How Americans' Time Use Patterns Have Changed From 2003 to 2017

Group July

26 November 2018

Introduction

The aim of this report is to discover how Americans' time use patterns have changed over the last 15 years, through discovering and analysing long-term trends over the years of 2003 to 2017. The datasets [1] used to tackle this comprise some of the results of the American Time Use Survey (ATUS) [2]. ATUS respondents were interviewed once about how they spent their time on the previous day, where they were, and whom they were with. The survey records the time spent on 431 different activities that are grouped into 17 different categories. In the analysis these categories will often be referred to by a variable beginning with `tu` followed by a two-digit *category number*. The meaning of each category number and identifications for the rest of the variables can be found in the lexicon and data dictionary files which are available on the ATUS website.

The Bureau of Labor Statistics website provides a number of charts looking at the 2017 annual averages for these categories. This report provides a more definitive conclusion on how time use patterns changed over the period, evidenced with statistical investigation of the data. The findings offer in-depth explanations as to where some of the biggest fluctuations are and suggests possible reasons for these fluctuations. Limitations to the analysis are noted where appropriate and these must be taken into account when performing any further analysis or summaries, and mentioned alongside any conclusions.

Usually when analysing data, statisticians will split the dataset into training data and validation data. This ensures that any conclusions drawn from exploratory data analysis (EDA) on the training data can be verified using the remaining "unseen" validation data. The analysis performed for this report follows this technique, splitting the data approximately in half to provide continuity and robustness in model training. The split was done by selecting even and odd months for training and validation data respectively, in order to ensure that seasonality did not cause incorrect rejections of hypotheses posed based upon the EDA, as well as ensuring there were no uneven gaps in the data which may lead to a loss of clarity in plotting. This split also allowed for conformity to one of the project requirements that the month of July was to be excluded initially and used for validation.

The Compelling Change in Caring for & Helping Non-HH Members

Exploratory Data Analysis

The ATUS dataset records time spent on activities in minutes. When we plot all 17 variables it appears that there is more variation in the proportion of American's partaking in activities than there is in time spent in minutes.Therefore, the initial concern of the EDA is to look at how the proportion of American's partaking in the 17 categories of activities has changed. The following table gives a summary of some of these changes. The activities included are those with a percentage change of over 10% and a variance greater than 0.5, i.e those which are deemed to be of interest for further analysis and exploration.

It is immediately clear that the most compelling change out of the 17 activity groups, over the period, is in *Caring for & Helping Non-household (Non-HH) Members* represented by `tu04` in the data. However, without some further exploration, this is arguably uninteresting.

When fitting a linear model, the method of best subsets with *Year* forced in as a variable tells us that the most significant variables to predict the proportion of American's spending time *Caring for & Helping*

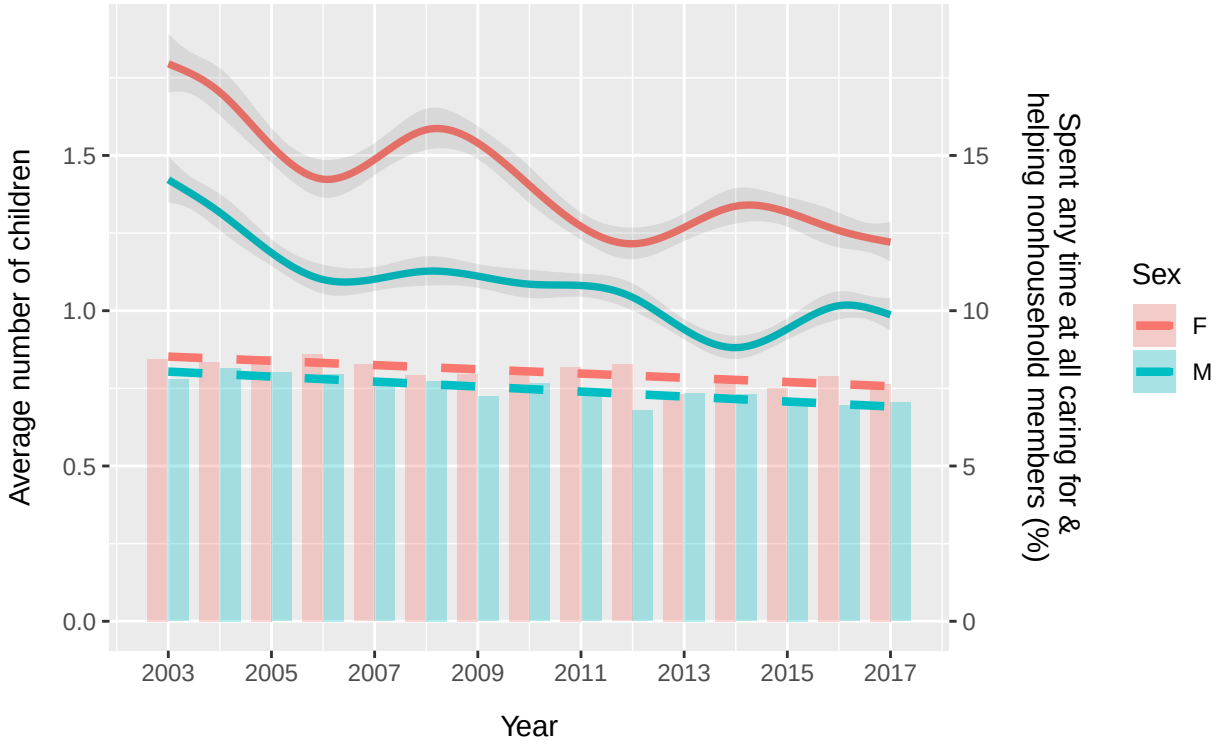
Table 1: Change in Participation of Activities

Measure	tu04	tu08	tu13	tu14	tu16
Variance	2.54	0.89	1.44	0.60	1.93
% Change	-32.11	-26.46	10.17	12.32	-24.08

Nonhousehold (NonHH) Members are *Sex* and *Number of Household Children*. Ths model is not sufficient to accurately predict `tu04` participation, however it does give a good idea as to what is causing the change over the period 2003 to 2017.

The non-negative, non-linear nature of the continuous activity data also means a more complex model would be suitable. An F-test shows such a complex model which incorporates *Sex* is a significant improvement over such a model which doesn't include it. This confirms the belief that is has a significant effect on participation in caring for & helping non-household members.

The plot below is a graphical interpretation of the model, built on the training data, with the average number of household children, for each group of the population, displayed as **bars**. Participation percentage in caring for & helping non-household members is displayed as smooth lines.



The key observation is that, as the table suggested, over the period the participation in caring for & helping non-household members has decreased for both men and women. Building a linear model using *Sex* and *Year* and performing a t-test confirms this. There also appears to be a link between participation in caring for & helping non-household members and the number of household children. Updating the spline curve model to include the number of household children and then again testing this hypothesis with a t-test confirms this is also the case. Looking at the graph, fluctuations in the average number of household children, on the whole seem to be followed by but if we check the correlations between them, they show this link is fairly weak (0.65 for Men and 0.49 for Women), suggesting there must be further reasons for this change; possibly not measured in this dataset. Therefore, further analysis must be done before a statistically sound conclusion on the causation of the changes can be made.

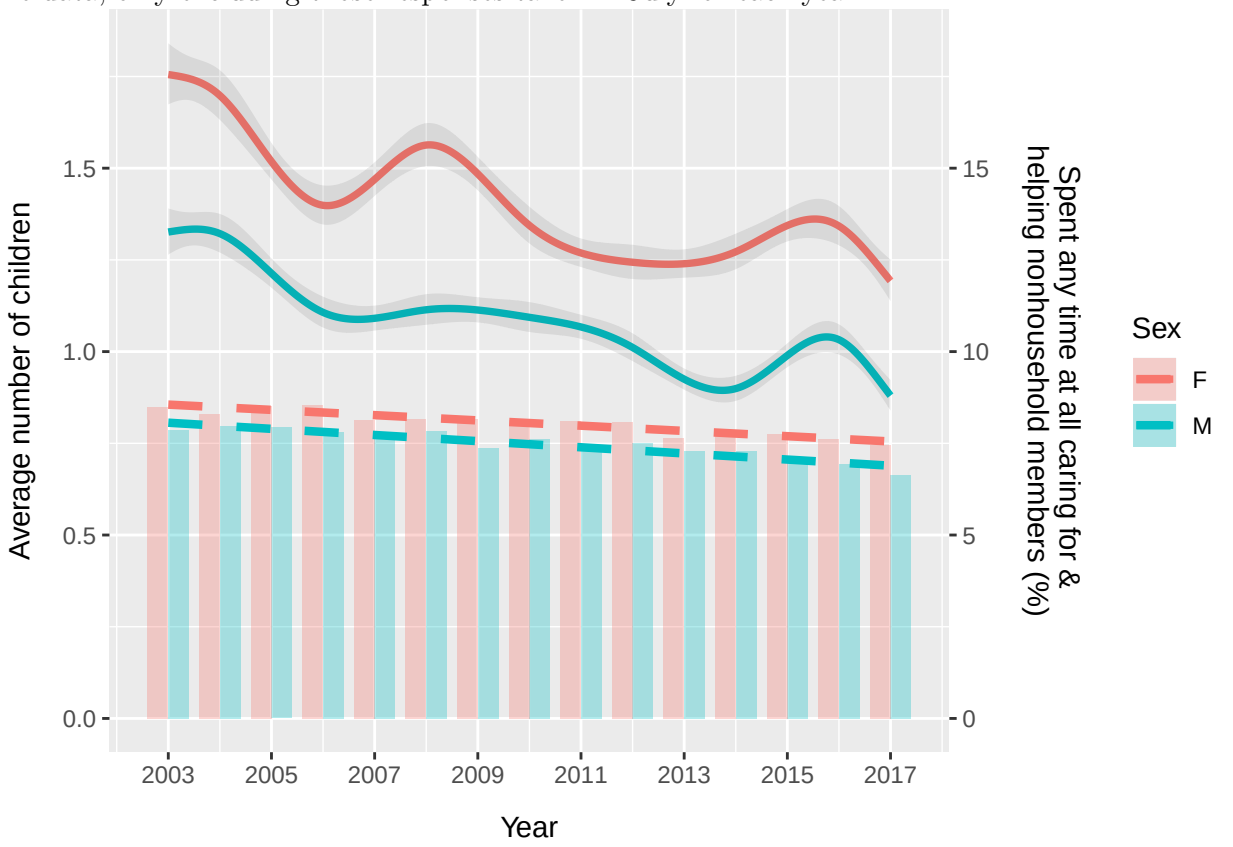
Validation

As with EDA, but this time formally, a linear model is built using *Sex* and *Year* and a t-test performed. It gives a p-value $<< 0.05$. This means there is sufficient evidence to conclude that the proportion of American's who participate in caring for & helping non-household members has infact decreased over the period 2003-2017 as suggested by the EDA.

The initial thoughts about the link with number of household children, from the EDA, again needed to be tested. A new model of the same structure was built. Another formal t-test could then be completed on an update of this model which includes the number of household children.

With a p-value $<< 0.05$, this test confirms this belief in the validation data.

We have now verified that the conclusions suggested by the EDA are correct: the percentage of American's who partake in caring for & helping non-household members has decreased over the period from 2003 to 2017. There also seems to be a link between the number of children in the household and the participation percentage. The plot below gives a graphical interpretation of changes using a model fit on all the data, only excluding those responses taken in July for each year.



How The Time Spent on Traditionally Gendered Activities has Converged as Gender Roles have Broken Down

```
fit.models.and.plot <- function(Data, variable, mu) {
    # Start building models for the data
    n <- nrow(Data)
    # model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE, knots = (c(as.Date
    model1 <- glm(as.formula(paste0('weighted.',variable,' ~ ns(TUDIARYDATE)')), data = Data, fam
    model2 <- update(model1, . ~ . + TESEX)
```

```
    # model3 <- update(model1, . ~ -1 + TESEX + TESEX:ns(TUDIARYDATE, knots = (c(as.Date('2006-01
    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(
    # 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(Data, TESEX == 'M'), se = TRUE, type
    female.prediction <- predict.glm(model5, newdata = filter(Data, TESEX == 'F'), se = TRUE, typ
    df <- data.frame(age = c(filter(Data, TESEX == 'M')$TEAGE, filter(Data, TESEX == 'F')$TEAGE),
    df<- transform(df, generation = factor(generation, levels = c("Millennials", "Generation X",

    if (variable == 'Working' | variable == 'Childcare') {
        print(ggplot(filter(df, generation != "Silent Generation"), aes(x=date, y=prediction, col
        geom_line() +
        geom_ribbon(aes(ymin = prediction - prediction_error, ymax = prediction + prediction_erro
    } else {
        print(ggplot(df, aes(x=date, y=prediction, colour=type)) +
        geom_line() +
        geom_ribbon(aes(ymin = prediction - prediction_error, ymax = prediction + prediction_erro
    }
}
Gender.Roles.Plotting <- filter(Gender.Roles.Data, month(TUDIARYDATE) != 7)
```

The discussion surrounding social division between genders in terms of roles and responsibilities has consistently been a major news topic in recent years. This debate has been fuelled by a new wave of feminism, as well as a number of different social media campaigns such as *'#MeToo'* and *'Times Up'*. Both of these movements aimed to address commonplace sexual harassment and discrimination, empowering women to take back control. Given this current social climate, it was decided that an interesting area of the ATUS data to focus on would be long-term changes in time use for different sexes. Specifically, areas of time use where pre-existing gender roles are present were examined.

Exploratory Data Analysis

Talcott Parsons' [3] study on gender roles - published in the 1950s - compared traditional gender roles with more liberal alternatives. A simplified version of the strictly traditional view of gender roles is detailed in the table below.

Table 2: Traditionalist Gender Actitivies

Male Activites	Female Activities
Working	Housework
House Maintenance	Cooking
Vehicle Maintenance	Childcare

The liberal viewpoint discussed by Parsons suggests an equal balance of time for the genders in these roles. Whilst this study was developed over 50 years ago, a preliminary look at the data confirmed that, in 2003, a complete transition to the liberal viewpoint had not occurred. Therefore, it was decided to investigate whether the time use of these specific activities was converging to equality or remaining stationary with noticeable differences between genders.

Table 3: Generations

Generation	Birth Years
Silent Generation	1928 - 1945
Baby Boomers	1946 - 1964
Generation X	1965 - 1980
Millennials	1981 - 1996

Following some initial exploration of the data and breaking down the respondents into different generations [4], another layer of complexity to the analysis was developed. Splitting the data by age allowed for further hypotheses regarding progression in converging time use by both genders in the areas discussed above; it would perhaps be assumed that younger people would be more progressive in this respect. Using this generational approach provided a way of tracking a population over time without having to arbitrarily pick age categories. A small number of participants fell outside of these ranges, however, these represented only a small number of participants and so were added to the next nearest generation (e.g. if a participant was born in 1998, they would be added to the *Millennials* generation).

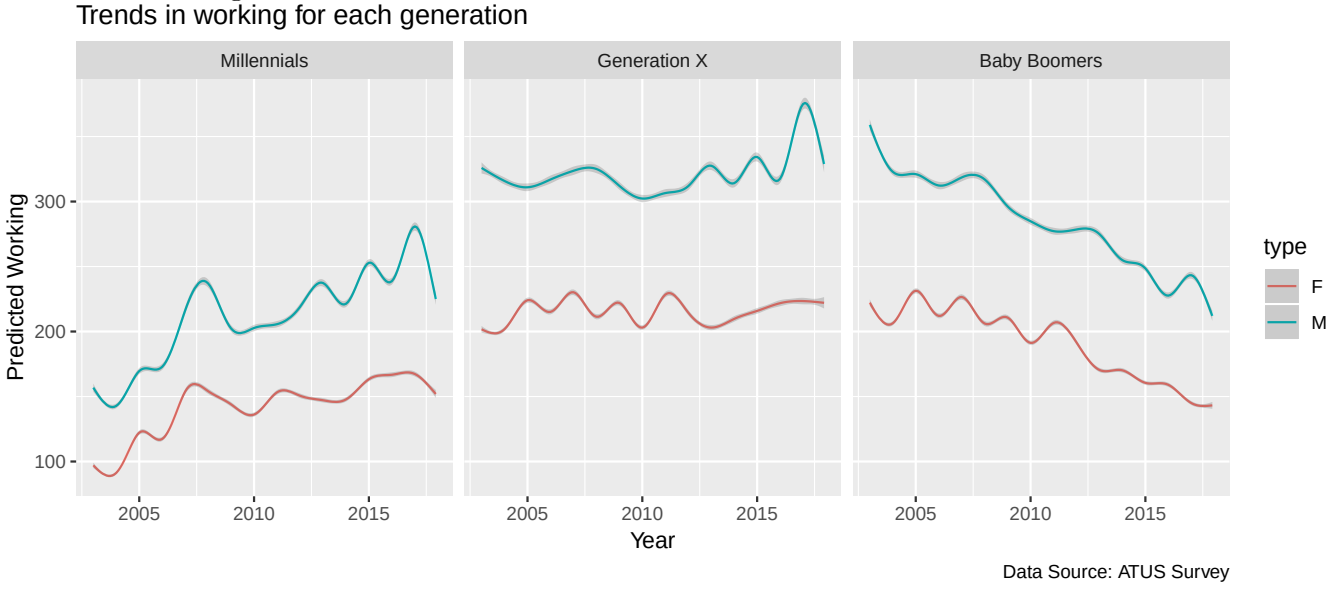
For each of the activities, a suitable generalized linear model was developed using sex, year and generation as parameters. The values used to predict each model were constructed using the weighted mean formula provided in the *ATUS User's Guide*, available online in the ATUS Survey Documentation. Splines were added to the model for each year and graphical representations were produced.

Validation

In order to validate the trends found in the EDA, the dataset was initially split

Working

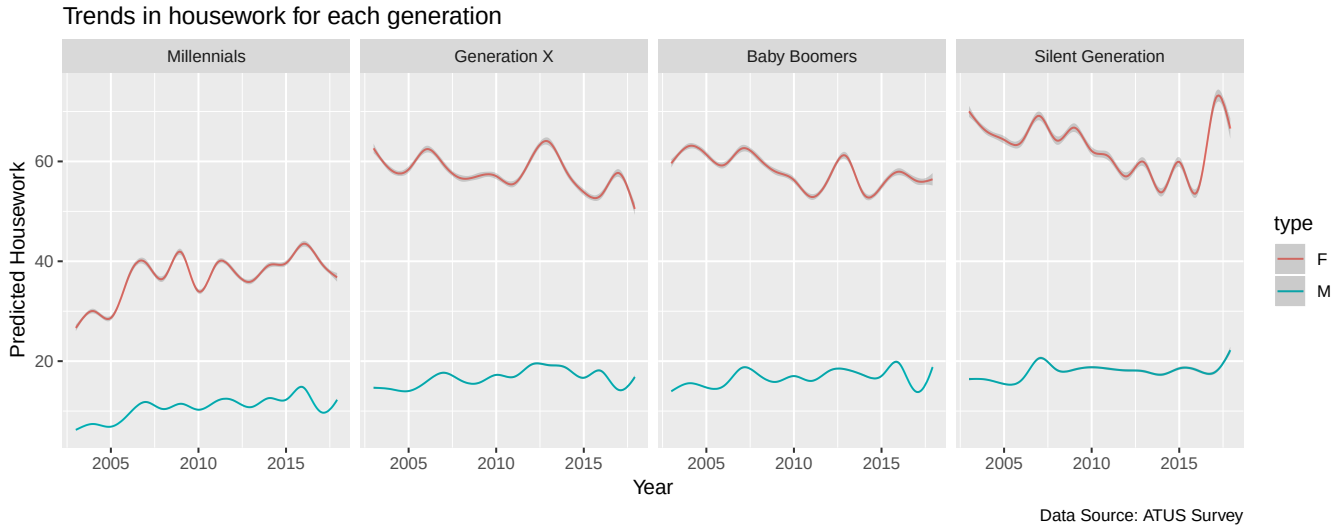
The plots below show the change in working patterns between three different generations. The *Silent Generation* have been excluded from this plot as, by 2010, the youngest of this generation will be 65. At this age, most will have retired and as such the plot is not very informative. Furthermore, the same scale is used for each graph (for purpose of easier comparison), so this affects the quality of comparison between the more informative generations.



House Maintenance and Vehicle Maintenance

For both of these variables, it was decided that the participation rate was too low to warrant deeper analysis. Whilst the findings represented that there existed a separation in gender, with men spending more time on both of these activities, the participation rates of around 3% for both reflected that these were more uncommon activities. As such, it was decided that there was not enough data to reflect the time spent on these activities in a suitable linear model.

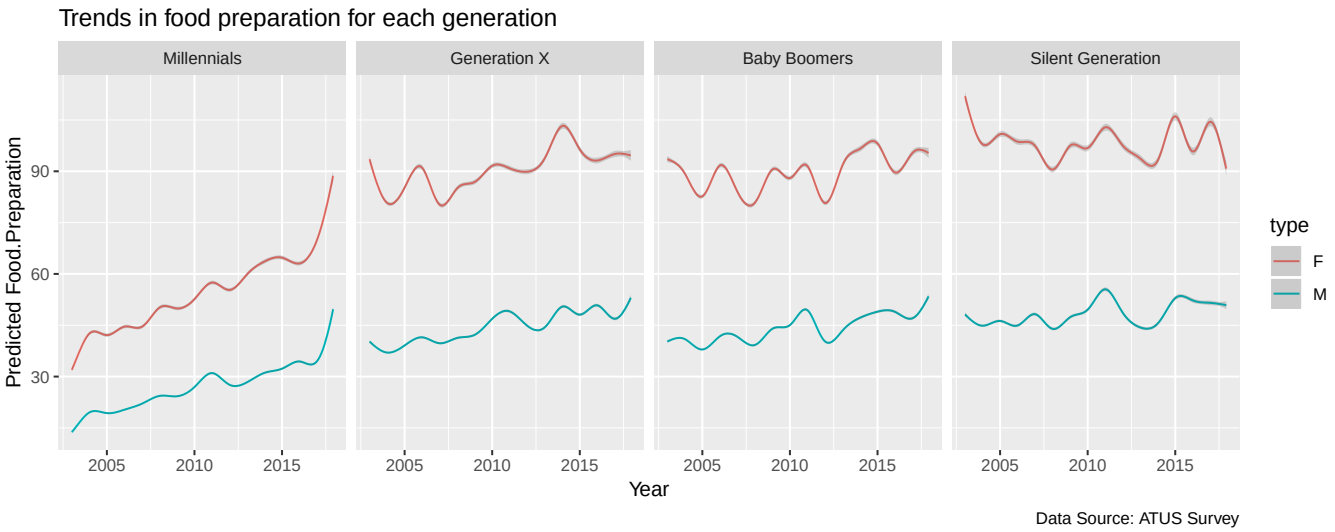
Housework



The graphs above show that for all generations, except for *Millennials*, there has been a decrease in the gap between men and women. This is as a result of men spending more time and women spending less time on housework, although the decrease for women is sharper than the increase for men. This trend is reflected strongest in *Generation X*, the second youngest generation, however, there is also strong evidence for this trend in the *Silent Generation* despite the slight rise in women in 2016. On the other hand, the gap seems to have increased slightly for *Millennials* - both sexes are increasing the amount of time spent on housework, but the increase is not as severe for men. Notably, *Millennials* also spend substantially less time doing housework than the other generations, with the highest point being around 45 minutes per day (women in 2016), compared to around 62 in *Baby Boomers* and *Generation X* and 72 in the *Silent Generation*. Therefore, whilst the gap is increasing for *Millennials*, it is still smaller than all of the other generations at around 25 minutes. Comparatively, the *Silent Generation's* gap started at over 50 minutes in 2003 and has decreased to just under 45.

transform(iris, Species=factor(Species,levels=c("virginica","setosa","versicolor")))) + geom_histogram(aes(Petal.W facet_grid(Species~.)

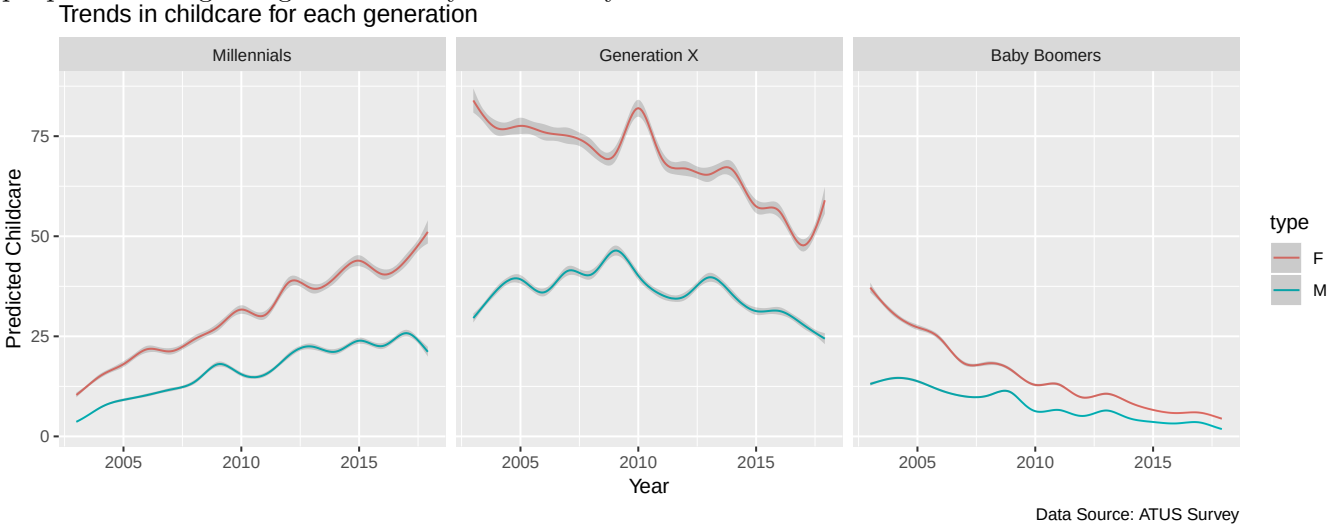
Cooking



The graphs above show a very similar trend to `Housework` in the sense that the gap is decreasing and that all generations, apart from *Millennials* spend a similar amount of time in general on this activity. However, the difference with this activity is that for *Baby Boomers* and *Generation X*, women are actually spending more time on this activity than they were previously - the decrease in the gap is as a result of a sharper increase for men. For *Millennials*, this is not true: both sexes are spending more time on this activity but the increase for both is at an equal rate and so the gap is remaining equal. The *Silent Generation* is the generation showing the biggest decrease in the gap between sexes; this is due to the decrease for women and an increase for men, although this generation started off with the highest values for both sexes.

Childcare

For similar reasons to working, the *Silent Generation* has been excluded from this section as most of the people in this age range are unlikely to have any household children.



These graphs show perhaps the greatest change in gender roles. The data for *Baby Boomers* and the *Silent Generation* shows very little, most likely as the variables included in childcare relate only to caring for household children, where a child is defined as someone below the age of 18. It is unlikely that many people within these generations would have children living with them at this age (in fact checking the data confirms that, in 2017, only 9% met this criteria). Furthermore at this age, their children are likely to be

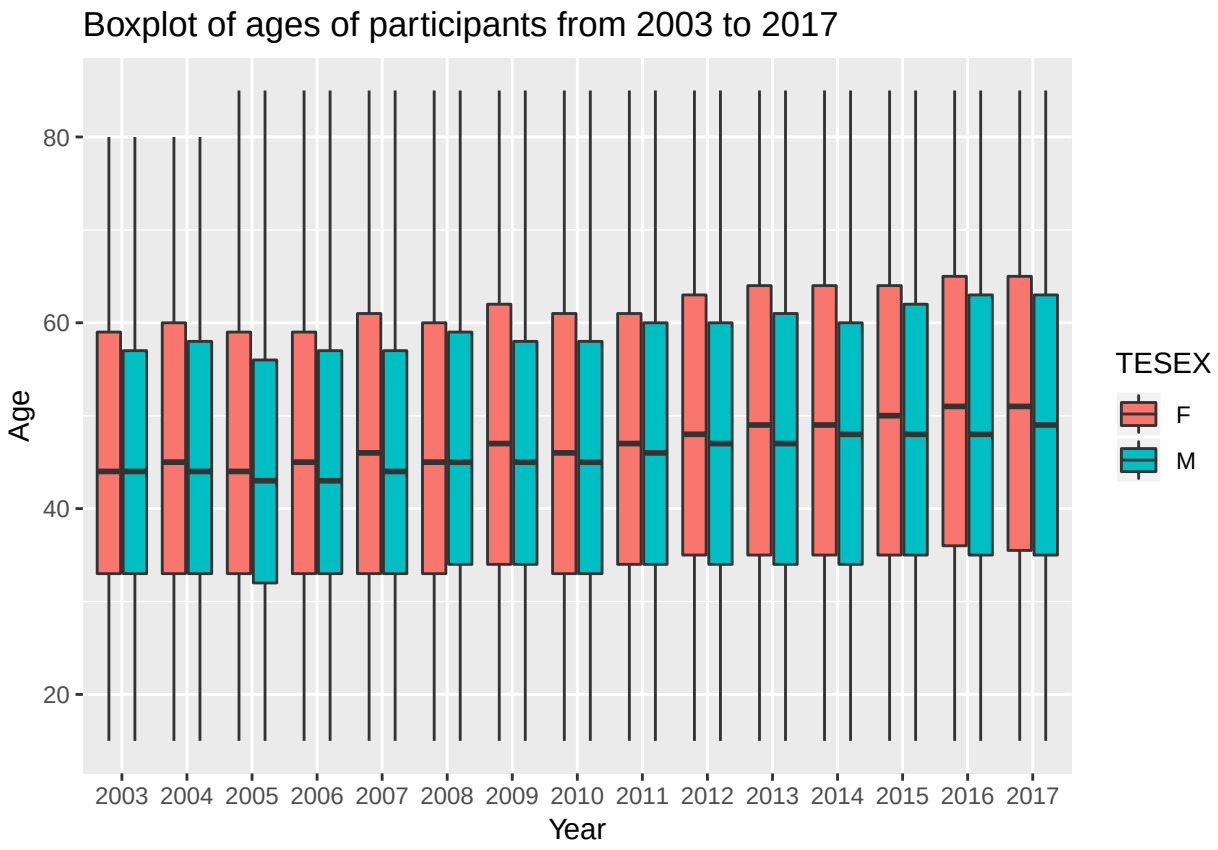
older and require less constant care as they become more independent. Therefore, the other generations, *Millennials* and *Generation X*, reflect more interesting trends. For reasons similar to the older generations, the drop off in childcare at the end of *Generation X* can likely be explained as children mature; however, the drop off for women is sharper than it is for men, leading to a convergence in the lines. The gap has decreased substantially, from 50 minutes per day in 2003 to less than half that, around 20 minutes in 2017. This could further be attributed to the decrease in working time for this generation, noted above. The *Millennials* on the other hand reflect an increase year on year for activities in childcare. Whilst the gap increases slightly over time, it doesn't ever reach a similar point to that of *Generation X*, with the biggest difference being around 40% compared to that of 60%.

Limitations of the Data

One of the major limitations imposed on this research is the length of the period within which the data was gathered. 15 years is a short period. Sporadic subset size due to the number of filters required to get down to the demographics is what actually need to investigate.

The code below is to check the distribution of age group in those 15 years. It would not be a problem if there is a participant who have joined the survey several times, as her/his data would be regarded as independent samples in different years.

```
#save Dan's tidy dataset as another dataset, so no need to run the code every time opening r stud
write.csv(Data_minus_columns,'Dataset.csv')
Dataset <- import('Dataset.csv')
library(dplyr)
require(ggplot2)
library(dplyr)
age <- select(Dataset, TUYEAR, TEAGE,TESEX)
age$TUYEAR <- as.factor(age$TUYEAR)
p <- ggplot(data = age , aes(x=TUYEAR, y=TEAGE,fill=TESEX)) + geom_boxplot(width=0.8)
p + ggtitle('Boxplot of ages of participants from 2003 to 2017') + xlab('Year') + ylab('Age')
```



It is clear that the amount of participants is not well distributed and should be considered in the analysis. However, it could lead to the birth rate and death rate of the country.

References

[1] "ATUS datasets." https://www.bls.gov/tus/datafiles_0317.htm.
[2] Bureau of Labor Statistics, "The american time use survey." https://www.bls.gov/tus/, 2017.
[3] T. Parsons, "Age and sex in the social structure of the united states," *American Sociological Review*, vol. 7, no. 5, pp. 604–616, 1942 [Online]. Available: http://www.jstor.org/stable/2085686
[4] "Millennials projected to overtake baby boomers as america's largest generation." http://www.pewresearch.org/fact-tank/2018/03/01/millennials-overtake-baby-boomers/.