Agenda for Today

- Announcements
- MoE: from count data to proportions (and percentages)
- The LEHD program
- Lab 4: LEHD data

Announcements

Piazza

Signup Link:
piazza.com/berkeley/summer2020/cp101

Why Piazza?

Democratic

- all students have access to questions and answers (vs. sending an email to instructors)
- Other students can answer questions

Why Piazza?

Privacy

- Comply with Family Educational Rights and Privacy Act (FERPA) regulations
- email addresses and course enrollment information are only accessible to instructors of the class
- Posts remains anonymous to classmates
- Only instructors and enrolled students can access posts

https://piazza.com/legal/ferpa

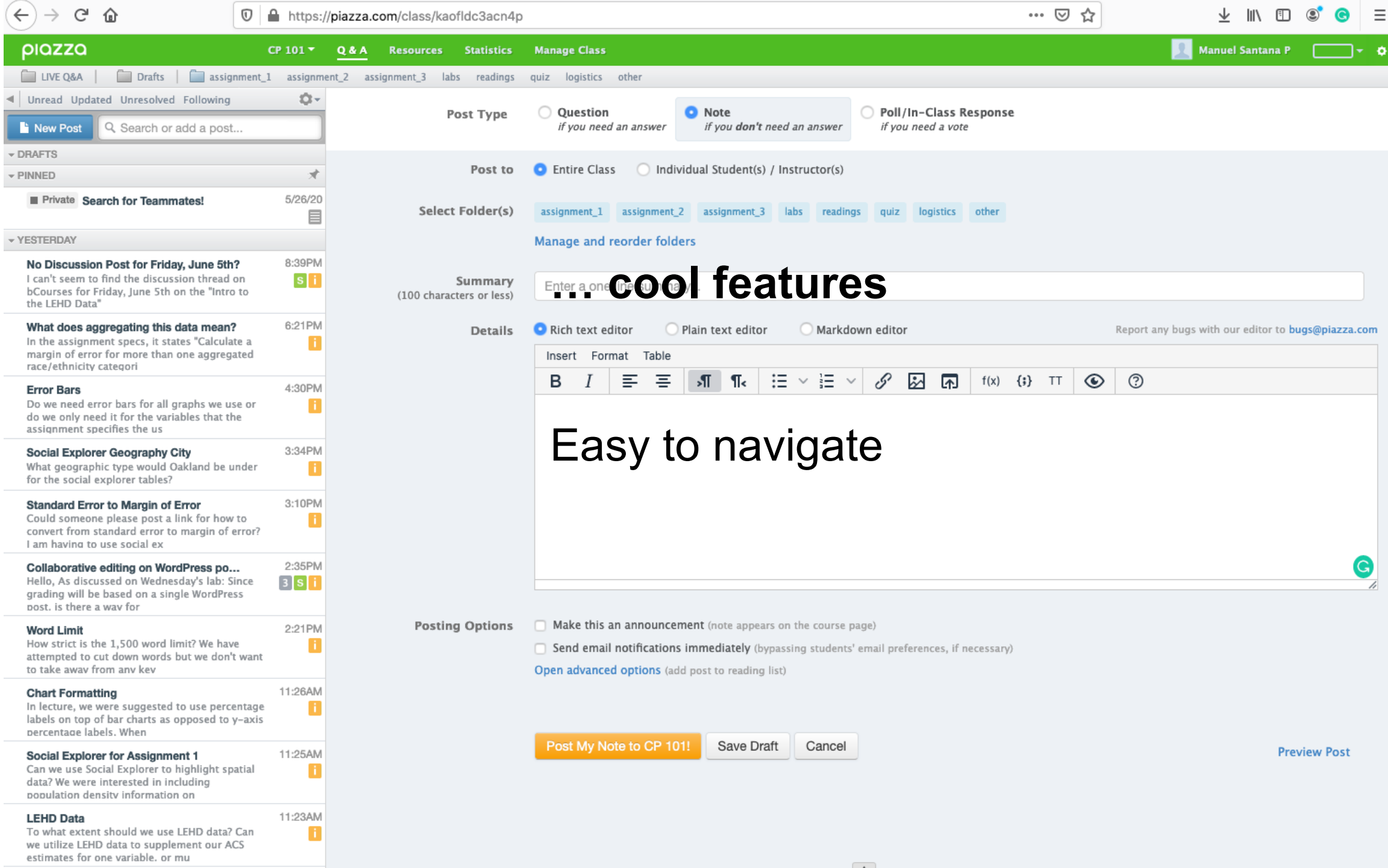
Why Piazza?

… instructors also benefit

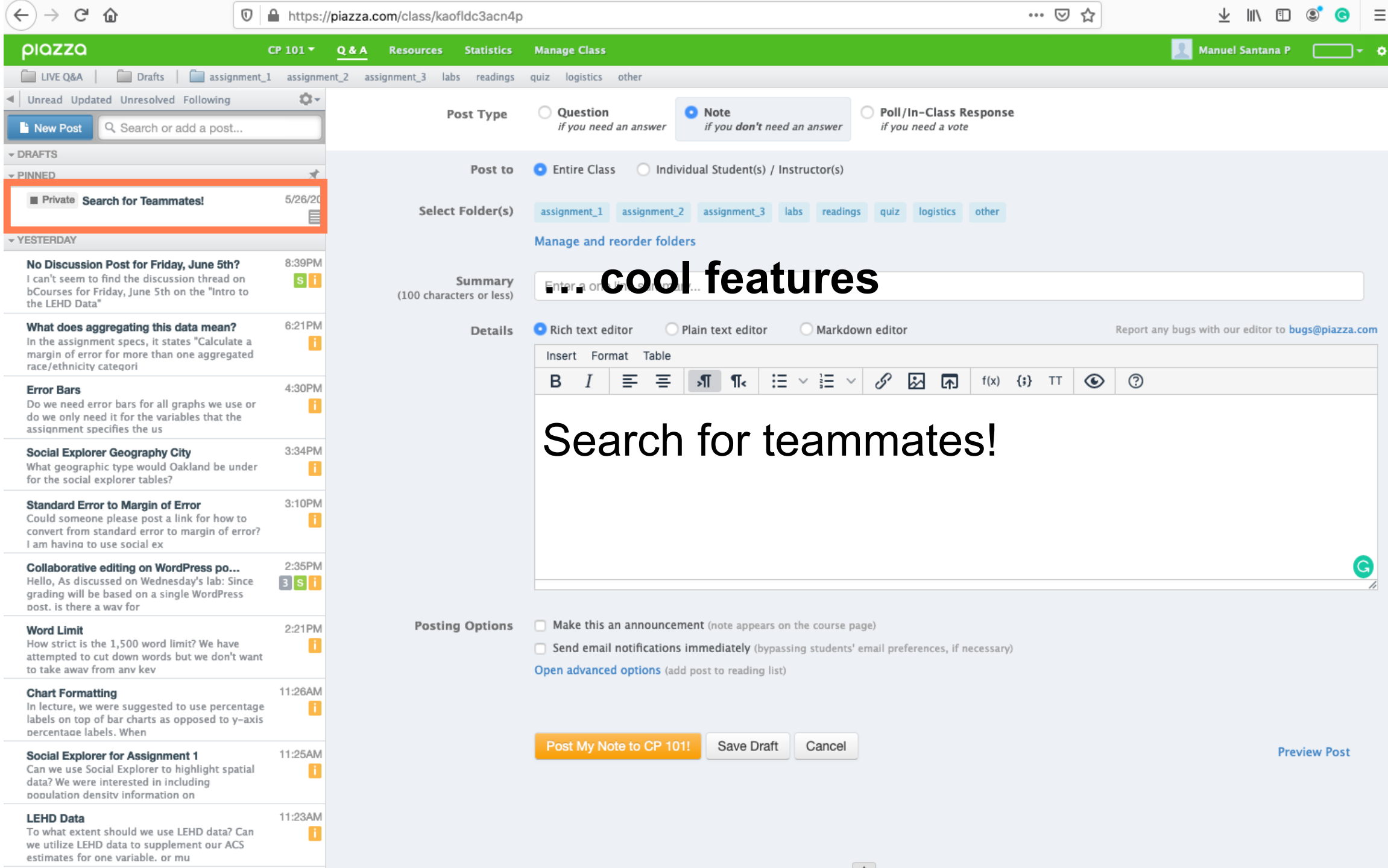
- Piazza helps instructors eliminate their over-crowded email inboxes

https://piazza.com/professors/show/ming_chow

Why Piazza?


https://piazza.com/class/kaofldc3acn4p
piazza
CP 101
Q & A
Resources
Statistics
Manage Class
Manuel Santana P
LIVE Q&A
Drafts
assignment_1
assignment_2
assignment_3
labs
readings
quiz
logistics
other
Unread
Updated
Unresolved
Following
New Post
Search or add a post...
DRAFTS
PINNED
Private
Search for Teammates!
5/26/20
YESTERDAY
No Discussion Post for Friday, June 5th?
I can't seem to find the discussion thread on bCourses for Friday, June 5th on the "Intro to the LEHD Data"
8:39PM
What does aggregating this data mean?
In the assignment specs, it states "Calculate a margin of error for more than one aggregated race/ethnicity categori
6:21PM
Error Bars
Do we need error bars for all graphs we use or do we only need it for the variables that the assignment specifies the us
4:30PM
Social Explorer Geography City
What geographic type would Oakland be under for the social explorer tables?
3:34PM
Standard Error to Margin of Error
Could someone please post a link for how to convert from standard error to margin of error? I am having to use social ex
3:10PM
Collaborative editing on WordPress po...
Hello, As discussed on Wednesday's lab: Since grading will be based on a single WordPress post, is there a way for
2:35PM
Word Limit
How strict is the 1,500 word limit? We have attempted to cut down words but we don't want to take away from any key
2:21PM
Chart Formatting
In lecture, we were suggested to use percentage labels on top of bar charts as opposed to y-axis percentage labels. When
11:26AM
Social Explorer for Assignment 1
Can we use Social Explorer to highlight spatial data? We were interested in including population density information on
11:25AM
LEHD Data
To what extent should we use LEHD data? Can we utilize LEHD data to supplement our ACS estimates for one variable, or mu
11:23AM
Post Type
Question
if you need an answer
Note
if you don't need an answer
Poll/In-Class Response
if you need a vote
Post to
Entire Class
Individual Student(s) / Instructor(s)
Select Folder(s)
assignment_1
assignment_2
assignment_3
labs
readings
quiz
logistics
other
Manage and reorder folders
Summary
(100 characters or less)
... cool features
Details
Rich text editor
Plain text editor
Markdown editor
Report any bugs with our editor to bugs@piazza.com
Insert
Format
Table
Easy to navigate
Posting Options
Make this an announcement (note appears on the course page)
Send email notifications immediately (bypassing students' email preferences, if necessary)
Open advanced options (add post to reading list)
Post My Note to CP 101!
Save Draft
Cancel
Preview Post

Why Piazza?


https://piazza.com/class/kaofldc3acn4p
piazza
CP 101
Q & A
Resources
Statistics
Manage Class
Manuel Santana P
LIVE Q&A
Drafts
assignment_1
assignment_2
assignment_3
labs
readings
quiz
logistics
other
Unread
Updated
Unresolved
Following
New Post
Search or add a post...
DRAFTS
PINNED
Private
Search for Teammates!
5/26/20
YESTERDAY
No Discussion Post for Friday, June 5th?
I can't seem to find the discussion thread on bCourses for Friday, June 5th on the "Intro to the LEHD Data"
8:39PM
What does aggregating this data mean?
In the assignment specs, it states "Calculate a margin of error for more than one aggregated race/ethnicity categori
6:21PM
Error Bars
Do we need error bars for all graphs we use or do we only need it for the variables that the assignment specifies the us
4:30PM
Social Explorer Geography City
What geographic type would Oakland be under for the social explorer tables?
3:34PM
Standard Error to Margin of Error
Could someone please post a link for how to convert from standard error to margin of error? I am having to use social ex
3:10PM
Collaborative editing on WordPress po...
Hello, As discussed on Wednesday's lab: Since grading will be based on a single WordPress post, is there a way for
2:35PM
Word Limit
How strict is the 1,500 word limit? We have attempted to cut down words but we don't want to take away from any key
2:21PM
Chart Formatting
In lecture, we were suggested to use percentage labels on top of bar charts as opposed to y-axis percentage labels. When
11:26AM
Social Explorer for Assignment 1
Can we use Social Explorer to highlight spatial data? We were interested in including population density information on
11:25AM
LEHD Data
To what extent should we use LEHD data? Can we utilize LEHD data to supplement our ACS estimates for one variable, or mu
11:23AM
Post Type
Question
if you need an answer
Note
if you don't need an answer
Poll/In-Class Response
if you need a vote
Post to
Entire Class
Individual Student(s) / Instructor(s)
Select Folder(s)
assignment_1
assignment_2
assignment_3
labs
readings
quiz
logistics
other
Manage and reorder folders
Summary
(100 characters or less)
...cool features
Details
Rich text editor
Plain text editor
Markdown editor
Report any bugs with our editor to bugs@piazza.com
Insert
Format
Table
Search for teammates!
Posting Options
Make this an announcement (note appears on the course page)
Send email notifications immediately (bypassing students' email preferences, if necessary)
Open advanced options (add post to reading list)
Post My Note to CP 101!
Save Draft
Cancel
Preview Post

… so please, enroll in piazza.com/berkeley/summer2020/cp101

https://piazza.com/professors/show/ming_chow

Reading Responses

Students are expected to respond to different sets of readings by submitting at least one question per session marked with an asterisk [*]

Assignment 1

- It is okay if you have more than 1,500 words. The word count **does not include references, citations, and sources**
- All data for one variable must come from the LEHD database – the only variable not found in the ACS is Commute Flows
- We will learn how to obtain commute flows data today during the lab
- WordPress does not allow for in-real-time collaboration

Assignment 1

- How to obtain MoE from Standard Errors?

$$\text{Margin of error} = MOE_\gamma = z_\gamma \times SE$$

Where z_γ denotes the *quantile* (also, commonly, a *z-score)*

The MoE reported by the Census are estimated with a 90 percent confidence level (z-score = 1.65)

Source: https://en.wikipedia.org/wiki/Margin_of_error

Assignment 1

- How to obtain MoE for percentages calculated from ACS count data?

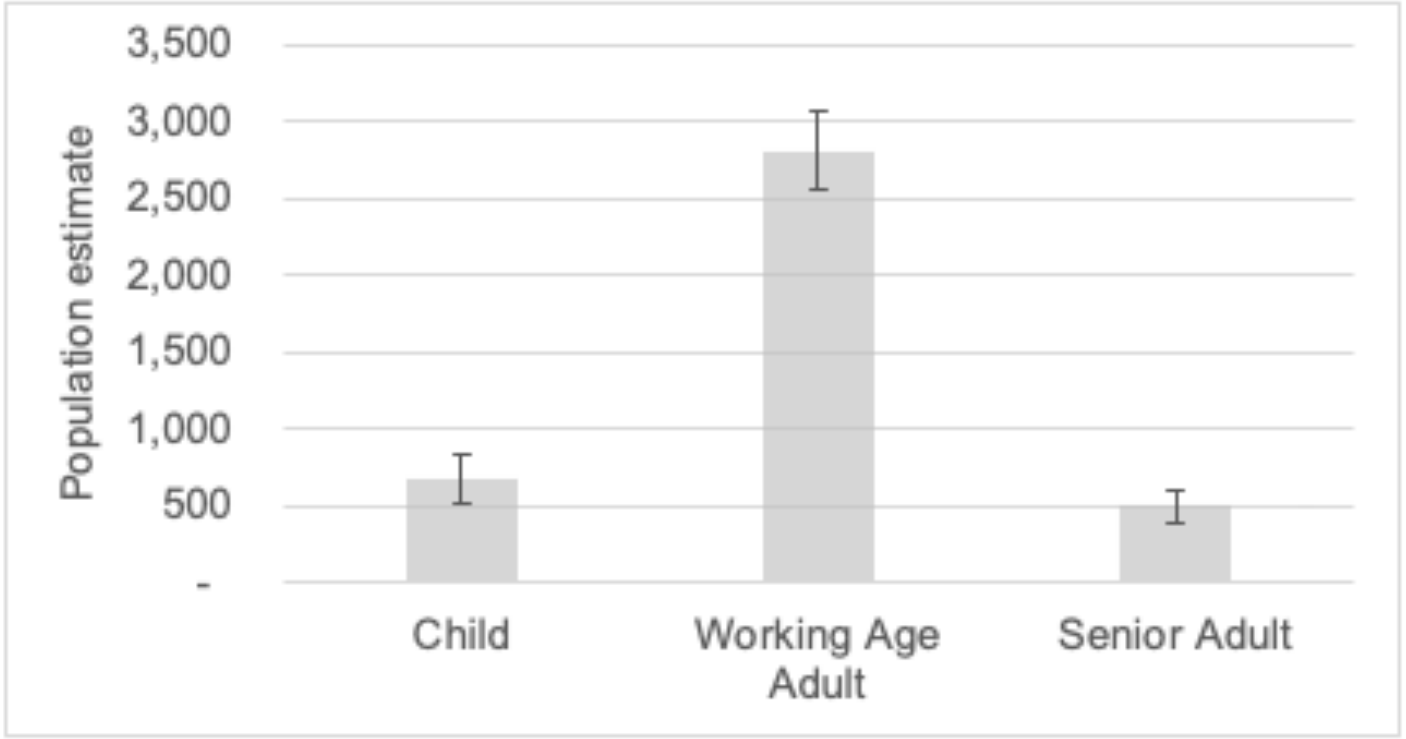
Table 1 - Population by Age Group (Tract 4004)

Age	Estimate	MoE
Child	678	161
Working Age Adult	2,810	247
Senior Adult	498	97
Total	**3,986**	**250**

Source: 2017 ACS 5-Year Estimates, Table S1810.

Note: Age summarized from Census table into three categories: "Child" (aged 17 and under), "Working Age Adult" (aged 18-64), and "Senior Adult" (aged 75 years and over).

Figure 1 - Population by Age Group (Tract 4004)



Source: 2017 ACS 5-Year Estimates, Table S1810.

Note: Age summarized from Census table into three categories: "Child" (aged 17 and under), "Working Age Adult" (aged 18-64), and "Senior Adult" (aged 75 years and over).

Understanding and Using American Community Survey Data

What All Data Users Need to Know

Issued July 2018

If we define the proportion as $\widehat{P} = \widehat{X}/\widehat{Y}$, then the MOE of this proportion is approximated as:

$$MOE(\widehat{P}) = \frac{1}{\widehat{Y}}\sqrt{[MOE(\widehat{X})]^2 - \left(\widehat{P}^2 * [MOE(\widehat{Y})]^2\right)}$$

Chapter 8, pg. 55

If we define the proportion as $\widehat{P} = \widehat{X}/\widehat{Y}$, then the MOE of this proportion is approximated as:

$$MOE(\widehat{P}) = \frac{1}{\widehat{Y}}\sqrt{[MOE(\widehat{X})]^2 - \left(\widehat{P}^2 * [MOE(\widehat{Y})]^2\right)}$$

$\hat{P}$ is the proportion of population under classified as children

Table 1 - Population by Age Group (Tract 4004)

Age	Estimate		MoE
Child	$\hat{X}$	678	161
Working Age Adult		2,810	247
Senior Adult		498	97
Total	$\hat{Y}$	**3,986**	**250**

If we define the proportion as $\widehat{P} = \widehat{X}/\widehat{Y}$, then the MOE of this proportion is approximated as:

$$MOE(\widehat{P}) = \frac{1}{\widehat{Y}}\sqrt{[MOE(\widehat{X})]^2 - \left(\widehat{P}^2 * [MOE(\widehat{Y})]^2\right)}$$

$\hat{P}$ is the proportion of population under classified as children

Table 1 - Population by Age Group (Tract 4004)

Age	Estimate		MoE	
Child	$\hat{X}$	678	161	$MoE(\hat{X})$
Working Age Adult		2,810	247	
Senior Adult		498	97	
Total	$\hat{Y}$	**3,986**	**250**	$MoE(\hat{Y})$

If we define the proportion as $\widehat{P} = \widehat{X}/\widehat{Y}$, then the MOE of this proportion is approximated as:

$$\mathrm{MOE}(\widehat{P}) = \frac{1}{\widehat{Y}}\sqrt{[\mathrm{MOE}(\widehat{X})]^2 - \left(\widehat{P}^2 * [\mathrm{MOE}(\widehat{Y})]^2\right)}$$

Table 1 - Population by Age Group (Tract 4004)

Age	Estimate	MoE	$\hat{P}$	$MoE(\hat{P})$
Child	678	161	17.0%	3.9%
Working Age Adult	2,810	247		
Senior Adult	498	97		
Total	**3,986**	**250**		

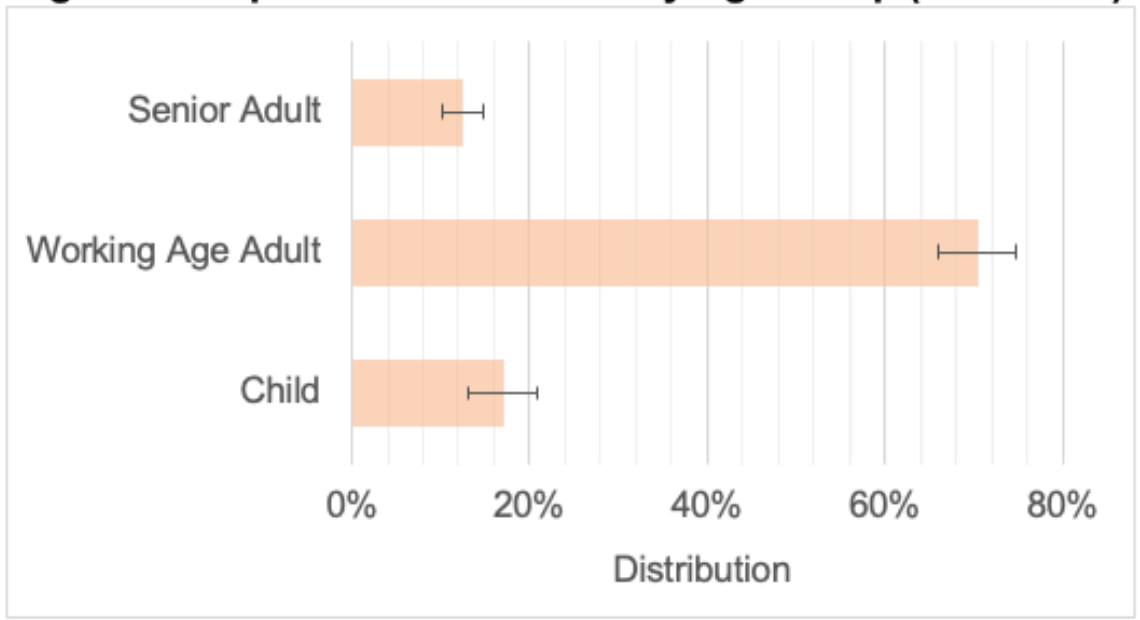
If we define the proportion as $\widehat{P} = \widehat{X}/\widehat{Y}$, then the MOE of this proportion is approximated as:

$$\mathrm{MOE}(\widehat{P}) = \frac{1}{\widehat{Y}}\sqrt{[\mathrm{MOE}(\widehat{X})]^2 - \left(\widehat{P}^2 * [\mathrm{MOE}(\widehat{Y})]^2\right)}$$

Table 1 - Population by Age Group (Tract 4004)

Age	Estimate	MoE	$\hat{P}$	$MoE(\hat{P})$
Child	678	161	17.0%	3.9%
Working Age Adult	2,810	247	70.5%	4.4%
Senior Adult	498	97	12.5%	2.3%
Total	**3,986**	**250**		

Figure 2 - Population Distribution by Age Group (Tract 4004)



Source: 2017 ACS 5-Year Estimates, Table S1810.

Notes: [a] **Base = 3,986**; [b] Age summarized from Census table into three categories: "Child" (aged 17 and under), "Working Age Adult" (aged 18-64), and "Senior Adult" (aged 75 years and over).

Table 2 - Population Distribution by Age Group (Tract 4004)

Age	Estimate	MoE
Child	17.0%	3.9%
Working Age Adult	70.5%	4.4%
Senior Adult	12.5%	2.3%

Source: 2017 ACS 5-Year Estimates, Table S1810.

Notes: [a] **Base = 3,986**; [b] Age summarized from Census table into three categories: "Child" (aged 17 and under), "Working Age Adult" (aged 18-64), and "Senior Adult" (aged 75 years and over).

The Longitudinal Employer-Households Dynamics Program

Berkeley
UNIVERSITY OF CALIFORNIA

Introduction to Urban Data Analytics

Manuel Santana Palacios

June 05, 2020

LEHD

Longitudinal Employer-Household Dynamics

Applications

- J2J Explorer
- LED Extraction Tool
- OnTheMap
- OnTheMap for Emergency Management
- PSEO Explorer
- QWI Explorer
- VEO Explorer

Useful Links

- Center for Economic Studies
- J2J Data
- LODES Data
- PSEO Data
- QWI Data
- VEO Data
- LED Workshop

Contact Information

Email us:

General

LODES/OnTheMap

QWI/QWI Explorer

J2J/J2J Explorer

or

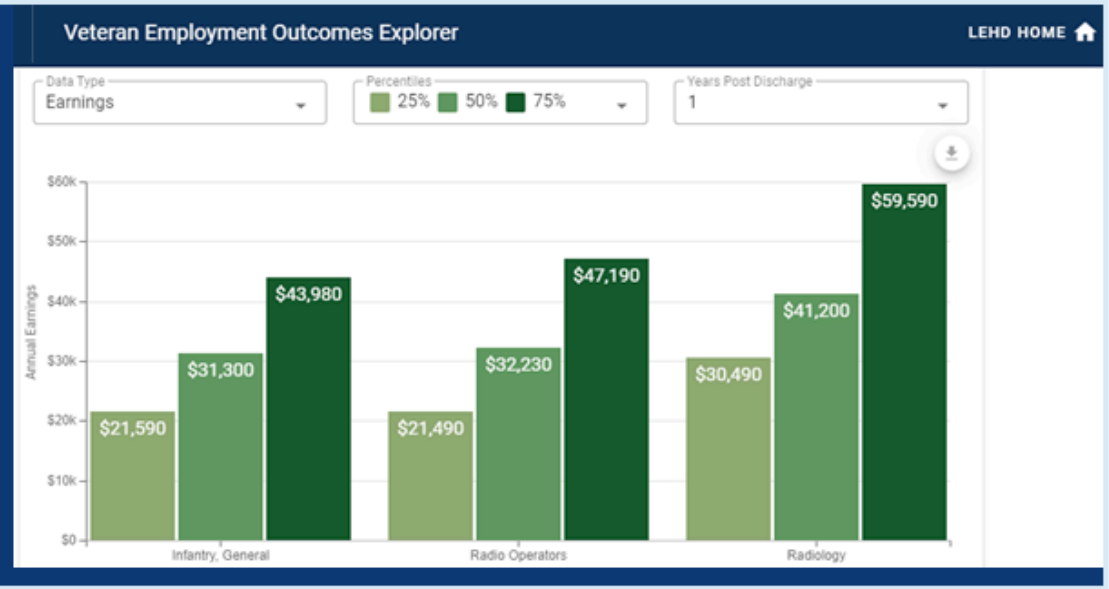
Veteran Outcomes Experimental Data Released

The U.S. Census Bureau announces the release of experimental Veteran Employment Outcomes (VEO) statistics. These tabulations show earnings and employment outcomes for U.S. Army veterans one, five, and 10 years after discharge, by military occupation, rank, demographics, industry and geography of employment. VEO data can be accessed via VEO Explorer, an interactive data tool.

View VEO Data

Start VEO Explorer

< 1 2 3 4 5 6 >



What's New?

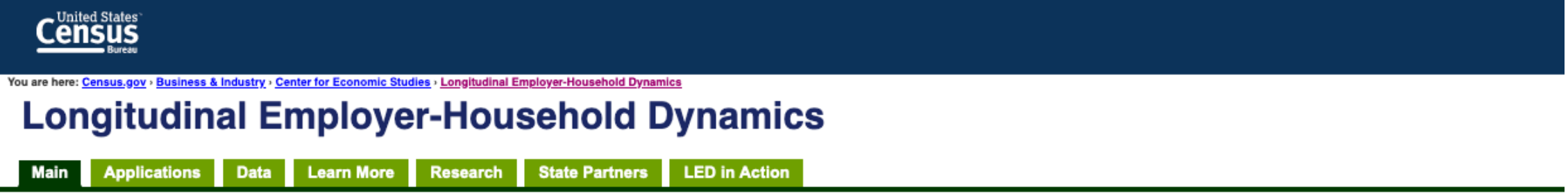
- 02/12/20: New Secure Electronic File Transfer Process for Local Employment Dynamics (LED) Partner File Submissions

View all announcements

About Us

The Longitudinal Employer-Household Dynamics (LEHD) program is part of the Center for Economic Studies at the U.S. Census Bureau. The LEHD program produces new, cost effective, public-use information combining federal, state and Census Bureau data on employers and employees under the Local Employment Dynamics (LED) Partnership. State and local authorities increasingly need detailed local information about their economies to make informed decisions. The LED Partnership works to fill critical data gaps and provide indicators needed by state and local authorities.

Under the LED Partnership, states agree to share Unemployment Insurance earnings data and the Quarterly Census of Employment and Wages (QCEW) data with the Census Bureau. The LEHD program combines these administrative data, additional administrative data and data from censuses and surveys. From these data, the program creates statistics on employment, earnings, and job flows at detailed levels of geography and industry and for different demographic groups. In addition, the LEHD program uses these data to create partially synthetic data on workers' residential patterns.



The LEHD program is **part of the Center for Economic Studies at the U.S. Census Bureau**. The program produces new, cost effective, **public-use** information combining **federal, state and Census Bureau data** on **employers** **and employees** under the Local Employment Dynamics (LED) Partnership. State and local authorities **increasingly need detailed local information about their economies to make informed decisions**.

Note

- The data released by LEHD are based on tabulated and modeled administrative data, which are subject to error
- Because the **estimates are not derived from a probability-based sample**, no sampling error measures are applicable.

Questions that can be answered with LEHD

- How has the distribution of women/men, race/ethnicity changed in certain employment sectors over time?

Questions that can be answered with LEHD

- How has the distribution of women/men, race/ethnicity changed in certain employment sectors over time?
- Where do young workers who work in San Francisco live?

Questions that can be answered with LEHD

- How has the distribution of women/men, race/ethnicity changed in certain employment sectors over time?
- Where do young workers who work in San Francisco live?
- What percentage of high-wage workers live in San Francisco but commute to San Jose (e.g. by Google bus)?

Questions that can be answered with LEHD

- How has the distribution of women/men, race/ethnicity changed in certain employment sectors over time?
- Where do young workers who work in San Francisco live?
- What percentage of high-wage workers live in San Francisco but commute to San Jose (e.g. by Google bus)?
- Where are the workplace destinations for workers living in a particular community or neighborhood?

Questions that can be answered with LEHD

- How has the distribution of women/men, race/ethnicity changed in certain employment sectors over time?
- Where do young workers who work in San Francisco live?
- What percentage of high-wage workers live in San Francisco but commute to San Jose (e.g. by Google bus)?
- Where are the workplace destinations for workers living in a particular community or neighborhood?
- How many jobs are there along the Oakland BRT line?
- How have average commute time in Oakland change over

Questions that can be answered with LEHD

- How has the distribution of women/men, race/ethnicity changed in certain employment sectors over time?
- Where do young workers who work in San Francisco live?
- What percentage of high-wage workers live in San Francisco but commute to San Jose (e.g. by Google bus)?
- Where are the workplace destinations for workers living in a particular community or neighborhood?
- How many jobs are there along the Oakland BRT line?
- How have average commute time in Oakland change over

Questions that can be answered with LEHD

- How has the distribution of women/men, race/ethnicity changed in certain employment sectors over time?
- Where do young workers who work in San Francisco live?
- What percentage of high-wage workers live in San Francisco but commute to San Jose (e.g. by Google bus)?
- Where are the workplace destinations for workers living in a particular community or neighborhood?
- How many jobs are there along the Oakland BRT line?
- How have average commute time in Oakland change over
- Who are the super-commuters in the Bay Area?

The Mercury News

Bay Area super-commuting growing: Here's where it's the worst

Alameda, Contra Costa counties continue to lead the way in the worst Bay Area commutes

Hopfenbeck, a Bay Area native, is one of an increasing number of super-commuters —people whose journey to and from work takes 90 minutes or more in each direction …

https://www.mercurynews.com/2019/09/11/supercommuting-is-not-just-for-central-valley-dwellers-map-shows-growth-in-bay-area-commutes/

United States Census Bureau

You are here: Census.gov › Business & Industry › Center for Economic Studies › Longitudinal Employer-Household Dynamics

Longitudinal Employer-Household Dynamics

Main | Applications | Data | Learn More | Research | State Partners | LED in Action

Applications

- J2J Explorer
- LED Extraction Tool
- OnTheMap
- OnTheMap for Emergency Management
- PSEO Explorer
- QWI Explorer
- VEO Explorer

Veteran Outcomes Experimental Data Released

The U.S. Census Bureau announces the release of experimental Veteran Employment Outcomes (VEO) statistics. These tabulations show earnings and employment outcomes for U.S. Army veterans one, five, and 10 years after discharge, by military occupation, rank, demographics, industry and geography of employment. VEO data can be accessed via VEO Explorer, an interactive data tool.

View VEO Data

Start VEO Explorer

Veteran Employment Outcom

Earnings

$43,050

$31,350

OnTheMap

Start | Base Map | Selection

Save | Load | Feedback | Previous Extent | Hide Tabs

Welcome to OnTheMap!

Start an analysis by using one of the tools below (Search, Import Geography, or Load .OTM file). Hover over the Help icons located throughout the application to see Help tips for using specific functionality. Sections in the control panel can be collapsed or opened by clicking the section title

2016 and 2017 data now available!
Read New Data Notice (08/29/2019)

Search

Search

Search All Names

Import Geography

Import from KML
Import from SHP
Import from GPS

Load .OTM File

Click the "Load" button below to load a .OTM file.

Load

OnTheMap

Worker

Home and Job

Locations

from state payroll tax records (unemployment insurance)

Household

Home and Job

Locations

from US Census Bureau

OnTheMap

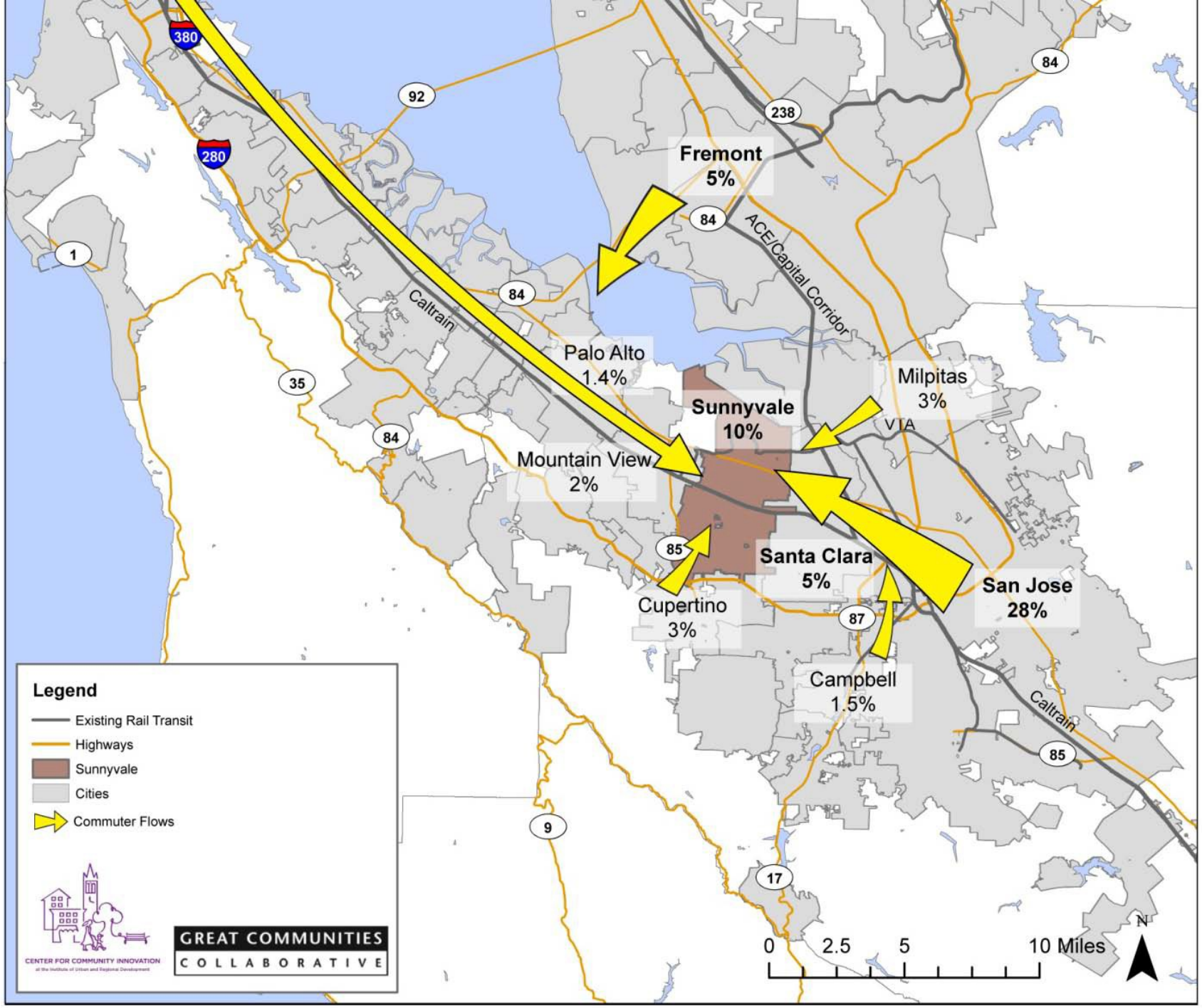
Commute Flows

- Age
- Income
- Industry

Do we need housing or office space near the Sunnyvale Caltrain stations?

Where do Sunnyvale workers live?

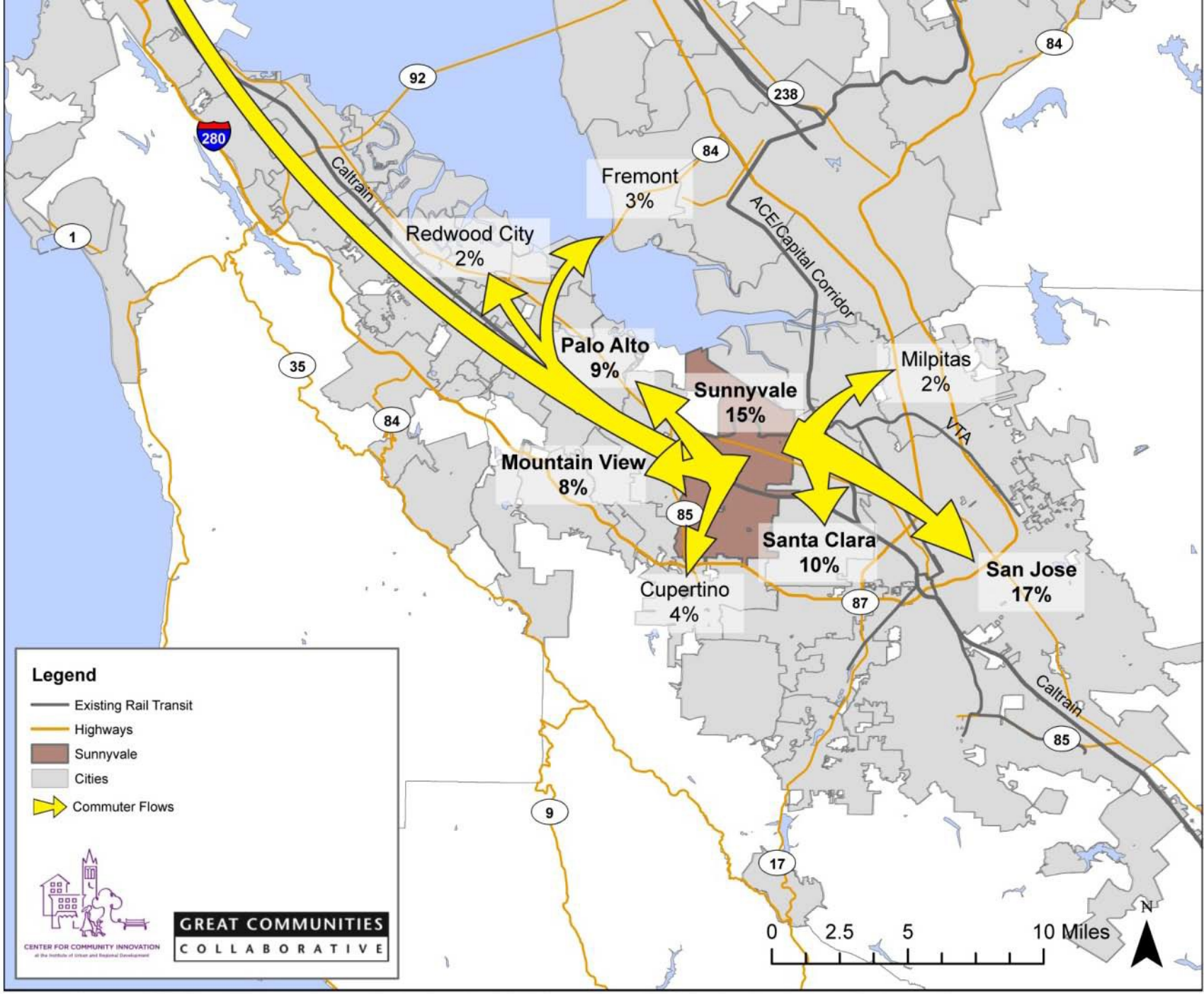
Do we need housing or office space near the Sunnyvale Caltrain stations?



Sources: U.S. Census, TIGER/Line, 2008; National Highway Planning Network, 2008; US Census Bureau, LED OnTheMap Origin-Destination Database, 2008.

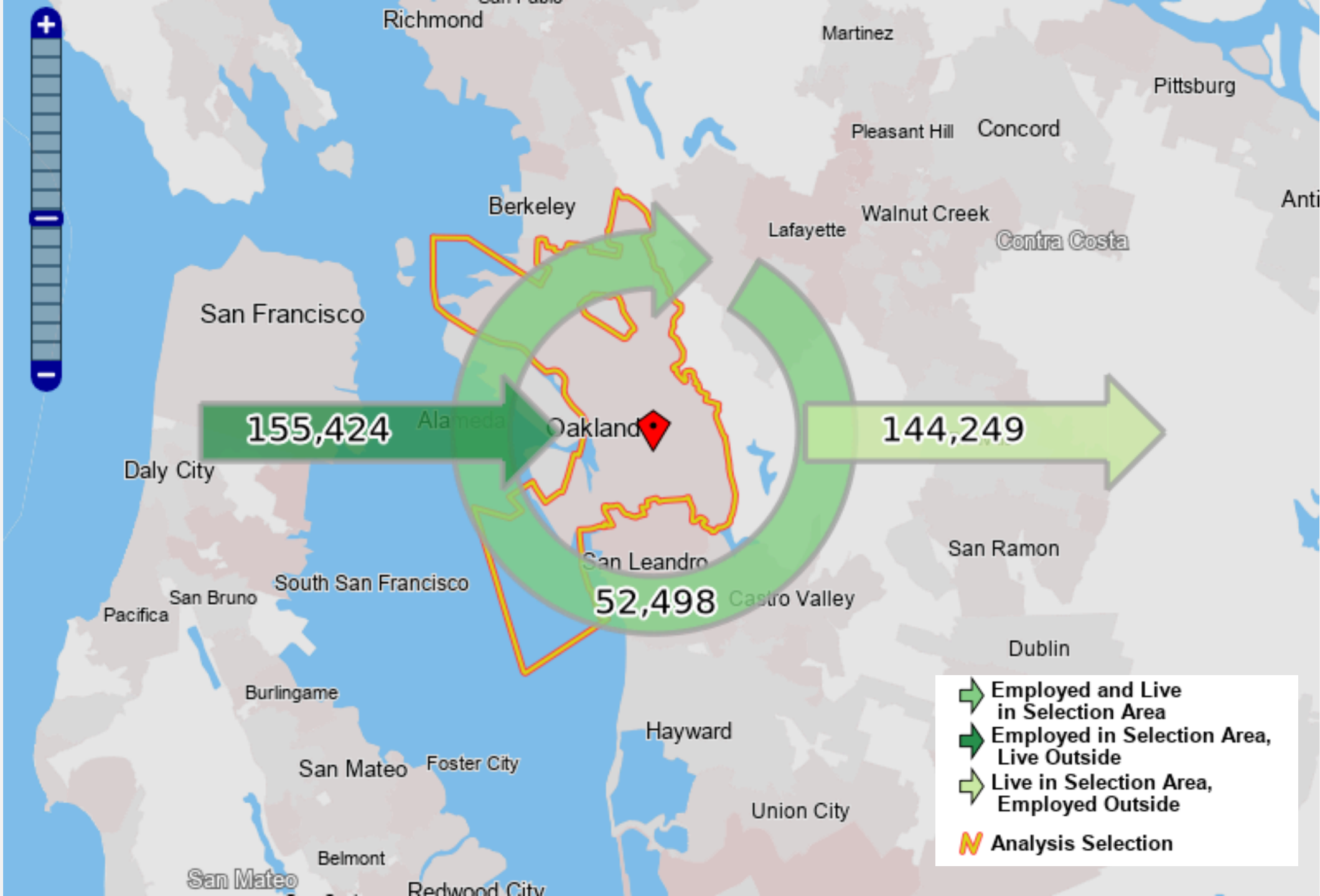
Where do Sunnyvale residents work?

Do we need housing or office space near the Sunnyvale Caltrain stations?



Sources: U.S. Census, TIGER/Line, 2008; National Highway Planning Network, 2008; US Census Bureau, LED OnTheMap Origin-Destination Database, 2008.

Oakland, CA
Commute Flows



Oakland, CA
Commute Flows

Inflow/Outflow Job Counts (All Jobs)

	2017	
	Count	Share
Employed in the Selection Area	207,922	100.0%
Employed in the Selection Area but Living Outside	155,424	74.8%
Employed and Living in the Selection Area	52,498	25.2%
Living in the Selection Area	196,747	100.0%
Living in the Selection Area but Employed Outside	144,249	73.3%
Living and Employed in the Selection Area	52,498	26.7%

OnTheMap Glossary - Report Terms:

Click here for the glossary of Mapping Terms

Note: For job definitions not otherwise specified, the reference period is the 2nd quarter of each calendar year (April-June). Throughout OnTheMap - unless otherwise stated - jobs are defined as 'Beginning of Quarter Employment'. Beginning of Quarter Employment is the total number of workers who were employed by the same employer in both the current (2nd) and previous (1st) quarter.

Total All Jobs A count of all jobs that meet a user's specification of geography, years, and/or labor market segments.

Total Primary Jobs A count of primary jobs that meet a user's specification of geography, years, and/or labor market segments. A primary job is the highest paying job for an individual worker for the year. The count of primary jobs is the same as the count of workers.

Total Private Jobs A count of private-sector jobs that meet a user's specification of geography, years, and/or labor market segments.

Total Private Primary Jobs A count of private-sector primary jobs that meet a user's specification of geography, years, and/or labor market segments. A primary job is the highest paying job (either public or private) for an individual worker for the year.

Jobs by Worker Age The categories are Age 30 or younger, Age 31 to 54 years, and Age 55 or older.

Jobs by Earnings The categories are $1,250 per month or less, $1,251 to $3,333 per month, and More than $3,333 per month.

Jobs by Industry Type (2-digit NAICS) The categories are all 2-digit NAICS codes, otherwise known as industry sectors. See the U.S. Census Bureau's documentation on NAICS for more information on these industry sectors.

Report Settings A section at the bottom of each report summarizing the settings and selection area(s) chosen in the Analysis Settings popup.

OnTheMap Glossary - Report Terms:

Click here for the glossary of Mapping Terms

Note: For job definitions not otherwise specified, the reference period is the 2nd quarter of each calendar year (April-June). Throughout OnTheMap - unless otherwise stated - jobs are defined as 'Beginning of Quarter Employment'. Beginning of Quarter Employment is the total number of workers who were employed by the same employer in both the current (2nd) and previous (1st) quarter.

Total All Jobs A count of all jobs that meet a user's specification of geography, years, and/or labor market segments.

Total Primary Jobs A count of primary jobs that meet a user's specification of geography, years, and/or labor market segments. A primary job is the highest paying job for an individual worker for the year. The count of primary jobs is the same as the count of workers.

Total Private Jobs A count of private-sector jobs that meet a user's specification of geography, years, and/or labor market segments.

Total Private Primary Jobs A count of private-sector primary jobs that meet a user's specification of geography, years, and/or labor market segments. A primary job is the highest paying job (either public or private) for an individual worker for the year.

Jobs by Worker Age The categories are Age 30 or younger, Age 31 to 54 years, and Age 55 or older.

Jobs by Earnings The categories are $1,250 per month or less, $1,251 to $3,333 per month, and More than $3,333 per month.

Jobs by Industry Type (2-digit NAICS) The categories are all 2-digit NAICS codes, otherwise known as industry sectors. See the U.S. Census Bureau's documentation on NAICS for more information on these industry sectors.

Report Settings A section at the bottom of each report summarizing the settings and selection area(s) chosen in the Analysis Settings popup.

Lab 4
LEHD OnTheMap web