# Draft Data & Methods Section

PA 604: Methods of Analysis in Public and Urban Affairs (Fall 2022)

15 points

Due: December 2, 2022

## Overview:

In this assignment, you'll be completing a draft version of your data and methods section for your final proposal.

The data and methods section is where you're introducing the reader to how you plan to answer your research question. You will discuss the data you propose to use – either data from qualitative observations or from a quantitative data set.

- *For quantitative projects, many researchers use secondary data (often culled from surveys or from administrative/official data). Please take a look through the ICPSR repository of secondary data sets, as a starting point for identifying a useful existing data set for your project. If you find a data set that could work for your research question, feel free to use it for this project.*

  *Alternatively, for quantitative projects, you can propose to do a survey of your own.*

Next, you must discuss your method or analytic strategy – the way you will analyze the data. If you're using qualitative data (e.g. interviews, ethnographic field notes, participant observation), analysis usually entails "grounded theory," where you code each interview, note, or observation into themes. If you're using a quantitative data set, analysis usually entails conducting bivariate (e.g. correlation, chi-square, t-test, ANOVA) or multivariate (e.g. regression) tests.

This assignment is designed to be edited, improved, and included as part of your final proposal. Since this assignment will be submitted for feedback, suggestions, and comments from me, you'll need to complete the assignment Google Docs, and follow these video instructions to submit your Doc on Canvas[1].

---

[1]You must share the document by granting editing permissions to anyone with the link, and copy and paste the sharing URL into Canvas

# Requirements

## Substance (10 pts)

- The data and methods section should be no more than <u>1.5 pages double-spaced</u>.

- The discussion of your data and methods differs based on whether you plan to use <u>qualitative</u> or <u>quantitative</u> data. This section should include these <u>5 main components</u> (*brief examples, by type, below each item in* <span style="color:blue">*blue*</span>/<span style="color:red">*red*</span>):

  1. A discussion covering the **characteristics of the data**

     - For Quantitative Data:

       * Secondary Data
         "*<span style="color:blue">Given my interest in how tactics influence policy change, I use data from the General Social Survey administered in 2014. The data set has 2,538 observations, with individuals as the unit of analysis. The data set includes an array of information, including demographics, behavioral, and attitudinal measures.</span>*"

       * Primary Data (e.g. your own survey)
         "*<span style="color:blue">Given my interest in how tactics influence policy change, I developed a survey, administered in 2021, to individuals in advocacy organizations. The survey asked questions about their organization's use of protests, lobbying, and advocacy tactics as well as their involvement in the policy process. The data set has 184 observations, with individuals as the unit of analysis. The data set includes an array of information, including demographics, socioeconomic, behavioral, and attitudinal measures.</span>*"

     - For Qualitative Data:
       "*<span style="color:red">Given my interest in how tactics influence policy change, I propose to conduct ninety interviews with activists from thirty on-campus, education-based organizations. This will entail interviewing 3 activists from each organization (the organization's leader/president, a member of their executive committee, and a rank-and-file member).</span>*"

  2. A discussion of the recruitment/**sampling strategy**

     - For Quantitative Data:

       * Secondary Data
         "*<span style="color:blue">The data were collected by NORC at the University of Chicago. Individuals were recruited for the study using a modified probability sampling technique. Specifically, NORC conducted probability sampling at the block level for all blocks in the United States. Next, individuals within each sampled block were selected using quota sampling – percentages to match the U.S. population – based on sex, age, and employment status.</span>*"

* Primary Data (e.g. your own survey)

  "*For my survey, individuals were selected from the population of all advocacy organizations in the United States. I obtained the sampling frame from Cengage's Encyclopedia of Associations: National Organizations of the U.S. Database. Next, I took a random sample of 200 organizations. From there, I visited the websites for all sampled organizations and retrieved the email addresses for all high-ranking members (e.g. Director, Executive Committee, etc.) in each organization, and emailed a link to the survey to all available emails. Because some organizations received more emails than others (given the number of available email addresses on an organization's website), I use the first completed survey from each organization, regardless of status.*"

- For Qualitative Data:

  "*Given my interest in how tactics influence policy change, I propose to conduct ninety interviews with activists from thirty on-campus, education-based organizations. This will entail interviewing 3 activists from each organization (the organization's leader/president, a member of their executive committee, and a rank-and-file member).* "

3. Discussion of your **variables of interest**

   - For Quantitative Data:

     "*The data set includes numerous variables to examine the relationship between tactics and policy change. I will use measures related to participation in social movement action (e.g. CIVDIS, OPPSEGOV, etc.) and those related to an individuals beliefs that they can have an impact on government (e.g. EFFGOV, GOVLIST, etc.). In order to get a broad picture of the effect of social movement participation and policy change, I will include demographic information about the participant (e.g. age, sex, race/ethnicity), and I will account for their political orientation and organizational involvement.*"

   - For Qualitative Data:

     "*The interviews will cover several topics to address why the organizations did or did not get involved campaigns to change educational policy. For those organizations that did get involved, I will ask participants about the types of action they engaged in (protest, lobbying, sit-ins, etc.), how they participated (alone or in a coalition), how the tactic changed over time, the immediate impact of their action, whether or not they plan to continue action, and what types of outcomes they wish to initiate.*"

4. A discussion of your method of (coding and) analyzing the data, or your **analytic strategy**

   - For Quantitative Data:

     "*I will use regression analysis to understand the effect of protest participation on the belief in policy change. Preliminary bivariate analysis,*

*using correlation, shows a strong positive and significant correlation between measures of movement action/participation (e.g. OPPSEGOV) and the belief that one can impact government policy (e.g. EFFGOV), r = .68, p < .001 This demonstrates that as participation increases, so too does belief that you can impact policy change. The regression analysis will build off this relationship by incorporating control variables (e.g. demographics) to account for spuriousness and build a causal model.*"

– For Qualitative Data:

"*I will use a grounded theory approach to the study of social movement action and its effect on policy change. I take a grounded approach because it is data-centric and allows me to pull out themes in the data (interviews) as they are collected and analyzed. Not having a predetermined set of themes or categories is beneficial because I will not be wed to categories or codes that don't necessarily appear in the responses of individuals. Moreover, the grounded theory approach allows for the expansion or contraction of categories, based on the themes generated in the interviews. I will compare these themes across interviews, and aim to create theory from themes that cut across responses of activist. Doing so will help paint a broad picture of how activists engage in action to impact policy change.*"

5. A **realistic timeline** for the completion of your project

– For Quantitative Data:

* Secondary Data

| 2023 Timeline | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Identify Data | ■ | | | | | | | | | | | |
| Download Data | | ■ | | | | | | | | | | |
| Clean Data | | | ■ | ■ | ■ | | | | | | | |
| Run Analyses | | | | | ■ | ■ | ■ | | | | | |
| Write Results | | | | | | ■ | ■ | ■ | | | | |
| Revise Results | | | | | | | ■ | ■ | ■ | | | |
| Submit to PSA Conference | | | | | | | | | | ■ | | |
| Submit to WSC Conference | | | | | | | | | | ■ | | |

| 2024 Timeline | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Present at WSC | | ■ | | | | | | | | | | |
| Present at PSA | | | ■ | | | | | | | | | |
| Revise Manuscript to Incorporate Feedback | | ■ | ■ | | | | | | | | | |
| Prepare Manuscript for Publication | | | | ■ | ■ | | | | | | | |
| Submit Manuscript to *Mobilization* Journal | | | | | | ■ | | | | | | |

4

* Primary Data (e.g. your own survey)

| 2023 Timeline | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| Develop Survey | X | X | | | | | | | | | | |
| Pilot Test Survey | | X | X | X | | | | | | | | |
| Revise Survey | | | | X | X | | | | | | | |
| Identify Participants (Sampling Frame) | | | | X | X | X | | | | | | |
| Select Sample | | | | | | X | X | | | | | |
| Recruit Participants / Send Survey | | | | | | | X | | | | | |
| Collect Responses | | | | | | | X | X | X | X | | |
| Wrap Data Collection | | | | | | | | | | X | | |
| Download, Clean, & Recode Data | | | | | | | | | | X | X | |
| Run Analyses | | | | | | | | | | X | X | X |
| Write Results | | | | | | | | | | X | X | |
| Submit to NPC Conference | | | | | | | | | | X | | |

| 2024 Timeline | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| Present at NPC | | | X | X | | | | | | | | |
| Revise Manuscript to Incorporate Feedback | | | | X | | | | | | | | |
| Prepare Manuscript for Publication | | | | X | X | X | | | | | | |
| Submit Manuscript to *Journal of Planning Practice and Research* Journal | | | | | | X | X | | | | | |

– For Qualitative Data:

### 2023 Timeline

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Develop Survey | ▓ | ▓ | | | | | | | | | | |
| Pilot Test Survey | | ▓ | ▓ | ▓ | | | | | | | | |
| Revise Survey | | | | ▓ | ▓ | | | | | | | |
| Identify Participants (Sampling Frame) | | | | ▓ | ▓ | ▓ | | | | | | |
| Select Sample | | | | | | ▓ | ▓ | | | | | |
| Recruit Participants / Send Survey | | | | | | | ▓ | | | | | |
| Collect Responses | | | | | | | ▓ | ▓ | ▓ | ▓ | | |
| Wrap Data Collection | | | | | | | | | | ▓ | | |
| Download, Clean, & Recode Data | | | | | | | | | | ▓ | ▓ | |
| Run Analyses | | | | | | | | | | ▓ | ▓ | ▓ |
| Write Results | | | | | | | | | | ▓ | ▓ | ▓ |
| Submit to NPC Conference | | | | | | | | | ▓ | | | |

### 2024 Timeline

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Present at NPC | | | ▓ | ▓ | | | | | | | | |
| Revise Manuscript to Incorporate Feedback | | | | ▓ | | | | | | | | |
| Prepare Manuscript for Publication | | | | ▓ | ▓ | ▓ | | | | | | |
| Submit Manuscript to *Journal of Planning Practice and Research* Journal | | | | | | ▓ | ▓ | | | | | |

## Format & Clarity (5 pts)

- 12-point font

- 1-inch margins on all sides

- First and Last Name on the first line of the header

- "Data & Methods" as the title, centered as the second line of your header or the first line of the paper

- Ensure that you're citations follow either ASA or APA format

- Clear grammar, punctuation, and flow of your writing

## Grading

You will be graded on 1) the clarity and format of your writing (including the title, name, font, margins, grammar, flow, etc), as well as 2) the presence and clarity of the 5 main components listed above.

# Example 1

The California Department of Corrections and Rehabilitation (CDCR) collects data on offenders that are released for a three-year follow-up period. I will submit a research application to obtain the historical reports from the Office of Research Data Concierge within the CDCR as well as a research proposal per their requirements. The 2019 recidivism report is based on offenders who were released in the 2014-15 fiscal year. They measure recidivism by arrests and returns to prison, however, the data analyzed will be limited to reincarceration. The data set includes offender demographics such as race and age, as well as characteristics including length of stay, and commitment offense.

Once I obtain the data sets, I will code the data and use regression analysis to understand the effect of correctional education on recidivism rates. I will use variables to identify the program participants, those who completed a program, and those who did not participate at all to measure whether they recidivated in the three-year post-release period. Upon completion of analyzing the data, I will write up the results and submit them to the CDCR per their requirements.

# Example 2

Given the lack of available data on museum mission statements, I propose to randomly sample 100 federally-funded anthropological museums in the United States that have Native American or Native Hawaiian artifacts in their collections. I will then clean the data collected to focus on the substantive words. The clean data will be run through a program that will look for mention of Native American, Native Hawaiian, and names of individual tribes, and coded as 0 (no reference) or 1 (reference). Those that received an initial score of 1 will be subject to further sentiment analysis to look for neutral and positive connotations. I will use the results of coding and sentiment analysis, and compare each mission statement to data on the funding, proximity to federally recognized tribal reservations, and repatriation disputes for each museum within the sample. I will then take the data and see if any patterns emerge between the type of reference (no reference, neutral sentiment, or positive sentiment) in the mission statement and whether the museum has engaged in repatriation disputes while accounting for differences in funding resources and proximity to tribal reservations.

# Example 3

First, I use California Department of Parks and Recreation administrative data to identify parks throughout the state that meet criteria needed to make observations regarding the impact of public restrooms on park visitation. Given the variation in parks, it is necessary to control for any factors outside of bathrooms that may influence park visitation. Therefore, I select 5 parks that are equal in acres (at least 2.8 acres), amenities (i.e. playgrounds, sidewalk, shade, gyms, active or recreational programs) as well as access (i.e. public transportation access, walkable access, parking access) and on-site bathrooms to compare them with 5 parks that are equal in acres, amenities as well as access without on-site bathrooms. I also ensure that the parks that I compare share similar to equal local population density and median income characteristics based on 2010 United States Census Data. I count bathrooms by the number of toilet stalls in a park. I use building codes, blueprints and personal observation to verify the number of toilet stalls in the parks studied. Then, I and research project staff will conduct ethnography by counting the number of people present at each selected park and tracking the length of their stay from 9am to 7pm based on each reviewer's line of sight. Data is collected on a structured research tool that is piloted twice for optimization before research use. Data is collected on the same weekday and weekend day in every selected park during each season during moderate weather from January 2020 to December 2020. Research project staff are stationed at each entrance of the park to ensure adequate vision. All research project staff data is gathered and collected after each observation study to confirm all data or remove duplicate data.

# Example 4

This investigation aims to explore the connection between parking requirements and transit ridership. Given my interest in the connection between parking availability and transit ridership, I propose to conduct a hundred interviews at ten station areas across the region. At each site, ten individuals will be randomly selected for interviews. The interviews will cover several topics to assess how proximity transit and parking availability factor into their commute patterns. I will ask participants about their general commute (origin and destination), how they arrived at the transit station (by public transit, in a personal vehicle, etc.), their reasons for choosing transit, their proximity to transit (workplace and residence), and how changes in parking availability at transit stations or proximity to transit would impact their travel patterns (reduction or addition).

Site selection is an important component of the interview process. The chosen sites include: (1) Beyer Boulevard, (2) H Street, (3) Old Town Transit Center, (4) Morena/Linda Vista, (5) Grossmont Transit Center, (6) El Cajon Transit Center, (7) Encanto/62nd Street, (8) Grantville, (9) Oceanside Transit Center, and (10) Escondido Transit Center. The interviews are intended to expand on existing research (Cervero 2007) by including stations with park-and-ride facilities. Given San Diego's sprawling spatial pattern, these facilities are seen as critical to the transit operations. In addition, stations were selected to reflect the diversity of communities in the region that depend on public transit and the various lines and systems serving those communities. Lastly, special consideration was given to recently redeveloped park-and-ride stations and stations where redevelopment is either planned or proposed.

I will use a grounded theory approach to the study of parking availability and its effect on transit ridership. I will take this approach because it is data-centric and allows me to pull out themes in the data as they are collected and analyzed. By not having a predetermined set of themes or categories, I will not be limited to categories or codes that do not accurately depict participants' responses. In addition, the grounded theory approach allows for flexibility in the number of categories, based on the themes generated in the interviews. Themes that emerge from the interviews will help illustrate how the availability of parking impacts transit ridership.