



EQUITABLE DATA COLLECTION

A TOOLKIT FOR DESIGNING & DISTRIBUTING SURVEYS

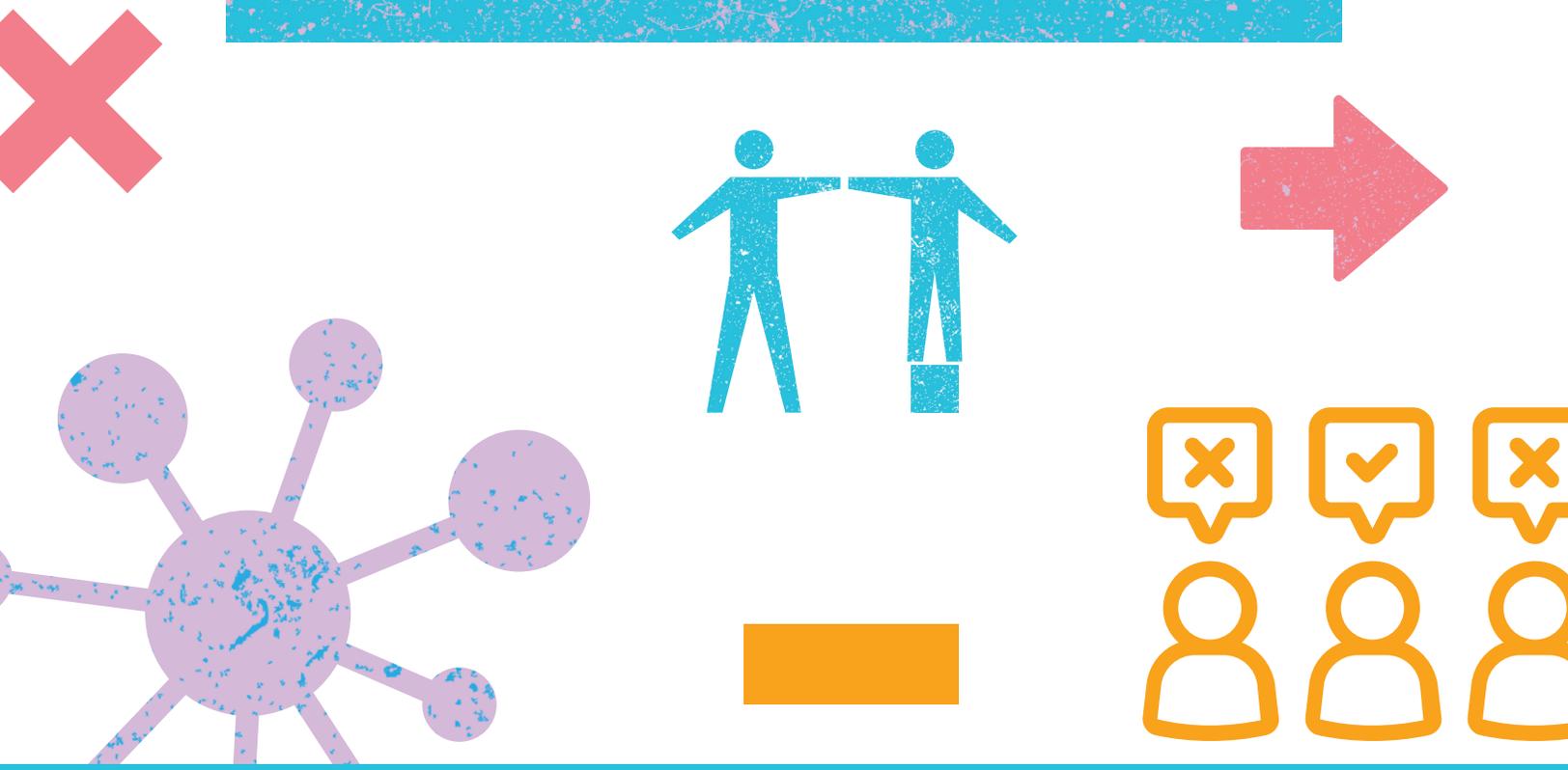


Table of Contents

<i>EQUITABLE DATA COLLECTION</i>	2
<i>BEFORE DESIGNING YOUR SURVEY</i>	3
<i>DESIGNING YOUR SURVEY</i>	3
<i>DISTRIBUTING YOUR SURVEY</i>	11
<i>SHARING YOUR SURVEY RESULTS</i>	11
<i>APPENDIX</i>	14

Equitable Data Collection: A Toolkit for Designing & Distributing Surveys

In 2020, the Long Beach City Council passed a resolution declaring racism a public health crisis. As a result, the City Council has adopted the [Racial Equity and Reconciliation report](#). The document lays out several actions necessary to address anti-Black racism and advance racial equity and in Long Beach. Significantly, the report envisions a future where race does not determine social and economic outcomes. These actions and this vision depend on accurate and specific data collection and data practices.

The City uses data to inform decisions about policy, programming, and other outcomes. City staff should be held accountable to using public information in an ethical manner that serves the public good. Staff may not use data collected from the public to reinforce existing biases or to make decisions that may exclude, harm, or criminalize low-income communities and communities of color. City staff must also work to include the diverse identities and experiences of Long Beach community members when designing city programs and policies.

This toolkit is meant to guide City employees as they design and distribute surveys, as well as to ensure that City staff equitably share data insights from surveys. This toolkit works in tandem with the [City's Communications Plan](#), [Equity Toolkit](#), [Data Privacy Guidelines](#), and [Digital Engagement Toolkit](#).

5 Principles for Equitable Data Collection

1. **Anonymized** – collecting non-identifiable data will help us in our work advancing racial equity
2. **Strategic** – only collect data where it is relevant and crucial
3. **Flexible** – understand that identities constantly shift
4. **Representative** – work to ensure that Long Beach is represented proportionally
5. **Trustworthy** – build and maintain relationships with community members via strong community engagement and transparency

Before Designing your Survey

Why Are We Collecting the Data?

Before designing a survey, we encourage you to challenge your assumptions and question what details the City needs to collect from residents. Recognize that asking for sensitive or private details about the community is a significant ask – make sure to justify City needs. The more questions staff ask and the more personal they are, the less likely residents are to complete a survey.

- How will this data advance policy goals?
- How might other departments utilize the data you collect?
- How will you protect any personally identifiable information in your data?
- Does potential exist for the data to be misused (unintended consequences)?
- How do you plan to share the results of your data with the community?
- Does this data already exist via administrative data? For example, could we get this same data from our existing systems like the Parks registration system?

Guidance on Whether to Collect Data

Race and gender impact aspects of our lives in ways we may not even consider. It is important to collect this information to understand if the data collected is representative of the population being surveyed, and also to determine whether our programs have unintended consequences for any stakeholders.

Designing your Survey

People-First Language

We encourage you to think of survey respondents from a people-first perspective. We often survey residents and stakeholders with the hope that data will improve our policies, programs, and services but also realize that as City employees it is imperative to build public trust. If a respondent feels uncomfortable or anxious while filling out a survey, they are unlikely to finish it. And the cumulative effect is that you will only have survey responses from people who felt comfortable taking the survey, which may disproportionately benefit one group over another.

Here are several approaches to designing a survey that prioritizes a comfortable experience.

1. Offer multi-select checkboxes, not single-select radio buttons

- This allows people to select multiple identities rather than selecting just one. This may mean that you need to do some data analysis work when your results are finalized so that you can gain insights from the group you are interested in studying.

2. Allow users to self-describe when relevant

- This allows survey responders to use their unique identity in the survey. Allowing responders to self-describe leads to messier data so you should only do this when relevant.

3. Write in personable, conversational language

- Using plain language makes it easier for people with low literacy skills or low knowledge of local government to take your survey. **This Hemingway app tests whether writing is clear enough.**
- Inclusive Strategies from the Communications **Equity Toolkit** (Pg.10 & 11)
 - Consider how different perspectives and contexts affect meaning and messages
 - Connect through storytelling, photos, art, and music
 - Use the City's Language Access services for translation or interpretation
- Employ inclusive language and terminology that resonates with communities of color. Convolutioned language can often discourage a community member from participating in a survey.

4. Do not require a response

- Do not make your demographics collection a smorgasbord of options, which can dehumanize and can offend respondents. Terminology in underrepresented communities evolves and it is our obligation as public servants to be flexible in matching this terminology.

5. For small sample sizes, do not ask identifiable questions

- The sample questions below should not be asked for surveys with a sample size that is so small that it would lead to identifiable results. For example, if you are surveying a group of 10 people, it would not make sense to ask a race question since you would know by the answer who the respondent was.

Highly Recommended Demographic Questions

After designing your survey and determining which demographic questions are relevant and necessary to collect, use these recommendations for how to structure demographic questions.

What is the primary language you speak at home? (select all that apply)

- English
- Spanish
- Tagalog
- Khmer
- Other _____

What is your zip code?

- | | | |
|--------------------------------|--------------------------------|--------------------------------|
| <input type="checkbox"/> 90802 | <input type="checkbox"/> 90807 | <input type="checkbox"/> 90814 |
| <input type="checkbox"/> 90803 | <input type="checkbox"/> 90808 | <input type="checkbox"/> 90815 |
| <input type="checkbox"/> 90804 | <input type="checkbox"/> 90809 | <input type="checkbox"/> 90822 |
| <input type="checkbox"/> 90805 | <input type="checkbox"/> 90810 | <input type="checkbox"/> 90831 |
| <input type="checkbox"/> 90806 | <input type="checkbox"/> 90813 | <input type="checkbox"/> 90802 |

Do you identify as someone who has a permanent or long-term disability?

- Yes
- No
- Prefer not to answer

Is this disability (select all that apply)

- | | |
|------------------------------------|---|
| <input type="checkbox"/> Visual | <input type="checkbox"/> Mental Health |
| <input type="checkbox"/> Hearing | <input type="checkbox"/> Prefer not to answer |
| <input type="checkbox"/> Cognitive | <input type="checkbox"/> Other _____ |
| <input type="checkbox"/> Physical | |

Race and Ethnicity Questions

I identify as (select all that apply):

Why are we asking this? We collect information to help us identify potential inequities or disparities in access to opportunities or outcomes for Long Beach residents.

- | | |
|---|---|
| <input type="checkbox"/> Asian/Pacific Islander | <input type="checkbox"/> Native American/Indigenous |
| <input type="checkbox"/> Black/African American | <input type="checkbox"/> Prefer not to answer |
| <input type="checkbox"/> White | <input type="checkbox"/> Identify as _____ |
| <input type="checkbox"/> Hispanic/Latinx | |
| <input type="checkbox"/> Middle Eastern/North African | |

Conditional formatting: If Asian/Pacific Islander, are you (select all that apply):

- | | |
|--------------------------------------|--|
| <input type="checkbox"/> Cambodian | <input type="checkbox"/> Taiwanese |
| <input type="checkbox"/> Hmong | <input type="checkbox"/> Chinese |
| <input type="checkbox"/> Laotian | <input type="checkbox"/> Korean |
| <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Japanese |
| <input type="checkbox"/> Indonesian | <input type="checkbox"/> Thai |
| <input type="checkbox"/> Singaporean | <input type="checkbox"/> Filipino/a |
| <input type="checkbox"/> Bangladeshi | <input type="checkbox"/> Sri Lankan |
| <input type="checkbox"/> Indian | <input type="checkbox"/> Samoan/Pacific Islander |
| <input type="checkbox"/> Pakistani | <input type="checkbox"/> Prefer not to answer |
| <input type="checkbox"/> Malaysian | <input type="checkbox"/> Other _____ |

Why are we asking this? Long Beach is proud of its diverse Asian community and understands that different people in this group experience distinctive outcomes.

Conditional formatting: If Hispanic/Latinx, are you (select all that apply):

- | | | |
|--------------------------------------|--------------------------------------|---|
| <input type="checkbox"/> Argentinian | <input type="checkbox"/> Dominican | <input type="checkbox"/> Panamanian |
| <input type="checkbox"/> Bolivian | <input type="checkbox"/> Ecuadorian | <input type="checkbox"/> Paraguayan |
| <input type="checkbox"/> Brazilian | <input type="checkbox"/> Salvadorian | <input type="checkbox"/> Peruvian |
| <input type="checkbox"/> Chilean | <input type="checkbox"/> Guatemalan | <input type="checkbox"/> Puerto Rican |
| <input type="checkbox"/> Colombian | <input type="checkbox"/> Honduran | <input type="checkbox"/> Uruguayan |
| <input type="checkbox"/> Costa Rican | <input type="checkbox"/> Mexican | <input type="checkbox"/> Venezuelan |
| <input type="checkbox"/> Cuban | <input type="checkbox"/> Nicaraguan | <input type="checkbox"/> Prefer not to answer |
| | | <input type="checkbox"/> Other _____ |

Why are we asking this? Long Beach is proud of its diverse Latinx community and understands that different people in this group experience distinctive outcomes.

This specific format avoids differentiating between race and ethnicity. The format of this question also avoids forcing Latinx respondents to identify their race as White. While this makes it easier for someone filling out this survey, it adds a layer of complication when conducting data analysis.

While this may change depending on what group you are interested in studying, we recommend categorizing anyone who answers Black in combination with another race as Black. No classification system is perfect but expanding this Black category in this way allows Afro-Latinos and other mixed-race people to be counted as Black, since society typically perceives them as Black.¹

When analyzing Latinos, consider that recent waves of immigration have seen increases in indigenous populations, especially in Long Beach where there is an established Chinanteco (Mexican indigenous) community. For more information, please click [here](#).

An important caveat: If you're comparing survey response data to Census data or American Community Survey data or any other large data set, you will need to make some assumptions to jointly analyze both data sets. A future version of this toolkit will explore sample assumptions to be made from this data.

For example, if your survey collected information on respondents who identify as Latinx and you wanted to compare it to data in the American Community Survey, you would need to reference the Hispanic or Latino origin category. People who are Hispanic are not necessarily Latinx; therefore, this is not exactly an apples to apples comparison. However, this assumption must be made to move forward with this analysis.

Gender and Sexual Identity

Gender and Sexual Identity²

Why are we asking this? We collect information to help us identify potential inequities or disparities in access to opportunities or outcomes for Long Beach residents.

Gender at Birth

- | | |
|------------------------------------|---|
| <input type="checkbox"/> Woman | <input type="checkbox"/> Prefer not to answer |
| <input type="checkbox"/> Man | <input type="checkbox"/> Other ____ |
| <input type="checkbox"/> Nonbinary | |

Current gender identity or expression

- | | |
|------------------------------------|---|
| <input type="checkbox"/> Woman | <input type="checkbox"/> Prefer not to answer |
| <input type="checkbox"/> Man | <input type="checkbox"/> Other ____ |
| <input type="checkbox"/> Nonbinary | |

I am:

- | | |
|-----------------------------------|---|
| <input type="checkbox"/> Lesbian | <input type="checkbox"/> Straight |
| <input type="checkbox"/> Gay | <input type="checkbox"/> Prefer not to answer |
| <input type="checkbox"/> Bisexual | <input type="checkbox"/> Other ____ |
| <input type="checkbox"/> Queer | |

Keep in mind - We can ask these survey questions when pertinent. Where people may experience distinct outcomes based on their sexual or gender identity.

The collection of data on sexual orientation and gender identity is very important given that data on LGBTQ communities (and intersectionality with other demographic indicators) is limited in scope.

¹Sanchez-Lopez, Pastor, Sanchez, et. al. "The State of Black Immigrants in California." Black Alliance for Just Immigration.

²GLAAD Media Reference Guide: 10th Edition. (2016, October). Retrieved from <https://www.glaad.org/sites/default/files/GLAAD-Media-Reference-Guide-Tenth-Edition.pdf>

Footnote - AB 677, Chiu. Data collection: sexual orientation. (2017, October). Retrieved from **Bill Text - AB-677 Data collection: sexual orientation. (ca.gov)**

Biases: Consider the Presentation and Influence of Your Survey

Various forms of cognitive bias can influence how we design survey and interview questions, as well as how we interpret responses. Below are some common pitfalls to consider and avoid:

Induction bias or leading questions occur when we, indirectly or deliberately, include the answer in the formulation of the question. For example, when asking “Your experience with the city’s Go Long Beach app was satisfactory, wasn’t it?”, this induces respondents to answer yes—even if their experience was frustrating.

In addition, to avoid induction bias when **analyzing** quantitative data, be sure you understand what is being measured and how and consider variation over time.

Confirmation bias occurs when we look exclusively for information that proves a personal argument or hypothesis. This tendency to support existing beliefs can lead to serious errors in data analysis and, subsequently, policymaking. For example, if City staff wrongly assume that older residents do not rely on internet access, this belief may bias a study meant to inform Long Beach’s digital inclusion efforts.

Selection bias occurs when a sample is skewed during the recruitment of study participants. When recruiting study participants—for surveys, interviews and focus group discussions—be sure to choose residents from across the City and from varied circumstances. Without diverse representation, the City cannot infer ideas about an entire population and may rely on false stereotypes. For example, if City staff assume that Long Beach Transit users are low-income—despite that bus riders are socio-economically diverse—and only include them in a study, the study findings could skew future city planning efforts.

Analytics bias occurs when data sets are incomplete or when data sets lack the context of other factors that impact your results. Additionally, understanding the data missing from the data set may tell as important a story as the data we do have. For example, suppose the City is attempting to answer the questions, “Did violent crime increase last year?” and the data clearly show that violent crime did increase. In order to draw conclusions, we would need to understand the other factors at play that may have caused an increase in violent crimes. The City would need to consider contextual factors such as unemployment rates and the COVID-19 pandemic in order to determine causation.

Question Types to Consider for Clear Data Analysis

Various forms of cognitive bias can influence how we design survey and interview questions, as well as how we interpret responses. Below are some common pitfalls to consider and avoid:

Likert scale questions comprise a psychometric response scale in which respondents specify their level of agreement to a statement typically in five points: (1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree.

Keep in mind - a Likert scale measures the strength/intensity of an attitude linearly. While Likert scales focus on agree/disagree, there are other quantitative scales that rank things like worst to best, rate satisfaction etc. The more categories you have in a Likert scale, the less likely people choose neutral.

Ranking questions ask survey respondents to compare a list of items with each other and arrange them in order of preference.

Open-ended questions add context to quantitative questions. If a survey asks the respondent to rate a City service and respondents rate it poorly, open-ended responses may explain the specific complaint. These types of questions are harder to analyze, as they require you to read through the responses and identify themes in the data to inform recommendations (also known as coding).

Storytelling - Stories are an important type of qualitative data. Often, they offer an easier way for people representing diverse cultures to share experiences with the City.

Qualitative options questions involve using personas based on data and asking survey respondents to select the persona that best fits them. For example, if you are interested in the needs and uses data related to people experiencing homelessness, you might interview stakeholders who have reached out to the City. From there, you might develop some personas such as Person Experiencing Homelessness, Service Provider, Concerned Resident, and Researcher. Ultimately, when you design the survey, you can ask respondents which persona they identify with and add relevant conditional questions.

Schema refers to who is responding to your survey. Sometimes for a survey, you are interested in a household rather than an individual. You may disseminate a survey to Long Beach parents soliciting data about their children. In these cases, it is important to differentiate the questions so that you obtain responses about the population you are truly interested in studying. For example, you may ask how the respondent identifies in terms of their gender, as well ask how their children identify.

Distributing your Survey

Translation

It is important to follow the City's Language Access Policy and ensure your survey is accessible in Spanish, Tagalog and Khmer. (The City will translate the sample survey.) Remember to write in plain language—using simple, community-friendly language in English means a better experience for respondents taking the survey in other languages.

For more information please visit: Tips for [a Successfully Translated Document](#).

Paper or Telephone Surveys

This toolkit refers mostly to digital surveys. However, it is important to consider stakeholders who lack access to the digital skills, technology, or Internet required to fill out your survey. In this case, distributing paper surveys and providing a space for community members to turn in these questionnaires—such as a community center or library—is a good way to reach people who do not use the internet. Administering a survey via telephone is also a way to reach people. Partnering with community-based organizations as trusted messengers to help administer the surveys via telephone is an equitable decision.

Paying for Survey Responses

Paying respondents to complete your survey is the most equitable decision as it compensates people for their time. People who may not otherwise be interested in completing a City survey may be incentivized with a \$10 gift card. You will collect richer data by reaching people who would have declined to take your survey without compensation.

Sharing your Survey Results

It is natural for data analyses to reflect your own perspective. Therefore, it is critical to be intentional in analyzing your survey results. Below are a variety of descriptive stats that you may apply to your data analysis:

Statistical Significance is the claim that a result from data generated by testing or experimentation is not likely to occur randomly or by chance. Rather, the result is likely attributable to a specific cause. Statistical significance is important for academic disciplines or practitioners who rely heavily on analyzing data and research.

Mean is the mathematical average of two or more numbers. The mean helps to estimate the central tendency or the “center” of a distribution of responses and may have many other uses.

Median is the middle number in a sorted, ascending or descending, list of numbers. It may be more descriptive of a data set than the average.

Mode is the value that appears most frequently in a data set.

Percentage is a rate, number, an amount of something, typically expressed as a number out of 100.

Frequency is the rate at which something occurs or is repeated over a period of time or in a given sample.

Range is the difference between the largest and smallest values.

Data visualizations such as graphs, charts and maps will help your audience understand the main takeaways from your research and may provide context.

Internal Data-Sharing

You can and should find ways to share in-depth raw data with fellow other City departments that may be able to use the data to improve their own programs, policies, and budgets. Often, we neglect to share data and needlessly reach out to the same residents with the same questions—resulting in community survey fatigue and wasted City resources.

External Data-Sharing

We must publicly show residents and stakeholders the results of the work to which they have contributed. Publishing a simple infographic lets people know what the City learned through their engagement, while protecting the anonymity of survey respondents.

Survey Fatigue

Survey fatigue happens when your audience becomes bored or uninterested in your surveys and it normally happens in two ways:

- Before taking the survey — overwhelmed by the sheer volume of requests for feedback, customers decide not to even begin your survey. The result is a drop in response rates as fewer customers decide to give you feedback.
- During the survey — this happens after someone has started the survey and is usually caused by poor survey design such as including too many questions, a high proportion of open text fields or asking the same question repeatedly. As a result, respondents can drop out midway through or lose interest and speed through, giving you inaccurate data.

Coordinating with other Departments prior to releasing a community to gather cross-departmental feedback, data, and responses can help mediate the fatigue communities often experience.

APPENDIX

Electronic Survey Template

<https://bit.ly/EquitableDataCollection>





City of Long Beach
411 W. Ocean Blvd.
Long Beach, CA 90802

Visit us at www.longbeach.gov

   @LongBeachCity

To request this information in an alternative format or to request a reasonable accommodation, please contact Technology & Innovation Department at Ryan.Kurtzman@longbeach.gov or (562) 570-6911. A minimum of three (3) business days is requested to ensure availability. Reasonable attempts will be made to accommodate request made within less than three (3) business days.