From September 2020 to May 2021, I performed various data analysis using R to assist with the Garcia v. DHCD class action litigation.

* From September – December 2020, the main focus of my work was calculating and modeling delays in disability accommodation. Scripts 1-3 were written to do this part of the project. For more information on this project, see the interactive web application I created here: <https://monica-chang.shinyapps.io/Garcia-vs-DHCD/>
* From January – May 2021, I performed various analyses that Laura asked me to in Scripts 4-10. These form less of a coherent project.
* Link to the Github repository (contains everything except the data for privacy reasons): <https://github.com/monica-chang/Garcia-vs-DHCD>.

This document is intended to explain what the various scripts, memos, and datasets in this folder contain. If you have any questions, please feel free to contact Monica Chang at [mychang07@gmail.com](mailto:mychang07@gmail.com) or 360-852-5609.

**SCRIPTS**

1. **clean\_ada\_data.Rmd:** This is an R Markdown file that cleans and merges the raw ADA request data from 2018 and 2020. This involved renaming columns for consistency, eliminating hyphens from the PEI, accounting for partial approvals, classifying certain requests as “unit type requests”, and adding any data about the zipcodes of service providers. The final export is approved\_adas\_complete.csv in the processed\_data folder.
2. **ada\_tables\_plots.Rmd:** This is an R Markdown file that generates some basic bar graphs showing the number of requests by accommodation type, the percentage of approved ADA requests by accommodation and reason type.
3. **ada\_delays.Rmd:** This is an R Markdown file that contains the bulk of analysis related to using shelter characteristics and zip codes to infer delays in ADA accommodation.
   1. **CONTEXT:** In the data that DHCD provided us, over 70% of values for the date an accommodation was met were missing – meaning that we could not easily calculate the average delay in accommodation. Because of this, we tried to be creative and use other methods to infer delays in accommodation:
      1. By looking at where individuals are transferred immediately after their request is approved and the characteristics of that transfer location, we can infer whether or not unit type accommodations (scattered site, first floor, wheelchair accessible) were met.
      2. By looking at where individuals are transferred immediately after their request is approved and the zipcode of their service provider, we can infer whether or not their request for closer proximity to a service provider was met.
   2. **INFERRING DELAYS IN ADA ACCOMMODATION FOR UNIT TYPE ACCOMMODATIONS**
      1. **Section 1 – Clean bed registry/shelter characteristics data:** I clean the dataset on shelter characteristics. This provides information on whether each location meets certain unit type accommodations (scattered site, first floor, wheelchair accessible, etc.)
      2. **Section 2 – Clean and merge transfer data:** I clean and merge the data on internal transfers (e.g. from Heading Home Congregate Shelter to a Heading Home Scattered Site) and external transfers (e.g. from Heading Home to JRI).
      3. **Section 3 – Merge approved ADAs for unit accommodations and transfer location:** For approved ADA requests that are for unit type accommodations (e.g. scattered site, first floor, wheelchair accessible), I add information about the location of all transfers that occurred after the request was approved. The final export is ada\_transfers.csv in the processed\_data folder.
      4. **Section 4 – Merge approved ADAs for unit accommodations and characteristics of transfer location:** I write a function that takes in the table generated from Section 3, the table generated from Section 1, and a TRUE/FALSE value indicating whether hotels are assumed to successfully meet unit accommodations. The table output contains a column indicating whether an individual’s unit accommodation was met by any of the transfers that occurred after their request was approved. The final exports are transfers\_chars\_hotel\_t.csv and transfers\_chars\_hotel\_f.csv in the processed\_data folder.
      5. **Section 5 – Keep the earliest transfer after approval that satisfies the request:** I write a function that takes in the table generated from Section 4 and supplements the missing date\_accommodation\_met values using the date of the transfer that satisfies the request and occurs most immediately after approval. The final exports are supplemented\_interesting\_ada\_transfers\_t.csv and supplemented\_interesting\_ada\_transfers\_f.csv in the processed\_data folder.
      6. **Section 6 – Plots showing delay in accommodation for approved ADAs:** This generates a series of Kaplan-Meier plots showing the percentage of approved ADA requests met over time for the following scenarios:
         1. Using only the data provided by DHCD.
         2. Using only the data provided by DHCD for unit accommodation requests.
         3. Inferring the date that a unit accommodation was met using transfer data and assuming hotels meet all unit accommodations.
         4. Inferring the date that a unit accommodation was met using transfer data and assuming hotels do NOT meet all unit accommodations.
      7. **Section 7 – Modeling delays in disability accommodation:** To compare whether requests that ask for certain accommodations or list certain reasons have longer or shorter delays in accommodation, I created a linear regression model. I use a Bayesian linear regression model to regress the length of delay on different predictors and calculate the average delay for requests with and without the predictor. Due to limitations of the data, the model is not reliable.
   3. **INFERRING DELAYS IN ADA ACCOMMODATION FOR PROXIMITY TO SERVICE PROVIDER ACCOMMODATION**
      1. **Section 9 – Calculating delays in accommodation for requests for proximity to service providers:** I go through a process similar to the one described above in Sections 1-7 to calculate the delays in accommodation for service provider requests. The main difference is that I use the zipcode of the transfer location and the zipcode of the provider to determine whether an individual’s accommodation is met. I use functions from the ggmap package to calculate the distance between the zipcode of transfer and the zipcode of provider. Temporarily, 20 miles is considered the cutoff for being close to a service provider.
4. **ada\_approval\_over\_time.Rmd:** This is an R Markdown file that generates some scatterplot showing the decline in the ADA approval rate, the change in the number of ADA requests, and the change in the number of approved ADA requests from 2015-2019.
5. **scat\_vacancy\_analysis:** This is an R Markdown file that generates a scatterplot that shows the change in the number of scattered site vacancies from 12/21/16 – 12/24/19.
6. **non\_compliance\_exit\_analysis.Rmd:** This is an R Markdown file that cleans 2018 and 2020 noncompliance data and merges the two datasets into issued\_ncs.csv in the processed\_data folder. This file also cleans 2018 and 2020 HUD exit data, attributes the appropriate text to the numerical HUD codes, and merges the two datasets into hud\_exit\_data.csv in the processed\_data folder. I then merge noncompliance data and TESI data with the dataset for approved ADA requests for unit accommodations to see how many of these individuals receive noncompliances/TESIs during their delay in accommodation. The final exports are ada\_ncs\_tesi\_exit\_final.csv
7. **interactive\_process\_analysis.Rmd:** This is an R Markdown file that takes in Erin’s spreadsheet and generates various scatterplots showing the change in the number of interactive process (IP) requests, the change in the number of IP requests that receives some form of follow-up, and change in the number of IP requests that receive some form of follow-up and accommodation.
8. **triage\_enrollment\_comparison.Rmd:** This is an R Markdown file that compares triage and enrollment dates and generates a few histograms showing the distribution of days from enrollment to triage from 07/02/2018 to 12/13/2019.
9. **clean\_legalserver\_export.Rmd:** This is an R Markdown file that takes the output of the Garcia-related LegalServer export and classifies the cases based on type (e.g. Garcia Related, Shelter Access, etc.), counts the number of cases under each advocate, and creates separate Excel files for each advocate. The exports are in the legalserver\_export subfolder of the processed\_data folder.If you want to re-run this script on a new LegalServer export, just upload the new export to the raw\_data folder and replace the name of the export in this script.
10. **miscellaneous.Rmd:** This is an R Markdown file that generates miscellaneous spreadsheets that Laura asked for, including: approved ADA requests that ask for proximity to a service provider, approved ADA requests that have NA for date\_received, individuals that underwent many transfers, etc.

**PROCESSED DATA**

* More important/useful files
  + **approved\_adas\_complete.csv**: This is a table with all approved ADA requests from 2015-08-04 to 2019-12-25. Each row is uniquely identified by PEI and the date the application was received. For each request, there is information about the date the application was received, the date of application decision, the date the accommodation was met (as provided by DHCD), the accommodations requested, the reasons for the request, and the zipcode of provider.
  + **ada\_geography.xlsx:** This is an Excel file of all the approved ADA requests for to be closer to a service provider. If you want to update the zipcodes of service providers, you should edit this spreadsheet and run clean\_garcia\_data.Rmd and ada\_delays.Rmd.
* Other ADA files
  + **ada\_request\_2018\_2020\_partials\_unprocessed.csv:** This is the merged table of 2018 and 2020 ADA request data where partial approvals are listed as such. Each row represents a unique accommodation requested (e.g. if someone requested both first floor and scattered site in their application, this would appear as 2 different rows).
  + **ada\_request\_2018\_2020\_partials\_processed.csv:** This is identical to ada\_request\_2018\_2020\_partials\_unprocessed.csv but with partial approvals accounted for (e.g. if a first floor request was approved but a scattered site was not, this is recorded) and frequently occurring answers in the “Other” category given their own accommodation category (e.g. AC Unit, additional bedrooms).
  + **all\_adas.csv**: This is ada\_request\_2018\_2020\_partials\_processed.csv pivoted so that each row represents a unique request made (i.e. if someone requested both first floor and scattered site in their application, this would appear as 1 row). Unit type requests (requests for scattered site placement, first floor, or wheelchair access) are identified as “interesting.”
  + **unit\_type\_ada\_transfers.csv:** This contains all approved ADA requests for unit type accommodations (e.g. scattered site, first floor, wheelchair accessible) including information about the location of all transfers that occurred after approval.
  + **transfers\_chars\_hotel\_t.csv:** This contains all approved ADA requests for unit type accommodations including information about whether any transfer that occurred after approval satisfied the requested accommodation. This dataset assumes that all hotel placements meet the requested unit type accommodations.
  + **transfers\_chars\_hotel\_f.csv:** This contains all approved ADA requests for unit type accommodations including information about whether any transfer that occurred after approval satisfied the requested accommodation. This dataset assumes that all hotel placements do NOT meet the requested unit type accommodations.
  + **supplemented\_unit\_type\_ada\_transfers\_t.csv:** This contains all approved ADA requests for unit type accommodations and supplements missing date\_accommodation\_met values using the method previously described. This dataset assumes that all hotel placements do meet the requested unit type accommodations.
  + **supplemented\_ unit\_type\_ada\_transfers\_f.csv:** This contains all approved ADA requests for unit type accommodations and supplements missing date\_accommodation\_met values using the method previously described. This dataset assumes that all hotel placements do NOT meet the requested unit type accommodations.
  + **issued\_ncs**.**csv:** This contains the PEI, date of noncompliance, type of noncompliance (e.g. termination, warning, TESI) and type of violation (e.g. health and safety, not participating in rehousing plan) for all non-compliances issued from 2014-2019.
  + **hud\_exit\_data.csv:** This contains the PEI, date of exit, exit destination (e.g. shelter, rental with subsidy) and exit reason (e.g. FAH, transferred, completed program) for all shelter exits from 2015-2019.
  + **ada\_ncs\_tesi\_exit\_final.csv:** This contains all approved ADA requests for unit type accommodations and the date of the first noncompliance after the ADA request is submitted, first TESI after the ADA request is submitted, first exit after the ADA request is submitted, and reason for first exit after the ADA request is submitted – if these events occurred.
* miscellaneous folder
  + **ada\_date\_acc\_met\_filled\_sample\_t.csv:** This contains a random sample of 50 individuals who have approved ADA requests and a filled out date when the accommodation was met and require a transfer and information about their subsequent transfers.
  + **ada\_date\_acc\_met\_filled\_sample.csv:** This contains a random sample of 50 individuals who have approved ADA requests and a filled out date when the accommodation was met and require a transfer.
  + **ada\_date\_acc\_met\_filled\_t.csv:** This contains all approved ADA requests that have a filled out date when the accommodation was met and require a transfer and information about their subsequent transfers.
  + **ada\_date\_acc\_met\_filled.csv:** This contains all approved ADA requests that have a filled out date when the accommodation was met and require a transfer.
  + **ada\_no\_transfers\_supplemented.csv:** This contains contains the date of application (date\_received), date of approval (date\_decision), and the date of internal and external transfers (date\_referral), transfer type (internal/external) and transfer location for a list of potential witnesses that Dan generated in 12/2020. Some individuals show up more than once because they underwent more than one transfer.
  + **approved\_adas\_many\_tranfers.csv:** This contains all approved ADA requests that have undergone a lot of transfers (10+).
  + **approved\_adas\_na\_date:** This contains all approved ADA requests that have “NA” for date\_received.
  + **service\_provider\_requests.csv:** This contains all approved ADA requests that include a request for proximity to service provider.
  + **transfer\_locations\_with\_zips.xlsx:** This contains the list of transfer locations (hotels/shelters) that do not have a corresponding zipcode – after another intern (Caitlin) found the zipcodes through Googling.
  + **transfer\_locations\_without\_zips.csv:** This contains the list of transfer locations (hotels/shelters) that do not have a corresponding zipcode.
* legalserver\_export folder
  + **Clean\_Garcia\_Related\_Cases\_ALL\_09\_24\_2020\_\_06\_25\_2021.xlsx**: Cases for all shelter advocates from 09/24/20 – 06/25/21 with separate tabs for each advocate. This Excel file was created manually by combining the Excel files for each advocate.
  + **Clean\_Garcia\_Related\_Cases.xlsx:** Cases for all shelter advocates from 09/24/20 – 06/25/21.
  + **Adam\_Hoole\_Garcia\_Related\_Cases.xlsx:** Adam’s cases from 09/24/20 – 06/25/21.
  + **Laticia\_Walker\_Simpson\_Garcia\_Related\_Cases.xlsx:** Laticia’s cases from 09/24/20 – 06/25/21.
  + **Linda\_Garcia\_Garcia\_Related\_Cases.xlsx:** Linda’s cases from 09/24/20 – 06/25/21.
  + **Other\_Advocates\_Garcia\_Related\_Cases.xlsx:** Other advocates’ cases from 09/24/20 – 06/25/21.

**RAW DATA**

All of the spreadsheets in this folder come directly from DHCD (unedited).

**SHINY**

All of these files are used to create the interactive Shiny web app: <https://monica-chang.shinyapps.io/Garcia-vs-DHCD/>.