



# Equity Toolkit

## User Guide

Version 1.1 | Published Dec 21, 2022

How to use data to  
advance equitable  
service at Metro

# Table of Contents

## Introduction and Background

[How does Metro define equity?](#)  
[What is the Equity Toolkit?](#)  
[What is in the Equity Toolkit?](#)  
[When should you consider using it?](#)  
[How do you get started?](#)  
[Who manages it? Who can you ask for help?](#)  
[How can you make equity a part of what we do?](#)  
[What is the Equity Toolkit trying to address?](#)  
[What is the Equity Working Group?](#)  
[How is Title VI applied to Bus/Rail service guidelines?](#)  
[How are “Title VI” and the toolkit different?](#)

## Equity Standards and Definitions

[Why are standards / definitions important?](#)  
[What are important definitions you need to know?](#)  
[When should you use rider vs. resident data?](#)  
[What should you know about using ACS data?](#)  
[A background on Metro survey data](#)  
[What should you know about Metro survey sample size?](#)  
[What should you know about Metro survey methodology?](#)  
[How can you use Metro survey data?](#)  
[What does current survey data tell us?](#)  
[Where can you get more detailed information?](#)

## The Equity Toolkit Data Hub

[What do you need to access the data hub?](#)  
[How do you get permissions to access the data hub?](#)  
[What operational data is available?](#)  
[What demographic data is available?](#)  
[How do you pull the data you need?](#)

## How to Conduct Analysis using the ETK

[Determine the objective](#)  
[Review the operating data available](#)  
[Factor in data limitations](#)  
[Determine the subject of the analysis](#)  
[Determine the level of detail needed from the data](#)  
[Determine Survey vs Census for demographics](#)  
[Determine your "unit of analysis"](#)  
[Pull the data](#)  
[Create a base case](#)  
[Compare the test group to the base](#)  
[Determine significance](#)  
[Document methodology and results](#)  
[QA/QC Overview](#)  
[QA/QC Peer Review](#)  
[QA/QC SME Review](#)

# Amending the User Guide

## Please help us make this guide better!

Questions? Concerns? Edits? Please email the Equity Toolkit Team: [equitytoolkit@wmata.com](mailto:equitytoolkit@wmata.com)

We will republish updated versions on a regular basis (to be determined). We look forward to hearing your thoughts and edits—big or small—so that we can continuously improve this User Guide.

Thanks!  
The Equity Toolkit Team





# Introduction

## What is Equity and the Equity Toolkit?

### **In this section:**

How does Metro define equity?

What is the Equity Toolkit?

What is in the Equity Toolkit?

When should you consider using it?

What should I do to get started?

Who manages it? Who can you ask for help?

How can you make equity a part of what we do?



# How does Metro define equity?

## Equitable Transit System

- Equity is an **outcome where anyone, regardless of identity, can use the transit system to access the region's opportunities and resources**; acknowledging differences in lived experience
- Some populations, particularly those of color, low-income and/or with disabilities, disproportionately experience injustice across several facets of life due to longstanding, structural challenges
- Metro recognizes that our investments and operational decisions change lives and therefore commit to intentionally prioritizing equity when designing, implementing, and evaluating our strategies, policies, practices, and investments

## Equitable Employer

- Equity at Metro is an **outcome where anyone, regardless of identity, can achieve success** in the workplace; acknowledging differences in lived experience
- Metro commits to providing and maintaining a work environment that recognizes, understands, respects, and encourages the unique contributions of each member of its workforce
- Policies, practices, and procedures are written and implemented to ensure fair treatment of all employees
- Decisions on promotion, training, professional development, and other opportunities as a result of employment are made based on merit and are free from discriminatory bias

# What does “equity” cover at Metro?

## Equitable Transit System

- Define / Measure\*
- Public Engagement\*
- Service\*
- Fares\*
- Infrastructure Investments
- Policing\*

## Equitable Employer

- Workforce Diversity
- Workforce Inclusion
- Contracting Diversity and Inclusion\*

\*Also identified in Board’s Transit Equity Framework

# What is the Equity Toolkit?

The Equity Toolkit is a virtual library of datasets, definitions, and guidance that help Metro factor transit equity into projects, reporting, and decision making, all with the goal of making our service more equitable.



The Equity Toolkit is focused on WMATA's efforts to make our **transit system** more equitable for our riders.

# Can you break down that definition?

The Equity Toolkit is a **virtual library**

(e-resources housed in a central SharePoint and data hub, accessible to all Metro employees)

of **datasets,**

(our own Metro data along with American Community Survey census data)

**definitions,**

(terms and concepts related to equity defined so you know what you're talking about)

and **guidance,**

(how-to's so you know what you're doing)

that helps us factor **equity**

(analyzing our own performance data, but also factoring in layers of census and rider data to better understand impacts across demographic groups)

into **projects,**

(integrating demographic data into projects to show how service delivery and quality varies across groups)

**reporting,**

(looking at preexisting and new performance measures through an equity lens)

and **decision-making,**

(advising leadership on decisions that improve equity in our service)

all with the goal of  
**making our service more equitable**

(a part of our mission at Metro)



# What is in the Equity Toolkit?



**In the Equity Toolkit:** A comprehensive, centralized set of relevant data tables that enables you to incorporate rider and regional demographics into your work

## In the Equity Toolkit SharePoint:

### DEFINITIONS

Explanations of agreed-upon **standards and definitions** to help you accurately communicate your results

### DATA

Details on how to access a centralized set of **data tables**

### GUIDE

This **User Guide** and other useful documentation

# When should you consider using it?

## Pursuing more equitable service | Two approaches:

1

Asking: Is our current service equitable?

2

Asking: What is the equity impact of a potential service or other operational change?



# When should you consider using it?

**Use this toolkit to add the "... and how does that differ for [specific communities]?" perspective to your work.**

This includes work that looks at our performance outcomes, our strategies, our policies, our processes, and our resource allocations. For example: The Performance Office regularly assess on-time-performance (OTP). The Performance Office should use the toolkit to assess OTP and ask how it differs for low-income Metro riders.

**Asking and answering this question will help ensure we are making decisions that factor in equity.**

You can also use this toolkit to get ideas for incorporating an equity lens into your work and learn how to define and get data on specific elements of our community. For example: Who "counts" as low-income and how do I find them?

# How do you get started?

- 1) **Join the Equity Toolkit Teams Group** – you can request access by emailing [EquityToolkit@wmata.com](mailto:EquityToolkit@wmata.com). In the Teams group, you can get help with your equity analyses, stay in the loop of current equity discussions at WMATA, and connect with other colleagues working to advance equity
- 2) **Visit the Equity Landing Page** – explore what Metro is doing to advance equity within our workforce and in the communities we serve
- 3) **Establish Access to the Data** – see the Data Hub section for how to do so.
- 4) **Know the Experts** – the list of Metro departments and employees below can serve as resources to you throughout your research – you should reach out to them early in your process with any questions.

Topics	Reach out to this Department:	Reach out to this Individual:
Survey data	RESR, PLAN	Kayleigh Campbell, Ray Yau, Matthew Zych
Census data	ETK QA/QC Teams Channel	Scott Traum, Diane Patterson
Ridership or operational data (i.e., homebrew, Trace, etc.)	PLAN API	Catherine Vanderwaart
Presentation to the Board, Executive Team, or public	EWG	Allison Davis (Board + Public) Jordan Holt (Exec Team)



# Who manages it? Who can you ask for help?

The below chart is a breakdown of responsibilities for all groups involved in the ETK:

Group	Responsibilities	How to contact
ETK Committee	<ul style="list-style-type: none"><li>• Toolkit processes</li><li>• Defining standards</li><li>• Maintaining repository of past work</li><li>• Granting access to data tables</li></ul>	<a href="mailto:EquityToolkit@wmata.com">EquityToolkit@wmata.com</a>
ETK Community	<ul style="list-style-type: none"><li>• General questions/discussions on equity</li><li>• Peer Reviews</li><li>• Coding questions</li></ul>	Equity Toolkit Teams Channel
Subject Matter Experts	<ul style="list-style-type: none"><li>• Questions specific to topic or area of expertise</li><li>• Methodology Review</li><li>• SME Reviews</li></ul>	See Slide 11
Equity Working Group	<ul style="list-style-type: none"><li>• Defining Equity</li><li>• Lead agency wide policy initiatives pertaining to equity</li><li>• Monthly meeting to discuss equity topics</li></ul>	Melissa Kim - <a href="mailto:MLKim@wmata.com">MLKim@wmata.com</a>

# How can you make equity a part of what we do?

Metro's annual performance management cycle provides an opportunity to ensure that your work on equity remains a priority and gets the visibility it deserves from your manager.

**Here are some ideas for Equity Toolkit-related actions to include in your Individual Performance Plans. They could be a good fit under “Personal Leadership”:**

- Use the Equity Toolkit to incorporate an equity lens into at least one existing report or analysis, and present recommendations to my leadership for actions Metro can take to improve. Make suggestions for how to improve the Equity Toolkit based on my use of it.
- Review at least one team process (SOP, OAP, Work Instructions, Checklist) to incorporate use of the Equity Toolkit.
- QA/QC at least two Equity Toolkit analyses.
- Serve as an Equity Toolkit subject matter expert and make continued updates to the Equity Toolkit (at least once per year).
- Present the Equity Toolkit to at least two groups of Metro staff and support their application of the Toolkit to their work.
- Proactively ask “what is the equity impact of this decision” when in meetings.



# Background

How does the Equity Toolkit align with Metro's mission?

## **In this section:**

What is the Equity Toolkit trying to address?

What is the Equity Working Group?

How is Title VI applied to Bus/Rail service guidelines?

How are "Title VI" and the toolkit different?

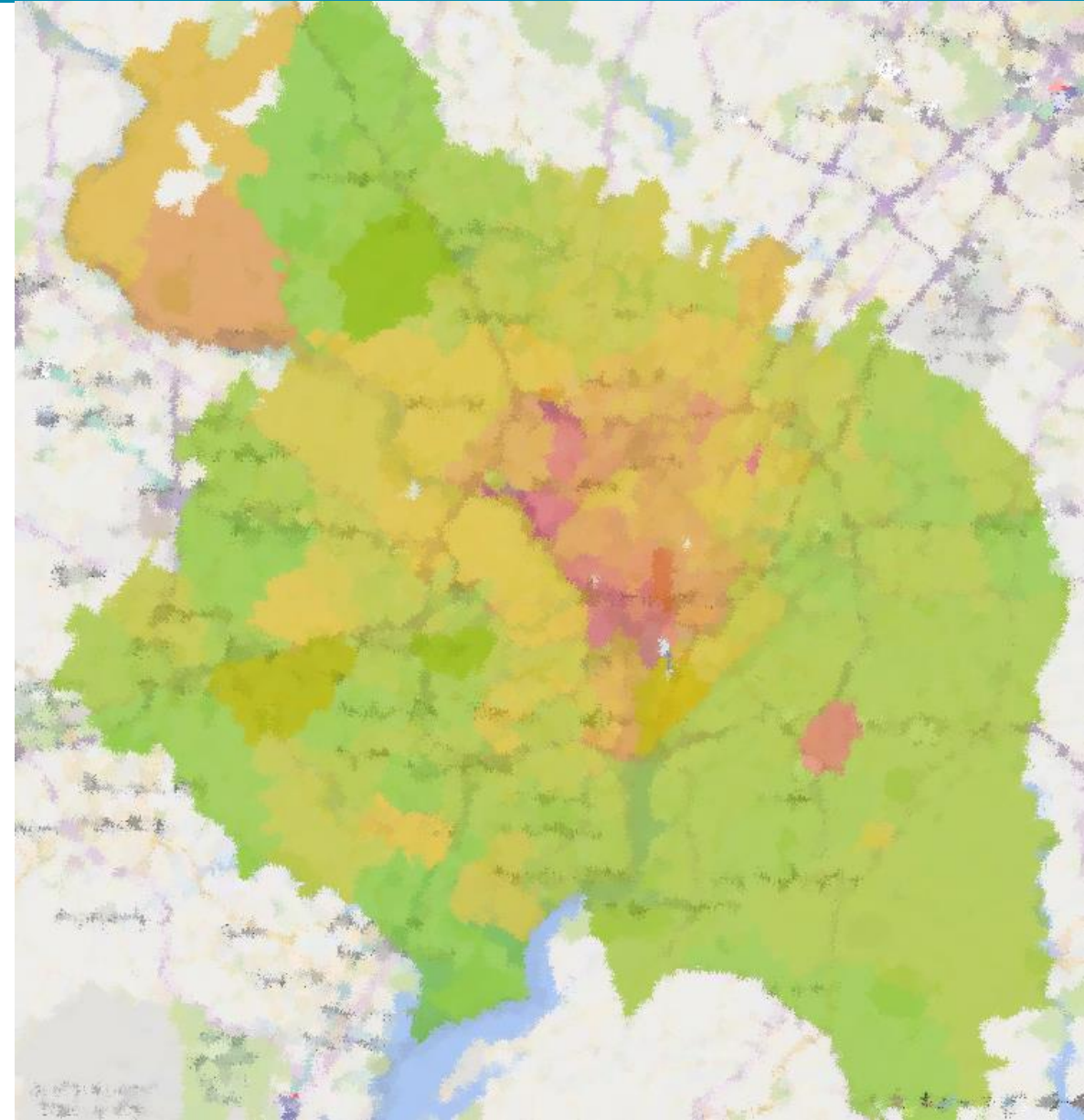
# What is the Equity Toolkit trying to enable?

## **This toolkit enables Metro staff to incorporate an equity lens in our work.**

It provides standard definitions for demographic and income variables, and ready access to a dataset that makes linking these variables to our service data easy.

It both expands access to the data needed to incorporate demographics with operational data AND ensures that staff are consistently applying the same equity definitions.

Metro has a long history of conducting Title VI analysis to document equity implications of major service changes, and to evaluate standards for service availability, vehicle load, vehicle headway, and on-time performance to ensure there is no significant difference in the service provided to people of color and low-income passengers. This Title VI work will continue outside of the Equity Toolkit.





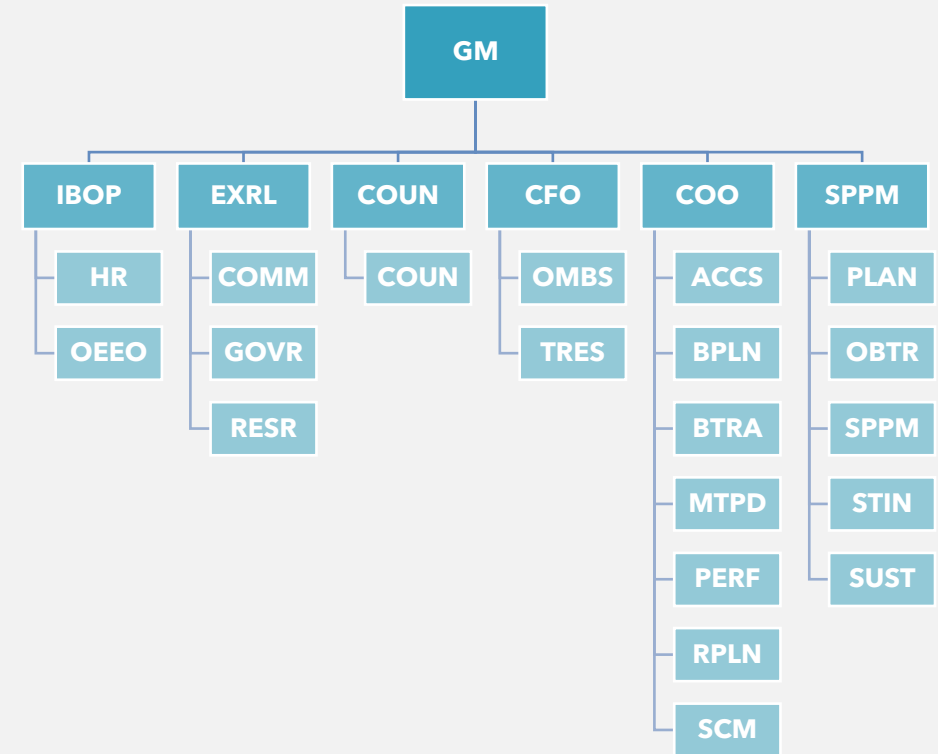
# What is the Equity Working Group?

**The Equity Working Group (EWG) is an interdepartmental working group** led by Office of Planning. The group is made up of over 30 members that represent 19 Metro offices and meets every four to six weeks.

**The purpose of the group is to develop proactive agency equity strategy that goes beyond Title VI, promote an inclusive transportation network, and facilitate cross-departmental communication related to equity.**

**The EWG is developing a Racial and Social Equity Policy Framework** focused on making Metro the region's equitable mobility provider of choice. One of the objectives within this framework is to build tools and processes to measure and prioritize as an input to Metro's decisions. The Equity Toolkit expands access to the data needed to conduct equity analyses and ensures that staff are consistently applying the same equity standards, thereby moving Metro toward meeting that objective.

**The EWG is made up of over 30 members representing 19 Metro offices**



# How is Title VI applied to Bus/Rail service guidelines?

## Bus or Rail Service Guidelines

- Comprehensive metrics to plan service and allocate resources
- Framework for recommending budget and other service changes

### Title VI Equity Impact Analyses

- Conduct for major service, fare, and access to fare media changes
- Evaluate disparate impact (DI) or disproportionate burden (DB) of proposed changes

### Title VI Monitoring

- Conducted every three years
- Monitors four standards during monitoring period
- Compares service for protected populations to mode average

# How are “Title VI” and the toolkit different?

**Analysts using the toolkit should not conduct a “Title VI” service equity analysis or service monitoring.**

Rather the toolkit is intended to enable analysts to incorporate demographic data into less formal work.

Title VI is a specific legal process that requires public outreach and sign-off from the Board of Directors. As mandated under law, the analysis is only conducted at certain times using a specific method established by the FTA.





# Equity Standards

What are the standards and definitions you need to know to accurately communicate your results?

## **In this section:**

The importance of standards

WMATA's definition of equity focus communities

Definition of key demographics



# Why are standards / definitions important?

It's important to remember that when we use the word "equity," we are talking about people, and specifically groups of people who experience current and historical structural oppression based on a particular trait.

In the first version of this toolkit, we chose to include six demographic variables that represent human traits for which people experience oppression in our society: race, ethnicity, income, gender, disability, and age. We prioritized these demographic variables because while many groups of people experience discrimination, the long history and continued violence of racism in this country is crucial to center in any discussion of equity. In particular, oppression along the other variables we included is inextricably connected to racism. For example, people of different races experience sexism differently.

The language that we use to talk about marginalized groups of people, particularly racially marginalized people, can be charged and complicated. And as social circumstances and culture change, language also changes. Therefore, as the data in this toolkit is refreshed annually, Metro will also review and update the language used to describe the people and variables included in the toolkit. Additionally, words have different meaning or connotation to different people. Therefore, we recommend that when using language to describe marginalized groups in an analysis based on the data in this toolkit, be as specific as possible and include a clear definition of how that language is used in the analysis.

# What are important definitions you need to know?

## Equity Focus Communities

**Definition:** Populations as defined by Metro's Equity Working Group (EWG) to capture the highest concentration of low-income, people of color, and people with disabilities in our service area.

*Equity Focus Communities are the Census block groups with the highest combined percentile rankings of % low-income households, % people of color, and % people with disabilities. They comprise about 30% of the population within the service area.*

# What are important definitions you need to know?

## Race and Ethnicity

**Definition:** In Metro rider surveys and the American Community Survey, respondents can choose to self-identify as one or more of the following racial groups: American Indian or Alaska Native, Asian, Black or African American, Hawaiian or other Pacific Islander, White, or Some Other Race. They can also choose to self-identify their ethnicity as Hispanic/Latinx or not.

**Terminology:** When referring to a specific racial or ethnic group, use the name of that group. When referring to a group of races or ethnicities, the toolkit uses people of color to include respondents who have identified as American Indian or Alaska Native, Asian, Black or African American, Hawaiian or other Pacific Islander, or Hispanic/Latinx. People of color is preferred to the terms minority or non-White.

**Reason for Including in the Toolkit:** Metro recognizes that public transportation providers have the responsibility and ability to address historical and ongoing racial discrimination and social inequity through engagement, service, fares, employment, and policing.

**Limitations:** The representativeness or lack of responses by people who do not speak English.

# What are important definitions you need to know?

## Poverty: Metro Title VI Definition

**Definition:** Standard for material deprivation set by WMATA that is to be used **only** when reporting on Title VI requirement.

**Threshold:** Household income less than \$30,000.

- **ACS Data:** Number of individuals with annual household income less than \$30,000 per year.
- **Ridership Survey Data:** Number of respondents with household income less than \$30,000 per year.

**Description:** This definition is based on the federal requirements to ensure non-discriminatory transportation and is required of transit agencies that receive federal funding. This definition is the standard that has traditionally been used by the agency as the poverty threshold for federal reporting and is included in the toolkit to ensure that WMATA continues to meet the federal requirement.

**Limitations:** Using only a static household income as a threshold does not consider other important factors in measuring economic disadvantage such as family size, cost of living, and inflation. Additionally, the threshold of \$30,000 is too low for the Metro region cost of living and thus excludes broad populations of individuals that require greater focus from the agency.



# What are important definitions you need to know?

## Poverty: Internal Standard Definition

**Definition:** Internal standard for poverty that is used **for all non-Title VI related policy decisions.**

**Threshold:** Household income less than 200% of federal poverty line.

- **ACS Data:** Number of households with annual income less than 200% of federal poverty line.
- **Ridership Survey Data:** Due to data availability, survey data approximates the metro poverty threshold as the number of respondents that fall below:
  - If household size is 2 or fewer: income under \$30,000 per year
  - If household 3 or more: income under \$50,000 per year

**Description:** This definition was created as a more inclusive standard for material deprivation that better reflects the economic realities of the region and allows the agency to better target communities that rely most heavily on public transportation. Tying the measure to the federal poverty line also considers important factors other than income such as family size and inflation. This definition will serve as WMATA's internal standard for poverty and should be used for all work not directly related to Title VI reporting.

**Limitations:** Currently not able to be measured in survey responses, so it must be approximated.

# What are important definitions you need to know?

## Disability Status

**Definition:** Registered MetroAccess riders who live in a zip code in the Metro service region.

**Reason for Including:** People with disabilities are often transit-dependent and face a myriad of societal barriers to living their full lives. We chose the universe of registered MetroAccess users because it is likely a more representative population of people with disabilities than data from the Metro Ridership Survey or the American Community Survey.

**Limitations:** Not all people with disabilities may be registered for MetroAccess, either because they solely utilize fixed-route services or they do not use any service provided by Metro.

# What are important definitions you need to know?

## Age

**Definition:** Expressed as an integer of the number of full 12-month cycles of life since birth; there is no rounding up to the next integer when in between 12-month cycles. Children less than 12 months old are expressed as age “zero.”

- Children: Residents aged 17 or under
- Adults: Residents aged 18-64
- Seniors: Residents aged 65 or older
- *Above Average % of Children:* % of residents aged 17 and under is over 25% (the average percentage of children in the DC area is 23, so a threshold over 25 would indicate a particularly high density of children)
- *Above Average % of Seniors:* % of residents aged 65 and over is over 15% (the average percentage of seniors in the DC area is 13, so a threshold over 15 would indicate a particularly high density of seniors)

**Reason for Including:** Age analysis can be helpful for internal service planning decisions; particularly given that seniors and children may generally be more heavily dependent on transit than those in other age groups.

**Limitations:** Broad age ranges within thresholds do not lend themselves to more age-granular analysis.

# What are important definitions you need to know?

## Gender

**Definition:** An individual's internal sense of being female, male, or something else.

**Reason for Including:** There is a wide body of evidence that shows patterns of travel differ by gender, due to a variety of reasons. Additionally, there are differences among other variables (such as poverty) based on gender.

**Limitations:** The Metro survey presents only "Female," "Male," and "Other Gender" as options and does not allow more specifics about other identities such as gender fluid or transgender. Additionally, the 2016 Rail Passenger Survey does not include a gender question. The gender data included for rail was estimated by matching the first name reported by a subsample of respondents to a database of names and sexes from U.S. social security records, as opposed to rider-reported gender.

The ACS only collects data on sex (the anatomical classification of a person, usually assigned at birth), and does not include an option for intersex. ACS sex data will likely not show a discernable difference between the geographic distribution of females and males across the region. This data is included to examine sex differences within other variables, such as race or income (i.e. to compare poverty levels between men and women).

Sex and gender are not the same variable, so they should not be considered equivalent across the two datasets.

# When should you use rider vs. resident data?

## To analyze riders | Use Metro Survey Data

- Every 4-5 years, WMATA conducts system-wide surveys of Metrorail riders and Metrobus riders. The last Metrorail Passenger Survey was conducted in 2016; the last Metrobus Passenger Survey in 2018. These are large-scale samples, with 10,000's of riders surveys.
- Survey data distinguishes people who actively use Metro from the people who live in the region
- Survey data differentiates bus and rail riders
- Time period, mode of access differences, fare price, and social-economic factors means that census data does not always match up with who is on board

## To analyze residents | Use American Community Survey (ACS) data

- The Equity Toolkit includes access to the US Census Bureau's American Community Survey, which collects data with relevant demographic characteristics of communities in the region. Unlike the decennial Census, the ACS is sent to a sample of addresses every month of every year, with results updated yearly.
- Data will be updated annually as new estimates are released so the toolkit will be up to date
- Data is aggregated on the Block Group level, the smallest geographical unit which the Census publishes data (600 to 3,000 people). Users can combine these groups to estimate characteristics of neighborhoods, wards, cities and regions and use this data to define areas of emphasis for the agency.

**Illustrative Example:** People of color make up 60 percent of our region's population but over 80 percent of Metrobus riders. The differences between the resident population and ridership can become stark at the route and line level, which is why it's important to consider both datasets.



# When should you use rider vs. resident data?

## Example: Say you want to understand who is *riding* at Pentagon Station?

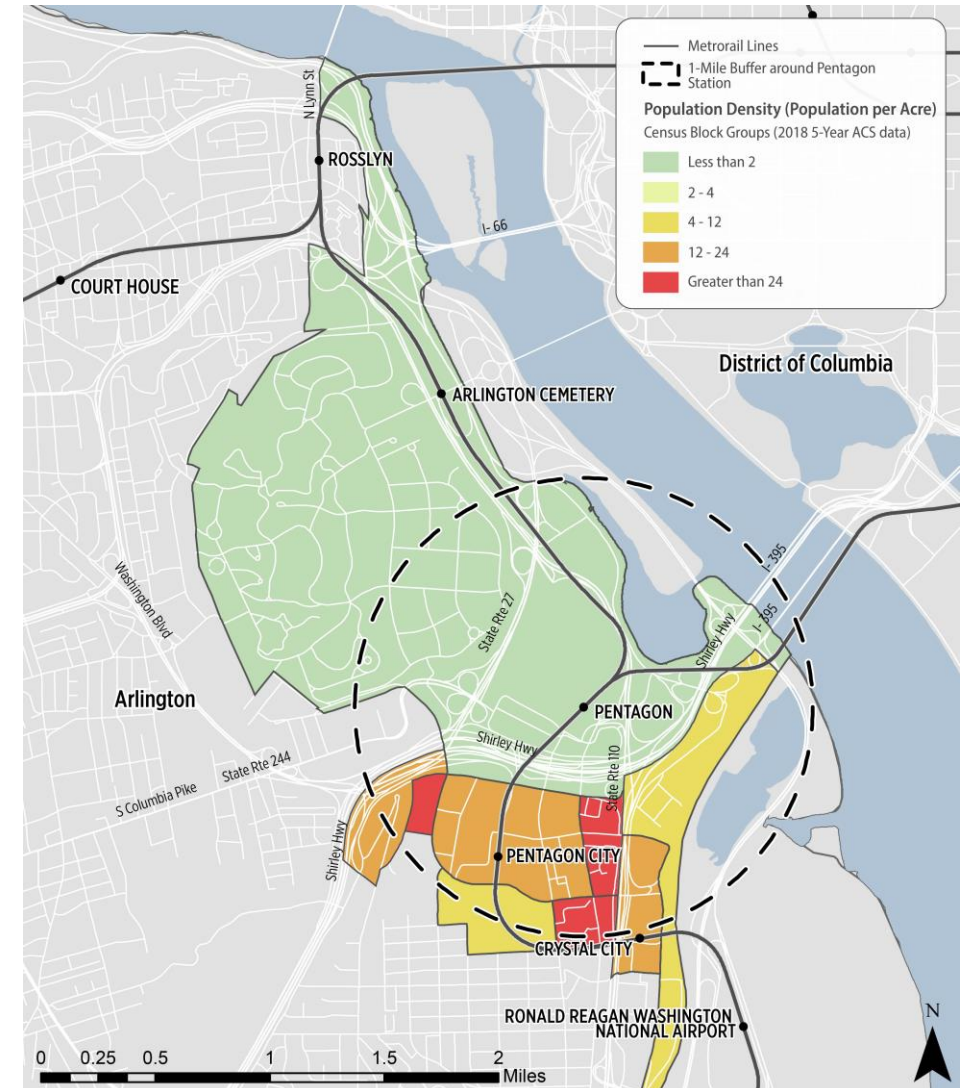
✗ Census data = those who live nearby (residents)

✓ Metro survey data = riders

### Why is survey data better?

- Most residents who live nearby use Pentagon City Metro Station, NOT Pentagon
- Most Pentagon riders are either Department of Defense employees or bus-to-rail transfers
- Fact: Pentagon Station is the most-used rail station for Alexandria residents, even though it is not in Alexandria

**Key: Understand what you are trying to analyze first: the riders or the surrounding community**

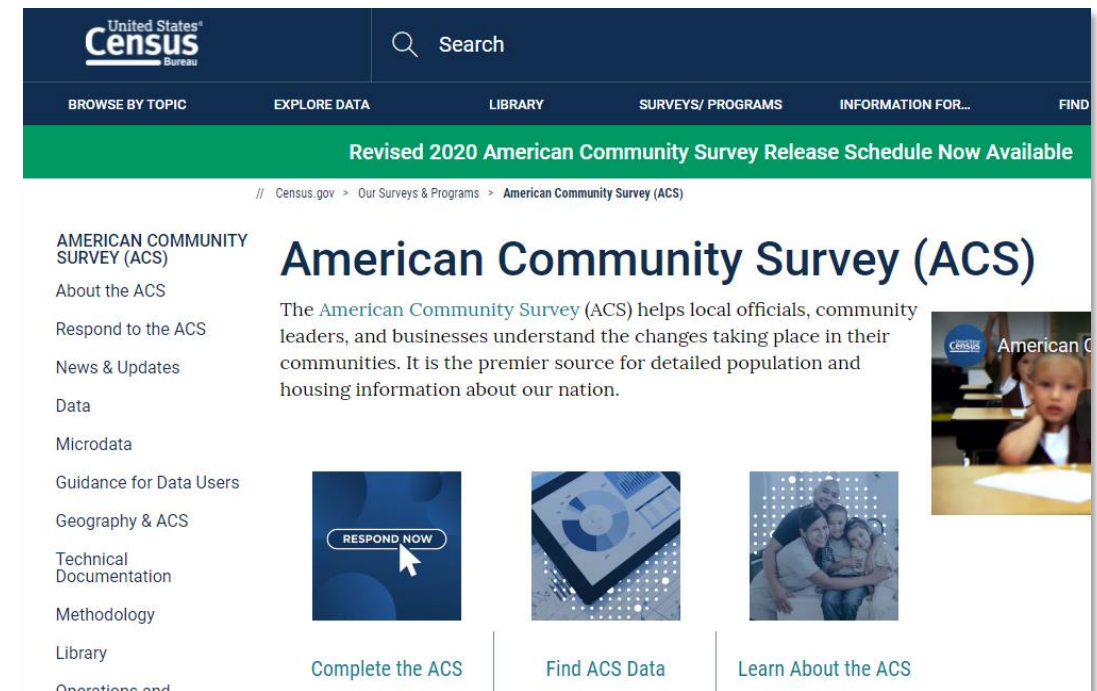


# What should you know about using ACS data?

While the demographics of an area may not always correlate with WMATA's riders, it is equally important to understand the characteristics of the region that we serve. Community characteristics allow WMATA to engage in more effective outreach to areas that we have historically missed and to identify mismatches in ridership and residents to provide better service.

**The Equity Toolkit includes access to the US Census Bureau's American Community Survey (ACS), which collects data relevant demographic characteristics of communities in the region.** Unlike the decennial Census, the ACS is sent to a sample of addresses every month of every year, with results updated yearly.

For the purposes of analyzing equity at WMATA, data is aggregated at the "Block Group" level, the smallest geographical unit which the Census publishes data (600 to 3,000 people). Users can combine these groups to estimate characteristics of neighborhoods, wards, cities and regions, and use this data to define areas of emphasis for the agency.



# What should you know about using ACS data?

## Methodology Overview

WMATA used GIS to join the ACS data to the location of WMATA rail stations, rail lines, bus stops, and bus routes. We assigned the demographic characteristics to WMATA assets to understand the residents in the surrounding geographic region.

Variables for each station and stop are determined by the population-weighted average of the block groups that are within an estimated walking distance of each location (0.5 miles for rail stations and 0.25 miles for bus stops).

Block groups that only partially fall within this radius are proportionally allocated based on the percentage of the area that falls inside of the specified radius.

## Three Different ACS Datasets for Rail and Bus

### Rail Datasets

- Station Demographic Data, with radii that overlap each other
- Line Demographic Data, with line radii that overlap each other
- System Demographic Data at Station level, with no overlap of radii

### Bus Datasets

- Stop Demographic Data, with radii that overlap each other
- Route Demographic Data, with line radii that overlap each other
- System Demographic Data at Stop level, with no overlap of radii

This is described in greater detail on the following page.

## Additional Datasets

- "Transit-shed" Demographic Data, using the merged rail station and bus stop radii, split by jurisdiction
- Zip Code Demographic Data, used primarily for MetroAccess analysis
- Block Group Demographic Data



# What should you know about using ACS data?

## Overlap Data | for stop/station- or line/route-level analysis

For stop or station level analysis, the Equity Toolkit has “Overlap Data”. **Overlap data can count the same block group for multiple stops, stations, routes, or lines.**

This is useful for an analysis that focuses on the demographic characteristics of a single location (i.e., What is the % poverty community with access to Stop X?). To avoid double-counting, it should not be used to aggregate at the route or line level.

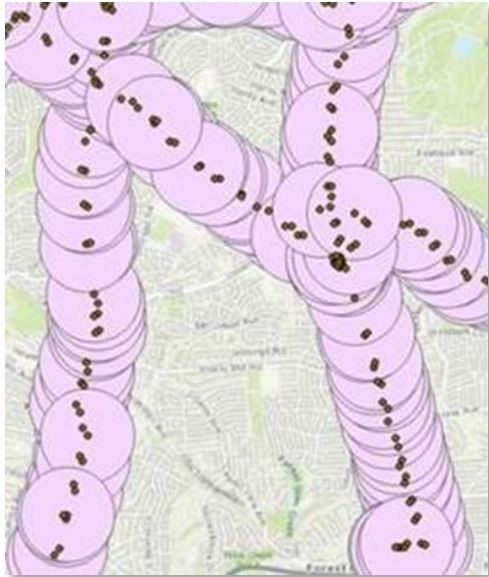


Image: Bus Stop Overlap Dataset

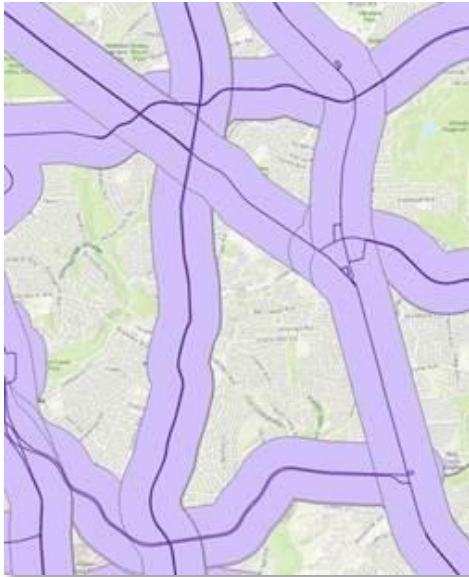


Image: Bus Route Overlap Dataset

## Non-Overlap Data | for system-, line-, or route-level analysis

For system, line, or route level analysis, the tool kit has “Non-Overlap Data”. **Non-overlap data does not count the same block group for multiple stops or stations.**

In this data set, an area is only assigned to one station or stop within its radius. This allows users to aggregate data of a line or segment and avoids double counting residents (i.e., what are the community characteristics of the X9 bus route).

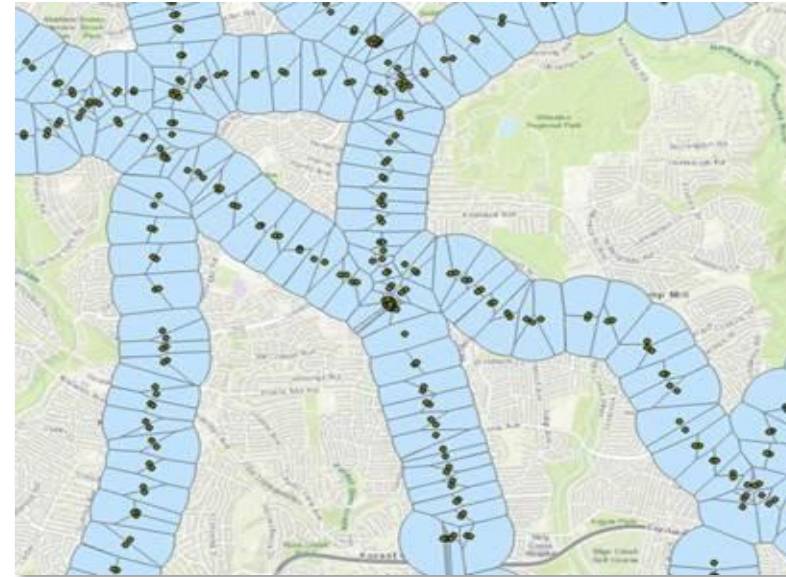


Image: Bus Stop Non-Overlap Dataset

# A background on Metro survey data

Every 4-5 years, WMATA conducts system-wide surveys of Metrorail riders and Metrobus riders. The last Metrorail Passenger Survey was conducted in 2016; the last Metrobus Passenger Survey in 2018. (Because of the pandemic, execution of the next Metrorail Passenger Survey has been postponed until Spring 2023.)

These passenger surveys sample trips, not individual riders—i.e., an individual rider may answer the survey multiple times, so long as they answer no more than one survey for each trip they take.

The Metrorail Passenger Survey samples all 118 station mezzanines and all time periods that Metrorail is open for business. All station mezzanines were sampled on weekdays, but only some station mezzanines were sampled on weekends (due to budget limitations and cost-benefit considerations).

Similarly, the Metrobus Passenger survey samples all Metrobus lines and all time periods. All Metrobus lines are sampled on weekdays, but only some are sampled on weekends.





# A background on Metro survey data

All WMATA staff who are interested in learning more about these surveys should contact the Office of Customer Research (RESR), the owner and custodian of the data files and related background information. Those who are unfamiliar with these datasets are strongly advised to consult with RESR and PLAN before analyzing them, as various weights and filters may need to be applied to yield results that accurately represent their populations of interest.

**Important note: The data files are proprietary to WMATA and can not be shared with anyone who is not a WMATA employee using them for WMATA project.** In some cases, access has been granted to outside parties contingent upon their completion of a non-disclosure agreement (NDA) with WMATA.

The screenshot displays the WMATA Customer Research website. The header features the WMATA logo, the title "Customer Research", and navigation links for Home, Employee Resources, Applications, News Room, Inside WMATA, and Departments. A search bar is located in the top right corner. Below the header, a breadcrumb trail shows "BROWSE" and "PAGE". The main content area is titled "Office of Customer Research (RESR)" and "Customer Research". It includes a description of RESR's role in building bridges between WMATA decision-makers, employees, and the public. A section titled "RESR accomplishes this by providing:" lists five services: Qualitative and Quantitative Research Services, Multi-Method Data Collection options, Questionnaire Design, Cross-Functional Team Leadership, and Effective Customer Experience Positioning. At the bottom, a paragraph states that RESR serves as an advisor to WMATA staff seeking guidance on quantitative and qualitative research, and provides contact information for further inquiries.

**Customer Research**

The Office of Customer Research (RESR) utilizes survey, opinion, and public policy research to build bridges between WMATA decision-makers, its employees and the public. RESR applies its energy and creativity to helping WMATA make better decisions by providing the highest quality research services.

RESR accomplishes this by providing:

- ✓ Qualitative and Quantitative Research Services
- ✓ Multi-Method Data Collection options
- ✓ Questionnaire Design
- ✓ Cross-Functional Team Leadership
- ✓ Effective Customer Experience Positioning

RESR serves as an advisor to WMATA staff who are seeking guidance on quantitative and qualitative research. A wide-range of research tactics are used, and RESR is the group to help you identify the appropriate tactic to answer your most pressing business and financial questions. Contact RESR with questions today about how to make your next project customer-centric.

# What should you know about Metro survey sample size?

**Sample size refers to the number of participants in a study, and influences two statistical properties: 1) the precision of our estimates and 2) the power of the study to draw conclusions.**

## Example

As an example, we might want to compare if trips by low-income riders are longer than trips by riders who are not low-income.

## Choosing the n

Since it would be impossible to survey every Metro rider, we have little choice but to focus on a segment of that larger population. This might mean randomly selecting only 100 riders for our study. The sample size, or  $n$ , in this scenario is 100. The survey describes the population of all riders based on the information obtained from the sample of 100 riders.

## Sampling error

No matter how careful we are about choosing our 100 riders, there will still be some margin of error in the study results. This is because we haven't talked to everyone in our population of interest. We can't be absolutely precise about how many riders are low-income because we can't look at every instance. This measure of error is known as sampling error. It influences the precision of our description of the population of all riders.

## Mitigating sampling error

Sampling error, though unavoidable, can be eased by sample size. Larger samples tend to be associated with a smaller margin of error. This makes sense. To get an accurate picture, we need plenty of examples to look at and compare.

However, there is a point at which increasing sample size no longer impacts the sampling error. This phenomenon is known as the law of diminishing returns.

## Calculating power

The second concept, power, refers to the probability of finding a statistically significant result. In our survey of riders, power is the probability of finding a difference in trip length across income level of riders.

We calculate power by specifying two alternative scenarios. The first, called the null hypothesis, is one that says there's nothing going on in the population of interest. In our study of riders, the null hypothesis might say that there's no difference in trip length for low-income riders.

The second is the alternative hypothesis, that trips by low-income riders are longer. A larger sample size gives more power, and better chance of finding a statistically significant result.

*Adapted from <https://www.iwh.on.ca/what-researchers-mean-by/sample-size-and-power>*

# What should you know about Metro survey sample size?

**What does it mean that the surveys were “sampled” at the station mezzanine, bus line, and time period level, and what are the implications for equity toolkit users?**

It means that demographic characteristics when reported at the station mezzanine or bus line level or at the time period level are statistically valid.

For the equity toolkit we’ve created two “datamarts”, one that displays rider survey demographics at the bus line or rail station level on a weekday, and a second that displays rider survey demographics at the time period level.

If the business question you are trying to answer is at a less aggregate level of analysis (for example bus route or bus stop), contact RESR or PLAN to determine if the appropriate way to use the survey data or if there are datasets that are more suitable for your project.

# What should you know about Metro survey methodology?

## 1. The percent of race and ethnicity variables in the equity toolkit are NOT additive

Both the ACS and Metro Ridership Survey allow respondents to select all race categories that apply. This means some respondents select two or more races

### Example

**Assume for simplicity that the survey has four total respondents**

- Rider 1 selects “White”
- Rider 2 selects “Black or African American”
- Rider 3 selects “Black or African American” and “Asian”
- Rider 4 selects “Asian”

### What races are the riders?

- 50% of riders are Asian
- 50% of riders are Black or African American
- 25% of riders are White

### How many riders are people of color?

- It would be INCORRECT to say, “Well if 50% of riders are Asian and 50% of riders are Black or African American then 100% of riders are people of color”
- Instead, we can see from the survey forms that 75% of riders are either Asian, or Black or African American

Rider #1's response:

What is your race? (Check all that apply)

<input type="checkbox"/> <sub>01</sub> American Indian or Alaska Native	<input checked="" type="checkbox"/> <sub>05</sub> White
<input type="checkbox"/> <sub>02</sub> Asian	<input type="checkbox"/> <sub>06</sub> Some other race (Specify):
<input type="checkbox"/> <sub>03</sub> Black or African American	
<input type="checkbox"/> <sub>04</sub> Hawaiian or other Pacific Islander	

Rider #2's response:

What is your race? (Check all that apply)

<input type="checkbox"/> <sub>01</sub> American Indian or Alaska Native	<input type="checkbox"/> <sub>05</sub> White
<input type="checkbox"/> <sub>02</sub> Asian	<input type="checkbox"/> <sub>06</sub> Some other race (Specify):
<input checked="" type="checkbox"/> <sub>03</sub> Black or African American	
<input type="checkbox"/> <sub>04</sub> Hawaiian or other Pacific Islander	

Rider #3's response:

What is your race? (Check all that apply)

<input type="checkbox"/> <sub>01</sub> American Indian or Alaska Native	<input type="checkbox"/> <sub>05</sub> White
<input checked="" type="checkbox"/> <sub>02</sub> Asian	<input type="checkbox"/> <sub>06</sub> Some other race (Specify):
<input checked="" type="checkbox"/> <sub>03</sub> Black or African American	
<input type="checkbox"/> <sub>04</sub> Hawaiian or other Pacific Islander	

Rider #4's response:

What is your race? (Check all that apply)

<input type="checkbox"/> <sub>01</sub> American Indian or Alaska Native	<input type="checkbox"/> <sub>05</sub> White
<input checked="" type="checkbox"/> <sub>02</sub> Asian	<input type="checkbox"/> <sub>06</sub> Some other race (Specify):
<input type="checkbox"/> <sub>03</sub> Black or African American	
<input type="checkbox"/> <sub>04</sub> Hawaiian or other Pacific Islander	

# What should you know about Metro survey methodology?

## **2. Some demographic data is unavailable for rail on Sundays**

- The Sunday rail survey did not ask gender, age, or household size so we have not calculated related variables. Caution should be taken in how you choose to aggregate and compare weekend values on rail and bus. Reach out to equity toolkit liaison with questions.

## **3. The weekday rail excludes the late-night period**

- The 2016 rail survey oversampled the Friday Late Night period, which is not likely to be representative of the typical weeknight rider so this data is excluded from the weekday rail demographic numbers. The bus data includes all time periods.

## **4. Using 2016/2018 surveys to estimate a different year**

- Take the percentages presented in the survey datamarts and weight by the ridership for the year/station/line/time period in question



# What should you know about Metro survey methodology?

## What does this mean for using the race and ethnicity variables?

- As presented in the equity toolkit, the % of each race includes the people who chose only one race, or selected two or more races and one of those is the % race being calculated
- Similarly, % race variables are not additive with % Hispanic or Latinx for the same reason-it could double count individuals
- The equity toolkit also includes a composite variable for “% people of color” which correctly adjusts for responses where two or more races were selected, as well as those who select Hispanic/Latinx . In other words, it doesn’t double count passengers.

# How can you use Metro survey data?

## Basic Approach

For the equity toolkit we've created two survey "datamarts", one that displays rider survey demographics at the bus line or rail station level on a **weekday**, and a second that displays rider survey demographics at the **time-period** level (e.g. AM Peak). What you are trying to analyze (something on a geographic basis or something for specific time period) will determine which data sets to use.

Typically, the survey data provide a percent (for example % low-income) for a given set of stations or lines. Multiply these percentages by the ridership to calculate the number of riders in a given demographic category for your business question.

### For example

Route 1: 50% low-Income \* 1000 daily riders = 500 low-income daily riders

Route 2: 60% low-Income \* 500 daily riders = 300 low-income daily riders

Total for Routes 1 and 2 = 800 daily low-income riders

Metro typically uses "**valid percent**" when reporting survey totals. Non-responses are not included in the denominator of the percentage

### For example

Say 120 people filled out the survey in total, but 20 people skipped the income question. Then, 50 people responded that they are low-income. The low-income percentage is 50%, based on 50/100 not 50/120

# How can you use Metro survey data?

Depending on the level of analysis, there may be categories (like bus lines or certain rail groupings) where **demographic (and other data) are not available**. As with valid percent referenced above, the analyst should remove these “null” values before calculating a percentage.

## Example

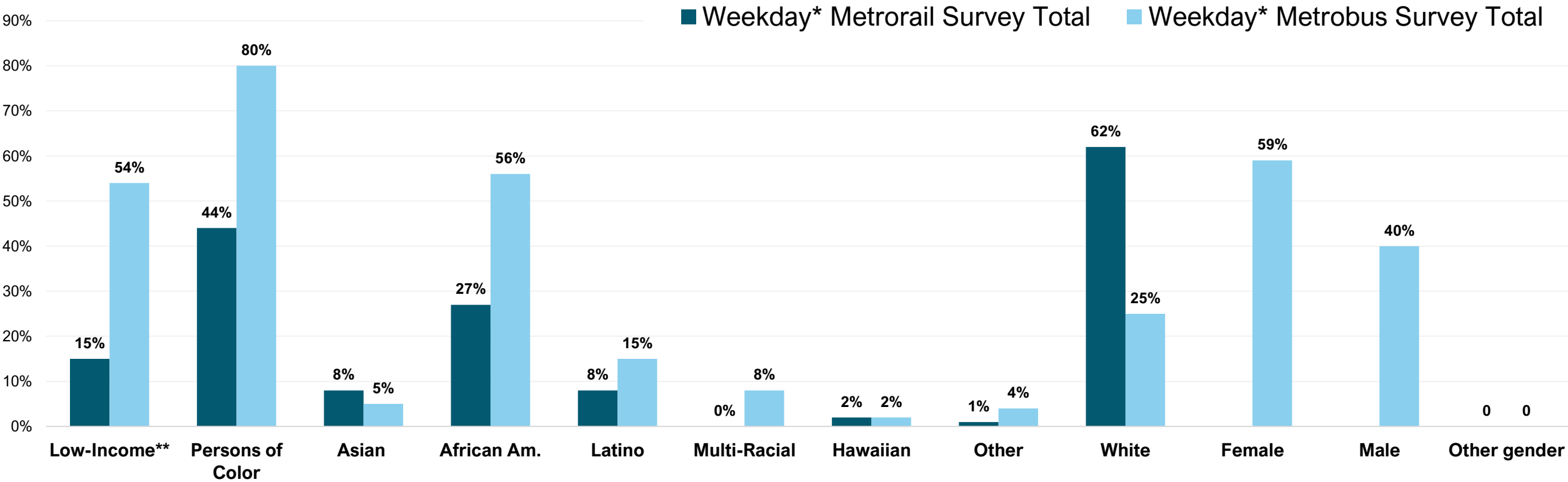
OTP data is available for 97 percent of riders. For this group, demographic data are available for 90 percent. The analysis should only include riders for which both OTP and demographic data are available. The analyst should identify and net out null cases.

## Talk to an SME first if the business question you are trying to answer is:

- At a less aggregate level of analysis (like a bus stop or only part of a bus line)
- If you want to analyze something about weekend ridership
- If you only look at a one or a few stations, one or a few bus lines, or at specific set of rail origin-destinations, as the sample sizes may not allow for such an analysis

# What does recent survey data tell us?

## What does recent Metro Rider Survey data tell us about our ridership demographics?



\*Weekday Only. Other reporting (such as that used in the Public Participation Plan) can include Saturday and Sunday

\*\*Low-income as defined for Equity Toolkit. Note this differs from the Title VI definition used in Public Participation Plan and other reporting

# Where can you get more detailed information?

## [\*\*Metrobus Annual Line Performance Report\*\*](#)

Rider demographic and income information by line and route

## [\*\*Metrobus Service Guidelines\*\*](#)

See how equity is factored into guidelines for the network and an assessment of the benefit each line provides

## [\*\*Title VI Plan\*\*](#)

Documents Metro's public engagement process, equity analyses, service standards, and results of monitoring service



# Equity Data Hub

A comprehensive, centralized set of relevant data tables to give you everything you need for conducting your analysis

## **In this section:**

How do you get permissions to access the data hub?

What do you need to access the data hub?

What operational data is available?

What demographic data is available?

How do you pull the data I need?

# How do you get permissions to access the data hub?

## Who can get access to the data hub?

Full time employees at Metro

## How do you request access?

Download a copy of the [Access Request Form](#) and fill out the unfilled fields. Have your manager fill out and sign the Requestor's Manager section. Then send the form to Scott Traum ([straum@wmata.com](mailto:straum@wmata.com)) who will work with IT to get the remaining signatures and provide access.

## Who can you share the data with?

Internal WMATA only, and you should write "draft" on all projects using the data

## Any caveats you should know about the data?

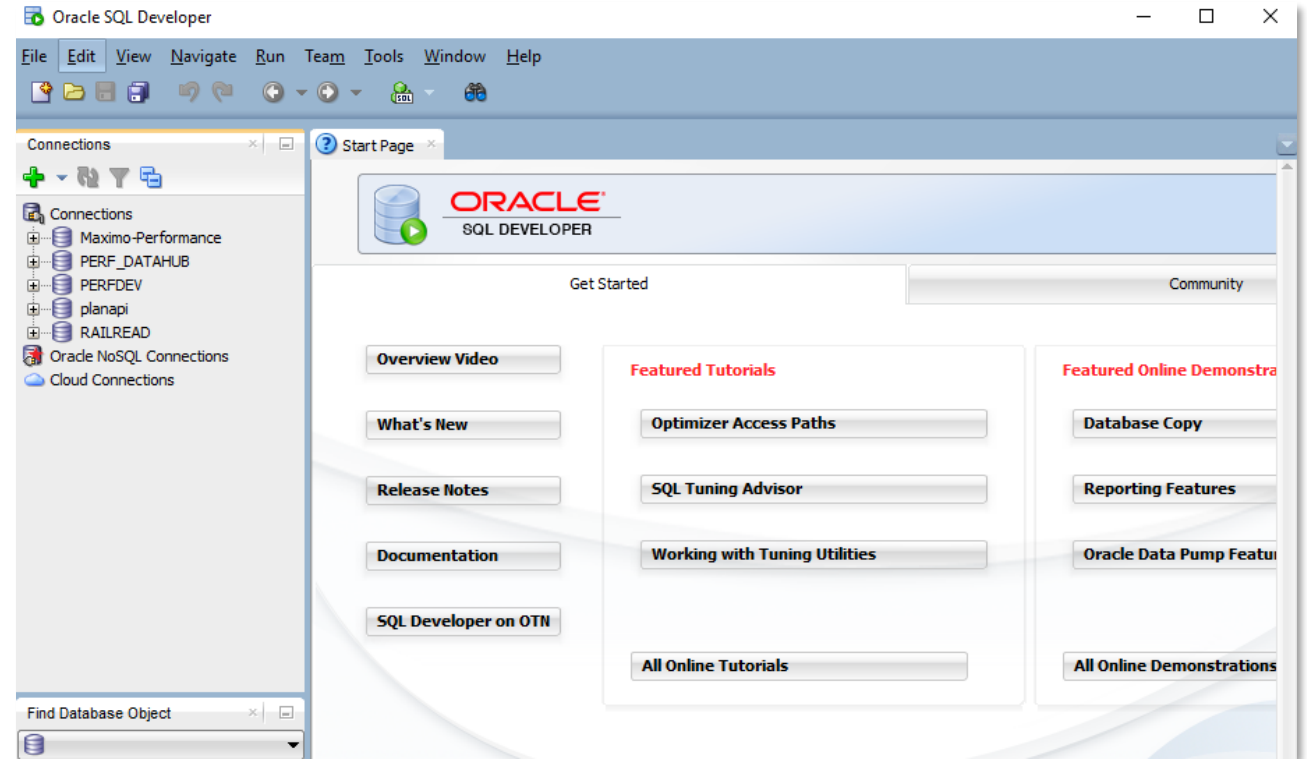
We have an entire section on how and when to use both the Census and Survey data. [Click to navigate to that section](#). See the [Action Section](#) for general guidance on verifying your results

Washington Metropolitan Area Transit Authority Oracle Database Access Form	
Requestor's Name:	Badge Number: E <input checked="" type="checkbox"/> Employee <input type="checkbox"/> Contractor
Department : Office Location:	Desk Phone No.: Cell Phone No. :
Date Submitted:	Date Required:
Duration of Access: <input checked="" type="checkbox"/> Permanent <input type="checkbox"/> Temporary	Database User Type : <input checked="" type="checkbox"/> Individual Account <input type="checkbox"/> Service Account <input type="checkbox"/> Schema Account
Application you need access to: CDMPROD.WMATA.COM	Database Name you need access to: WMATA Equity Toolkit Role
Type of Access: <input type="checkbox"/> Read Only – All Tables <input checked="" type="checkbox"/> Read Only – Specific Tables <input type="checkbox"/> Read/Write – Specific Tables <input type="checkbox"/> DBA <input type="checkbox"/> Schema Privileges	
Provide Details of Access Needed (Attach additional sheets if required): This user needs to be granted access to the WMATA Equity Toolkit via the WMATA Equity Toolkit Role	
Purpose of Access: This user will be using WMATA Equity Toolkit to conduct equity analyses for Metro.	
Approvals:	
Requestor's Manager Name:	Employee ID:

# What do you need to access the data hub?

The data is stored in Oracle SQL on CDMPROD, which means you will need to download, install, and use SQL Developer to view, query, and manipulate the ETK datasets.

1. Download and Install SQL Developer ([instructions](#))
2. Complete ETK SQL Access form and send to Scott Traum (see previous slide for link)
3. Once you have access to the ETK SQL Role, create a new connection in SQL Developer ([instructions](#))



# What operational data is available?

Mode	Available data	The formal name of the table in the data hub
<b>Metrorail</b>	train schedule adherence	RAILOPS.RAIL_TRIP_STOPS_MASTER
	customer on-time performance	Ridership_Portal.TT_F_RAIL_OTP_SUMMARY SMARTTRIP.TT_F_ORIGIN_DEST_PAIR_DTL_HIST
	ridership	SMARTTRIP.TT_F_METRO_RIDERSHIP_EACHDAY
<b>Metrobus</b>	bus on-time performance	SMARTTRIP.SCHLD_SMRY_AND_TRIP_STOPS
	customer on-time performance	SMARTTRIP.BUS_OTP_V
	ridership (APC)	SMARTTRIP.EQUITY_BUS_Ridership_RideCheck
<b>MetroAccess</b>	trip on-time performance	SMARTTRIP.EQUITY_MA_F_TRIP_DETAIL
	ridership	SMARTTRIP.EQUITY_MA_F_TRIP_DETAIL
	rider information (name, address, etc.)	SMARTTRIP.EQUITY_MA_D_riders

# What demographic data is available?

1

## Residents in the vicinity of Metrorail and Metrobus operations

Demographic data on people who live within Metro's service region from the American Community Survey (ACS). ACS datasets are organized into three levels for RAIL and BUS assets, and one level for ACCS

### For all rail stations and bus stops (level analysis with buffer)

**Rail station** → smarttrip.equity\_rail\_stn\_ovrlp\_census

**Bus stop** → smarttrip.equity\_bus\_stops\_ovrlp\_census

### For regular routes (level analysis with buffer)

**Rail line** → smarttrip.equity\_rail\_lin\_ovrlp\_census

**Bus route** → smarttrip.equity\_bus\_route\_ovrlp\_census

### For all rail lines and bus routes with nonoverlapping polygons

**Rail lines** → smarttrip.equity\_rail\_stn\_unovrlp\_census

**Bus routes** → smarttrip.equity\_bus\_stop\_unovrlp\_census

Rail spatial analysis uses a ½ mile buffer around stations. Bus spatial analysis uses a ¼ mile buffer around stops.

### For MetroAccess

ACS data is aggregated by zipcode, which will map to the trip origin zipcodes in the ACCS operational datasets.

smarttrip.equity\_census\_by\_zipcode

### Administrative

**Block group data** → equity\_block\_group\_census

**Transitshed by jurisdiction** → equity\_transitshed\_juris\_census

**For more information on these datasets, navigate to [slides 32-33](#)**

2

## Metrorail and Metrobus riders

Demographic data on people who actively use Metro services (more detailed information in standards section)

Rail Customer 2016 Survey; Bus Customer 2018 Survey (Table names below)

smarttrip.equity\_survey\_bylinestationadj

smarttrip.equity\_survey\_system\_byperiod

# How do you pull the data you need?

## Querying the Equity Toolkit Data

1. Open SQL Developer and load the WMATA Equity Toolkit connection (should open a new worksheet in SQL Developer for that connection)
2. Navigate to the tables in the ETK using the following steps:
  - Expand connection in Connections pane, expand Other Users, and find ETK tables in the SMARTTRIP / RAILOPS / Ridership\_Portal folders.
3. Using the table list on the previous slides of the ETK, identify which tables you'd like to utilize
4. Decide whether to export data from SQL to Excel to analyze, use pre-written queries to generate common metrics and tables, or write your own queries to join datasets and analyze equity metrics.

## Other

- Alternatively, you can contact Elijah Evans (eevans2@wmata.com) as a SQL subject matter expert during the soft launch phase of the DataHub.
- Teams across Metro can consider assigning a "SQL SME" to satisfy requests for data from team members.

test\_query\_ETK.sql

```
1  /*Select All Data from an Equity Census table*/
2
3  Select
4  *
5  from
6  --insert schema/table name below
7  SMARTTRIP.equity_survey_system_byperiod
8
```



A map of the Washington Metropolitan Area, showing various cities and transit routes. The map is color-coded with green for parks and open spaces, yellow for urban areas, and blue for water bodies. Major highways are shown in red and orange. The map includes labels for cities such as Ellicott City, Columbia, Gaithersburg, Rockville, Aspen Hill, Wheaton, Silver Spring, Bethesda, Potomac, Washington, Providence, Annandale, Mason, Alexandria, Springfield, Rose Hill, Woodlawn, Mount Vernon, Fort Washington, Waldorf, and Occoquan. The map also shows the Potomac River and the Chesapeake Bay.

# Conducting Analysis with the ETK

## In this section:

How do you determine the scope of your analysis?

What key questions should you ask when conducting your analysis?

How can you ensure that your analysis is correct?

# Conducting your analysis



With the wealth of data available through the Equity Toolkit, spending time **defining the scope** of your analysis and **answering some key questions** will ensure that we’re conducting our analyses throughout the agency in a consistent and meaningful way.

In this section, we’ll walk through the key steps to approach, conduct, and quality-check your analysis.

As we walk through each step, we’ll apply it to three “use cases” – vetted examples of the ETK in action. You can refer to the use cases files linked below as review this section:

**Use Case 1** | [MetroAccess OTP](#)

**Use Case 2** | [BUS Log-on Rates](#)

**Use Case 3** | [BUS All-Day Service Plan](#)

# 1 | Determine the objective

Before starting your analysis, it is critical to determine the objective of the analysis. What question(s) are you trying to answer? How does answering this question provide value to Metro?

## **Use case 1 | Does MetroAccess On-Time Pick-up Performance (OTP) differ by riders' race and income?**

The findings provide a data-driven basis for improving equitable quality of service for MetroAccess riders. Leaders may act on the results by prioritizing an equity emphasis in trip scheduling, holding service delivery providers accountable to on-time pick-ups, particularly first pick-ups of the day, and monitoring changes in OTP over time to quickly act on trends indicating less equitable service is being provided.

## **Use case 2 | Are all residents provided the same level of real-time bus information at the stop-level?**

The findings resulting from this analysis will provide useful information for decision-makers as they improve Bus Logon capabilities.

## **Use case 3 | Does the Frequent Service Plan improve access to all-day frequent transit service (defined as every 12 minute or better, 7am to 9pm Monday through Friday) for the region's low-income residents and residents of color?**

The answer for this question will provide useful information for decision-makers as they consider moving the Frequent Service Plan proposal forward.

## 2 | Review the operating data available

Match that question to the relevant Metro operating data available. What data support this analysis? How are these data summarized; by rider, by vehicle, by bus route, etc.? Also, what do the data report? For example, rider on-time performance and vehicle schedule adherence are two different ways to report on-time performance data. Which best relates to the business question you are trying to answer?

**Use case 1** | Ideally, we would be able to use rider demographic information to compare against OTP by zip code, but ACCS does not collect rider demographic data by trip origin location, so we used the ACS data by zip code to compare demographic information about zip code residents against OTP.

Datasets (1) MetroAccess trip origin location zip code & OTP data (at the stop level) for specified date range from Trapeze; (2) Demographic data from American Community Survey data by Census Block Group; (3) List of zip codes within each Census Block Group

**Use case 2** | We used three datasets in this analysis.

- Bus Stop Logon Data (bs\_v\_logon\_stop\_level) for April 2021
- ACS Block Group Data for Bus Stops (equity\_bus\_stop\_unovrlp\_census)
- Ridecheck Data for Bus Stop Lat / Long (equity\_bus\_ridership\_ridecheck)

**Use case 3** | While both current and proposed service level information was used to inform this effort, no operating data from the toolkit was needed for this analysis.

# 3 | Factor in data limitations

What will the data not be able to answer? Very rarely are data sources tailored specifically to a research question, and most data has limits on its usability when you first obtain it. Analysts may find datasets to be incomplete (i.e., missing data) or vary in quality (i.e., unusable formats), and should plan for these issues before beginning their research. For survey data such as the Ridership Surveys and the Census data, problems with data collection and sampling prevent the use of certain categories of data. Be sure to check in with the SME for each data source if you are ever unsure about using a specific data point.

**Use case 1 |** Some zip codes are not entered in the standard ##### zip code format, making it impossible to determine the trip origin location zip code; these data points were excluded from the analysis.

Since ACCS does not collect rider demographic data by trip origin location, the analysis is limited to comparing the demographic characteristics of the zip code residents near the pick-up location vs. actual riders.

**Use case 2 |** N/A

**Use case 3 |** N/A

## 4 | Determine the subject of the analysis

What am I trying to determine? Is it equity for riders? Is it access to service for a specific geographic community? Ridership data allows for focus on a specific set of bus lines or rail stations. However, Census data tell us about who lives in a specific region, regardless of whether they currently use Metro.

**Use case 1** | Does MetroAccess On-Time Pick-up Performance (OTP) differ for trips originating in areas with an above average proportion of residents of color and low-income residents?

**Use case 2** | Is real-time Bus information equitable for residents around stops in areas with an above average proportion of residents of color and low-income residents?

**Use case 3** | Does the Frequent Service Plan benefit low-income residents and residents of color, whether or not they currently use Metro?

This analysis differs from a Title VI analysis, because this analysis looks at the change for all residents, whether they take Metro or not. Title VI, on the other hand, looks only at riders.



# 5 | Determine the level of detail needed from data

Match the correct level of detail for your business question to the ridership or number of households (from Census data). For example, is rail station level OK or do you need origin and destination level?

**Use case 1** | Stop-level data based on zip code of riders' trip origin locations is tied to Census data.

**Use case 2** | Stop-level Bus logon / logoff data is joined to Census data based on bus stop location.

**Use case 3** | For this analysis, Census data assigned to rail stations and bus stops are the appropriate level of detail.

As described on page 33 of this document, the toolkit includes Census demographic data assigned to Metro rail stations and bus stops. Since the analysis is at a system level, "Non-Overlap Data" was used, rather than the "Overlap Data."

## 6 | Determine Survey vs Census for demographics

Please review the guidance in Section 2. Survey data should be used for questions about riders. Census data should be used for questions about people who live in specific community. There are trade-offs of using either dataset. Survey data are critical in understanding differences in demographics between bus and rail riders, and between the time of day and distance that a rail rider travels. However, it does not provide much insight into who lives in a community, especially people not currently served by Metro.

**Use case 1** | Census data is used for this analysis because survey data on the demographics for MetroAccess riders based on their trip origin location does not exist, although other general demographic survey data for ACCS riders is available. Therefore, the demographics are for the residents of the zip codes where MetroAccess trips originate.

**Use case 2** | Census data is used for this analysis to understand the demographics of residents near bus stops.

**Use case 3** | Census data is the appropriate data source. Census data provides insight into how all residents are impacted by this plan. This analysis will provide insight into the implications this proposal has on all of the region's most vulnerable residents, not just Metro's current riders.

# 7 | Determine your “unit of analysis”

What Metro Measure will help best answer the research question? Analysts should pick units of analysis that helps them provide insight on the service or policy they are researching. For example, if the focus of the analysis is reliability of bus service, one might want to focus on OTP or run times, but if the focus of analysis is accessibility, the focus should be on metrics such as distance from stops, hours of service, and headway length. It is recommended that analysts choose more than one unit of analysis for their research in order to understand nuances in the data, understand discrepancies better, and explain the results the best way possible.

**Use case 1** | On-Time Pick-up Performance for areas with above and below average concentrations of people of color and low-income residents.

**Use case 2** | Bus Stop Logoff Percentage for locations with above and below average concentrations of people of color and low-income residents.

**Use case 3** | Availability of frequent service (defined as 12 minute frequency or better, from 7am to 9pm Monday through Sunday) by community group.

# 8 | Pull the data

Using the SQL scripts in the datahub as a guide, pull the necessary operations, ridership and demographic data together for analysis.

**Use case 1** | Used a SQL script to pull the #On-Time Stops, #Total Stops by trip origin zip code for the designated time period. Used a SQL script to pull demographic data by zip code. Used a lookup function to bring together OTP by zip code with demographic data by zip code.

**Use case 2** | Used a SQL script to join the Stop Logon data and the Census data, as well as calculate the Stop Logon % for each stop. Within the Excel file, used a lookup function to join the Bus Stop regional ID #'s with the Lat / Lon coordinates from the Ridecheck table in the Datahub. Then used an Excel formula to calculate the quartiles for each bus stop based on the persons of color and low-income measures.

**Use case 3** | Used a SQL script to pull demographics for bus stops and rail stations with all day frequent service (1) prior to the All Day Service Plan proposal and (2) under the All Day Service Plan proposal.

*Note: An analysis of bus stops receiving all day frequent service had previously been conducted. The resulting list of Regional Bus Stop IDs for bus stops receiving all day frequent service was used in the query.*

## 9 | Create a base case

What is the basis for comparison for your analysis? Quantifying equity means comparing the distribution of benefits between socio-economic groups and determining the differences in impact. Analysts should establish a broader population (i.e., all Riders/HHs) to provide an accurate measure of Metro's service for emphasis groups. For example, an OTP rate of 65% for bus riders of color might appear to be acceptable, but if the system wide average for all riders is 85%, then bus performance is inequitable. Analysts should pull the data for their base case from the same source used for the rest of their analysis to ensure similarity for comparison.

**Use case 1** | The base case for this analysis is overall MetroAccess OTP and the average percent of low-income residents and residents of color by zip code for the zip codes included in the trip origin location data. These averages were calculated using the data from the SQL queries described in Step 8.

**Use case 2** | The base case for this analysis is overall Bus Stop Log-Off percentage for the entire month of April, system-wide. The baseline was calculated using the data from the SQL queries described in Step 8.

**Use case 3** | The base case for this analysis is all of the region's residents with access to frequent service.

# 10 | Compare the test group to the base

Which method will I use to make the comparison between what is being tested and the base case? There are multiple different methodologies an analyst can use to test results such as comparison of means, difference in differences, and regression analysis.

**A Note on Regression Analysis:** While this toolkit provides you with the data necessary to run a regression, it is generally recommended to use other methodologies. Regression analyses tend to be difficult to communicate to broader audiences and other methods can communicate similar findings with simpler messaging.

**Use case 1** | Compare the OTP for trips originating in areas with above average concentrations of low-income and residents of color to overall OTP and trips originating in areas with below average concentrations of low-income and residents of color.

**Use case 2** | We compare the base case (system-wide result) to the system-wide result for stops in areas with the 4<sup>th</sup> quartile of residents of color and low-income residents.

**Use case 3** | Compare the percent increase in access to all-day frequent service for (1) low-income residents and (2) residents of color to the increase in access for all residents.



# 11 | Determine significance

Is the result of the research meaningful? Determine whether your test case is different from the base case in a way that is inequitable. Then, determine what difference would be considered significant. For example, a difference of 0.1% is not the same of one that is 30%. **There is no formal standard for what is considered significant.** In some cases, a 5% difference is negligible and in others 5% can be indicative of a serious issue in need of urgent action. Analysts should rely data and subject matter experts to decide whether findings are meaningful, and then build a case to leadership for a recommended action that is tailored to the results.

**Use case 1** | Started off by using a t-test to compare the average OTP for riders with trips originating in areas with above average concentrations of low-income and residents of color to riders with trips originating in areas with below average low-income and residents of color. Then plotted OTP & % POC & % low-income residents on a scatterplot to identify whether there was a significant relationship. Both tests resulted in finding that there was not a major difference in OTP between the populations. The results of this analysis are intended to provide additional information to ACCS leadership regarding communities and geographic focus areas for targeted OTP improvement efforts.

**Use case 2** | As may be the case for Equity Toolkit analyses that focus on Census demographic data, this use case does not include a statistical analysis. The results of this analysis are intended to provide additional information to an effort to increase Bus Logons.

**Use case 3** | As may be the case for Equity Toolkit analyses that focus on Census demographic data, this use case does not include a statistical analysis. The results of this analysis are intended to provide additional information about the Frequent Service Proposal, rather than determine with significance whether it is equitable or inequitable. As such, this step is not applicable for this use case.

# 12 | Document methodology and results

Document your findings as well as your data sources and methodology used so they can be referenced by analysts looking to review, replicate, or expand upon your work in the future.

**Use case 1** | Overall MetroAccess service reliability as measured by OTP is equitable. OTP was slightly lower for trips originating in areas with above average concentrations of low-income residents compared to overall OTP and especially compared to areas with below average concentrations of low-income residents; however with a standard deviation of ~3 percentage points, these differences in averages are negligible. The findings focused on identifying specific geographic areas for improvement where OTP was less than one standard deviation below the average. Methodology and results were documented in a PowerPoint format with supporting data in an Excel file.

**Use case 2** | All residents are not provided the same level of real-time bus information, but the difference is still within the measure's performance band set by leadership in BusSTAT. Methodology and results were documented in a PowerPoint format with supporting data in an Excel file.

**Use case 3** | The increase in frequent rail access for all of the region's residents was 23%, compared to 26% for low-income residents and 34% for people of color. The increase in frequent bus access for all of the region's residents was 512%, compared to 582% for low-income residents and 759% for people of color.

In both cases, the percent increase is higher for low-income residents and people of color in the region than for all of the region's residents.

Methodology and results were documented in a PowerPoint format with supporting data in an Excel file.



# Quality Assurance/ Quality Control

How can you ensure that my analysis is correct?

## **In this section:**

Overview

Getting started

Peer review

SME review

Every analysis performed using the Equity Toolkit **must be** reviewed before findings are made available outside of the ETK team. The QA/QC process is in place to make sure that Metro is producing consistent and accurate equity analyses the first time, every time. Having someone review your analysis helps catch errors in code, correct terminology, and simply produces a better product. The following lays out the two-step review process for all analyses:

**Note: This process is not meant to replace other QA's or reviews that exist outside of the ETK; this is an additional review to ensure that the Tool Kit is being used correctly**

### 1. Peer Review (Slide 63-64)

- Review by **fellow analyst** that also uses ETK
- Checks the following:
  - Methodology
  - Math/Data
  - Terminology
  - Presentation

### 2. SME Review (Slide 65-66)

- Review by **select individuals** in the agency **that are experienced** in the topic of analysis
- Checks the following:
  - Data used correctly and methodology
  - Findings don't contradict previous analysis
  - Deliverables presented at right level

# QA/QC | Getting Started

The Equity Toolkit Microsoft Teams Channel serves as the main collaboration space for all things ETK. Users can share research methodologies, team up on code, or solicit a QA/QC from their peers.

## How to join

Email [equitytoolkit@wmata.com](mailto:equitytoolkit@wmata.com) with your request. The team will authorize access, and you'll receive an automated email inviting you into the Team. You'll also be asked to fill out a short form outlining your areas of expertise and where you might be able to provide QA/QC support. You will also automatically be granted access to the data tables in Oracle.

## How to use

### Channels

General	Code Chat	QA / QC
General questions and discussion + sharing resources and ideas	Questions about conducting analysis when using the ETK	For QA/QC coordination and correspondence



See next few slides for more details of how you'll use this channel during your QA/QC

Once you have conducted the analysis, have a peer or your manager check your work. Are the queries capturing the data? If analysis is in Excel, do the formulas work? Does your peer/ manager accept the calculations and methodology? Does he or she accept the finding/ test for significance?

**Use case 1** | Laura Moeini performed the peer review. She did not find any errors in the calculations or methodology. She provided feedback to enhance the clarity of presenting key findings in the results briefing deck.

**Use case 2** | Xin Zhou performed the peer review of Use Case 2. She did not find any material errors with the use case but did recommend several ways to clarify the presentation. These changes were implemented, and the peer review was successfully finished.

**Use case 3** | [enter response here]



# QA/QC | Peer Review: Analyst Steps

**Before you share your analysis and findings outside of your department, all analysts must complete the following steps:**

**Step 1** – Join the Equity Toolkit Team on Microsoft Teams

**Step 2** – Post a request for a review in the QA/QC Channel on the Equity Tool Kit Page. Requests for review should include:

- A short summary of your project
- Prerequisite knowledge and skills needed from reviewer (i.e., knows Oracle SQL, has experience with ridership data)
  - Note: Try to find a reviewer who's knowledgeable enough about the work they don't have to start from zero, but who is far enough away from the work that they can be objective. You can look at the reviewer list and @mention specific reviewers who you think might be a good match for your analysis
- A timeline for the analysis and any important deadlines

**Step 3** – After you find a reviewer, arrange a short handoff meeting to provide the reviewer the following:

- A more detailed summary of your analysis and methodology
- Present initial findings
- Share locations of data, analysis, and deliverables
- Go over the [checklist](#)
- Provide any other information needed for review

**Step 4** – Email the peer reviewer with all files and documentation pertaining to this project including datasets, code or Excel files, final deliverables, and the checklist. **Be sure to cc [EquityToolkit@wmata.com](mailto:EquityToolkit@wmata.com)**

**Step 5** – After the review is complete, make all changes necessary to satisfy the checklist and provide updates to reviewer. The reviewer may also have other suggestions outside of the check list – these can be considered “nice to haves” and are not necessary to proceed. Once the all the checklist items are satisfied, the review will send you a notice of approval via email with a completed checklist – save this with your documentation. You may now proceed to the SME review.

# QA/QC | SME Review: Analyst Steps

## SME QA/QC process for analysts

1. After you have completed the peer QA/QC process, review the list on the right. Determine if your assessments meet any of the criteria
2. Reach out to the relevant contact(s) with your findings, methodology, and intended audience and cc [EquityToolkit@wmata.com](mailto:EquityToolkit@wmata.com). Coordinate with the SME to determine how they would like to proceed – only one SME in each section is required to review your assessment. Some SME's may only need minimal information for their review and others may want to dive deep.
3. Allow the individual a minimum of **5 business days** to review your analysis. They will provide you with feedback and requests for changes in that time frame if they are necessary.
4. Once all changes are incorporated and the SME has given their approval, your analysis is complete and ready to be shared with your desired audience
5. Send your final assessment to [EquityToolkit@wmata.com](mailto:EquityToolkit@wmata.com) so it can be included in the repository of equity analysis on our SharePoint site

Topics	It should go to this group for review:	Reach out to this Individual:
<b>Survey data</b>	Customer Research/ Planning	Kayleigh Campbell, Ray Yau, Matthew Zych
<b>Census data</b>	ETK QA/QC Teams Channel	Scott Traum, Diane Patterson
<b>Includes ridership or operational data (i.e., homebrew, Trace, etc.)</b>	Planning	Catherine Vanderwaart
<b>Is being presented to the Board, EMT, or public</b>	EWG	Allison Davis (Board + Public) Jordan Holt (EMT)

# QA/QC | SME Review: Sample Handoff Email

**Subj: Equity Toolkit SME Review – [Include Topic]**

Hello **[SME Name]**,

I am completing an equity analysis on **[topic of your report]** where I analyze **[provide a quick summary of your report]**. This report utilizes elements of the Equity Tool Kit, and part of the process is to solicit a review from a Subject Matter Expert (SME). The goal of this review is to ensure that all work stemming from the Tool Kit is aligned with previous work, to ensure that data is being used/interpreted correctly and deliverables are designed for the correct audience.

My analysis specifically [topic that you need reviewed (i.e., includes Census data, is going to the Board, etc.)] and the Equity Toolkit Committee has identified you as a SME for this topic. Would you have time to provide a review of this assessment in the next 5 business days?

As a reminder, the main objective of this review is to ensure that a SME is comfortable with the work that is produced so the review is largely left up to your discretion. There is no specific standard for a review, and you can provide whatever feedback you would like before approving an assessment; however your approval is required for work to be published.

If you would like to discuss further, I am available to meet at **[provide times]**, otherwise the relevant materials are attached for your review.

Please let me know if you have any questions, and I look forward to hearing back from you.

Best  
**[Your Name]**

# QA/QC | SME Review: SME Steps

## SME QA/QC process for Subject Matter Experts

1. If you are contacted by an analyst, review their findings and deliverables and determine the level of review that will be required for the assessment. Feel free to set up a meeting with the analyst to further discuss findings
2. Share any insight you have on the subject area and ask any questions you may have about analysis. Be sure to share any common pitfalls upfront so they can be addressed quickly
3. Review the analysis using the questions to the right as a guide. **The purpose of this review is to ensure all assessments can be reviewed by an expert. You may decide to delegate the review to a colleague or determine the assessment does not need a review.** In general, we would like the ensure all assessments have done the following:
  - The data is used and interpreted correctly
  - The findings are accurately represented and don't conflict with things that we have already presented
  - The audience is considered, and the findings are presented at the right level
4. Provide feedback to analyst within 5 business days and coordinate with analyst to ensure necessary changes are incorporated into deliverables

## Questions to Consider During SME Review

What is the research about?

What will be done with the assessment?

What changes is it hoping to inform?

What has been done before on this topic?

Do these results contradict any of our previous reports?

Who's the final audience and does the analysis have the right level of detail?

Who will have access to the assessment?

Is the data used correctly?

What is the methodology?

Has the analyst considered all the data limitations?

What is the sensitivity of that data and is it being shared properly?

# QA/QC | SME Review: Further Guidance

## **What does it mean to be a SME Reviewer?**

Your name would be listed as a point of contact next to your area of expertise. All analysts conducting analysis that includes survey data would be required to reach out to you and gain your approval before publishing their results. Analysts would provide you with findings, methodology, and deliverables for your review. When you are comfortable with the results being shared, you can provide your approval.

## **What should a SME be reviewing for?**

The main objective is to ensure the SMEs are comfortable with the work that is produced by the ETK, so the review is largely left up to your discretion. In general, we want to make sure that analyses are using the data correctly, findings are not contradicting work that you or someone else in the agency has produced, and presentations are geared toward the right audience. However, there is no specific standard for a review, and you can provide whatever feedback you would like before approving an assessment. Your approval is required for work to be published.

## **What level of review is required from a SME?**

Again, I want to stress that there is no specific standard for a review. In fact, you may read a description of an analysis and decide that you can delegate the task to a colleague or that the assessment requires no review at all. That is ok! Our goal is to create a process where a SME has an opportunity to review and can exercise discretion on a case-by-case basis.

# Please help us make this guide better!

Questions? Concerns? Edits? Please email the Equity Toolkit Team: [equitytoolkit@wmata.com](mailto:equitytoolkit@wmata.com)

We will republish updated versions on a regular basis (to be determined). We look forward to hearing your thoughts and edits—big or small—so that we can continuously improve this User Guide.

Thanks!  
The Equity Toolkit Team





