

Swipe Write to Find Mr. Miss/ Right

Understanding How Different Demographics
Describe Themselves on Online Dating Platforms

Why Bother?



Who Would You Go Out With?



looking for romance in all the right places...talented technically,well read, loves the out doors and a good symphony is always a treat.

i am a mystery wrapped inside of an enigma.some might call that a conundrum burrito.if that's you, we should be friends.



Overview

Objective 1.0:

Find patterns in how different demographics use language to market themselves online.

Components

Getting to Know the Demographics

Demographics and Text Together

Doc2Vec and Word2Vec Together

About the Data

- OkCupid Data from 2012
- Approximately 60,000 profiles from the San Francisco area
- Profiles were anonymously submitted to the study with consent
- Average age: 32 years
- Average height: 5 foot 6 inches
- Average time since online: 40 days

About the Data



Part 2: Clustering With Demographics and Text

Cluster 1

1. I'm looking for a good solid relationship..I'm ready to settle down and eventually get married..I'm a single mom of 3..would like to be with someone who would love my kids; go to church..
2. I am a passionate, forthright, sarcastic, witty, and fancy myself an undiscovered comedian. Attend church regularlyhave beautiful three-year old twin boys who are full of life and questions.
3. Divorced at the end of 08 and ...a working dad with 2 great, teen boys who live with me half-time. Loves bicycling, snow skiing, hiking, travel. exploring the great outdoors and attending church

Cluster 1

1. I'm looking for a good solid relationship..I'm ready to settle down and eventually get married..I'm a **single mom of 3**..would like to be with someone who would love my kids; **go to church**..
2. I am a passionate, forthright, sarcastic, witty, and fancy myself an undiscovered comedian. **Attend church** regularlyhave beautiful **three-year old twin boys** who are full of life and questions.
3. Divorced at the end of 08 and ...a working dad with 2 great, **teen boys** who live with me half-time. Loves bicycling, snow skiing, hiking, travel. exploring the great outdoors and **attending church**

Cluster 2

1. I'm 27, 5 foot 2, I have a shitty construction job , I like to smoke weed, drink, and of course have fun. I like to go out on the weekends and sometimes weekdays, usually to a bar or two...
2. I'm just a city boy, ...I used to be a singer in a smokey room. i enjoy the smell of wine and cheap perfume. for a smile we can share the night.
3. I'm just a college student trying to figure out what life has in store. ...I'm definitely a geek and, although i hate to admit this, probably a little bit hipster. but nonetheless; die hipster, die. do i smoke weed? Yup . do i drink? yup. other than that, i just like to have fun with fun people.

Cluster 2

1. I'm **27**, 5 foot 2, I have a shitty construction job , I like to **smoke weed, drink**, and of course have fun. I like to go out on the weekends and sometimes weekdays, usually to a bar or two...
2. I'm just a **city boy**, ...I used to be a singer in a **smokey room**. i enjoy the smell of **wine** and cheap perfume. for a smile we can share the night.
3. I'm just a **college student** trying to figure out what life has in store. ...I'm definitely a **geek** and, although i hate to admit this, probably a little bit hipster. but nonetheless; die hipster, die. do i **smoke weed**? Yup . do i **drink**? yup. other than that, i just like to have fun with fun people.

Cluster 3

1. I'm not good with writing these so here goes nothing... I'm working on education at the moment. I'm funny, random, crazy, and fun. I don't smoke don't drink and don't do any drugs.
2. Im born and raised in sf mostly and some time in ireland. lets see im irish and asian. i'm sarcastic .. like to go out with my friends which sometimes involve drink and definate randomness, don't smoke but i don't care if you do..
3. I'm a transplant from ohio been here ten years now, i'm home.having been raised in the midwest with traditional values, love of god, country, and respect for women. Although i don't smoke drink or do drugs. i love life and you should to

Cluster 3

1. I'm not good with writing these so here goes nothing... I'm working on education at the moment. I'm funny, random, crazy, and fun. I **don't smoke don't drink and don't do any drugs.**
2. Im born and raised in sf mostly and some time in ireland. lets see im irish and asian. i'm sarcastic .. like to go out with my friends which sometimes involve drink and definate randomness, **don't smoke** but i don't care if you do..
3. I'm a transplant from ohio been here ten years now, i'm home.having been raised in the midwest with traditional values, love of god, country, and respect for women. Although i **don't smoke drink or do drugs.** i love life and you should to

Do we agree?

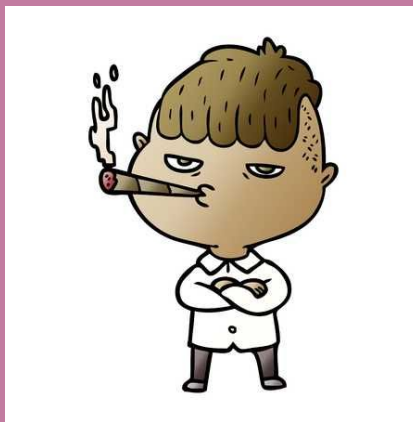
Cluster 1

- have kids
- older (average 41)
- live far
- more religious



Cluster 2

- do drugs and smoke
- younger (average 29)
- most active online
- local

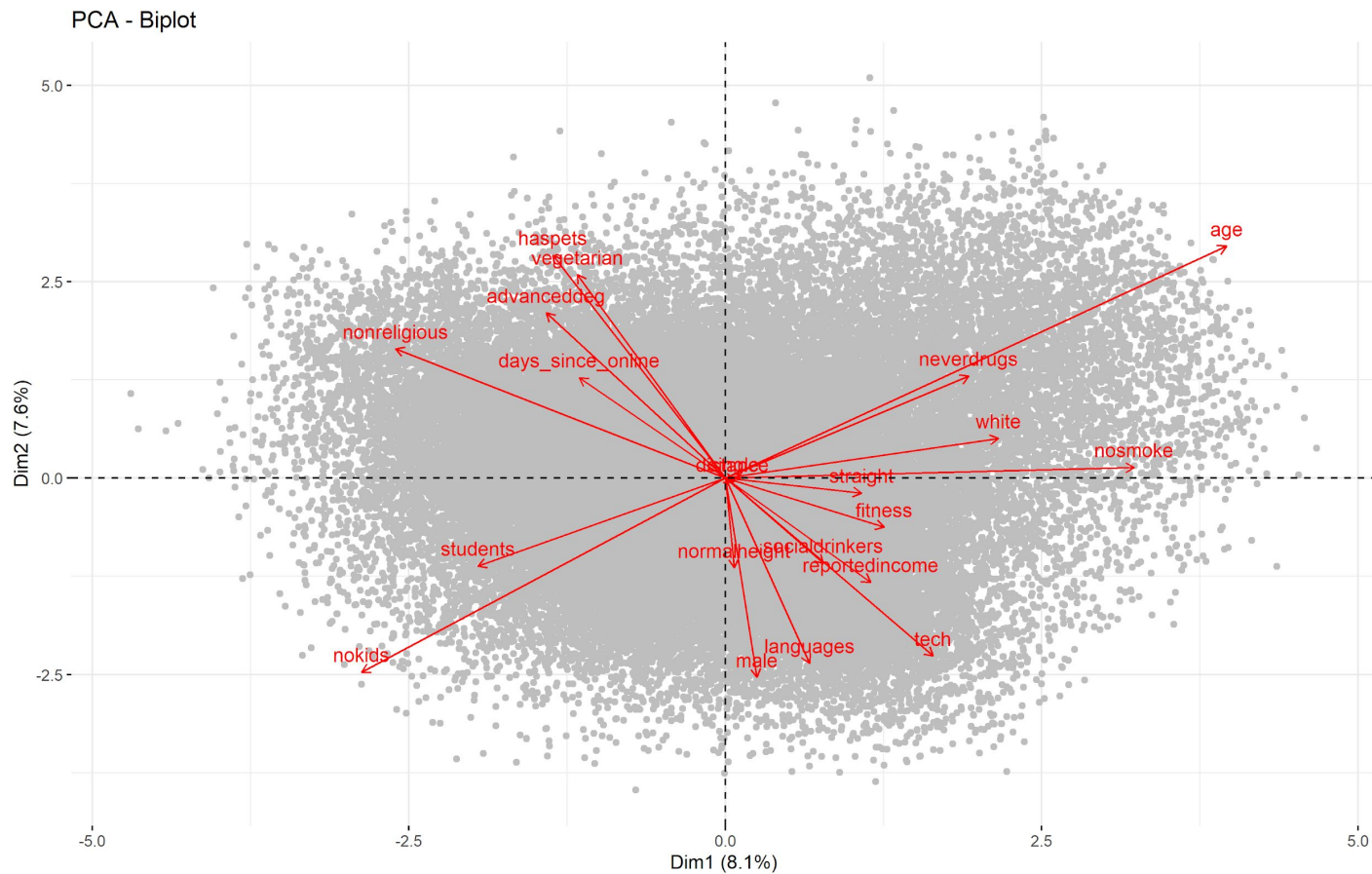


Cluster 3

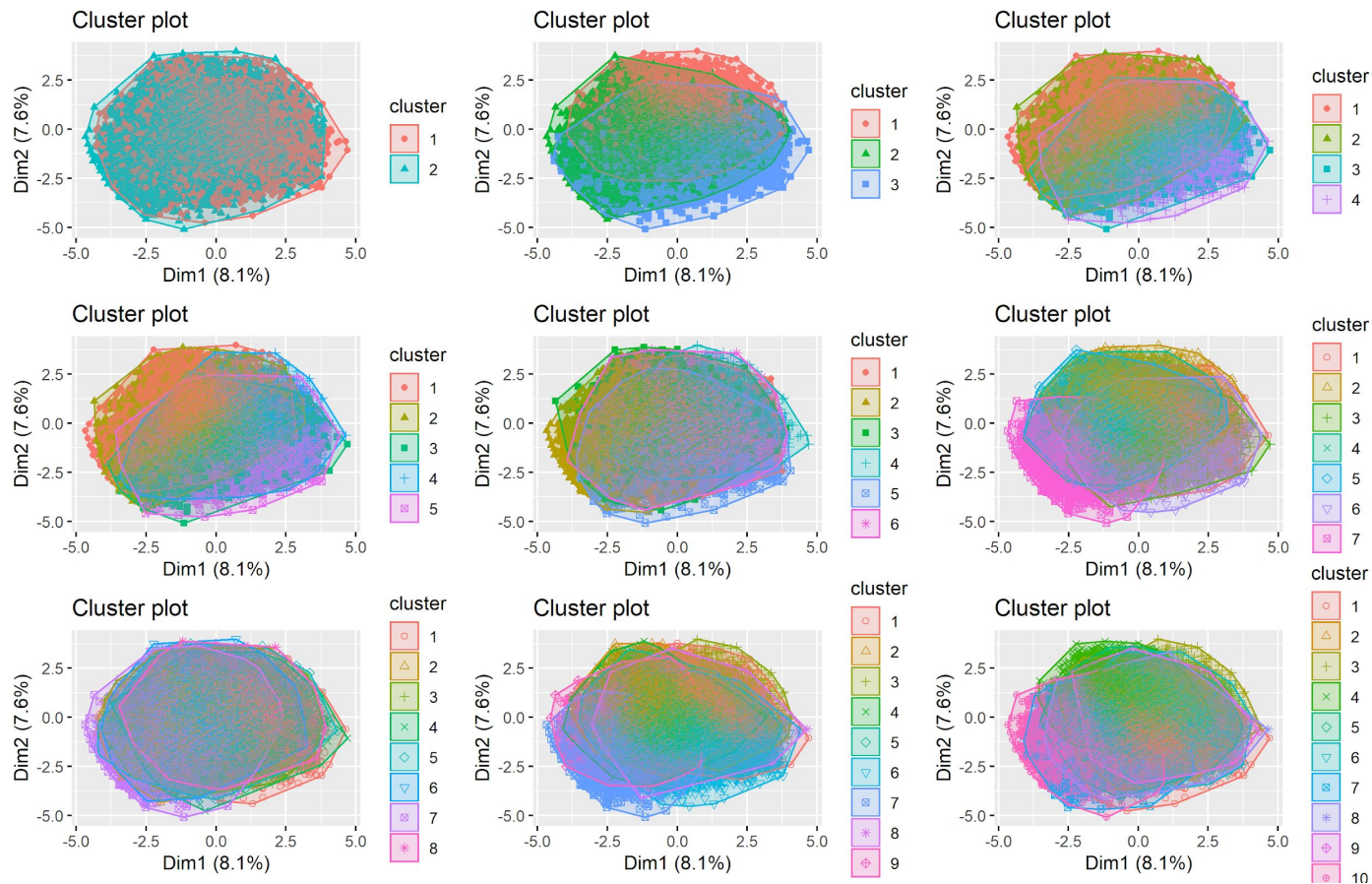
- never do drugs or smoke
- average 31



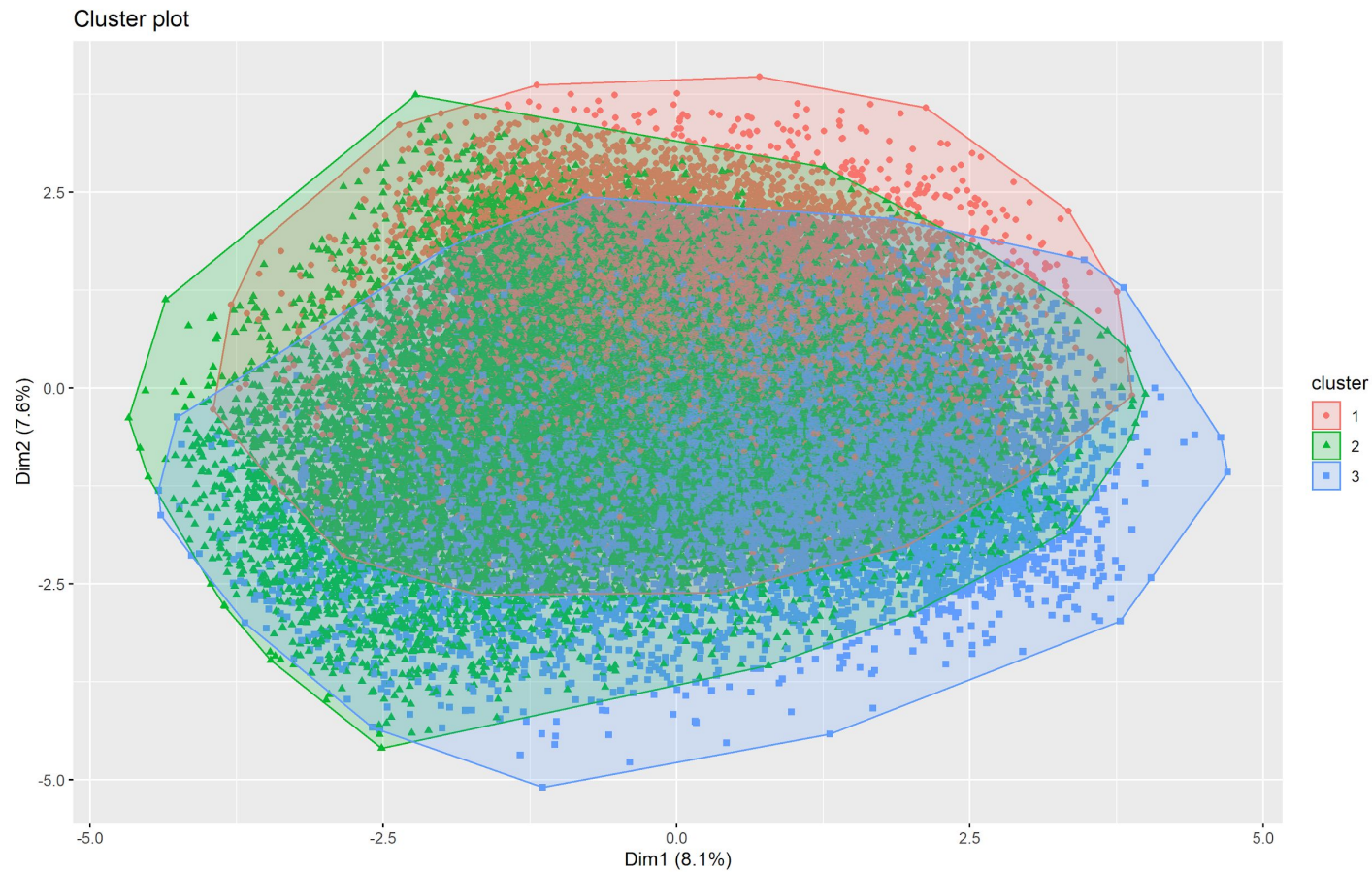
PCA



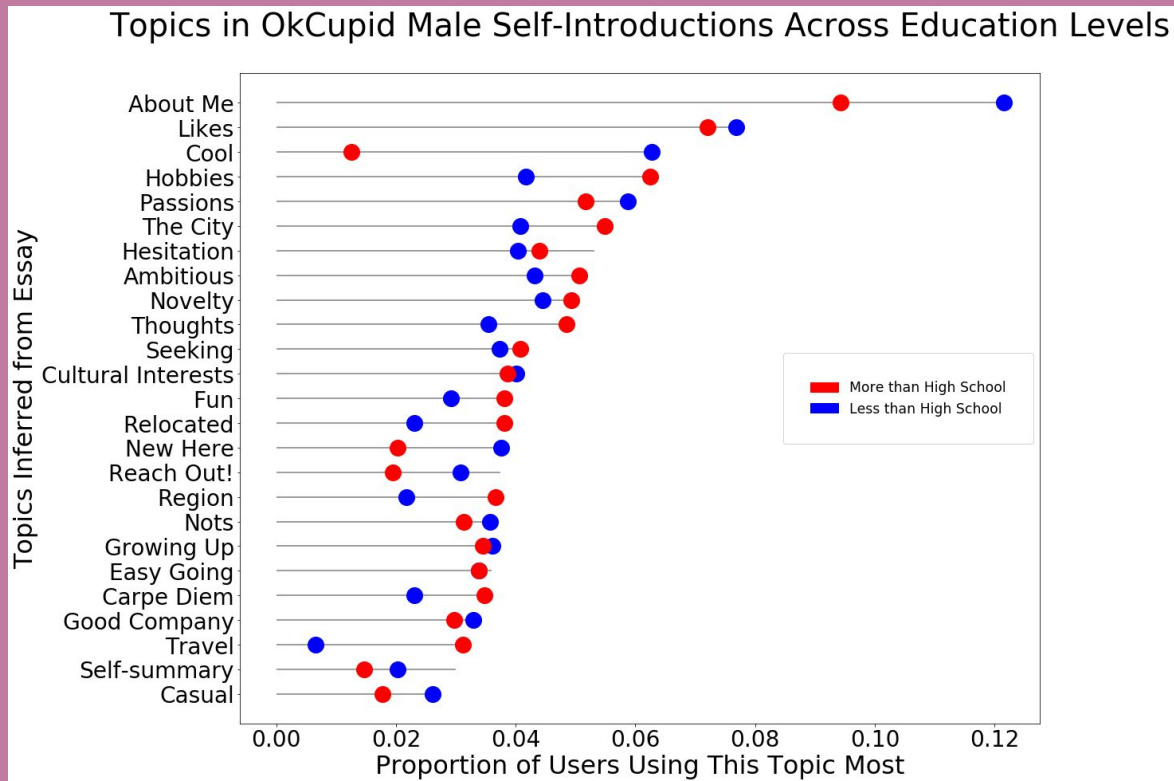
CLARA



CLARA

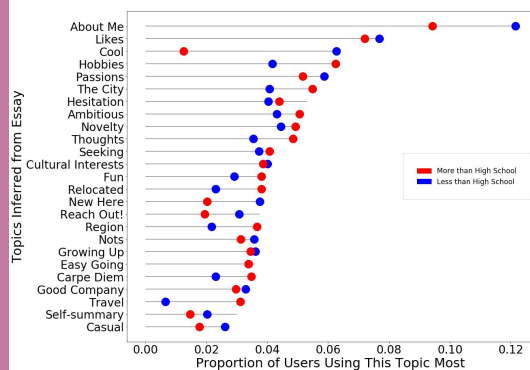


What do people talk about- Topic Models

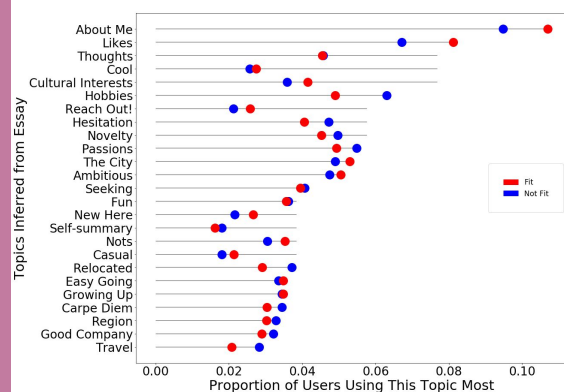


Big Difference!?

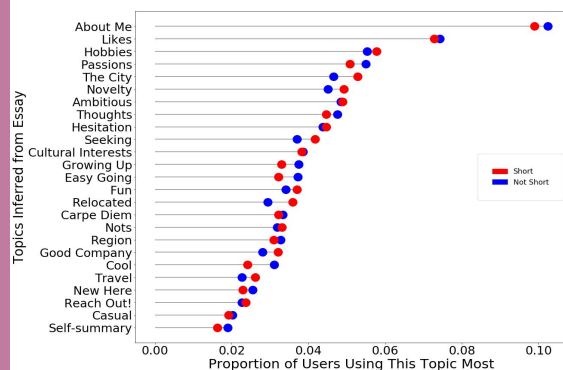
Topics in OkCupid Male Self-Introductions Across Education Levels



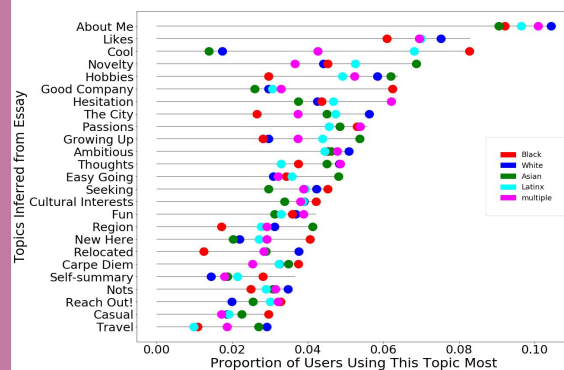
Topics in OkCupid Male Self-Introductions Across Fitness Levels



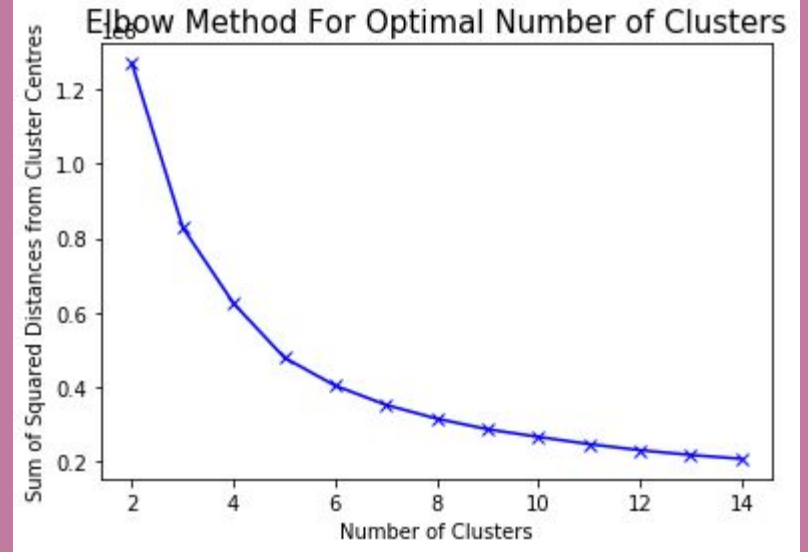
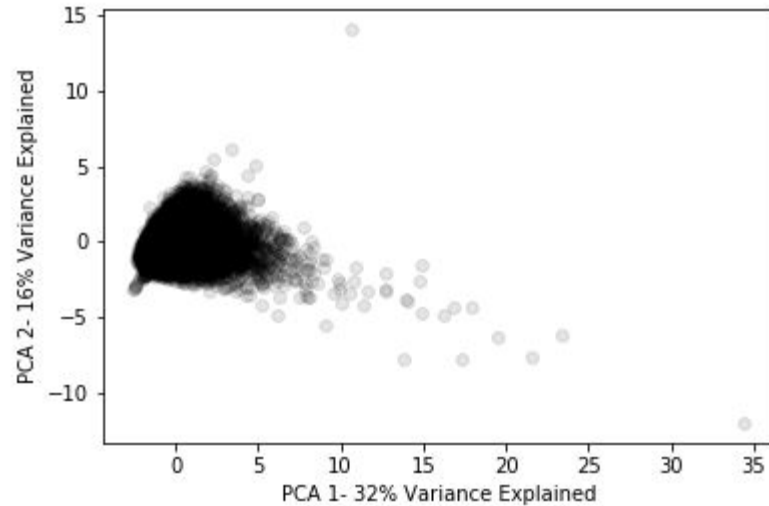
Topics in OkCupid Male Self-Introductions Across Height Groups



Topics in OkCupid Male Self-Introductions Across Racial Groups



Doc2Vec Models



What We've Learnt So Far

Profiles are fairly (unfairly!?) homogeneous

Focus more on distinctiveness and the outliers

Future Steps

Use Agglomerative Clustering (more bottom-up)

Work more on interpreting what clusters mean, especially with Doc2Vec

Look for memorable/ unusual words rather than common topics