

A Long Look at Love

(San Francisco Edition)

Rodrigo Arguello

Project Overview

The goal of this project is to get to know the OKCupid population of San Francisco to decide whether it is worth using the app for dating there. To do so, this project will find out the kinds of people that populate the scene as well as whether they reflect the population around them.



Data Overview

OKCupid data for 59,946 users in
San Francisco, June 2012.

Dataset cleaned from the OkCupid Profile Data for
Introductory Statistics and Data Science Courses (Journal
of Statistics Education 2015).

The dataset contains typical user data, lifestyle
habits, declared income and education, and answers
to 10 essay questions.



Data Overview



Sex	Age	Body Type	Diet	Drinks	Drugs	Ethnicity	Income	Religion	Orientation
Male / Female	18+	Average, fit, thin, etc.	Vegan, eat anything, etc.	Never, sometimes, socially, etc.	Never, sometimes, often	White, asian, black, etc.	0+	Christianity, atheism, hinduism, other, etc.	Straight, gay, bisexual

The analysis will focus on the users' physical attributes, several habits, ethnicity, income, religion, and orientation.

Questions

First few questions reveal:

- Distribution of age and gender
- Most common physical attributes

Next few questions reveal:

- Drug, drinking, and diet habits of users

Last few questions show:

- How OKCupid users compare to the average San Francisco resident or, in the absence of that information, American

What kinds of people populate the San Francisco dating scene?

— — —

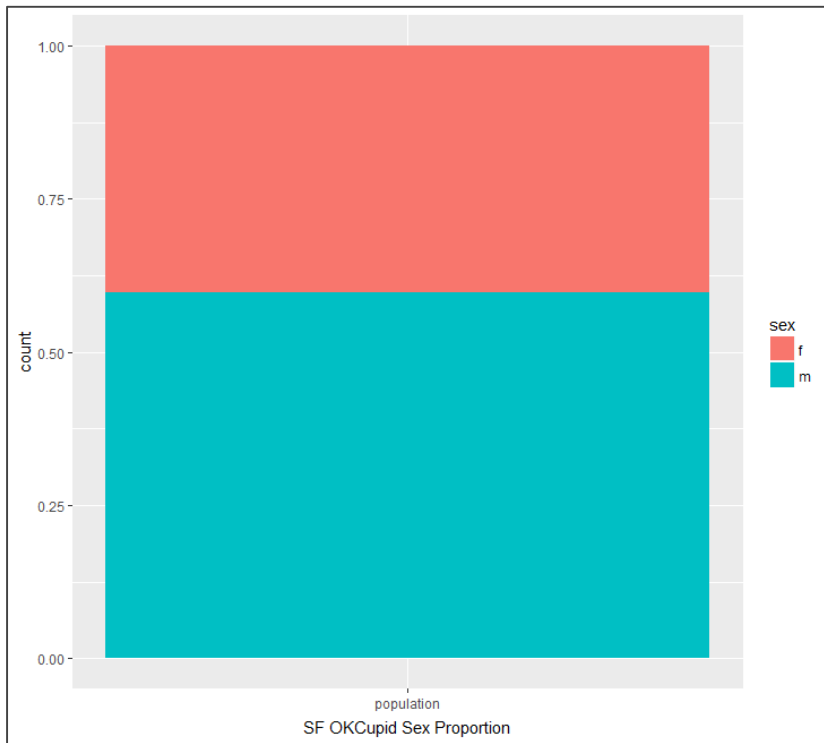
For both genders...

- What's the main age range?
- How fit are users?
- Which religions are the most common?
- Which ethnicities are most prevalent?
- Is there a high proportion of LGBTQ users?

Answering these questions will shed some light on the most common traits OKCupid users see. This will help give us an idea of the kinds of people we'd expect to find in SF OKCupid.

What kinds of people populate the San Francisco dating scene?

-Sex



Sex selected.

X axis is a single category.

Y axis shows proportion as fraction of 1.

Color shows Sex (M/F)

- More Males than Females
(6:4)

What kinds of people populate the San Francisco dating scene?

-Age and Sex

— — —



Age and Sex selected.

X axis shows Age by single years.

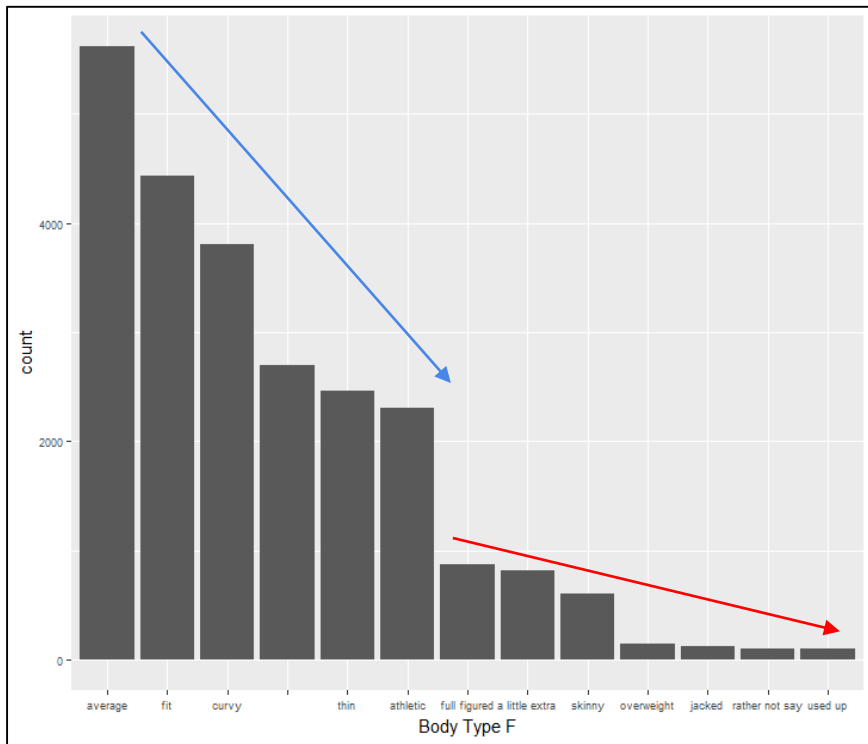
Y axis shows frequency for each year.

Color shows Sex (M/F)

- Most users in the 22-33 age range
- Most common age is 26

What kinds of people populate the San Francisco dating scene?

-Body Type by Sex = Female



Body Type and Sex selected.

Filtered Sex by Female.

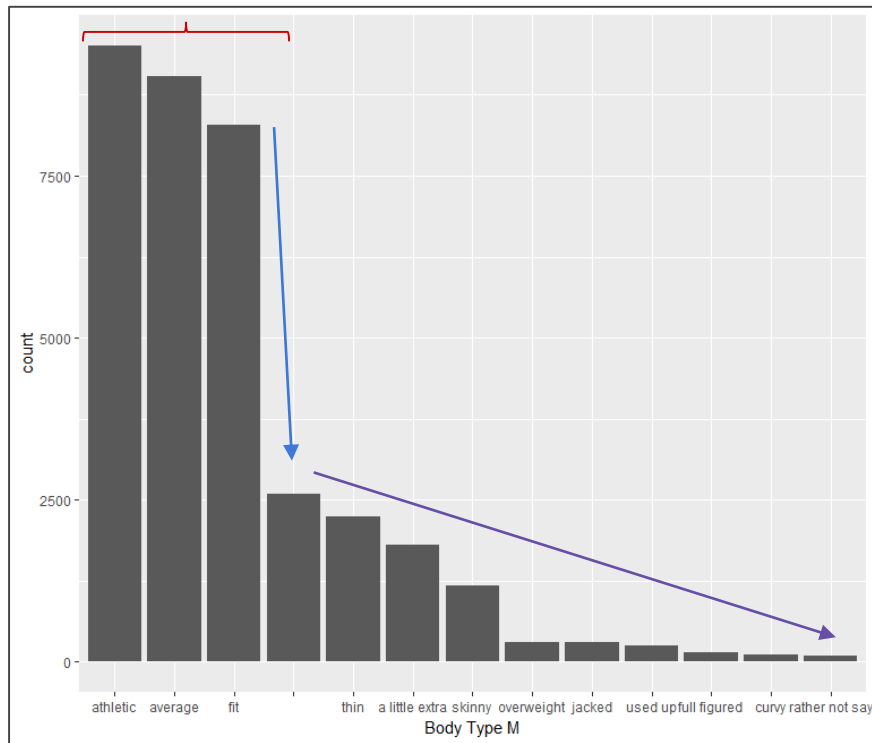
X axis shows options for body type.

Y axis shows frequency for each body type.

- Noticeable right-skew
- Relatively steep incline until last seven body types.

What kinds of people populate the San Francisco dating scene?

-Body Type by Sex = Male



Body Type and Sex selected.

Filtered Sex by Male.

X axis shows options for body type.

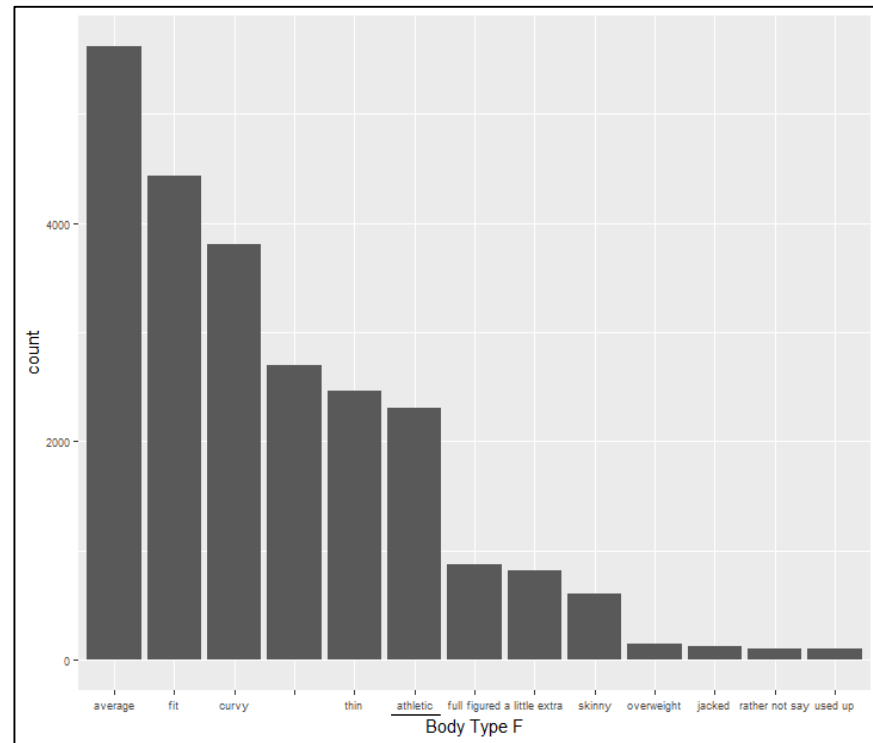
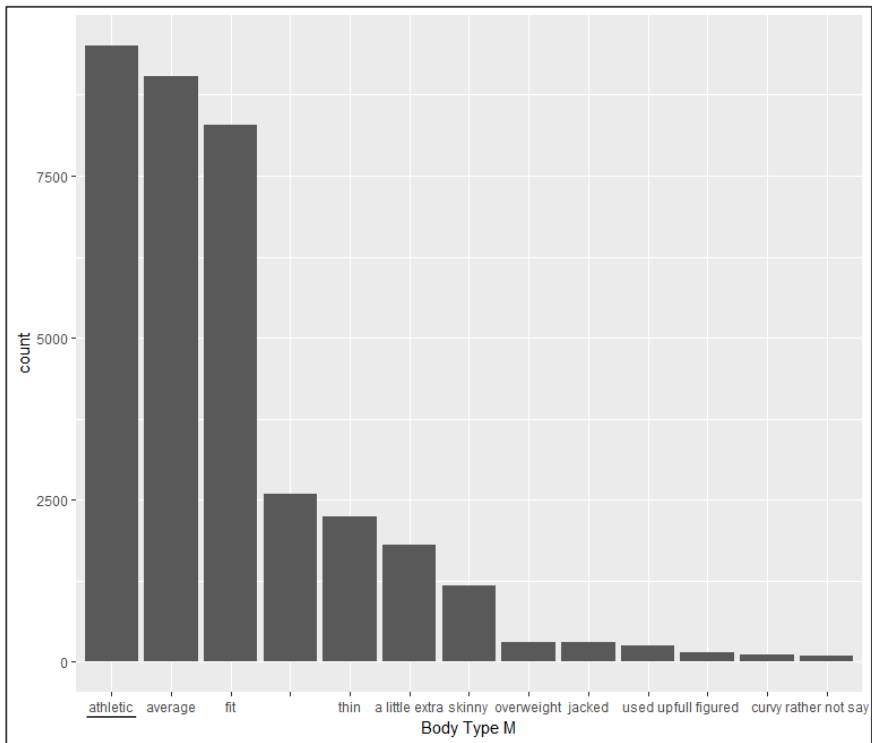
Y axis shows frequency for body types.

- Strong right-skew
- Top 3 results far above the rest
- Sharp drop after “fit”
- Gentle slope afterwards

What kinds of people populate the San Francisco dating scene?

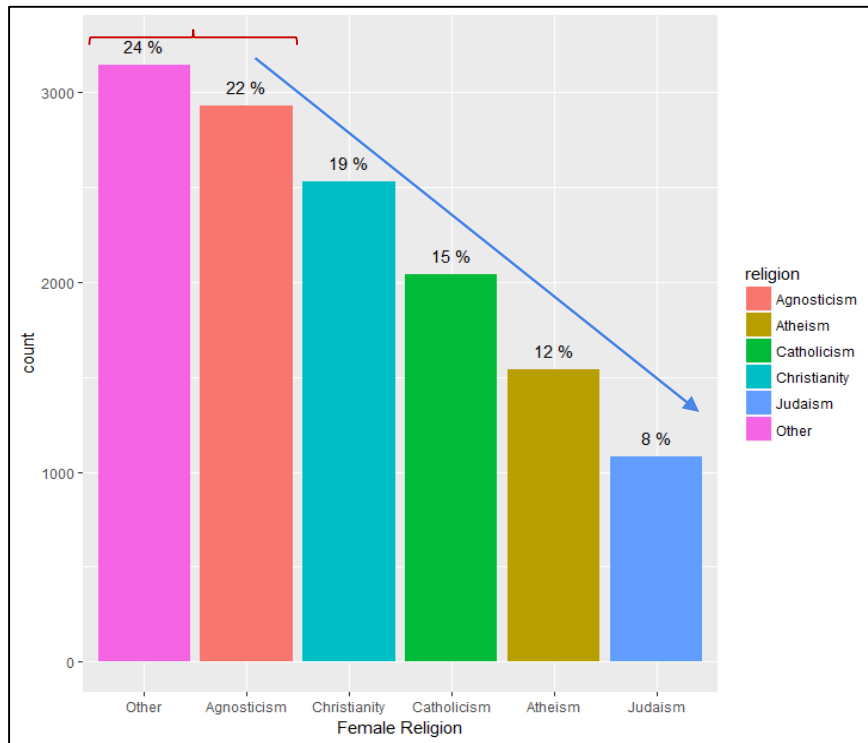
-Body Type by Both Sexes

-- --



What kinds of people populate the San Francisco dating scene?

-Religion by Sex = Female



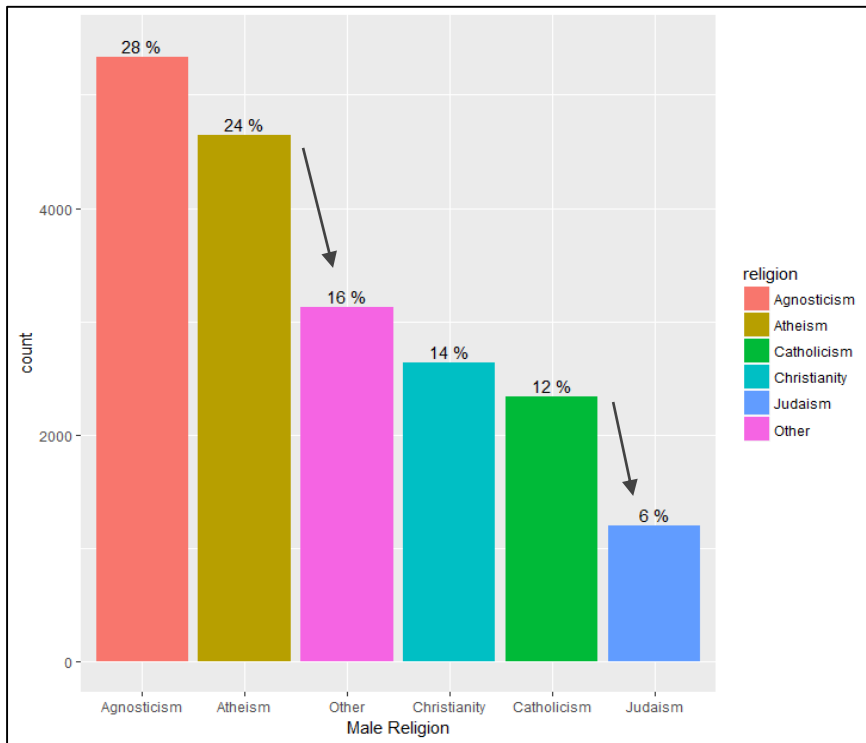
Religion and Sex selected.

Data sorted by religion count, top 20 religions looked at and grouped.

- Top 2 religions very close
- Fairly consistent and gentle slope

What kinds of people populate the San Francisco dating scene?

-Religion by Sex = Male



Religion and Sex selected.

Data sorted by religion count, top 20 religions looked at and grouped.

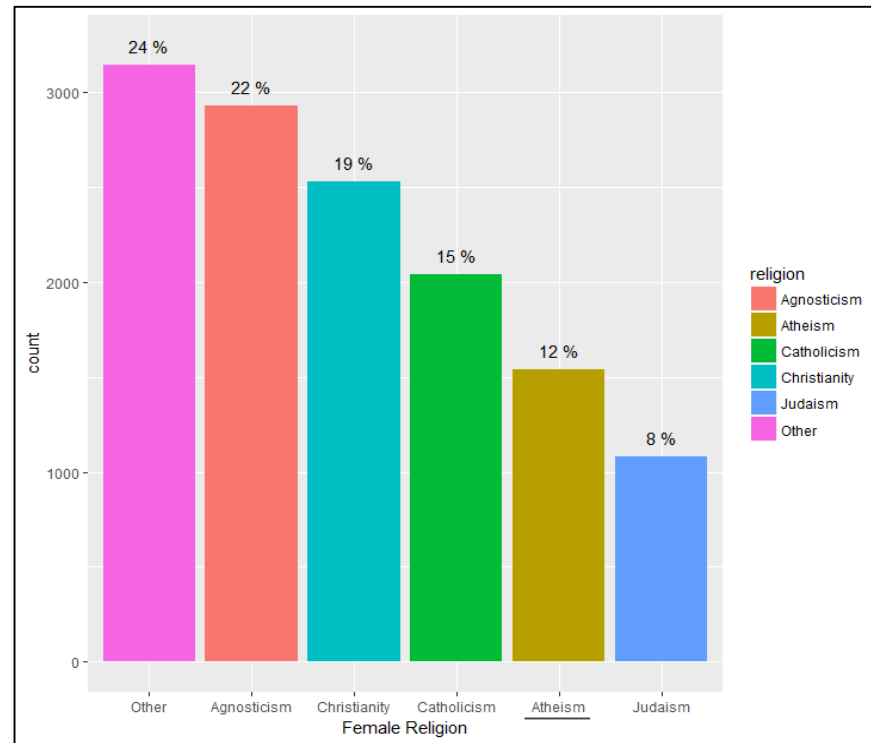
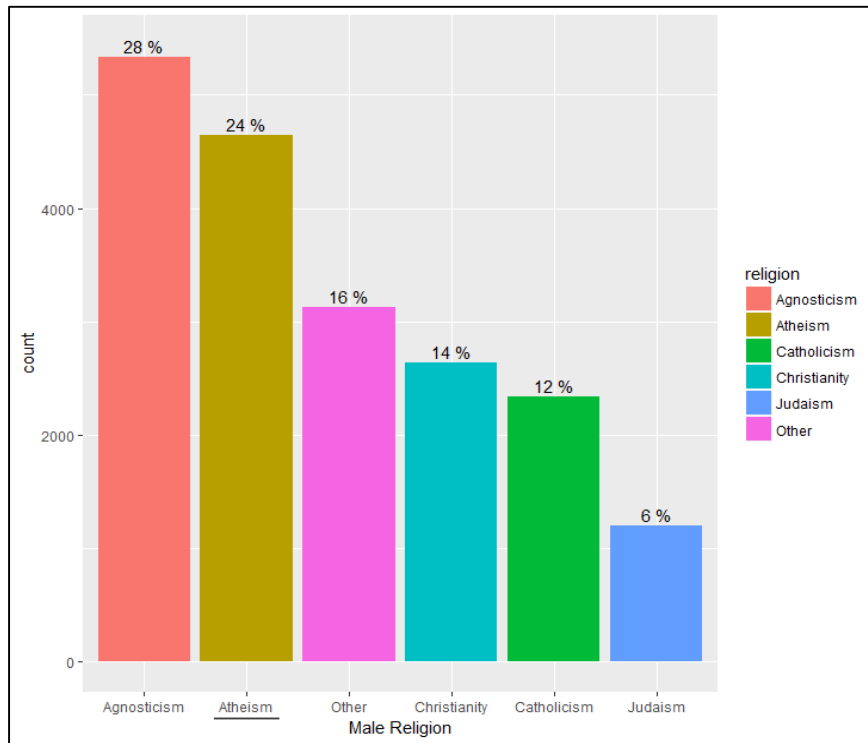
X axis shows religion.

Y axis shows frequency.

- Noticeable drops after top 2 and middle 3 religions.

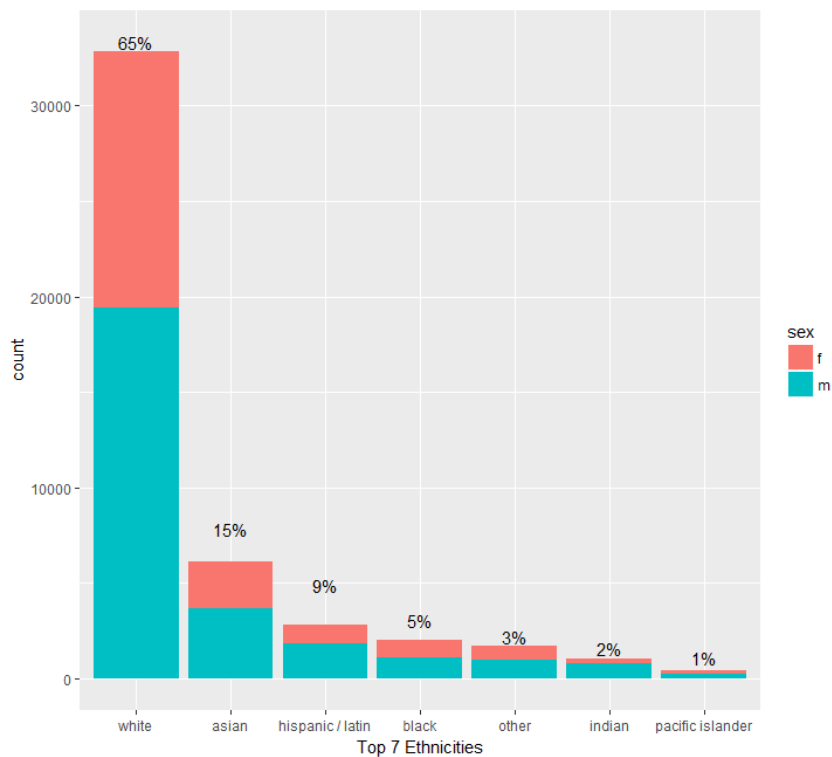
What kinds of people populate the San Francisco dating scene?

-Religion by Both Sexes



What kinds of people populate the San Francisco dating scene?

-Ethnicity and Sex



Ethnicity and Sex selected.

Top 20 ethnicities looked at and grouped into top 5 ethnicities + other. Count for each ethnicity graphed while using sex as fill.

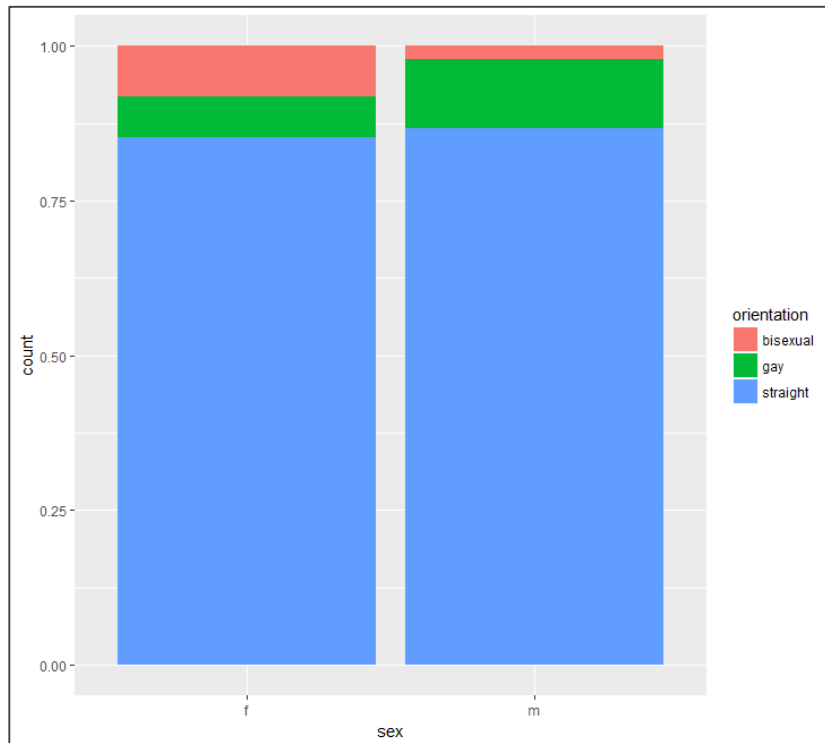
X axis shows top 6 ethnicities.

Y axis shows frequency for each ethnicity.

- White much more prevalent than other ethnicities.

What kinds of people populate the San Francisco dating scene?

-Orientation by Sex



Orientation and Sex selected.

Orientation chosen as fill for Sex as x axis. Proportion instead of count shown.

X axis shows the two sexes.

Y axis shows proportion of orientation.

- Men more likely to be **gay** than women
- Women more likely to be **bisexual** than men

What kind of habits should we expect to find amongst SF OKCupid users?

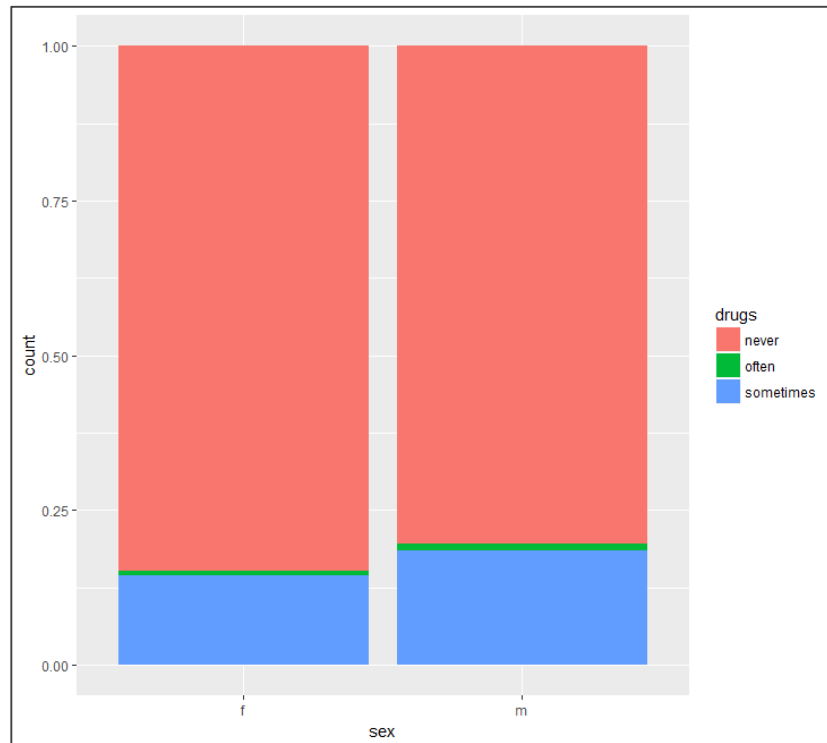
For each gender, what habits can we expect to find for...

- Drug use?
- Drinking?
- Diets?

Figuring out the most common habits for such categories will shed some light into San Francisco culture and give us an idea of what each sex is up to.

What kind of habits should we expect to find amongst SF OKCupid users?

-Drugs by Sex



Drugs and Sex selected.

Drugs chosen as fill for Sex as x axis. Proportion instead of count shown.

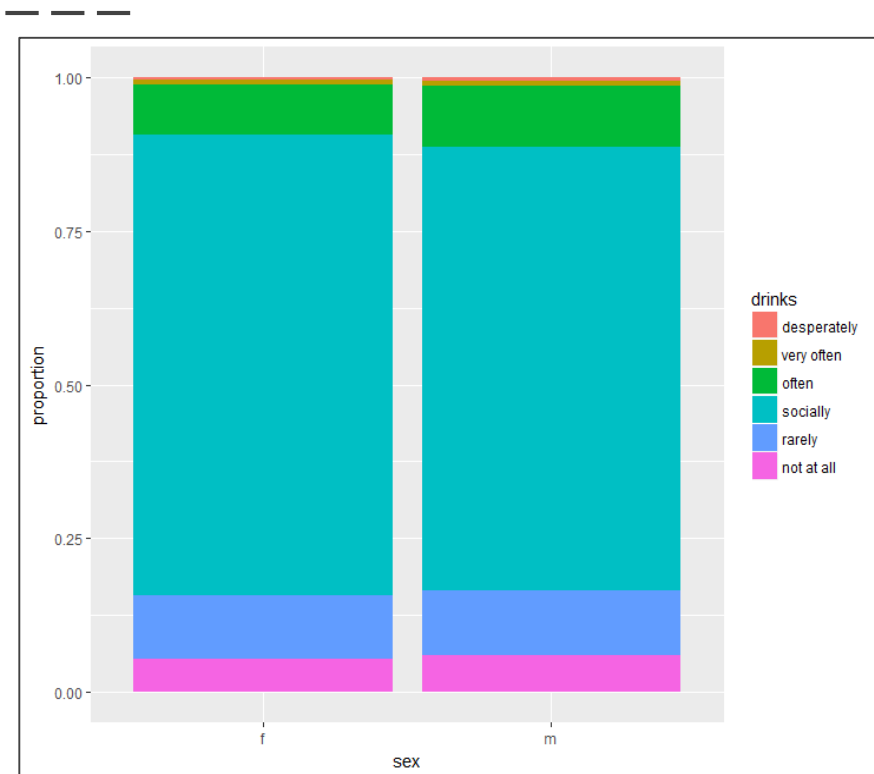
X axis shows the two sexes.

Y axis shows proportion of drug habits.

- Men are very slightly more likely to use drugs “often”
- Men are more likely to use drugs “sometimes”

What kind of habits should we expect to find amongst SF OKCupid users?

-Drinking by Sex



Drinks and Sex selected.

Drinks chosen as fill for Sex as x axis. Proportion instead of count shown.

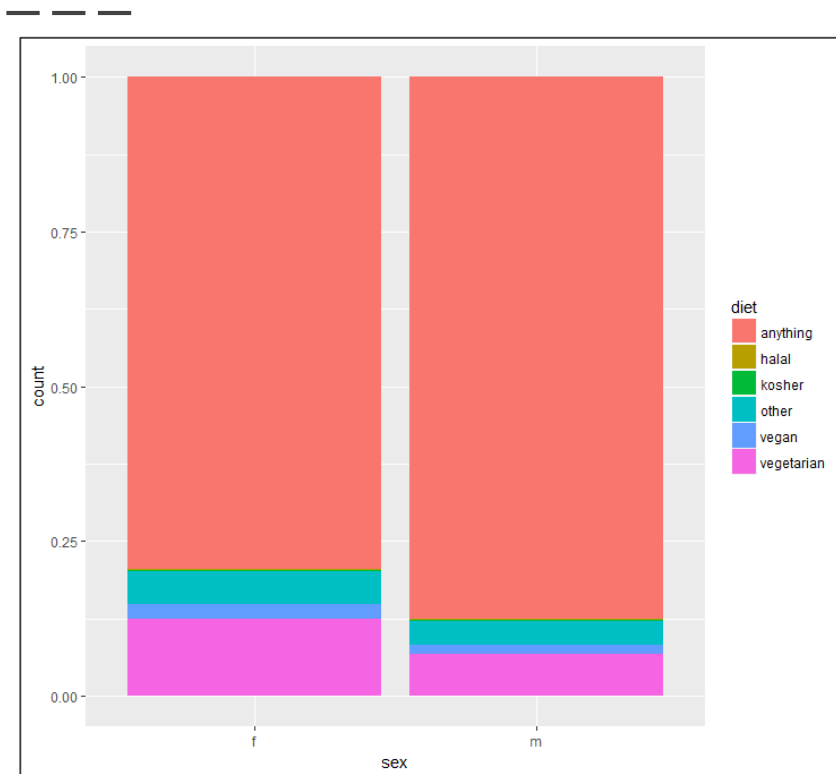
X axis shows the two sexes.

Y axis shows proportion of each drinking habit.

- Women are slightly more likely to drink “socially” than men
- Men slightly more likely to drink “not at all” or “often” than women
- Proportions relatively consistent otherwise

What kind of habits should we expect to find amongst SF OKCupid users?

-Diet by Sex



Diet and Sex selected.

Diet chosen as fill for Sex as x axis. Proportion instead of count shown.

X axis shows the two sexes.

Y axis shows the proportion of each diet habit.

- Men are significantly more likely of eating “**anything**”
- Women are significantly more likely to be “**vegetarian**”
- Women are more likely to follow “**other**” diets or be “**vegan**”

How do OKCupid users compare to the average American/San Francisco resident?

— — —

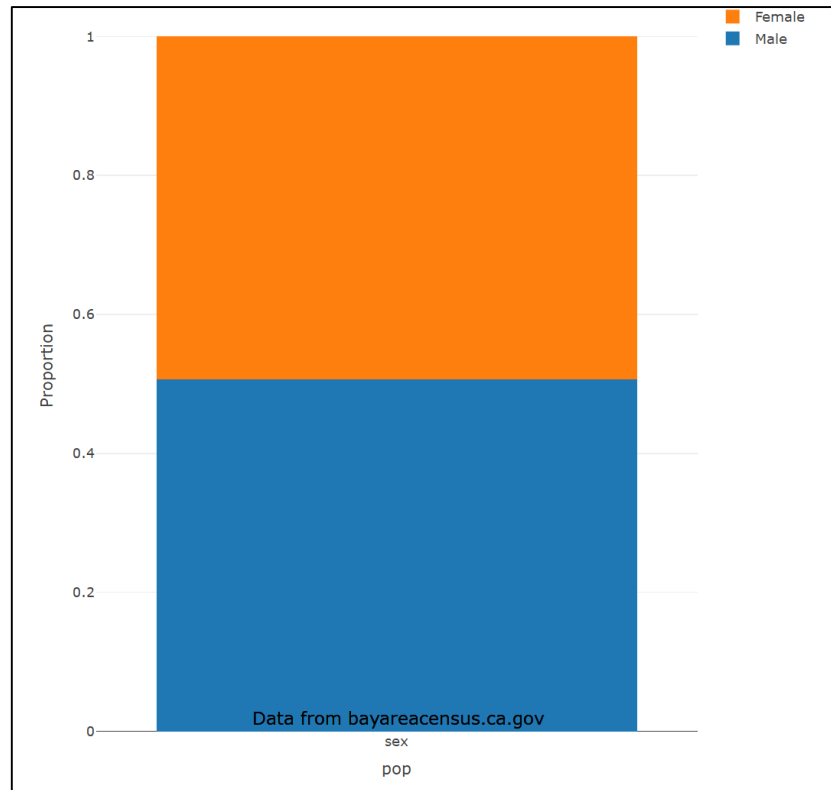
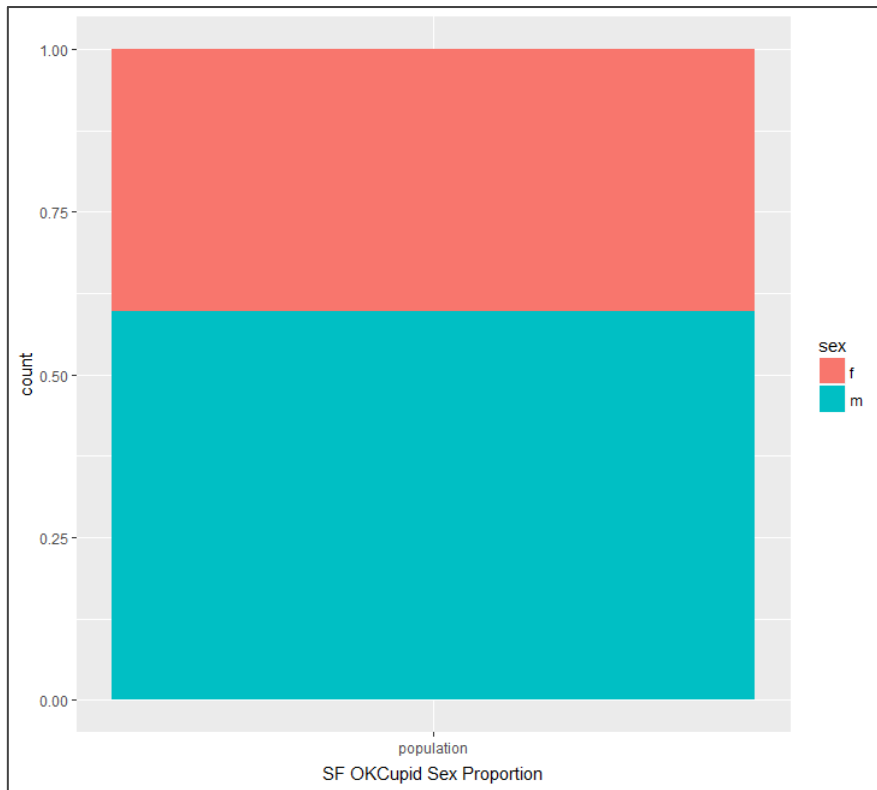
How does the average American/San Francisco resident compare to OKCupid users when it comes to...

- Sex?
- Height?
- Income?
- Ethnicity?

Answering these questions will allow us to discern whether SF OKCupid users tend to be above or below average when it comes to certain characteristics. Perhaps the virtual dating landscape doesn't reflect San Francisco very well.

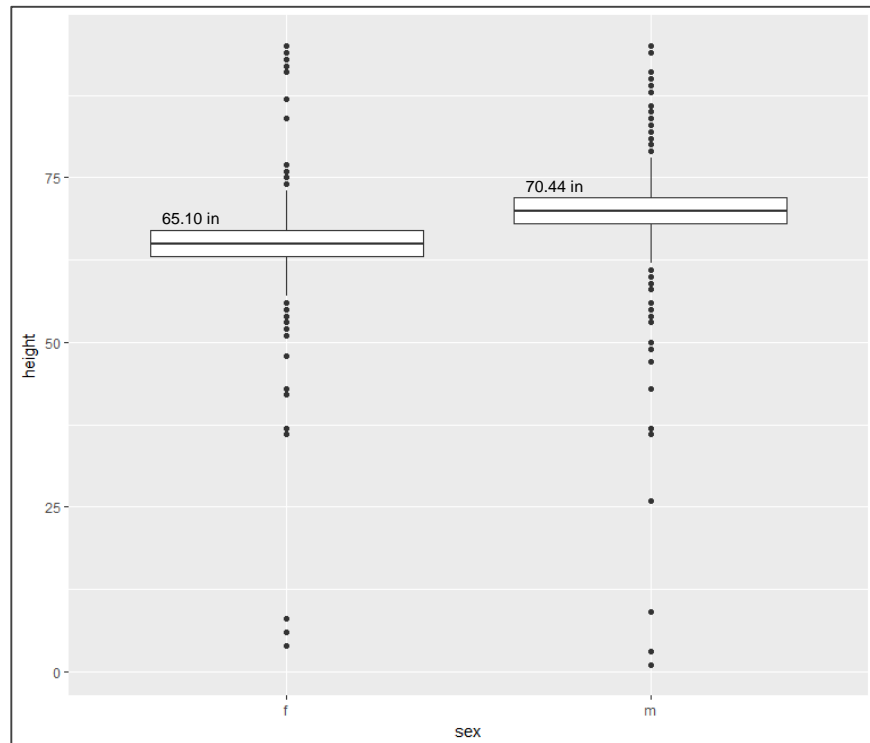
How do OKCupid users compare to the average American/San Francisco resident?

-Sex



How do OKCupid users compare to the average American/San Francisco resident?

-Height by Sex



Height and Sex selected.

Height filtered to only include data > 0 . Data summarized showing median height for each sex.

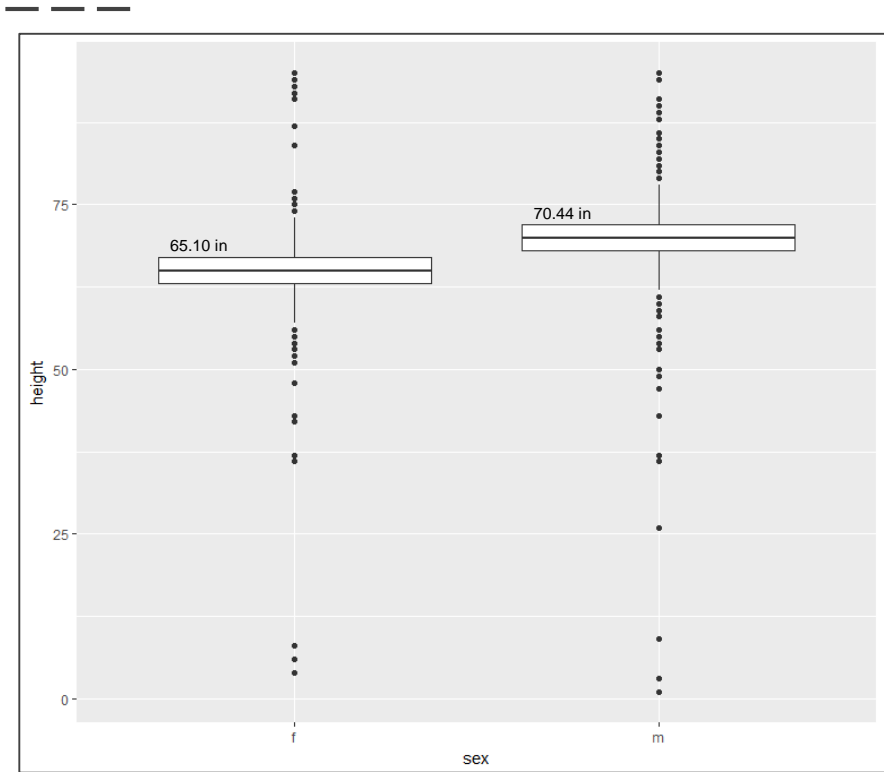
X axis shows the two sexes.

Y axis shows height in inches.

- Male median is 70 in. (5'10)
 - Mean is 70.44 in. (5'10.4)
- Female median is 65 in. (5'5)
 - Mean is 65.10 in. (5'5)

How do OKCupid users compare to the average American/San Francisco resident?

-Height by Sex

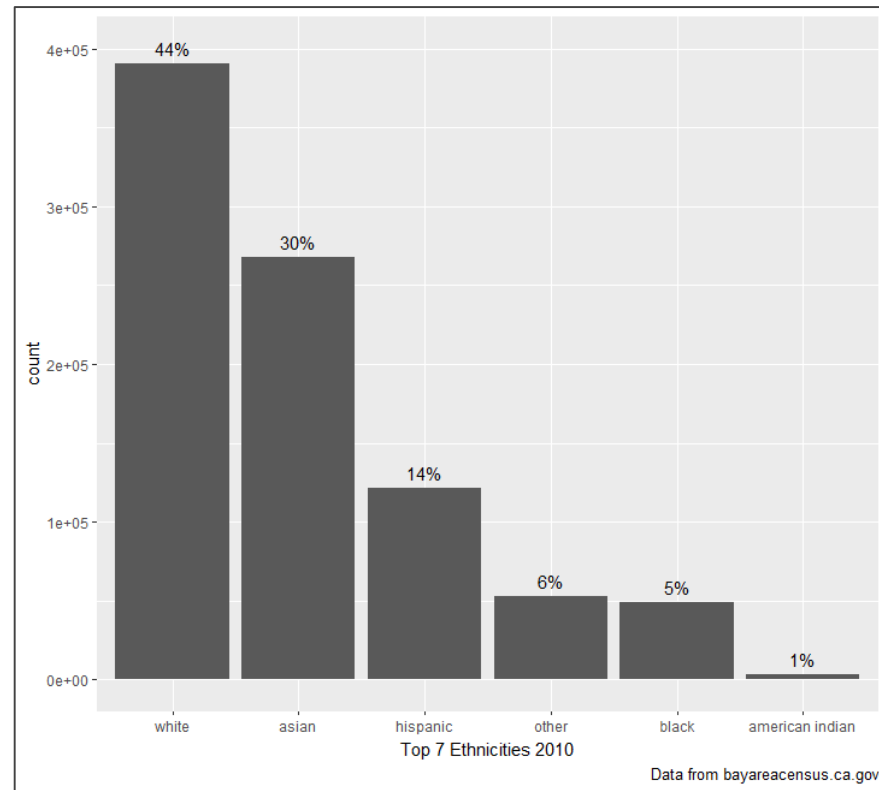
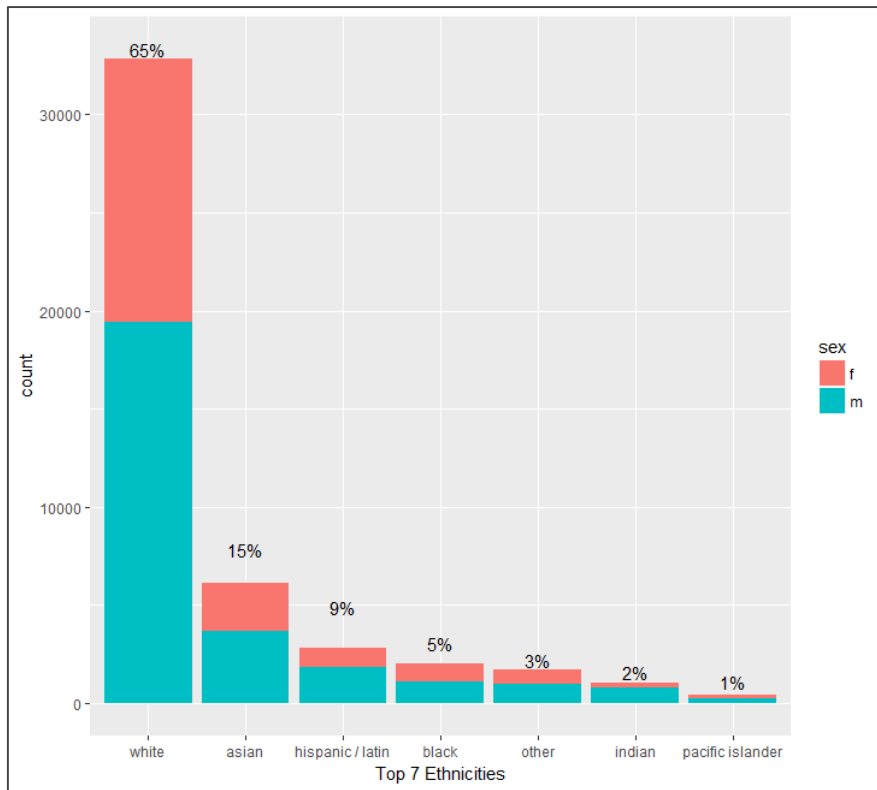


U.S.A. Average Height (in.)	
Men	Women
69.5 (5'9.5)	63.8 (5'4)

Source: livestrong.com

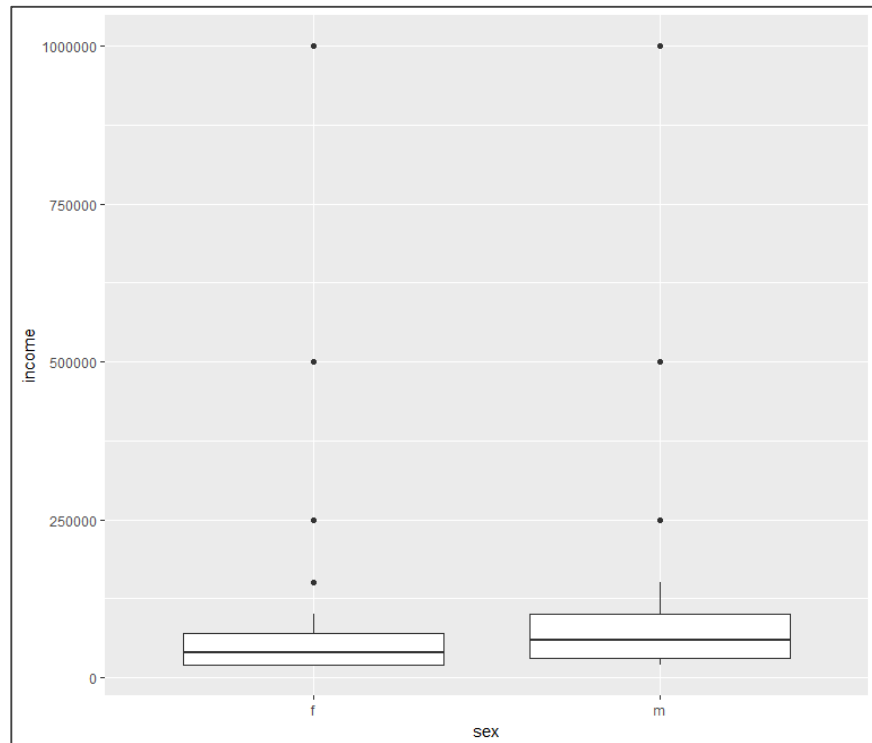
How do OKCupid users compare to the average American/San Francisco resident?

-Ethnicity



How do OKCupid users compare to the average American/San Francisco resident?

-Income by Sex



Income and Sex selected.

Income filtered to only include data > 0 . Data summarized showing median income for each sex.

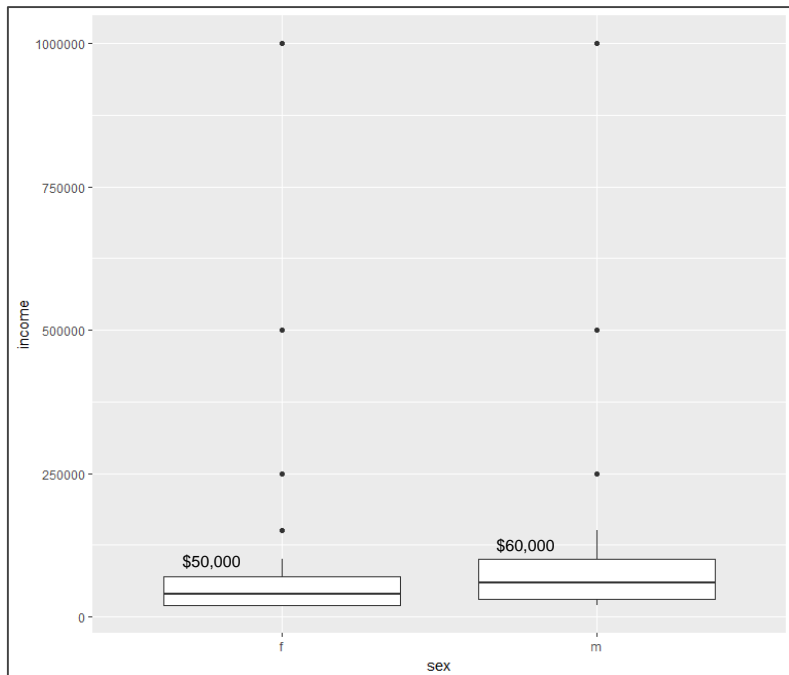
X axis shows sex the two sexes.

Y axis shows income (in \$).

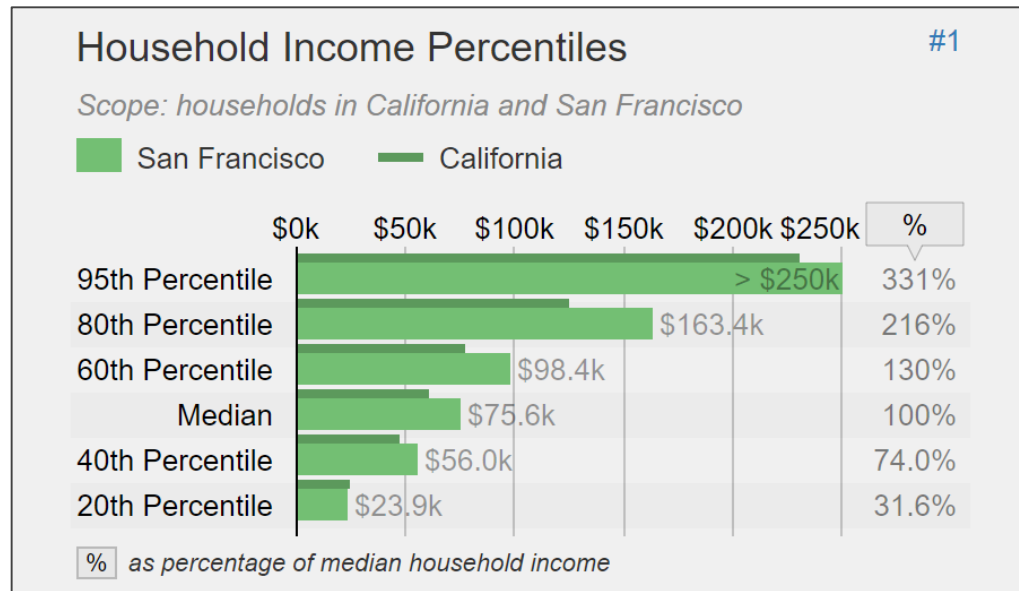
- Male median is \$60,000
 - Mean is \$110,984.4
- Female median is \$50,000
 - Mean is \$104,395

How do OKCupid users compare to the average American/San Francisco resident?

-Income



OKCupid Individual Income



Source: statisticalatlas.com

San Francisco Household Income

Conclusions

- Sex

- 6:4 Male-Female split
 - Women have a much easier time finding men

- Age

- Most users in the 22-35 age range

- Body Type

- Men tend to be Athletic, Average, Fit, and not much else
 - Suspicious
- Women tend to be Average, Fit, and Curvy

- Ethnicity

- Whites strongly overrepresented

- Orientation

- Men likelier to be gay than women
- Women likelier to be bisexual than men



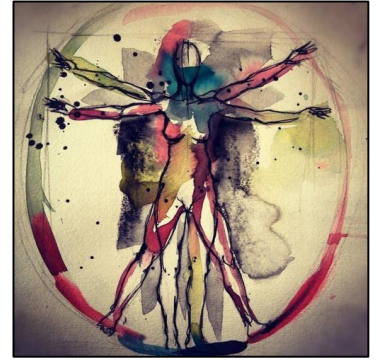
Conclusions

- Drugs
 - Men more likely to use drugs sometimes and often
- Drinking
 - Women more likely to drink socially than men
 - Men more likely to not drink or to drink often
 - A gender of extremes
- Diet
 - Men more likely to not have a strict diet
 - Women more likely to be vegan or vegetarian



Conclusions

- Sex – OKCupid vs SF
 - OKC 6:4 Male-Female split vs SF 5:5 Male-Female Split
 - Males overrepresented
 - Men might be more likely to look for dates in general
 - Men might be more likely to use online dating
- Height – OKCupid vs America
 - Men and women are around an inch taller than the average American
 - San Francisco residents might receive better exercise and nutrition than the average American
 - Online daters might be boosting their height slightly to look more appealing



Conclusions

● Ethnicity - OKCupid vs SF

- Whites overrepresented in both populations but significantly more so in OKCupid
 - White people might be more likely to search for dates online
 - OKCupid might appeal to white people more
 - Perhaps East meets East and such draw certain ethnicities away from OKCupid



● Income - OKCupid vs SF

- Median income for men is 20% more than women
 - Men might be more likely to disclose income
- Mean income for men and women is nearly double the median
 - Men and women with high incomes are more likely to disclose it
- Median income for men and women is below San Francisco median household income
 - Not a completely fair metric, but relative closeness makes OKCupid seem truthful



Conclusions

— — —

People sometimes fear that online daters aren't very truthful in their profiles, but OKCupid users in San Francisco don't seem to be lying about their attributes.

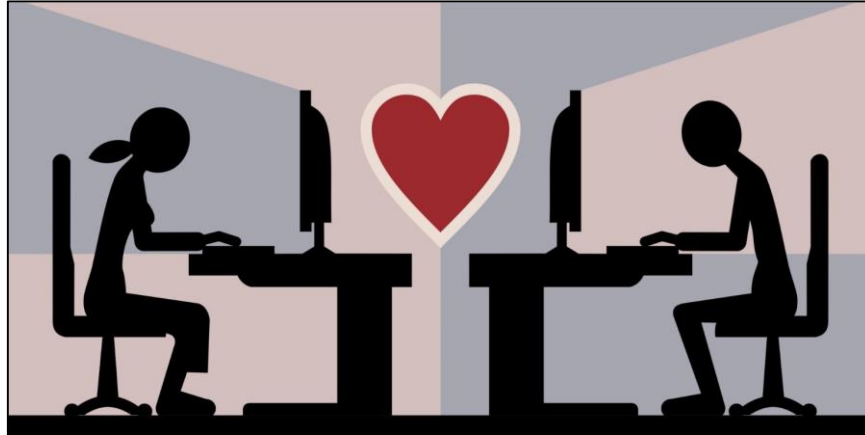
If you plan on using OKCupid in San Francisco, you're likely to be pleasantly surprised by the physical characteristics of the people you meet.

However, if you're looking for more variety in when it comes to ethnicities or orientation, then OKCupid might not be the app for you.

Conclusions

— — —
If you're looking for an athletic, white, agnostic, straight man or an average, white, agnostic, straight woman, then you just might be in luck.

Turns out finding people online isn't that different from finding people offline.



A Long Look at Love

(San Francisco Edition)

Rodrigo Arguello