

Data Wrangling for Coworking Space Market Feasibility Project

What data do we need?

We are attempting to put together a demographic sample of citizens of our market areas, South Coast Santa Barbara and pertinent Ventura County areas, in order to gauge if demographics and information about the area show evidence of a strong likelihood /potential for locals to want to join our coworking space as members. However, glancing at the data dictionary for the PUMS data provided by the US Government, we see that there are 285 variables entered about these folks - which I doubt we will even come close to needing all of this information.

2016 Global Coworking Survey - What statistics are we searching for?

As a result of this, instead we will look to another important piece of this puzzle - the 2016 Global Coworking Survey Demographic. Please note that this is data extrapolated only for the United States, which is more helpful to us compared to global information.

Distance to home

Starting on this page at slide 47, we see a list of the top ten reasons why folks choose a coworking space. Most of these things are factors that we have to take control of in terms of our environment, but there is one that translates toward a demographic market feasibility: A close distance to home.

Continuing on to slide 57, we see the breakdown of how folks get to their coworking spaces:

- Driving: 51%
- Cycling: 22%
- Walking: 19%
- Public Transit: 8%

In addition, we see that in an area with the market size we're looking at, 81% of respondents will arrive to their coworking space within 20 minutes of departing their house.

Age of Member

Moving on to slide 49, we see the age breakdown of participants of the survey which is clearly important:

- Age 18-29: 28%
- Age 30-39: 39%
- Age 40-49: 19%
- Age 50-59: 13%
- Age 60+: 1%

Job Status

On slide 51 we can see that job status could play a role in who the makeup of members are. According to survey results (as referenced on slide 76), as an owner of a coworking space you want to prioritize the acceptance of freelancers and entrepreneurs as members. The reasoning for this is that many folks join these

spaces for networking and collaboration - people want to use the space to find contracts, job opportunities, or to see if there are any folks around who want to collaborate on a project with them (I must admit - sounds pretty exciting!). We see the following:

- Employee: 51%
- Entrepreneur: 12%
- Freelancer: 32%
- Other: 5%

Moving over to slide 55, we further see the following breakdown of Job type:

- IT (Software Engineer, Web Developer): 27%
- Consulting: 15%
- PR, Marketing, Sales, Advertising: 5%
- Design: 7%
- Project Management: 4%
- Research: 4%
- Writing: 5%
- Business Development (includes entrepreneurs): 3%
- Education: 5%
- Higher Management: 5%
- Art: 3%
- Other: 16%

Gender

As reported by coworking spaces, the average share of female members sits at 38%; as reported by the individual member respondents, 41% of them were female. Let's sit at a happy medium of 40%.

Relationship Status

In the US, relationship status of survey respondents were as follows:

- Single: 19%
- In a relationship, unmarried: 22%
- Married: 47%
- Separated, Divorced, Widowed: 7%
- Other or NA: 5%

In addition to marital status, The 2017 Global Coworking Survey contains global information regarding number of children of respondents:

- 0 children: 64%
- 1 child: 11%
- 2 children: 18%
- 3 children: 7%

Health Insurance of members

In the US, the following breakdown of health insurance per respondent was as follows:

- Enhanced Health Insurance: 35%
- Basic Health Insurance: 52%
- No Health Insurance: 9%
- Other: 1%

- NA: 3%

Last place of work - before starting coworking

In the US, the following breakdown illustrates where respondents worked before deciding to go to a coworking space:

- Home Office: 44%
- Traditional Office: 37%
- Coffee Shop: 11%
- Small Shared Office Community: 2%
- No Fixed Location: 6%

Highest Level of School Education

As seen in The 2017 Global Coworking Survey, we also have information on education levels of survey respondents on slide 5:

- Doctoral or higher: 4%
- Master: 41%
- Bachelor: 41%
- High School: 10%
- No Education: 1%
- NA: 3%

Relative Income

The survey contains information regarding how the respondent gauges their own level of income vs. their cost of living. I really wish this would be represented better, as straight income data, but we'll see what we can do with this later on. As seen on slide 60, respondents rated their income compared to cost of living as:

- Very High: 5%
- Rather High: 34%
- Somewhere in the Middle: 45%
- Rather Low: 9%
- Very Low: 1%
- NA: 6%

What data do we need?

In summary, we need to be on the lookout for the following in our PUMS Data:

- Does the person Drive, Bike, Walk, or Transit as their commute?
- How long would that person's mode of transportation take to get to our coworking space?
- What is the age of the person?
- What is the person's job type?
- What field does the person work in?
- What is the person's gender?
- What is the person's marital status?
- How many children does the person have?
- What type of health insurance does the person have, if any?
- What is the person's current workspace?
- What is that person's highest level of education?

How to get this data

We can get a sample of data from our market areas by going to the PUMS Data Website and following these steps:

- In the “topic name” search bar, type “PUMS” and hit search. Wait for your search to load.
- Click “2016 ACS 1-year Public Use Microdata Samples (PUMS) - CSV format”. You should be directed to a new page.
- Select “California Population Records”. Your browser should begin downloading a CSV file. It should be very large. Save this to the folder “Starting Point Data Files” in the repository for this R Project. It is recommended that you delete this file once you have run the script below due to its very large size.
- Navigate to the list of PUMA codes on the US Census website and search for the PUMA codes you’re looking for. In my case, I need Santa Barbara South Coast (8303), Ventura City (11104), Oxnard/Port Hueneme (11103), and Moorpark area for good measure (11106).
- Run the R script below in order to get only the market areas you want from this file. This script will also save your filtered result in a file that will take much less time to read upon subsequent analyses, as long as you have begun this activity from the designated R Project instead of just this file alone.

```
Census <- read.csv("Starting Point Data Files/ss16pca.csv")
MarketCensus <- Census[Census$PUMA == 08303 | Census$PUMA == 11104 | Census$PUMA == 11103 | Census$PUMA == 11106, ]
write.csv(MarketCensus, file = "PUMS Data SB-Ventura County Market Areas.csv")
```