

Investigation of the CDC's NSFG Data 1973-2015: Results, Lessons Learned, and Future Directions

Julianne; Julianne.Ammirati@emory.edu

April 25, 2017

The code used to generate this report is available from my GitHub Account:

https://github.com/JulianneA/N741_Milestone1

Please Note: The CDC required a data use agreement to obtain .dat files, so data will not be uploaded to github. They are available to download from the CDC websites cited in the references section.

The National Survey of Family Growth (NSFG)

Dating back to 1973, the CDC has been surveying women regarding their sexual and reproductive health using a cycle-standardized National Survey of Family Growth Interview. The most recent iteration of the survey occurred in 2015, and this survey marks the ninth cycle of data to be collected since the project's inception. Over the course of this project, over 76,000 female respondents have been surveyed on items including demographics, contraceptive use, pregnancy history and outcomes, and sexual education history. The information collected under this survey has been used by the U.S. Department of Health and Human Services to direct funding and plan the implementation of health services throughout the country [1].

Note: Cycles 1-6 are officially referred to a cycle number, but surveys from 2006-present are solely referred to by their year of collection. For the purposes of this project, these three surveys will be referred to as Cycles 7-9.

Proposed Analysis

I did not intend to analyze the entire data set, but rather, had planned to conduct this exploratory analysis by focusing mostly on family planning variables (coded as Section E of the data set in the three most recent iterations). A lot of previous publications have come out of this data set, but these publications have all covered information collected in a single cycle. Given the increasing interest in women's contraceptive utilization in the current political climate, I believe that the data could be better visualized for public understanding. Additionally, by combining the data across cycles and breaking it down into various groups based on race/age/demographics, I think the data can tell a better story about the groups that are being left behind as contraceptive methods improve.

Initially, I planned to reduce the number of variables to strictly sociodemographic variables and contraceptive variables, while examining cases collected throughout all cycles of the survey (1973-2015). My analysis plan would have allowed me to evaluate trends in a

subset of variables over time, to examine how certain contraceptive options have become more or less utilized.

Theoretical Context

Healthy People 2020 identified improved access to effective family planning as an objective [2]. According to the CDC, the most reliable forms of contraception are sterilization or long acting reversible contraceptives (LARCs), a category that includes both copper and hormonal intrauterine devices (IUDs) and hormonal implants [3]. Current LARC utilization rates are lower in the US relative to other developed nations, especially among women with of lower age and socioeconomic status [4], but it is unclear when the divergence in utilization rates occurred and what additional factors contribute to the lower utilization. The CDC has collected data through the NSFG that could provide information regarding when this divergence occurred and who is most impacted, which is information that could be used to identify causes of differential utilization and steps that could be taken to ensure that all women are provided the information and opportunities to choose the best contraception for them. Currently, it is theorized that differential utilization is due to personal choice, misinformation on the part of providers or patients, or accessibility, and without first understanding the traits of women who access various birth control types, it is difficult to assert ways to promote safe, effective birth control for all women.

Data Acquisition

The results of the NSFG projects are available in a .dat file with setup programs that can be run in SAS, STATA and SPSS to properly import and code variables. In order to import the data into R, I chose to download both the setup program (.SAS) and the data file (.dat) from the CDC to my personal drive at Emory. I opened the .SAS program and then updated the directory to read in the .dat file. Finally, I exported to a CSV file using the following command in SAS:

```
dm "dexport new 'H:\SAS 9.4 Temporary Files\2013_2015_FemRespData.csv' ";
```

I attempted this process with the most all NSFG data cycles (1973-2015), and was *mostly* successful (details to follow on issues that arose). This was a particularly arduous way of importing the data set, but it was chosen because the CDC provided all necessary files to compile the data through SAS.

Data Wrangling

When the Real and the Ideal Quickly Diverged

Upon further inspection of code books, it became evident that the 1973 and 1976 files could not be used in analysis. Throughout these cycles, the CDC interviewed only married women or unmarried women who had borne children for the NSFG. From 1982 onward, unmarried and childless women were included in the NSFG, therefore, the initial two cycles of the NSFG were excluded here because the differences in sampling populations introduces too great a threat to data validity and generalizability.

Year of NSFG	Number of Respondents	Variables	Notes
2013-2015	5699	3207	Dubbed Cycle 9 for this project. [5]
2011-2013	5601	3094	Dubbed Cycle 8 for this project. [6]
2006-2010	12279	3315	Dubbed Cycle 7 for this project. [7]
2002	7643	3087	Cycle 6, 3 SAS setup files. [8]
1995	10847	5753	Cycle 5 [9]
1988	8450	1437	Cycle 4 [10]
1982	7969	742	Cycle 3 [11]
1976	8611	470	Excluded: Did not include unmarried, childless women. [12]
1973	9797	451	Excluded: Did not include unmarried, childless women. [13]
All Respondents	76896		
1982-2015	58488		

Preparing for Merge

Sixty-three variables were extracted from the NSFG data sets (Appendix A). These variables capture information regarding respondent demographics, contraceptive utilization, sexual education history, and reproductive history. These variables were selected with special consideration for variables present in the most recent three cycles of the NSFG (2006, 2011-2013, and 2013-2015), which are most similar in their data cleaning methodology and actual variables and information collected from women. Two variables were added prior to data merge: CYCLE and YEAR which both indicate which NSFG survey data set that respondent had been included.

Expected Complications: The Known Unknowns

There were many issues I expected to arise when developing my data analysis plan, including issues with coding schemes and information regarding differences in information being collected overtime. While the scope of these changes and the r code they required to be remedied were not accurately foreseen, these issues were at least expected prior to analysis.

Birth Control Has Changed

Contraceptive methods have been developed or fallen out of use in the last 35 years, so the contraceptive use variables has changed greatly overtime. Previous iterations of the survey included information regarding contraceptives such as women's condoms, today sponge, and cervical caps, but these methods were not included as distinctly coded variables in the most recent cycles. While many methods seem to have fallen out of popularity or have been discovered and then used through to the most recent cycle, thus appearing in only the first few cycles or the last few cycles, this pattern was not always consistent. Some methods such as Lunelle, a hormone injection, was not included on many of the cycles, as Lunelle did

not receive FDA approval until 2000 and then quickly fell out of use. This method appears as a distinct variable only in the 2002 and 2006 Cycles.

Information Collected has Changed

Since its inception in 1973, there has been a seven-fold increase in the number of variables maintained in the data set. While many of these are recodes and imputed variables, some are useful for further examining reproductive health history or sexual education experiences. The final three cycles (2006-2015) have the most consistency between them relative to earlier iterations, but even these cycles experience inconsistencies due to the changing types of birth control utilized by American women.

Naming Conventions Have Changed and Variable Inconsistencies

This survey has been collected in large cohorts of women over the past 34 years. While providing the user with enormous amounts of data, the CDC has been *consistently inconsistent* in their variable coding schemes. These differences in coding were mostly expected, but the scale with which variable names changed and coding was altered was much higher than expected. Single letter changes have been made to the variable names overtime, with each requiring a recode of the variable name prior to merging. Earlier cycles of the survey seem to have used a non-descriptive naming scheme for some variables, with an alphanumeric qualifier, such as "C_12" or "D20P3" rather than the descriptive variable naming conventions used in later cycles. Additionally, some iterations of the survey utilize name brands for some methods (Lunelle and Depo rather than a category called "hormone injection"), while other methods are lumped together as a class, such as IUDs or sterilization. It is unclear what prompted the CDC to classify certain contraceptive options by a single brand name rather than in groups, but this makes it increasingly difficult to track changes over time, as it would require a variable recode so that all methods are categorized as a type (hormonal injection, IUD, etc.) rather than just as a singular formulation.

Processing Capacity of My Mac Struggled with the Amount of Data

Once all CSV files had been read into R, the project took an immense amount of time to save and even longer to knit. Since the document can still fail to knit even though r-chunks all successfully run, I had decided to frequently knit to HTML to ensure that changes made did not cause knitr to fail. This process took a lot of time to read in all CSV files, add and cut variables, and remerge, so I needed to find a way to reduce the toll the procedure took on my computer to increase the speed at which I could knit and test out new code. This fix was fairly simple, requiring me to save the subset of variables I was using to an R-data file, write to csv then import those csv files prior to merging across NSFG cycles. Once this process ran successfully, all code that imported the original CSV files, pulled and renamed subsets of variables, and exported to an R-data file and CSV were muted, allowing the code to remain in the markdown (to promote reproducibility) while still allowing the markdown to work only with the smaller subset of variables (65 per cycle versus approximately 22,000 across all cycles).

Unforeseen Complications: The (previously) Unknown Unknowns

While I anticipated data wrangling to take up a large portion of my project plan, there were many unexpected issues that arose over time that were often unanticipated until an r-

chunk crashed and I had to look for a source of the problem. While a long data wrangling process would be acceptable given the fact that the majority of my analysis requirements focused on visualization rather than modeling or labor intensive statistics, these unanticipated delays were an overwhelming setback that took up an immense amount of time and resources while also detracted from the quality of visualization.

Not all SAS Setup Programs Are Created Equal

Unfortunately, not all the SAS setup routines worked. The setup .sas files ran to completion on all cycles, but these routines did not read in all the headers. This complication was entirely unexpected, and taught me a valuable lesson: I took for granted the setup routines and the .dat files and did not look through the sas log to ensure the scripts ran *properly*. This complication went undiscovered until the data merge continued to fail, as each individual cycle had a differing number of variables. While I retried the setup programs, they continued to generate csv files without complete header rows. This led me to question the integrity of the data, and forced me to go through the relevant sections of the code book to ensure that the columns filled with data I was utilizing properly matched the correct name.

Illogical Data Entry Method

Not all iterations of the survey follow basic data coding etiquette. While I did not anticipate the data to be perfect from the first few cycles, I had expected the data to at least be *coded*, which it was not always. While the code books assert that the data is numeric, with 1/0 standing in for values, the file itself has "(1) Uses Contraception" in the field, rather than just "1". I am not certain why the CDC coded the data if they wound up not using the codes, but they began regularly using codes as of Cycle 5 (1995). This issue was fairly easy to correct, but were not discovered until after the data was merged and the data class of certain variables were not as expected.

Difficulty Understanding Codes

The code books for the NSFG are extensive, often totaling hundreds of pages per cycle and including the coding schemes for each variable as well as information about recoded and computed variables and the survey questions themselves. While I had expected the process of learning the coding to be time consuming, I did not anticipate it to be confusing. I had been thankful to discover the code books and had assumed that the information provided within the code books were extensive enough to be able to adequately understand the data enough to recode them. However, the coding schemes were often more unnecessarily nuanced than I had expected, containing codes that represent the same thing without adequately explaining what the differences in the codes signify. One major example of this comes from one cycle in which female and male sterilization were represented by a total of 5 codes, with little information to indicate why one code was selected over another outside of whether the respondent or her male partner were sterile.

Protected Data

The CDC has opened access to this data, but certain variables are still protected, such as information regarding the region or zip code of the respondent. While this information was collected by the researchers, the CDC would only allow access to an independent researcher if additional approvals and data-use agreements were obtained.

Techniques Used

Data recoding and computing became the bulk of my project, as I worked to resolve unforeseen issues that arose in the data cleaning process. The pulling of a subset of variables and renaming the variables to ensure that the files for each cycle would successfully run required fairly straightforward coding, examples of which appear below. Variables selected for analysis were imported into a "NSFGyear_subset" data set using the select command from the dplyr package. Variables that were not collected in that cycle were added to that cycle's subset of necessary variables with "NA" entered in the field to ensure that each cycle contained all necessary variables prior to merge (second line of code below). If information was available in the data set under a different variable name, that variable was renamed and imported into the subset (third line of code below)

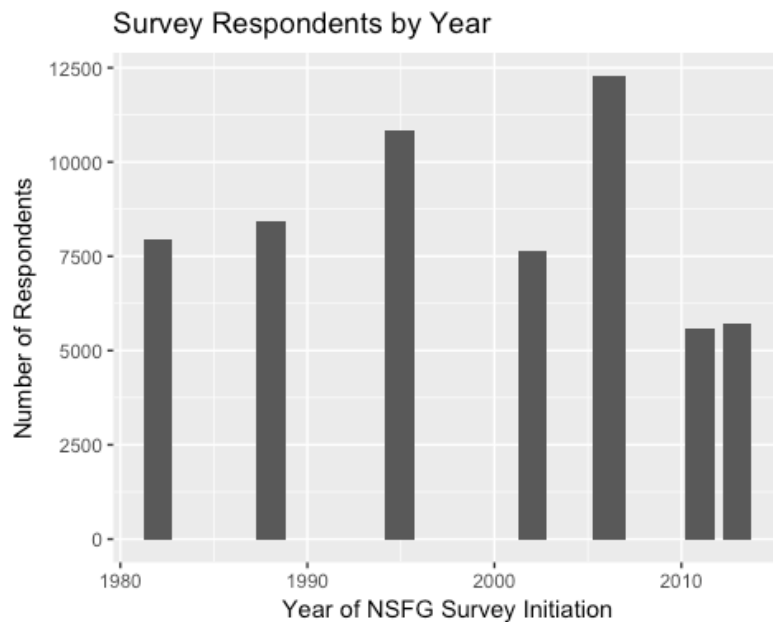
```
NSFG1995_subset <- select(NSFG1995, CASEID, HIGRADE, PREGNOWQ, MAYBPREG)
NSFG1995_subset$SEDWHBC <- NA
NSFG1995_subset$CURRMETH2 <- NSFG1995$MTHLSTS1
```

Coding schemes from the most recent three cycles (2006-present) were established as the standard coding scheme, and all prior years were forced to match these cycles. If variables were collected but the coding differed from the most recent cycles, recoding procedures were undertaken. This recoding was required to alter numeric values that referred to certain birth control methods (as is the case in the code example below) and also to change the meaning of codes to indicate whether a respondent ever used a method of birth control, refused to answer the survey question, or the data was not collected.

```
#Recode 1982
NSFG1982sub$RecMeth[NSFG1982sub$CURRMETH1==1] <- 6 #female ster
NSFG1982sub$RecMeth[NSFG1982sub$CURRMETH1==2] <- 5 #male ster
NSFG1982sub$RecMeth[NSFG1982sub$CURRMETH1==3] <- 3 #Pill
NSFG1982sub$RecMeth[NSFG1982sub$CURRMETH1==4] <- 19 #IUD
NSFG1982sub$RecMeth[NSFG1982sub$CURRMETH1==5] <- 12 #diaphragm
```

Once recoding was completed, all cycles had sixty-three variables plus 2 additional variables to identify the year and cycle the respondent's data is from. These 65 variables were identical across all the cycles and an rbind command was used to merge the data into a single data frame with approximately 60,000 respondents.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Data Analysis

Sample Demographics

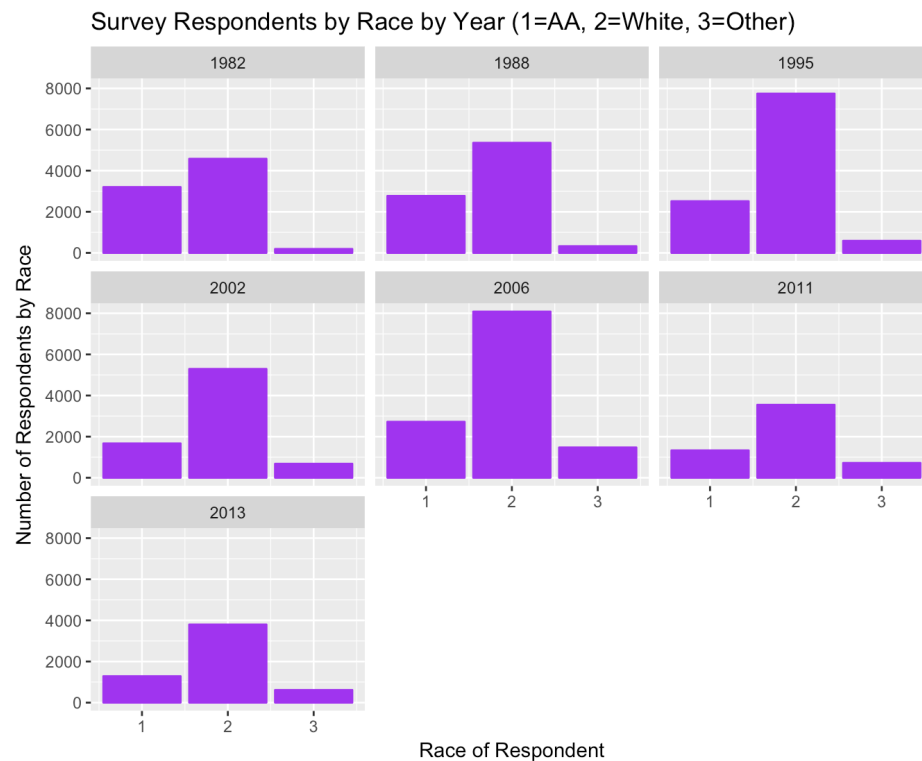
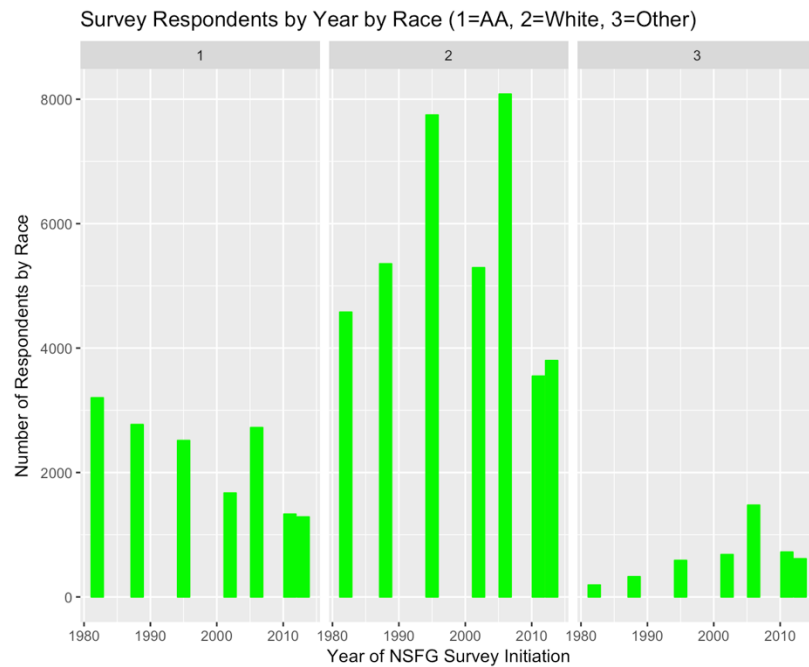
Following the completed merge of NSFG cycles 3-9, preliminary analyses were conducted to calculate demographic frequencies and percentage breakdowns for variables such as age, race, ethnicity, education levels, and marital status. These values were calculated using filtering and "summarise" along with basic statistical commands (mean, standard deviation) for continuous variables and the "prop.table(table())" for categorical variables in order to tabulate percentages. ANOVAs were run on continuous variables.

Table 1: Basic Demographics

Variable / Year	1982	1988	1995	200 2	2006- 2010	201 1- 201 3	201 3- 201 5	All Responde nts	p- value
Age	27.26 +/- 8.18	29.54 +/- 8.13	30.59 +/- 8.31	29.5 +/- 8.43	28.64 +/- 8.45	28.6 1 +/- 8.38	28.8 8 +/- 8.4	29.08 +/- 8.39	0.16 56
Highest Grade of Education	12.71 +/- 4.48	12.99 +/- 4.14	13.03 +/- 2.82	13.3 1 +/- 3.13	13.3 +/- 3.25	13.4 6 +/- 3.67	13.5 9 +/- 3.33	13.17 +/- 3.55	0
% Hispanic	6.24	7.59	14.32	20.7 9	22.18	26.0 5	24.4 6	16.85	
% Black	40.17	32.79	23.18	21.8 5	22.17	23.7 6	22.5 7	26.49	
% White	57.44	63.36	71.41	69.2 4	65.82	63.3 6	66.6 6	65.65	

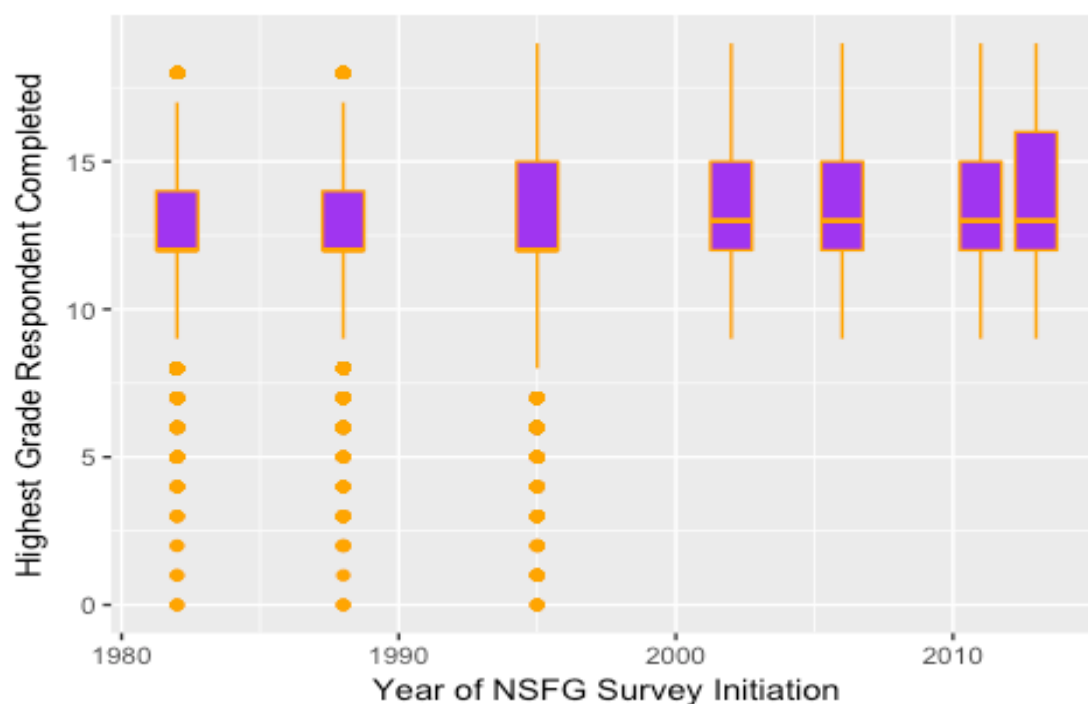
% Currently Married	44.56	47.7	48.78	40.3	32.34	30.55	29.81	39.9
% Never Married	39.5	34.52	36.9	38.55	45.58	45.21	46.89	40.72
% Residing in Metro Areas	39.65	29.34	Data Unavailable	47.23	Data Unavailable	41.9	41.05	39.06
% Preferring English/Spanish	Data Unavailable	Data Unavailable	97.09; 2.91	92.59; 7.41	92.33; 7.58	94.25; 5.73	93.58; 6.37	94.03; 5.94
% Private Insured	Data Unavailable	Data Unavailable	Data Unavailable	63.27	55.41	48.49	52.48	55.56
% Medicaid Insured	Data Unavailable	Data Unavailable	Data Unavailable	12.84	20.07	23.62	26.32	20.08
% Uninsured	Data Unavailable	Data Unavailable	Data Unavailable	6.8	3.66	6.09	4	4.93
% Receiving Public Assistance	Data Unavailable	Data Unavailable	Data Unavailable	10.24	0.35	0.67	87.99	1.11
% Employed Full Time	35.1	46.47	42.09	43.06	39.67	37.14	38.78	40.59

The following graphs depict how many participants of each racial category were recruited for the various iterations of the NSFG. The first plot uses race as a variable to split the plots into 3 (African American, White, Other) which then chart the count of women of each racial category across the years. The second series of plots uses Survey Year as a variable to split the plots into 7 different graphs (one for each survey year) which then charts the count of women of each racial category within that year. Both are useful to show the changing diversity seen in each iteration of the survey. For example, it is clear that in 2002, a higher percentage of the respondents were white relative to 1982. This conclusion is supported by the table above, but these graphs depict number of women, rather than percents.

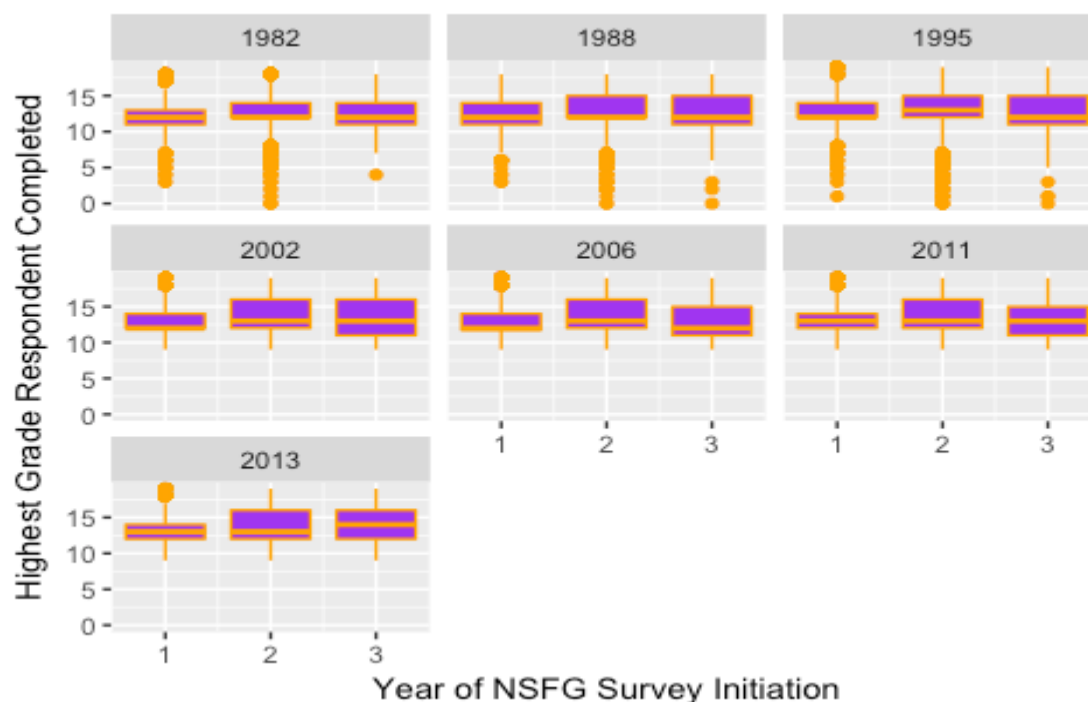


The following graphs depict the changes in highest level of education achieved for all respondents across survey years. This plot was generated because an anova found the differences between years to be significant, but post hoc tests were not performed to determine which years were significantly different from the others. While the first plot depicts the mean highest grade completed for all respondents in each cohort, the second graph contains one plot per survey year and splits the cohort by race. It is clear that women of color and other racial minorities typically have a lower level of education relative to the Caucasian group, but it is unclear if this difference across races is statistically significant.

Educational Attainment by Survey Year



Educational Attainment by Race for each Survey Year



Exploratory Analysis of Contraceptive Utilization

This section contains information pertaining to contraceptive utilization across the survey cycles. These values were calculated using the "prop.table(table())" as well as various data reformatting methods such as "data.frame(matrix())", "table()", and "cbind()". I struggled greatly with figuring out the best way to cross-tabulate this information, and while correct, the method used is inefficient and time consuming.

Below is a selection of contraception methods and the percentage of women in each survey who *have ever* tried that method.

Year of NFGS	Pill	Condom	IUD	Withdrawal	Rhythm	NFP	Patch	Ring
1982	76.07	48.32	18.78	20.4	13.63	1.8	NaN	NaN
1988	77.84	39.16	14.49	11.38	9.15	1.48	NaN	NaN
1995	76.29	80.99	10.29	39.12	23.71	4.11	NaN	NaN
2002	73.13	88.99	5.66	55.48	16.29	3.39	0.79	NaN
2006-2010	70.42	92.9	NaN	59.03	17.11	3.89	10.88	5.98
2011-2013	69.49	94.5	NaN	64.2	16.88	3.15	11.14	8.96
2013-2015	67.92	94.27	13.18	64.66	14.39	2.93	9.46	9.46

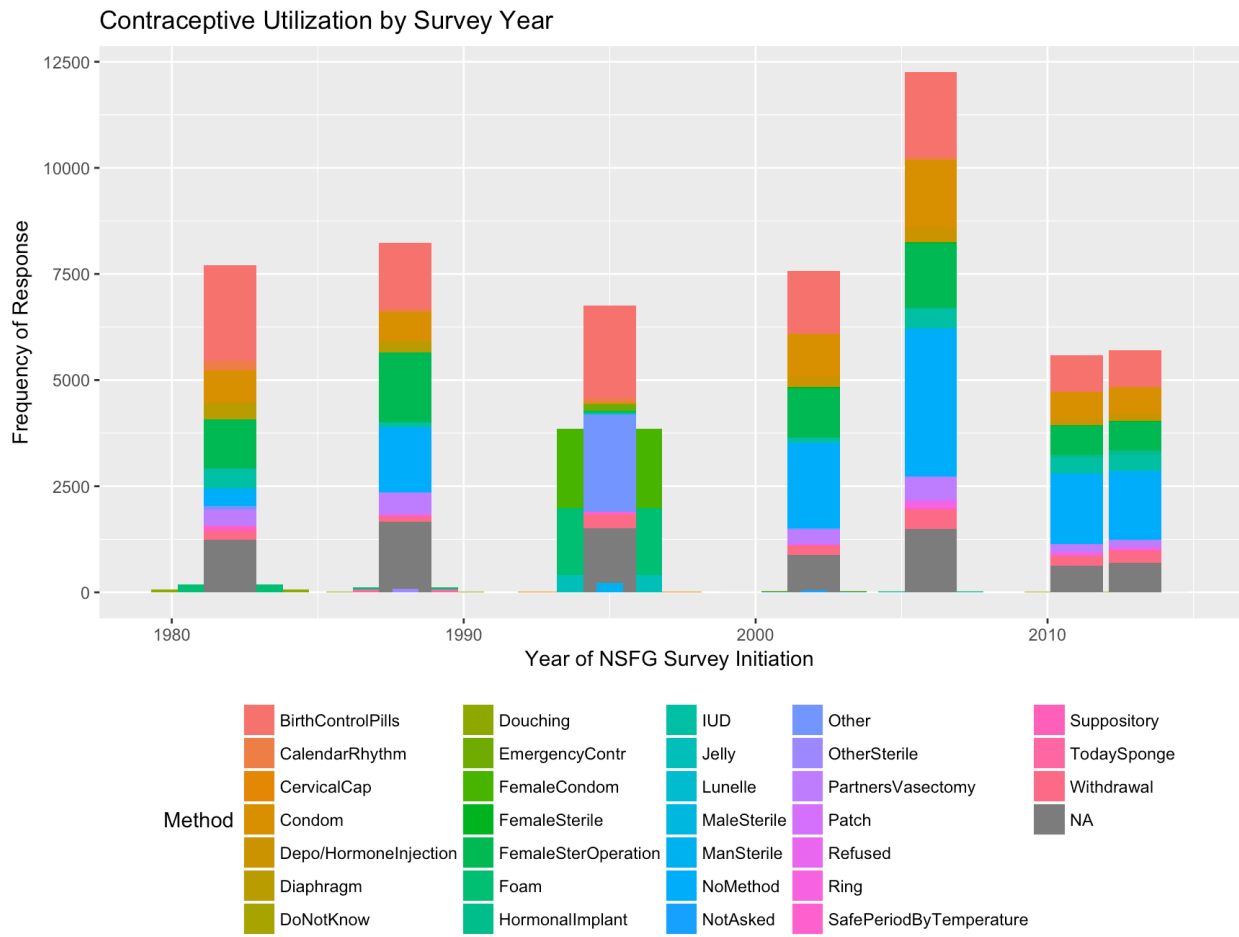
Data Visualization

These graphs are visualizations of the current contraceptive method utilized by women in this survey.

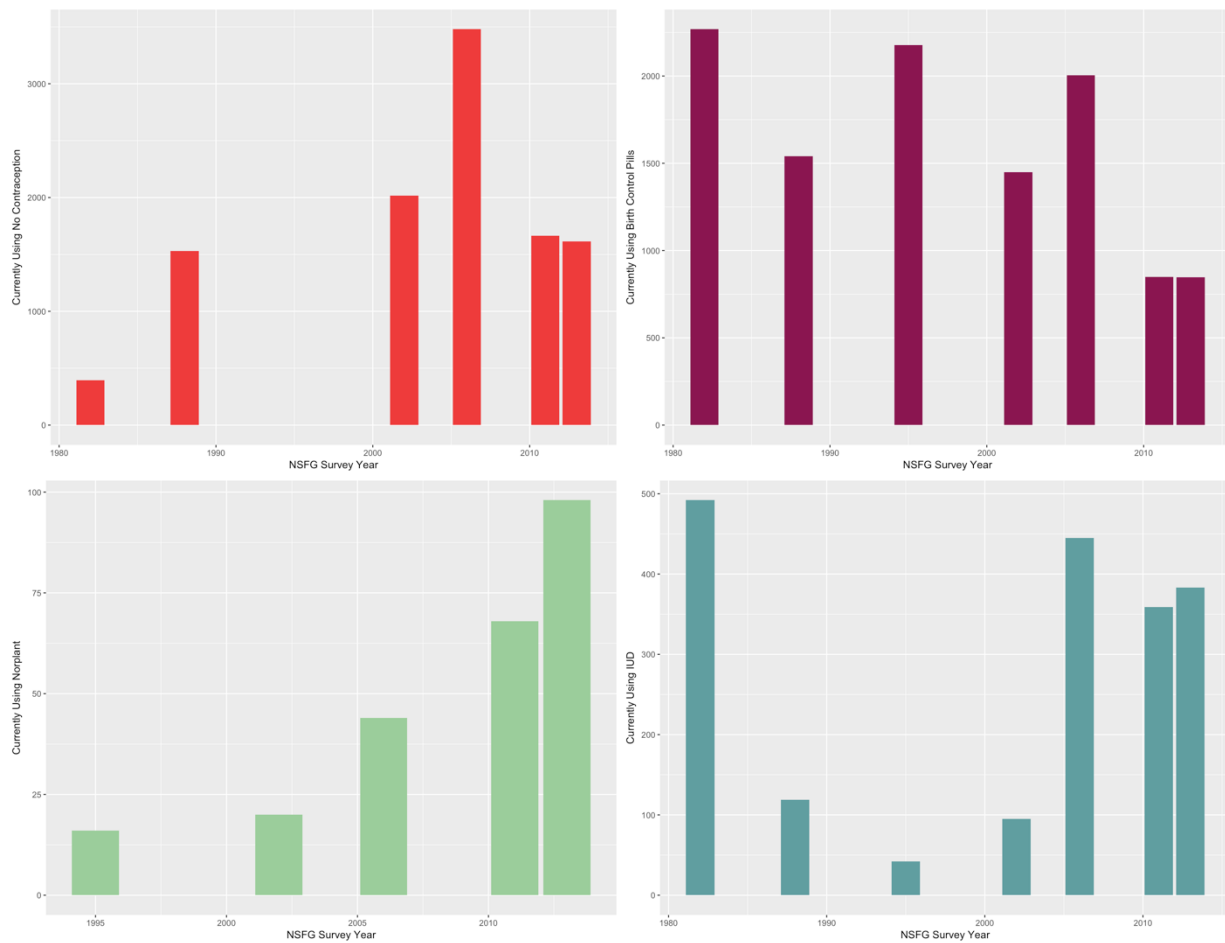
GGPlot was utilized in the making of all but one visualization (qplot was used), but there are clear ways to improve these graphics. The addition of animation would alleviate the difficulty of seeing certain variables within the graphs, especially when dealing with longitudinal changes. Issues with data recoding have made it difficult to develop visualizations that capture a complete story, since so many data quality issues have arisen, and data recoding should be continued to ensure that the data is perfectly cleaned prior to building out visualizations for publication.

This graph displays the count of women who utilize each birth control method. Because each cycle had differing number of female respondents, this graphic does not indicate the presence of any trends over time. Additionally, the stacking of methods in some years remains a mystery to me. This does not appear in the other stacked graphs, and I'm not sure how its occurring.

```
## Warning: position_stack requires non-overlapping x intervals
```

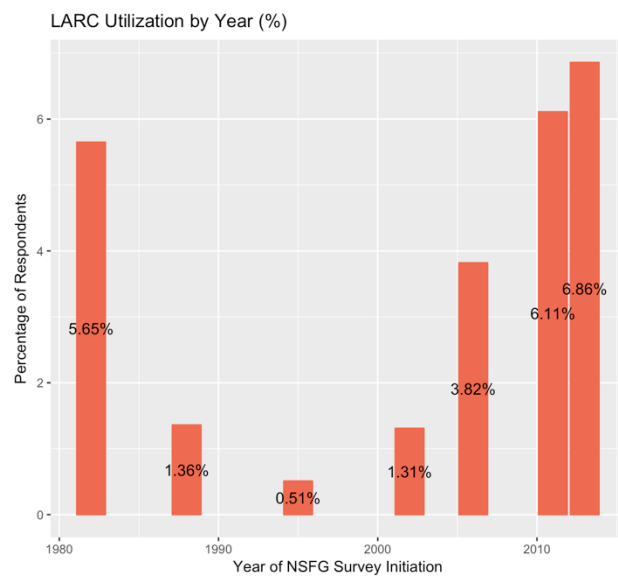


The following graphics indicate the number of women in each cycle that utilize particular forms of popular contraceptive methods (Pills, Norplant, or IUD) or use no contraception. Again, these values are frequencies, not percentages, so they are influenced by the amount of women recruited into the cycle and do not truly indicate trends of increased or decreased utilization over time. Notably, Norplant was not available prior to the 1995 NSFG survey.

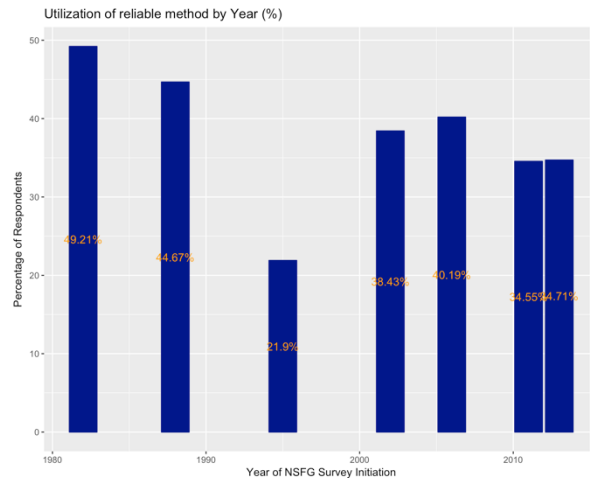


Basic Year By Current Contraceptive Method Use Plots

The following plot depicts the utilization of LARCs as a percentage of total respondents for each survey cycle.

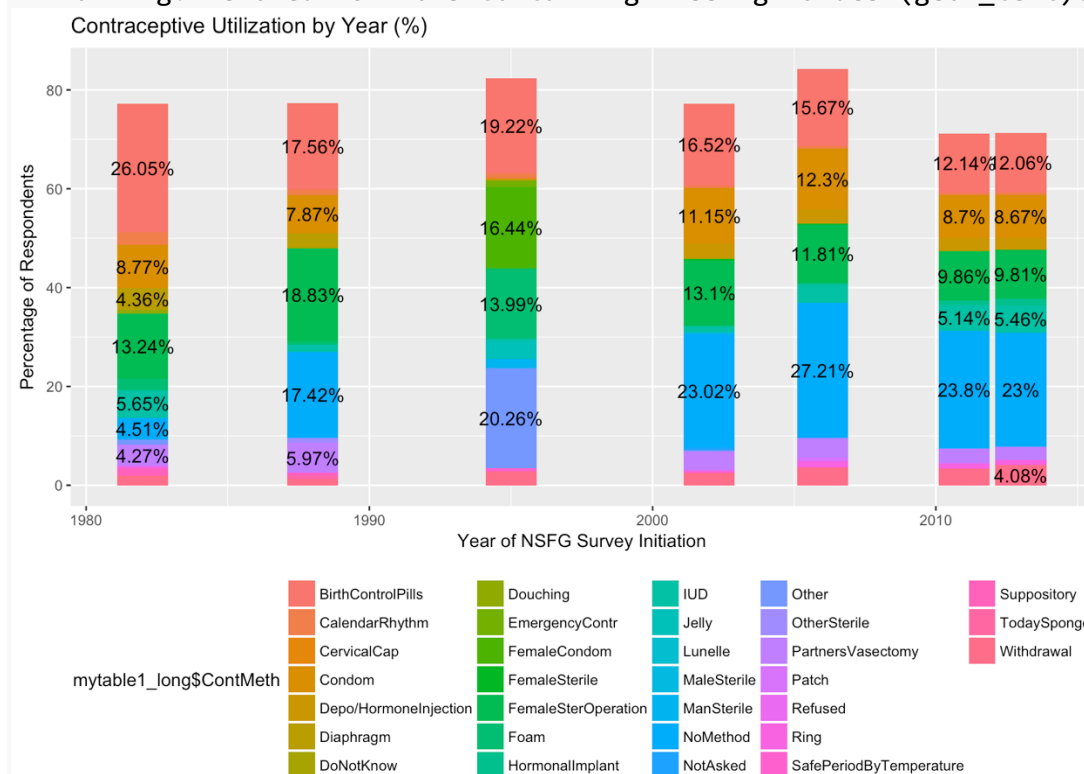


The following plot depicts the utilization of RELIABLE contraception, defined here as <10% failure rate, as a percentage of total respondents for each survey cycle.



The following graphic depicts contraceptive utilization as a percent of respondents in the cycle. Any method with a utilization % of over 4% has the corresponding percentage value overlaid. This visualization is much more useful than prior graphics that just depict the utilization as a frequency, rather than a percentage. None of the cycles reach 100% because of missing data.

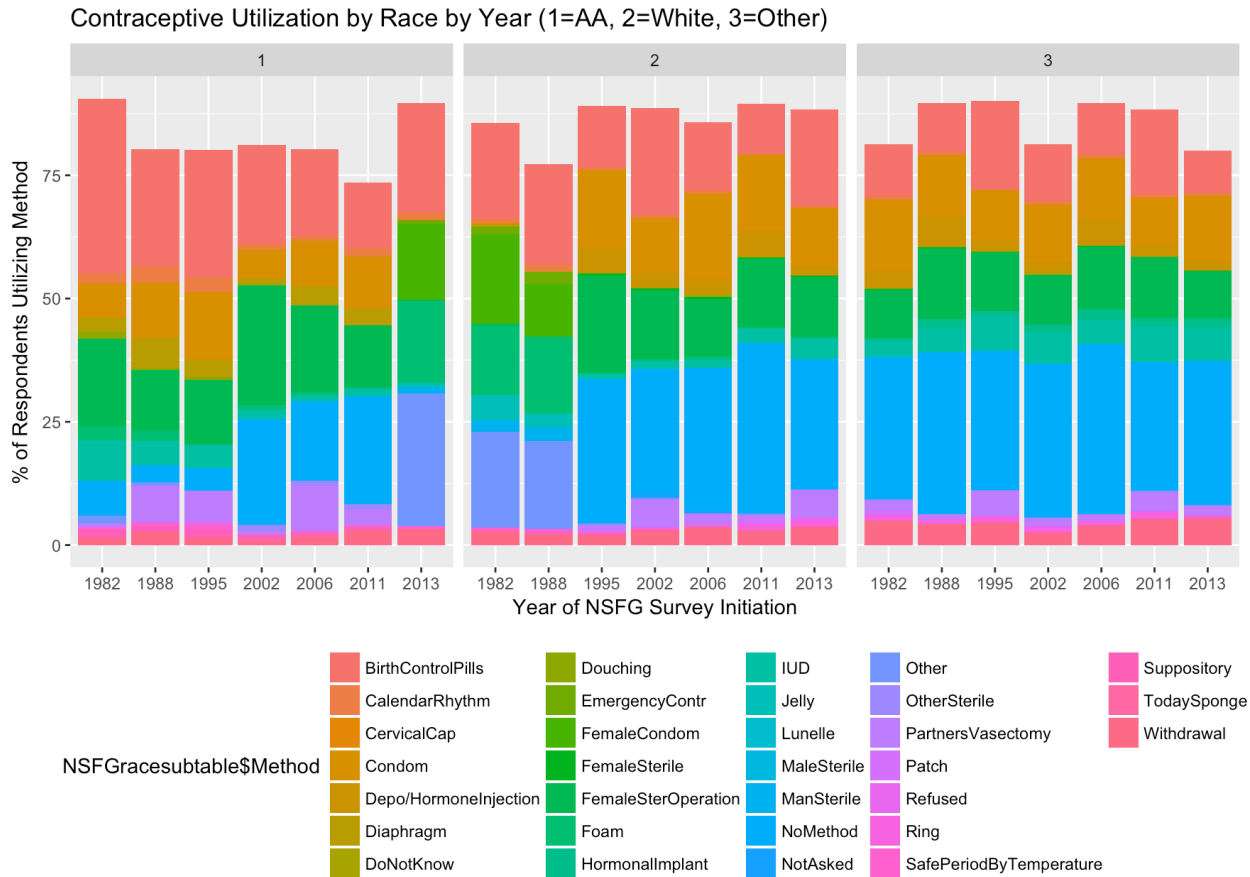
Warning: Removed 182 rows containing missing values (geom_text).

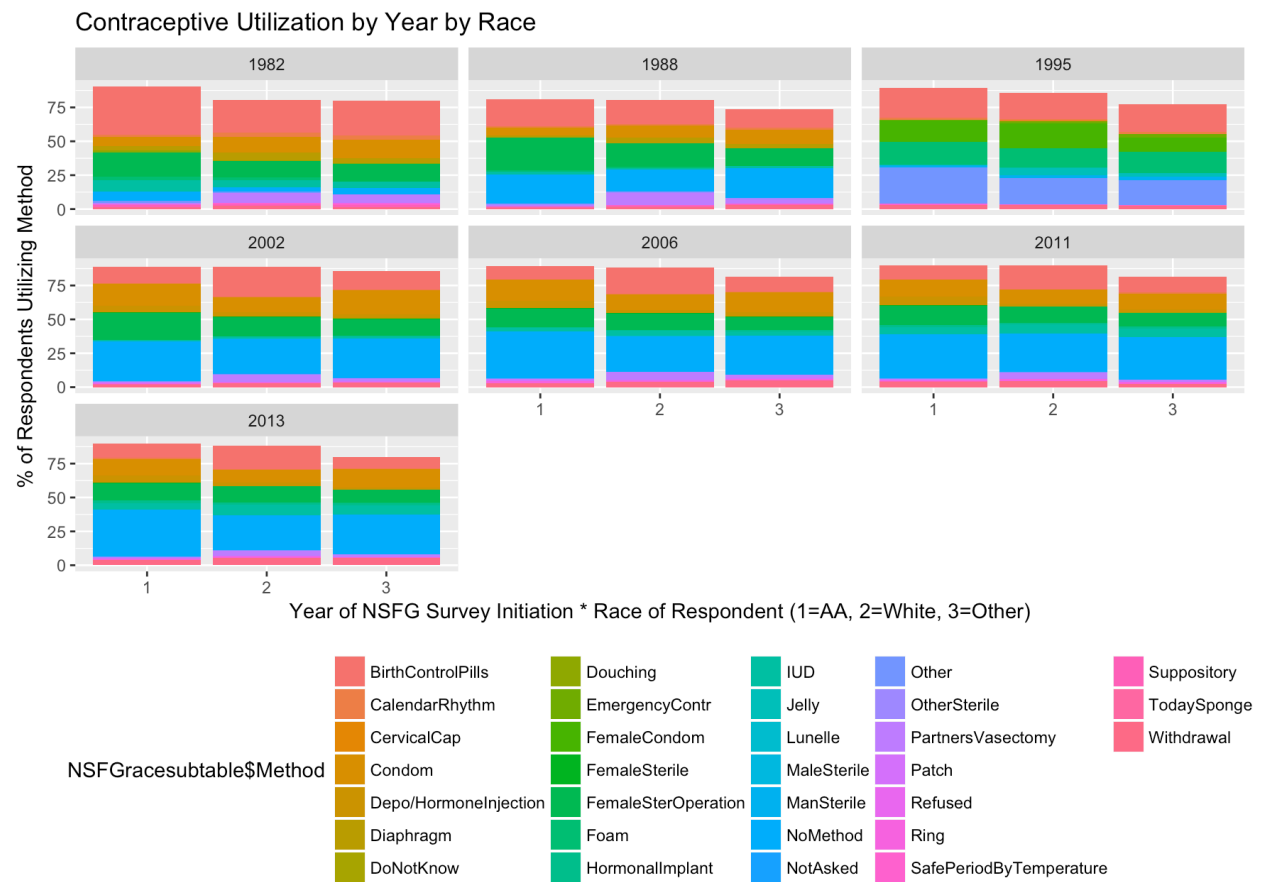


Basic Year By Current Contraceptive Method By Race

The following graphics display the same information as the graph immediately above, but this information is stratified by Race (in addition to stratification by survey year). In the first

graph, the graphs are first split by race than the x-axes are split based on survey year. This visualization can show trends in utilization within a race over time. In the second, survey year is used to split of the graphs into a series to determine what the racial differences in utilization were by year. This visualization is useful for seeing differences between individuals of different races within the same survey cycle.





Final Thoughts and Future Directions

The recoding and reformatting of the data sets was the most time consuming part of this project. My methods changed with each recode and rearrangement of data, as the data was never in the proper format for the desired analysis.

Since I do intend on using R in my future work, it was important for me to spend time on data cleaning, rather than animations, since data wrangling will be a necessary part of any dissertation-required work that I must complete. While I eventually chose to consider the data "clean enough" to conduct analyses and generate graphical representations, I would not feel comfortable trying to publish findings without more thorough investigations into the data. This project has helped me to better understand the layering required for ggplot and the usefulness of dplyer and piping (which would have ideally been used more from the start). Most notably, I have solved various problems throughout the process that have empowered me to feel confident working with a data set with fewer variable discrepancies.

References

- [1] CDC, National Center for Health Statistics: National Survey of Family Growth. Retrieved from <https://www.cdc.gov/nchs/nsfg/>
- [2] Healthy People 2020: Family Planning. Retrieved from <https://www.healthypeople.gov/2020/topics-objectives/topic/family-planning>
- [3] https://www.cdc.gov/reproductivehealth/unintendedpregnancy/pdf/contraceptive_methods_508.pdf
- [4] Sonfield A. Popularity Disparity: Attitudes About the IUD in Europe and the United States, Guttmacher Policy Review, Fall 2007 10(4).
- [5] https://www.cdc.gov/nchs/nsfg/nsfg_2013_2015_puf.htm
- [6] https://www.cdc.gov/nchs/nsfg/nsfg_2011_2013_puf.htm
- [7] https://www.cdc.gov/nchs/nsfg/nsfg_2006_2010_puf.htm
- [8] https://www.cdc.gov/nchs/nsfg/nsfg_cycle6.htm
- [9] https://www.cdc.gov/nchs/nsfg/nsfg_cycle5.htm
- [10] https://www.cdc.gov/nchs/nsfg/nsfg_cycle4.htm
- [11] https://www.cdc.gov/nchs/nsfg/nsfg_cycle3.htm
- [12] https://www.cdc.gov/nchs/nsfg/nsfg_cycle2.htm
- [13] https://www.cdc.gov/nchs/nsfg/nsfg_cycle1.htm

Appendix A

Below is a summary of the sixty-three variables chosen for extraction:

- MAINTENANCE
 - CASEID: Respondent's (R's) Study ID.
 - CMINTVW: Century Month of Interview.
 - INTVWYEAR: Year of Interview.
- DEMOGRAPHICS
 - AGE_R: Age of R
 - MARSTAT: Marital Status of R
 - HIGRADE: R's current grade in school or highest grade/year attended
 - DEGREES: R's highest college or university degree received
 - PARMARR: Whether R's biological parents married at R's birth
 - ACASILANG: Language to be used in ACASI
 - AGER: Recode of R's Age
 - HISPANIC: Whether R identifies as Hispanic
 - RACE: R's Race
 - CURR_INS: Current health insurance status
 - METRO: Metro versus Non-Metro as place of residence
 - RELIGION: R's current religions affiliation
 - LABORFOR: R's employment status
 - POVERTY: R's income as a percentage of poverty level
 - TOTINCR: R's total household income
 - PUBASSIS: Whether R currently receives public assistance
- CONTRACEPTIVE USE
 - PILL- Whether R has ever used birth control pills?
 - NORPLANT- Whether R has ever used hormonal implants?
 - CONDOM- Whether R has ever used condoms?
 - VASECTMY- Whether R has ever depended on partner's vasectomy?
 - DIAFRAGM- Whether R has ever used a diaphragm?
 - DEPOPROV- Whether R has ever used depo-provera or other injectable?
 - WIDRAWAL- Whether R has ever used withdrawal?
 - RHYTHM- Whether R has ever used rhythm method?
 - SDAYCBDS- Whether R has ever used Standard Days or CycleBeads?
 - TEMPSAFE- Whether R has ever used natural family planning (safe temperature)?
 - WOCONDOM- Whether R has ever used a women's condom?
 - FOAMALON- Whether R has ever used contraceptive foam?
 - JELCRMAL- Whether R has ever used spermicide jelly or cream?
 - CERVLCAP- Whether R has ever used a cervical cap?
 - SUPPOSIT- Whether R has ever used suppositories?

- TODAYSPG- Whether R has ever used the Today Sponge?
- IUD- Whether R has ever had an IUD?
- LUNELLE- Whether R has ever used Lunelle injection?
- PATCH- Whether R has ever used contraceptive patch?
- RING- Whether R has ever used the contraceptive ring?
- OTHRMETH01- Has R used another method?
- ECTIMESX- Number of times R used emergency contraception.
- FIRSMETH1- What was first method R used?
- CURRMETH1- What is the current method R uses (1)?
- CURRMETH2- What is the current method R uses (1)?
- CURRMETH3- What is the current method R uses (1)?
- CURRMETH4- What is the current method R uses (1)?
- REPRO HEALTH
 - MAYBPREG- Currently, does R believe she could be pregnant?
 - NPREGS_S- Total number of pregnancies
 - PARITY- Total number of live births
 - ABORTION- Total number of induced abortions
 - HYST- Has R had a hysterectomy?
 - OVAREM- Has R had an ovariectomy?
 - OTHR- Is R surgically sterile due to another operation?
 - PREGNOWQ- Is R currently pregnant?
- SEXUAL HX
 - EVERSEX- Has R ever had heterosexual intercourse?
 - RHADSEX- Has R ever had heterosexual intercourse (recoded)?
 - YNOSEX- Main reason R has not had heterosexual intercourse?
- SEX ED
 - SEDBC- Did R receive formal Sex Ed Before 18: Methods of Birth Control
 - SEDBCG- R's Grade When Received Instruction on Birth Control
 - SEDBCSX- Received Sex Ed on Birth Control Methods Before/After 1st Sex
 - SEDWHBC- Did R receive formal Sex Ed Before 18: Where to Get Birth Control
 - SEDWHBCG- R's Grade When Received Instruction on Where to Get Birth Control
 - SEDWBCSX- Received Sex Ed on Where to get Birth Control Before/After 1st Sex

Appendix B: Larger Version of Contraceptive Utilization Graphic

