**Katie’s dataset notes**

* 01.Overview\_Geospatial\_Data.pdf: Changes recommended.
  + Prose notes because I don’t have the original slide files.
  + Title: I’d make the period in the title into a “\_”.
  + Slides 17-18
    - Location example looks like an all-white ranch/4H program (<https://asotincountyfairandrodeo.org/4-h/>)
    - Attribute explanation not clear (add example attributes?)
    - Possibly use this location instead? <https://commons.wikimedia.org/wiki/File:Coyote_Point_Trail_at_Whitewater_State_Park,_Minnesota_(44136078811).jpg> (photo has a CC 2.0 license)
  + Slide 31: CA maps look great (local, social science-y)
  + Slide 34: Given the racialization of what counts as “crime” in the United States (and even more overt racism in police presence), I’d replace crime locations with something else, e.g., motor vehicle accident locations or point environmental pollutant emissions locations from the Toxics Release Inventory (e.g., <https://www.epa.gov/toxics-release-inventory-tri-program/tri-basic-data-files-calendar-years-1987-2019>)
  + Rest of slides: Good data!
* 02\_Introduction\_to\_GeoPandas.ipynb: datasets updated/links checked/done
  + Datasets used (all from United States Census):
    - California counties shapefile ('notebook\_data/california\_counties/CaliforniaCounties.shp')
    - California places cartographic boundaries for 2019 from the [United States Census Bureau](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb\_2019\_06\_place\_500k.zip) ('notebook\_data/census/Places/cb\_2019\_06\_place\_500k.zip')
* 03\_CRS\_Map\_Projections.ipynb: datasets updated/links checked/done
  + Datasets used
    - California counties shapefile (‘notebook\_data/california\_counties/CaliforniaCounties.shp’)
    - United States states shapefile (‘notebook\_data/us\_states/us\_states.shp’)
    - California places cartographic boundaries for 2019 from the [United States Census Bureau](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb\_2019\_06\_place\_500k.zip) ('notebook\_data/census/Places/cb\_2019\_06\_place\_500k.zip')
  + 04\_More\_Data\_More\_Maps.ipynb: datasets updated/links checked/done Datasets used
    - 'notebook\_data/transportation/BerkeleyBikeBlvds.geojson' (Berkeley bike boulevards)
      * Social science-relevant
      * Local
    - 'notebook\_data/alco\_schools.csv' (Alameda school locations)
      * Social science-relevant
      * Local
    - 'notebook\_data/parcels/parcel\_pts\_rand30pct.geojson'
      * Social science: Housing parcel data is definitely social science-relevant!
      * Local
      * Wasn’t clear on first read what a parcel was or the origin of the (I’d say “housing” or “real estate” parcel data and clarify the origin of the dataset in the workshop text)
    - ‘notebook\_data/berkeley/BerkeleyCityLimits.shp’
      * Social science-relevant
      * Local
* 05\_Data-Driven\_Mapping.ipynb: datasets updated/links checked/done
  + Datasets used
    - California counties shapefile ('notebook\_data/california\_counties/CaliforniaCounties.shp')
    - Alameda County schools ('notebook\_data/alco\_schools.csv')
    - Berkeley bike boulevards ('notebook\_data/transportation/BerkeleyBikeBlvds.geojson')
* 06\_Spatial\_Queries.ipynb: datasets updated/links checked/done
  + Datasets used
  + California census tract cartographic boundaries for 2019 from the [United States Census Bureau](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb\_2019\_06\_place\_500k.zip) ('notebook\_data/census/Tracts/cb\_2019\_06\_tract\_500k.zip')
  + California protected areas ('notebook\_data/protected\_areas/CPAD\_2020a\_Units.shp')
  + Berkeley city limits ('notebook\_data/berkeley/BerkeleyCityLimits.shp')
  + Alameda County schools ('notebook\_data/alco\_schools.csv')
  + Berkeley bike boulevards ('notebook\_data/transportation/BerkeleyBikeBlvds.geojson')
  + Bay Area Rapid Transit (BART) stations ('notebook\_data/transportation/bart.csv')
* 07\_Joins\_and\_Aggregation.ipynb datasets updated/links checked/done
  + Datasets used
    - California American Community Survey (ACS) data from the United States Census Bureau by tract ('notebook\_data/census/ACS5yr/census\_variables\_CA.csv') (developer note: could use updating at a later date)
    - California census tract cartographic boundaries for 2019 from the [United States Census Bureau](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb_2019_06_place_500k.zip) ('notebook\_data/census/Tracts/cb\_2019\_06\_tract\_500k.zip')
    - Alameda County schools ('notebook\_data/alco\_schools.csv')
* 08\_Pulling\_It\_All\_Together.ipynb datasets updated/links checked/done
  + Datasets used
    - Census tracts in Alameda County joined to American Community Survey data: 'outdata/tracts\_acs\_gdf\_ac.json'
    - California 2019 grocery store data: 'notebook\_data/other/ca\_grocery\_stores\_2019\_wgs84.csv'
* 09\_ON\_YOUR\_OWN\_A\_Full\_Workflow.ipynb datasets updated/links checked/done
  + Datasets used, no changes recommended, but not listed at beginning
* - Alameda County polling locations ('notebook\_data/ac\_voting\_locations.csv')
* - California census tract cartographic boundaries for 2019 from the [United States Census Bureau](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb\_2019\_06\_place\_500k.zip) ('notebook\_data/census/Tracts/cb\_2019\_06\_tract\_500k.zip')
* - Census place boundaries for Berkeley ('outdata/berkeley\_places.shp')
* - Alameda County bus routes for fall 2020 ('notebook\_data/transportation/Fall20Routeshape.zip')

10\_OPTIONAL\_Fetching\_Data.ipynb datasets updated/links checked/done

* + Datasets used:
    - - California census tract cartographic [boundaries](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb\_2019\_06\_place\_500k.zip) for 2019 from the United States Census Bureau
    - - United States county cartographic [boundaries](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb\_2019\_us\_county\_500k.zip) for 2019 from the United States Census Bureau
    - - San Francisco [bikeway routes](https://data.sfgov.org/api/geospatial/ygmz-vaxd?method=export&format=GeoJSON)
    - - Berkeley [bikeway routes](https://data.cityofberkeley.info/api/geospatial/fgw9-98ic?method=export&format=GeoJSON)
    - - Open Street Map San Francisco and Berkeley street networks obtained with the Python OSMNX package

**11\_OPTIONAL\_Basemap\_with\_Contextily.ipynb updated links/checked/done**

* + **Datasets used (**not listed at beginning)
    - OpenStreetmap Mapnik basemap (via contextily)
    - 2019 cartographic boundary (cb\_) file of California (06) places: <https://www2.census.gov/geo/tiger/GENZ2019/shp/cb_2019_06_place_500k.zip>
    - BART station locations: 'https://raw.githubusercontent.com/dlab-berkeley/Geospatial-Fundamentals-in-Python/master/notebook\_data/transportation/bart.csv'
    - ax, source=cx.providers.NASAGIBS.ModisTerraTrueColorCR (via contextily)
* **12\_OPTIONAL\_Interactive\_Mapping\_with\_Folium.ipynb updated links/checked/done**
  + **Datasets used (**not listed at beginning)
    - OpenStreetMap and CartoDB Positron basemaps
    - California counties shapefile ('notebook\_data/california\_counties/CaliforniaCounties.shp')
    - California places cartographic [boundaries](https://www2.census.gov/geo/tiger/GENZ2019/shp/cb_2019_06_place_500k.zip" \t "_blank) for 2019 from the United States Census Bureau ('notebook\_data/census/Places/cb\_2019\_06\_place\_500k.zip')
    - California passenger rail [stations](http://opendata.mtc.ca.gov/datasets/passenger-rail-stations-2019" \t "_blank) subsetted to Bay Area Rapid Transit (BART) locations) from the Metropolitan Transportation Commission (MTC) Open Data portal ('notebook\_data/transportation/Passenger\_Rail\_Stations\_2019.zip')
* **13\_OPTIONAL\_geocoding.ipynb**
  + **Datasets used (**not listed at beginning)
    - **Alameda County schools: "./notebook\_data/alco\_schools.csv"**
    - **Contextily basemap ‘ax’**
    - **CartoDB Positron tileset**
* **14\_OPTIONAL\_Plotting\_and\_Mapping\_with\_Altair.ipynb**
  + **Datasets used (**not listed at beginning)
    - ACS 5-year data: notebook\_data/census/ACS5yr/ (especially Alameda County)
    - Census tract shapefiles: 'zip://./notebook\_data/census/Tracts/cb\_2018\_06\_tract\_500k.zip'
* **15\_OPTIONAL\_Voronoi\_Tessellation.ipynb**
  + **Datasets used (**not listed at beginning)
    - 2020 General Election voting locations for Alameda County: 'notebook\_data/ac\_voting\_locations.csv'
    - Census tracts for Alameda County: "zip://notebook\_data/census/Tracts/cb\_2013\_06\_tract\_500k.zip"
* **16\_OPTIONAL\_Introduction\_to\_Raster\_Data.ipynb**
  + **Datasets used (**not listed at beginning)
    - Missing dataset list at the top
* **99\_Questions\_Answers.md: no datasets!**

**Dataset versions:**

**Dataset versions**

**02**

**- California counties shapefile: 'notebook\_data/california\_counties/CaliforniaCounties.shp'**

**- California places 2018 shapefile from the United States Census Bureau: 'notebook\_data/census/Places/cb\_2018\_06\_place\_500k.zip'**

**03**

* **California counties shapefile: ‘notebook\_data/california\_counties/CaliforniaCounties.shp’**
* **United States states shapefile: ‘notebook\_data/us\_states/us\_states.shp’**
* **California places 2018 shapefile (from the United States Census Bureau): ‘notebook\_data/census/Places/cb\_2018\_06\_place\_500k.zip’**

**04**

**Berkeley bike boulevards: 'notebook\_data/transportation/BerkeleyBikeBlvds.geojson'**

**Alameda County schools: 'notebook\_data/alco\_schools.csv'**

**Housing parcels: 'notebook\_data/parcels/parcel\_pts\_rand30pct.geojson'**

**Berkeley city limits: ‘notebook\_data/berkeley/BerkeleyCityLimits.shp’**

**05**

**California counties shapefile: 'notebook\_data/california\_counties/CaliforniaCounties.shp'**

**Alameda County schools: 'notebook\_data/alco\_schools.csv'**

**Berkeley bike boulevards: 'notebook\_data/transportation/BerkeleyBikeBlvds.geojson'**

**06**

**Census tracts: 'notebook\_data/census/Tracts/cb\_2013\_06\_tract\_500k.zip'**

**California protected areas: 'notebook\_data/protected\_areas/CPAD\_2020a\_Units.shp'**

**Berkeley city limits: 'notebook\_data/berkeley/BerkeleyCityLimits.shp'**

**Alameda County schools: 'notebook\_data/alco\_schools.csv'**

**Berkeley bike boulevards: 'notebook\_data/transportation/BerkeleyBikeBlvds.geojson'**

**Bay Area Rapid Transit (BART) stations: 'notebook\_data/transportation/bart.csv'**

**07**

**American Community Survey (ACS) data for California: 'notebook\_data/census/ACS5yr/census\_variables\_CA.csv'**

**Census tracts: 'notebook\_data/census/Tracts/cb\_2013\_06\_tract\_500k.zip'**

**Alameda County schools: 'notebook\_data/alco\_schools.csv'**

**08**

**Census tracts in Alameda County joined to American Community Survey data: 'outdata/tracts\_acs\_gdf\_ac.json'**

**California grocery store data: 'notebook\_data/other/ca\_grocery\_stores\_2019\_wgs84.csv'**

**09**

**Alameda County polling locations: 'notebook\_data/ac\_voting\_locations.csv'**

**Census tract boundaries for California: 'notebook\_data/census/Tracts/cb\_2013\_06\_tract\_500k.zip'**

**Census place boundary shapefile for Berkeley: 'outdata/berkeley\_places.shp'**

**Alameda County bus routes for fall 2020: 'notebook\_data/transportation/Fall20Routeshape.zip'**

**10**

**Census tract shapefiles for California: https://www2.census.gov/geo/tiger/GENZ2018/shp/cb\_2018\_06\_tract\_500k.zip**

**San Francisco bikeway routes: https://data.sfgov.org/api/geospatial/ygmz-vaxd?method=export&format=GeoJSON**

**Berkeley bikeway routes: https://data.cityofberkeley.info/api/geospatial/fgw9-98ic?method=export&format=GeoJSON**

**Open Street Map San Francisco and Berkeley street networks obtained with the Python OSMNX package**

**11**

**OpenStreetmap Mapnik and CartoDB basemaps (via contextily)**

**Cartographic boundary file of California (06) places in 2019 from the United States Census: https://www2.census.gov/geo/tiger/GENZ2019/shp/cb\_2019\_06\_place\_500k.zip**

**Bay Area Rapid Transit (BART) station locations: https://raw.githubusercontent.com/dlab-berkeley/Geospatial-Fundamentals-in-Python/master/notebook\_data/transportation/bart.csv**

**12**

**OpenStreetMap and CartoDB Positron basemaps**

**California counties shapefile: https://raw.githubusercontent.com/dlab-berkeley/Geospatial-Fundamentals-in-Python/master/notebook\_data/california\_counties/CaliforniaCounties.shp**

**California census places shapefile for 2018: https://www2.census.gov/geo/tiger/GENZ2018/shp/cb\_2018\_06\_place\_500k.zip**

**California passenger rail stations (subsetted to Bay Area Rapid Transit (BART) locations) from the Metropolitan Transportation Commission (MTC) Open Data portal: http://opendata.mtc.ca.gov/datasets/passenger-rail-stations-2019**

**13**

**Alameda County schools: 'notebook\_data/alco\_schools.csv'**

**CartoDB Positron basemap**

**14**

**Alameda County polling locations: 'notebook\_data/ac\_voting\_locations.csv'**

**California census tracts: 'zip://notebook\_data/census/Tracts/cb\_2018\_06\_tract\_500k.zip'**

**15**

**Alameda County polling locations: 'notebook\_data/ac\_voting\_locations.csv'**

**California census tract cartographic boundaries in 2013: 'zip://notebook\_data/census/Tracts/cb\_2018\_06\_tract\_500k.zip'**

**16 (links checked)**

**National Land Cover Database (NLCD) land cover from 2019 downloaded from the Multi-Resolution Land Characteristics (MRLC) Consortium ('notebook\_data/raster/nlcd\_2019.tiff')**

**California census tract cartographic boundaries in 2019 from the United States Census Bureau ('notebook\_data/census/Tracts/cb\_2019\_06\_tract\_500k.zip').**