
Data Project Guidelines

The Big Picture

This data project is one of the central components of the course. In class, we are covering a wide range of empirical approaches to measuring mobility and inequality for the United States as a whole over the past two centuries. We are going to examine how structural changes to the US economy and major political movements shaped mobility and inequality patterns nationally. The goal of this data project is to do the same thing, but with a narrow focus on Williamsburg. As a class, we will construct inequality and intergenerational mobility data for Williamsburg from the times of the Civil War to the present. These data will then form the basis your final research papers.

The data project consists of two distinct components. The first is the creation of an intergenerational dataset of Williamsburg families, constructed by linking those families across federal censuses. These intergenerational data will provide insights into the geographic and occupational mobility of Williamsburg residents over time. The second is a dataset of property histories. These data will capture ownership histories and changes in covenants across the different neighborhoods of Williamsburg. Details on the construction of each of the datasets are provided below.

An Intergenerational Dataset of Williamsburg Families

Each student will be responsible for linking a sample of approximately 20 Williamsburg residents forward from the 1870 federal census to the 1880, 1900, 1910, 1920, 1930 and 1940 censuses.¹ This methodology is similar to that used in several of the papers discussed in class, including Long & Ferrie (2007), Long & Ferrie (2013), Feigenbaum (2014) and Collins & Wanamaker (2015).

General Instructions

You will be assigned a spreadsheet containing approximately twenty individuals from the 1870 federal census living in Williamsburg. The spreadsheet will contain a unique id number for each individual as well as the individual's name, place of residence in 1870, year of birth, place of birth, and gender. Each individual will also have a link to a page of Ancestry.com search results.

¹The original 1890 census manuscripts were destroyed. The 1940 census is the last federal census that is publicly available. Census data becomes fully public after 72 years.

1. Open the link to the Ancestry.com 1870 census search results. Note that you may first have to log in through Swem's link to Ancestry available [here](#). Given the individual's characteristics, you should be able to find a unique match in this first page of search results.
2. Once you have identified the match, copy the url for the match's 'View Record' link, paste it into the appropriate column in your spreadsheet, and then click on the match's 'View Record' link, opening the detailed 1870 census information for the individual. In your spreadsheet, record the race and household member names in the appropriate columns in your spreadsheet. These details will help you link the individual to subsequent federal censuses.
3. Now that you have obtained all of the relevant information for linking your 1870 individuals, search for them in the 1880 federal census. The Ancestry.com search page for the 1880 census can be accessed [here](#). Use all of your information about the individual to try to find them in 1880 (name, birth year, birth place, family member names, etc.).
4. Once you determine the best match, record the match quality in the relevant columns on your spread sheet. The first aspect of match quality is whether the match represents a good, unique match. Enter the most appropriate of the following designations in the match type column:
 - **Good and unique** - In this case, the match looks like a good match for your individual and there should be no other matches that look like a good match.
 - **Good and nonunique** - In this case, the match looks like a good match for your individual but there is at least one other individual that looks like a good match.
 - **Bad** - In this case, even the best match does not look like a good match.

The second aspect of match quality is your assessment of how good the match is on a zero to ten scale, with zero being a terrible match and ten being a perfect match. This is a subjective score and your scores may differ from those of other students. The key is that you be consistent throughout your dataset with how you assign your match quality scores. Enter your score in the match quality column.

5. If you found a good match (whether unique or nonunique), copy the link to the relevant 1880 census record (the 'View Record') link in the relevant column in your spreadsheet.
6. Now view the 1880 census record for your match. If the individual has any children in the 1880 census who are not already in your spreadsheet, create new rows in the spreadsheet and enter these individuals' details. These details include the individual's name, birth year, birth place, gender, race and the url for the individual's 1880 census record. You should also create a new id number for the individual using the following rules:

- The first part of the new individual's id number should be the same as the original individual's number.
 - Add a period to this number and then add a '1' after the period if it is the first new individual from that household added to the dataset.
 - Add a '2' after the period if it is the second new individual from that household added to the dataset.
 - Keep increasing this last digit of the new id number for each additional individual from the household added to the dataset.
 - As an example, suppose that you find individual 7 from the original dataset in the 1880 federal census. In 1880, this individual has three children who were not in your original sample. You would create three new rows in your dataset for these individuals, assigning them id numbers 7.1, 7.2 and 7.3. If you then find individual 7.2 in the 1900 census with two new children of his own, you would create two new rows in your dataset for these children, assigning them id numbers 7.2.1 and 7.2.2.
7. Now repeat steps 3 through 6 for all of your individuals from 1880, including the original individuals that you linked from the 1870 census and the new individuals you added to the spreadsheet in step 6, searching the 1900 census. The Ancestry.com search page for the 1900 census can be accessed [here](#).
 8. Continue to repeat the steps to link your individuals, both the original individuals and everyone added with each additional census, to the [1910](#), [1920](#), [1930](#), and [1940](#) federal censuses.

Once everyone has completed their samples, I will combine the samples into a single dataset and then transcribe additional details from all of the linked census records. I will then post the complete dataset online so that everyone can use it for their final projects.

A History of Williamsburg Neighborhoods

The second part of the data project requires each student to explore the history of two neighborhoods in Williamsburg. You will investigate the history of a single residential property in each neighborhood, tracing its price and ownership history and identifying any changes in covenants and zoning regulations restricting the use of the property. These data overlap with the types of evidence explored in Aaronson et al. (2017) and Rothstein (2017).

General Instructions

There will be a Doodle poll posted online listing several Williamsburg neighborhoods. You will sign up for two different neighborhoods. For each of these two neighborhoods, you will need to do the following:

1. Identify a single family dwelling in the neighborhood that is at least ten years old. This can be easily accomplished using a real estate search website like [Zillow](#). Note that there is a spreadsheet posted online giving location information for the neighborhoods in the Doodle pool.
2. If your property is located in the City of Williamsburg, use the city's online property lookup tool, available [here](#), to find details on the ownership history and assessed value of the property. If your property is located in James City County, the same information can be obtained [here](#). Record any and all relevant information about the property and note the parcel ID number. This number will be useful for searching earlier records at the clerk's office.
3. Once you have exhausted the information available through the online property information systems, you need to go to the district court clerk's office to trace the property's history further back. The clerk's office is located in the courthouse at 5201 Monticello Avenue, Williamsburg VA 23188.² At the clerk's office, you can use the computer terminals to search for earlier deeds related to your property. Trace the property's deeds back as early as you can, reading through each deed to uncover who was buying and selling the property and whether any restrictive covenants were attached to the property.³
4. Create a summary of the property history, citing any relevant deeds or other property records and including images of records you deem particularly important. Submit this summary to me by email.

Once everyone has submitted their property summaries to me, I will combine them into a single dataset for use by the entire class when working on the final project.

References

- Aaronson, D., Hartley, D., & Mazumder, B. (2017). The effects of the 1930s HOLC redlining maps. Tech. rep., Federal Reserve Bank of Chicago.
- Collins, W. J., & Wanamaker, M. H. (2015). The Great Migration in black and white: New evidence on the geographic mobility of American Southerners. *Journal of Economic History*, 75, 947–992.
- Feigenbaum, J. J. (2014). A new old measure of intergenerational mobility: Iowa 1915 to 1940. *Unpublished Manuscript*.

²Note that the clerk's office is open weekdays during business hours (exact hours can be found on their website). If you are unable to go to the courthouse due to schedule constraints or some other reason, please contact me and we will work out an alternative assignment for you.

³The clerk's office charges 50 cents per page for any printouts or copies of deeds. I will gladly reimburse you for any pages you print at the clerk's office.

Long, J., & Ferrie, J. (2007). The path to convergence: Intergenerational occupational mobility in Britain and the US in three eras. *The Economic Journal*, 117(519).

Long, J., & Ferrie, J. (2013). Intergenerational occupational mobility in Britain and the US since 1850. *American Economic Review*, 103(4), 1109–1137.

Rothstein, R. (2017). The color of law: A forgotten history of how our government segregated America. *New York, NY: Liveright*.