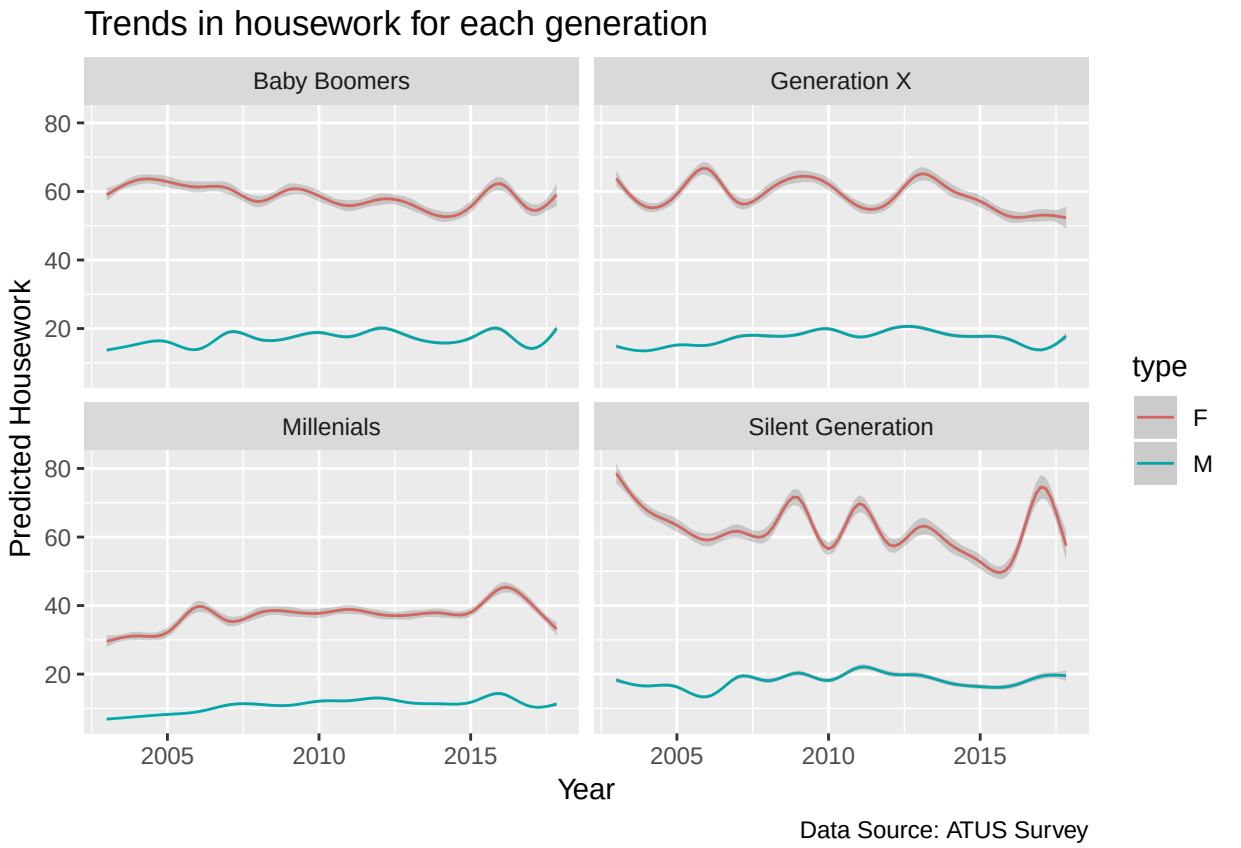
```
## Warning: glm.fit: algorithm did not converge
```

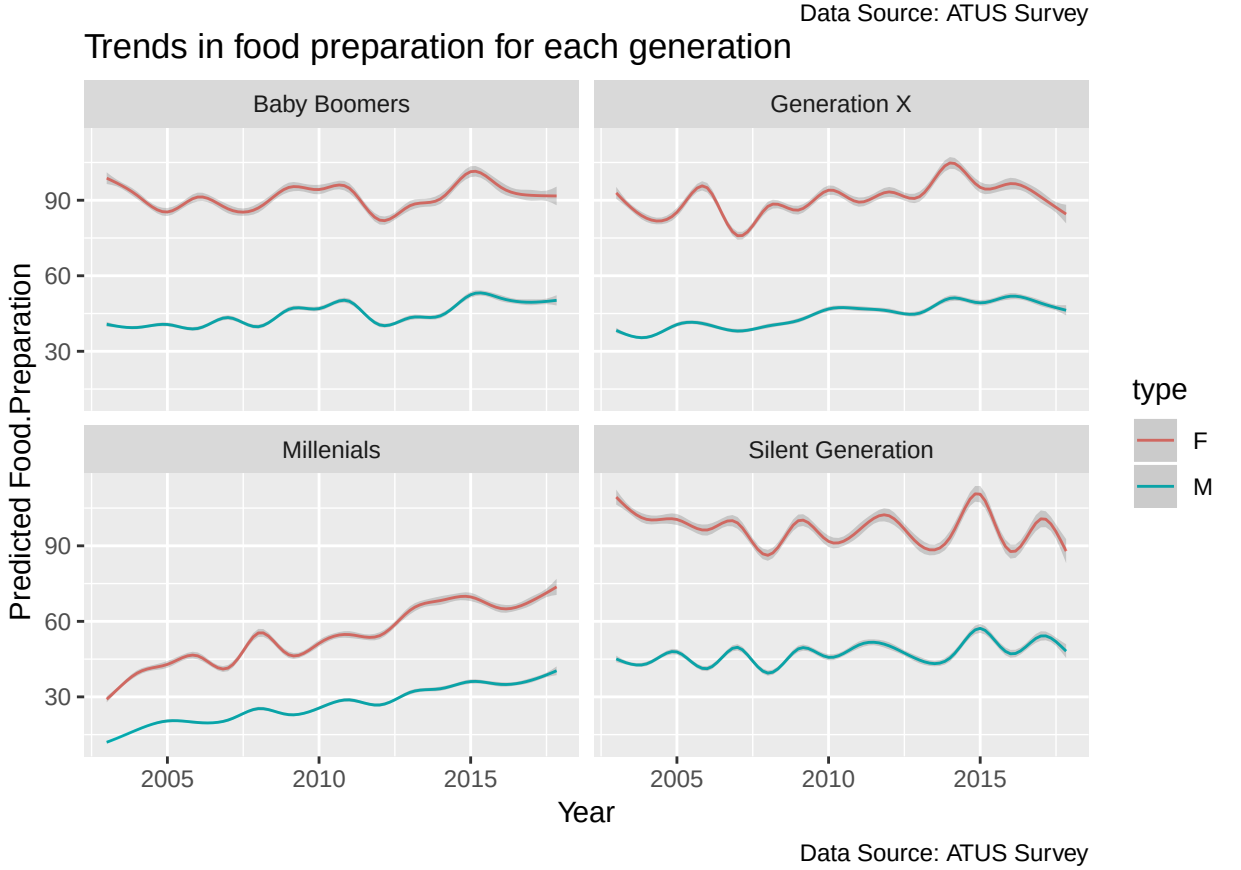


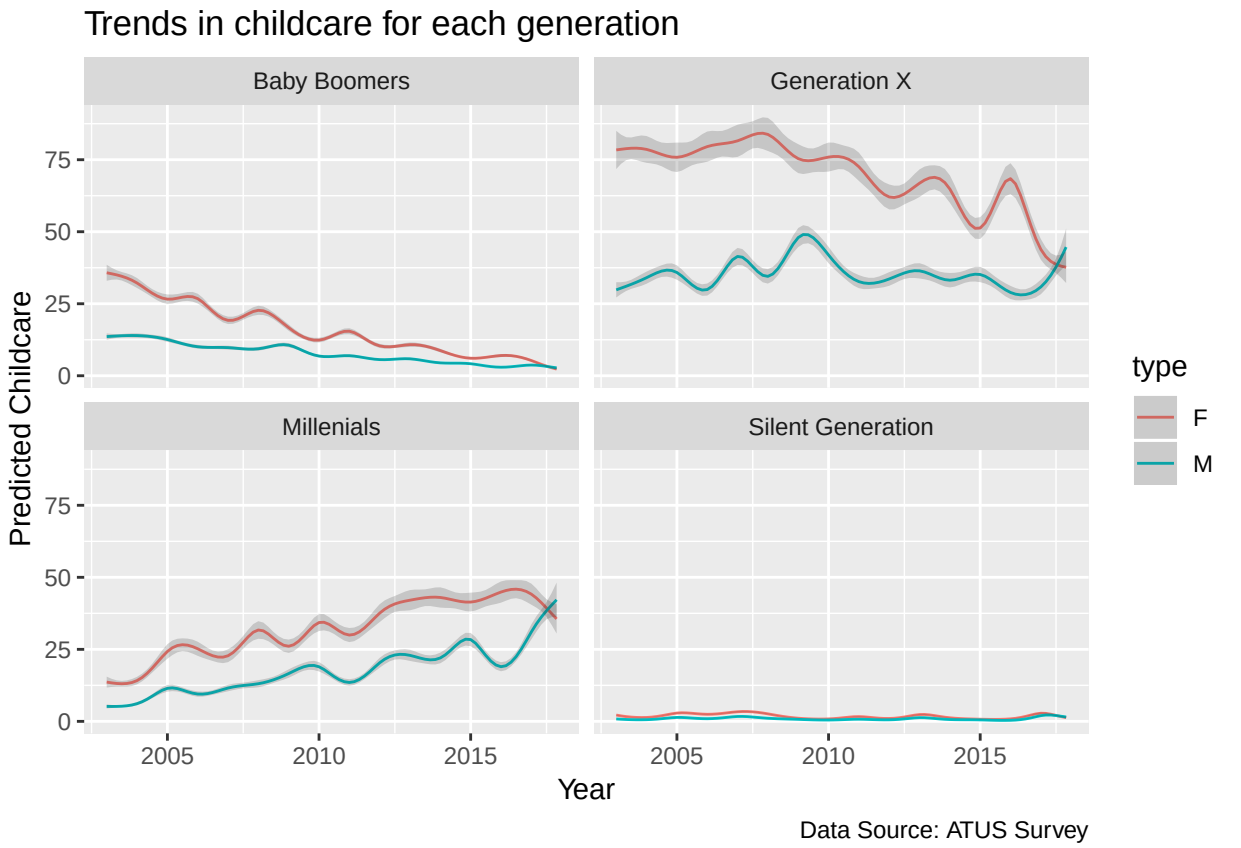

```
## Warning: glm.fit: algorithm did not converge
```


Trends in house maintenance for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted House.Maintenance
0
10
20
30
40
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in vehicle maintenance for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Vehicle.Maintenance
0
5
10
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in housework for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Housework
80
60
40
20
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in food preparation for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Food.Preparation
90
60
30
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey

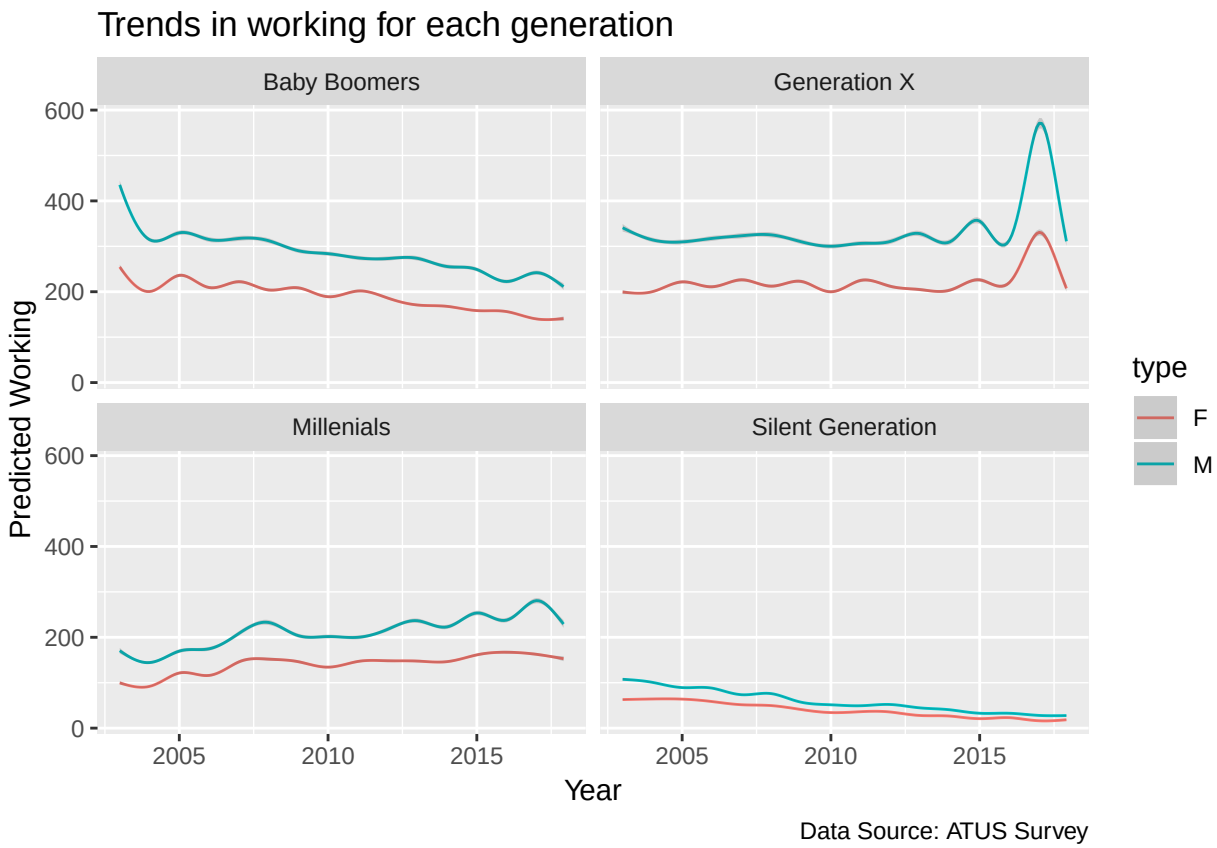


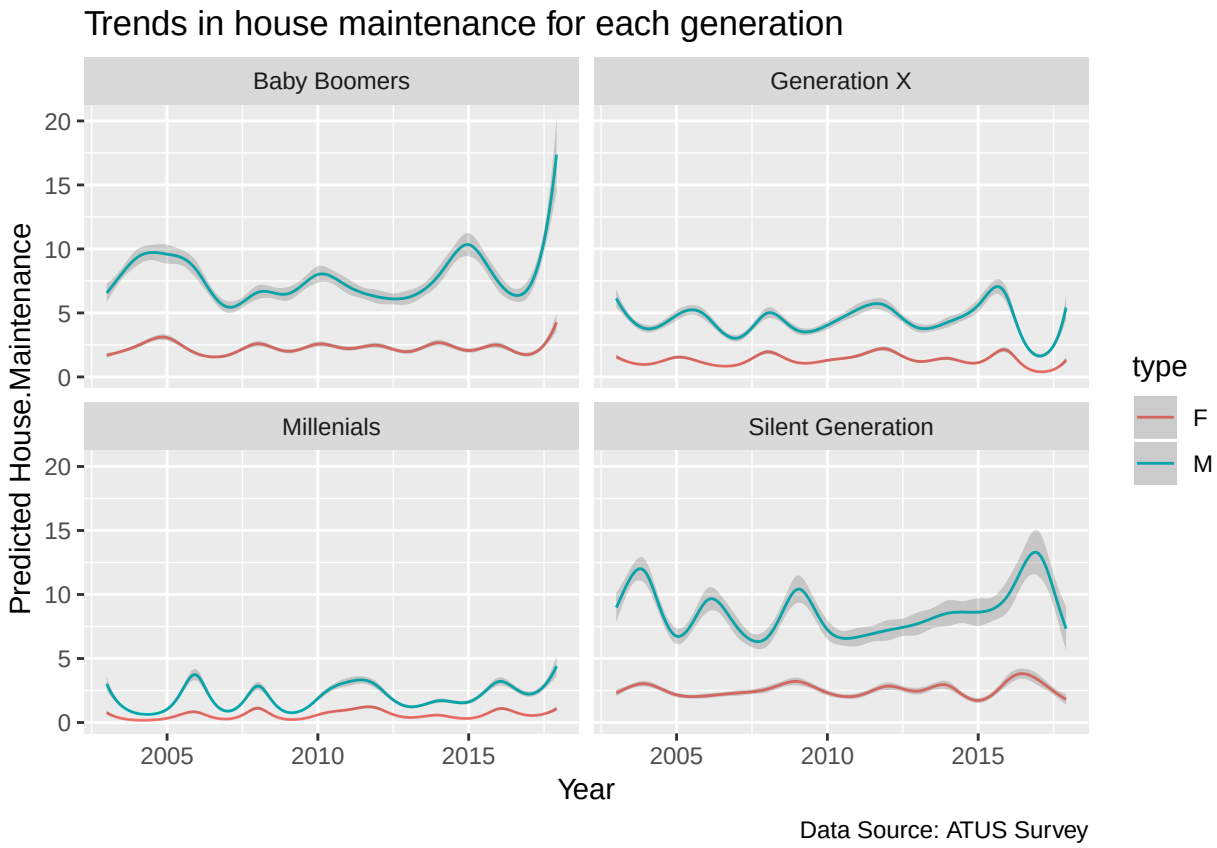
Do on entire dataset

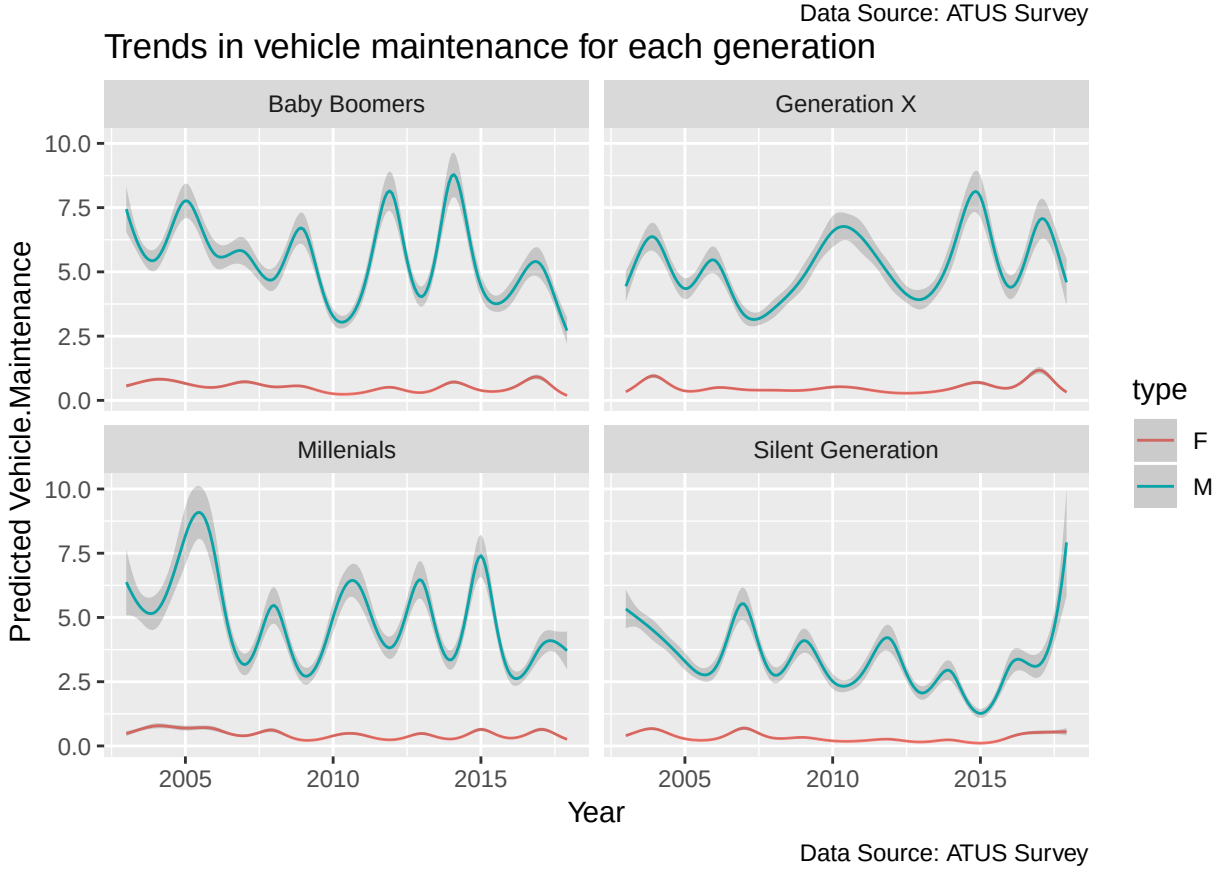
```
# Make a variable for all years.
Dates <- c()
for (i in 4:17) {
    if (i < 10) {
        Dates[i-3] <- paste(paste0("200", i), "01", "01", sep = "-")
    } else {
        Dates[i-3] <- paste(paste0("20", i), "01", "01", sep = "-")
    }
}
# Commented out lines show how to introduce knots to the models
for (variable in c("Working", "House.Maintenance", "Vehicle.Maintenance", "Housework", "Food.Preparatio
    # Start building models for the data
    n <- nrow(Gender.Roles.Data)
    # model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE, knots = (c(as.Date("2006
    model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE)')), data = Gender.Roles.Da
    model2 <- update(model1, . ~ . + TESEX)
    # model3 <- update(model1, . ~ -1 + TESEX + TESEX:ns(TUDIARYDATE, knots = (c(as.Date('2006-01-01'),
    model3 <- update(model1, . ~ -1 + TESEX + TESEX:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    model4 <- update(model3, .~. + GEREG + GEREG:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    model5 <- update(model3, .~. + TUGENERATION + TUGENERATION:ns(TUDIARYDATE, knots = c(as.Date(Dates)
    # model5 <- update(model3, .~. + GEREG + GEREG:ns(TUDIARYDATE, knots = c(as.Date(Dates))))
    # Use same method as in lab 5 to plot them
    # change the model number in the predict functions below for different models
    male.prediction <- predict.glm(model5, newdata = filter(Gender.Roles.Data, TESEX == 'M'), se = TRUE
    female.prediction <- predict.glm(model5, newdata = filter(Gender.Roles.Data, TESEX == 'F'), se = TRU
    df <- data.frame(age = c(filter(Gender.Roles.Data, TESEX == 'M')$TEAGE, filter(Gender.Roles.Data, TE
```

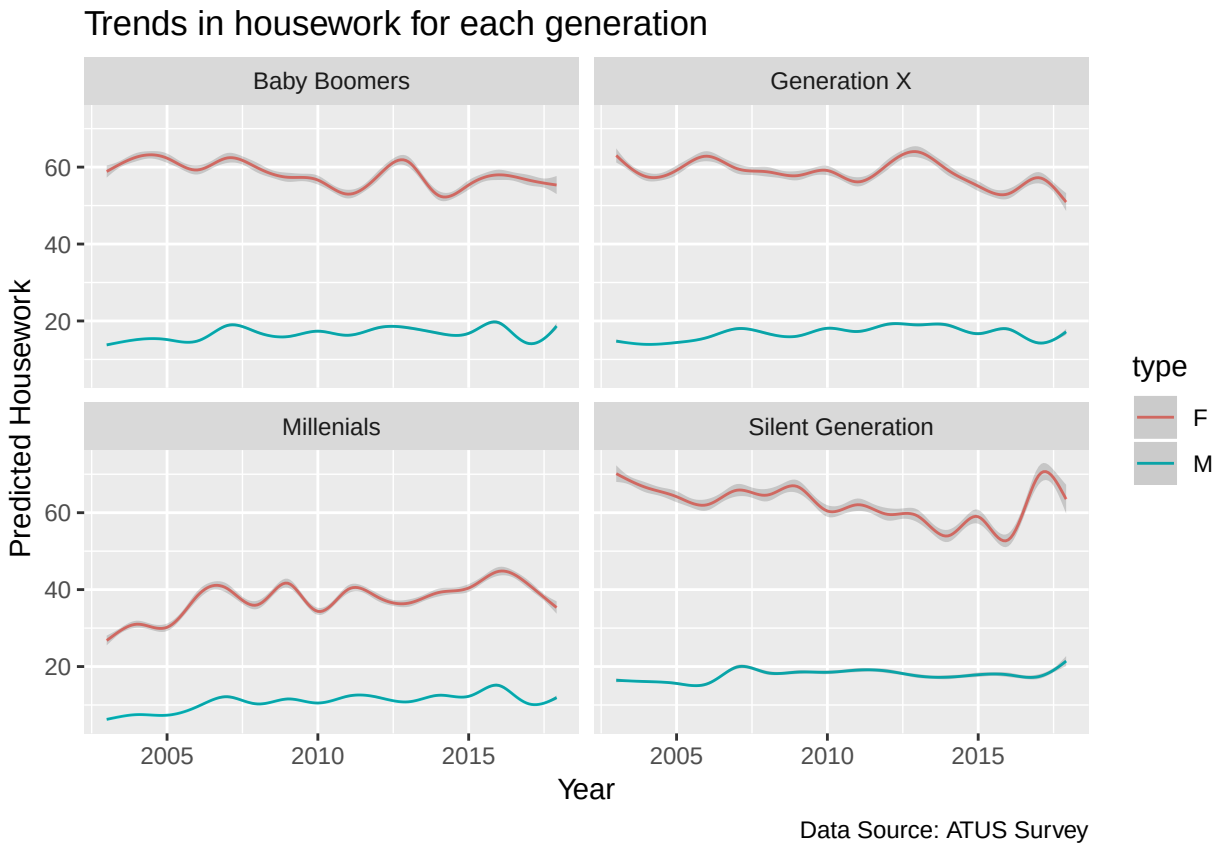
```
    model3_plot <- ggplot(df, aes(x=date, y=prediction, colour=type)) +
    geom_line() +
    geom_ribbon(aes(ymin = prediction - prediction_error, ymax = prediction + prediction_error), alpha =
    print(model3_plot)
}
```

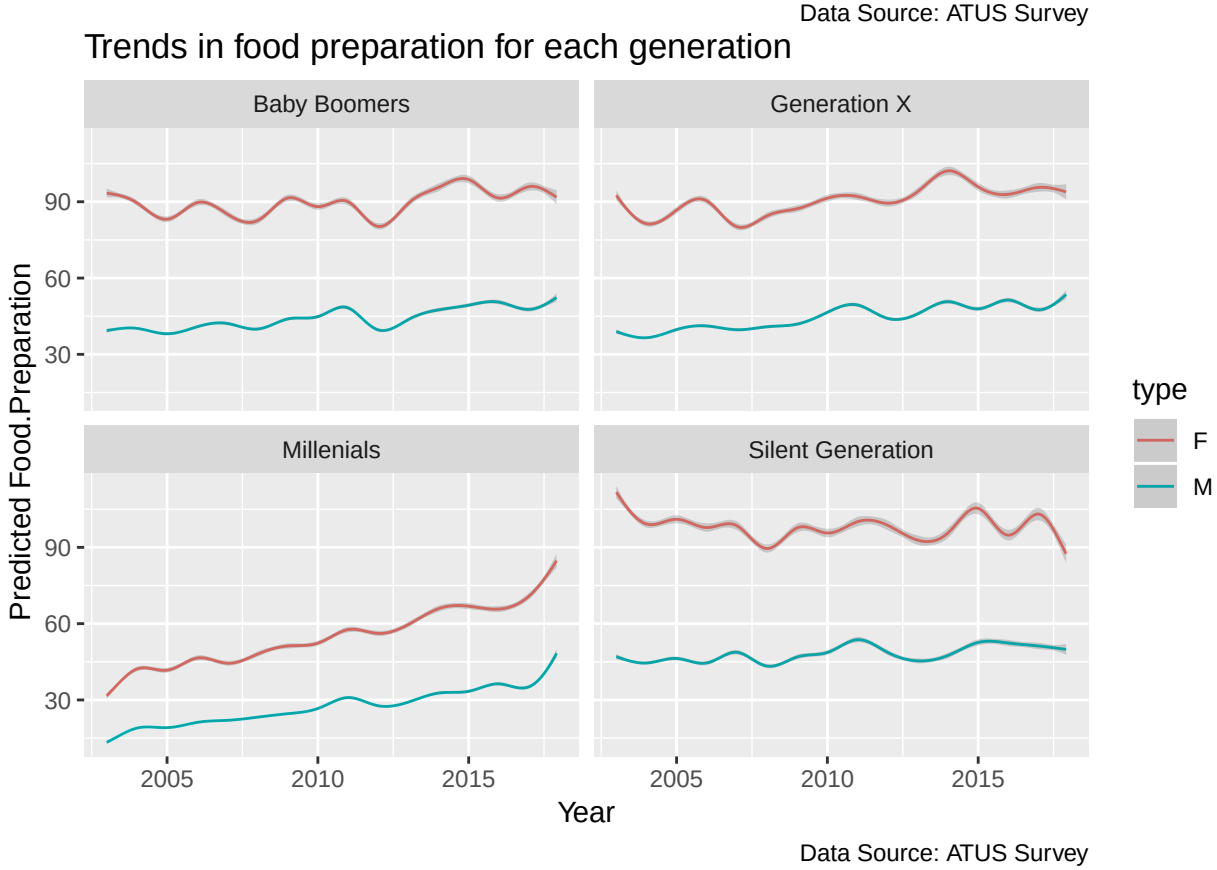
```
## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge
```

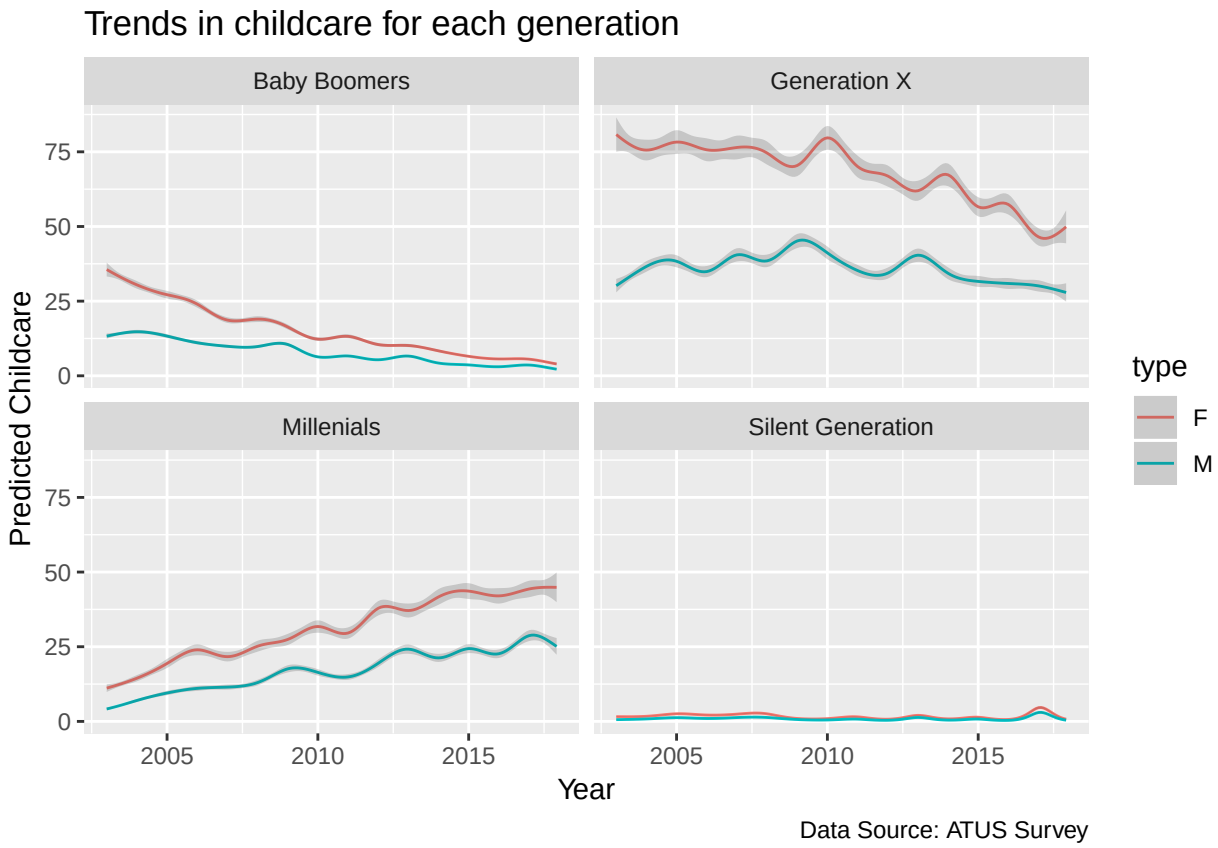



Trends in house maintenance for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted House.Maintenance
Year
type
F
M
Data Source: ATUS Survey


Trends in vehicle maintenance for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Vehicle.Maintenance
Year
type
F
M
Data Source: ATUS Survey


Trends in housework for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Housework
60
40
20
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in food preparation for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Food.Preparation
90
60
30
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey


Trends in childcare for each generation
Baby Boomers
Generation X
Millenials
Silent Generation
Predicted Childcare
0
25
50
75
2005
2010
2015
Year
type
F
M
Data Source: ATUS Survey