

Simon Fong
Brian Aguirre
Mulin Yang
Justine Lee

ECE 143
Spring 2019

Overview

1. Context
2. Collecting and Cleaning
3. Interesting Points
4. Deeper Analysis

1. Context

About the data

- <http://subtleasiandating.org>
- 37 question quiz that calculates your 'dateability' score as a guy.
- Quiz posted Subtle Asian Dating group.
- About 1000 responses recorded, but had to stop recording because database was getting full.

1. Height:

0	Centimeters
0	Feet 0

Inches

2. ☐ STEM, law, or medicine major.

3. ☐ Graduated.

4. ☐ Masters in progress or received.

5. ☐ PhD in progress or received.

6. ☐ Asian parents think your degree is a waste of time.

7. IQ score or estimate.

IQ Score Meanings from this [site](#):

- Over 140 - Genius or near genius
- 120 - 140 - Very superior intelligence
- 110 - 119 - Superior intelligence
- 90 - 109 - Normal or average intelligence
- 80 - 89 - Dullness
- 70 - 79 - Borderline deficiency
- Under 70 - Definite feeble-mindedness

Rubric

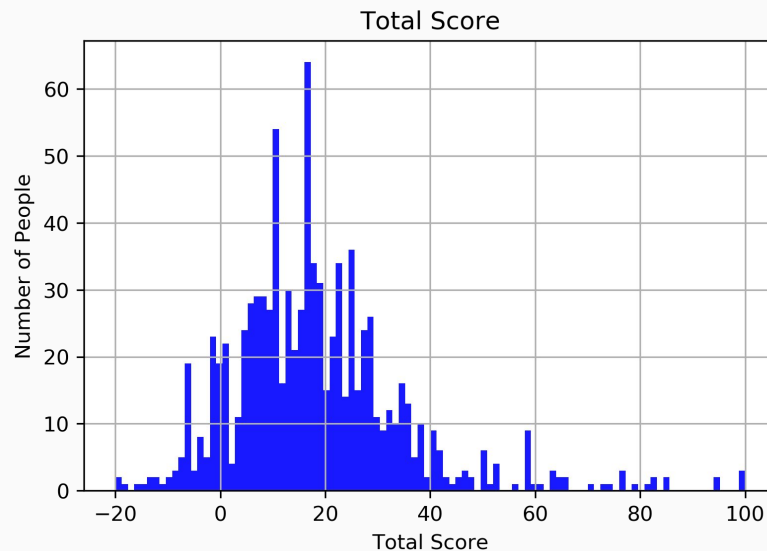
- Rubric inspiration from a popular Facebook post in Subtle Asian Dating.

Start with 0 points.

- Add 1 point for every inch over 5'10.
- Subtract 1 point for every inch below 5'7.
- Add 5 points if 6 figure salary.
- Subtract 1 point for every tattoo.
- And so on for 37 questions...

Quiz Scores

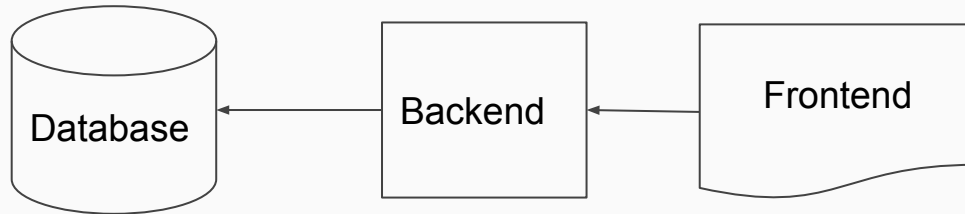
- Average: 18
- Median: 16
- Standard Deviation: 16.840
- Min: -20
- Max: 100



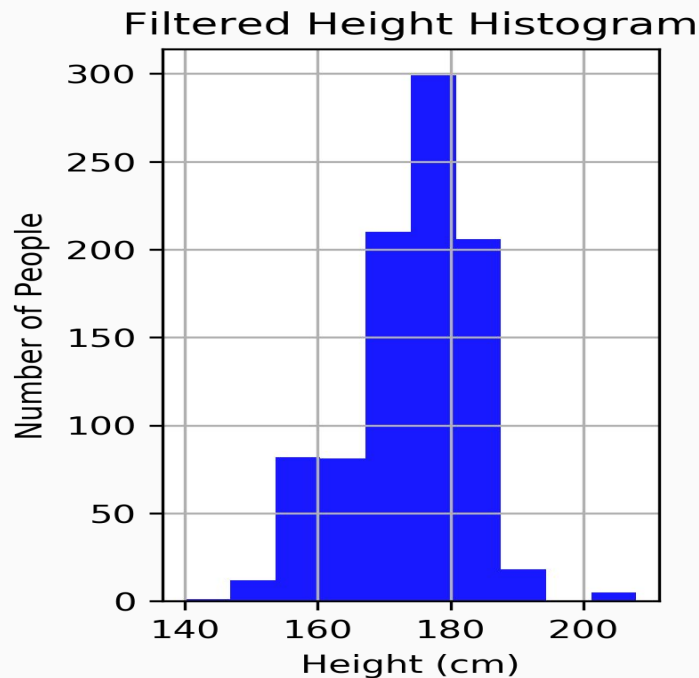
2. Collection/Cleaning

How we collected the data

- HTML/CSS/Javascript for frontend form
- Python/Flask for backend web application
- MongoDB for database



Removing outliers



We had some outliers in our data such as this height:

Height: 1.26×10^{29} cm

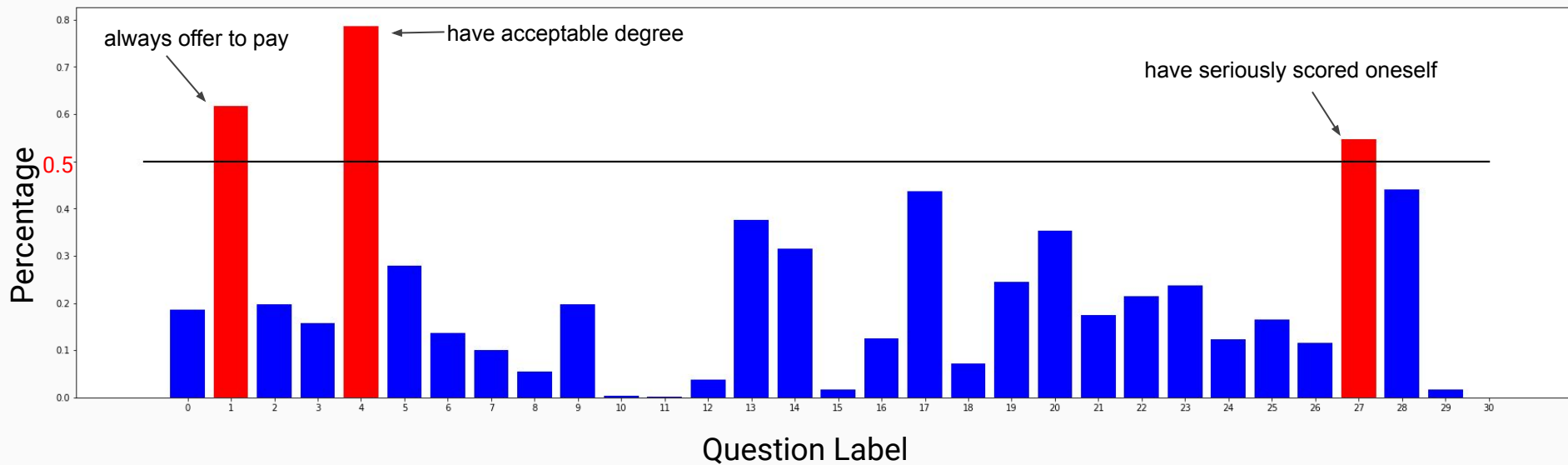
Basically way too tall.

So we removed some height outliers such as above **8 ft** and below **4 ft**.

We did the same for the rest.

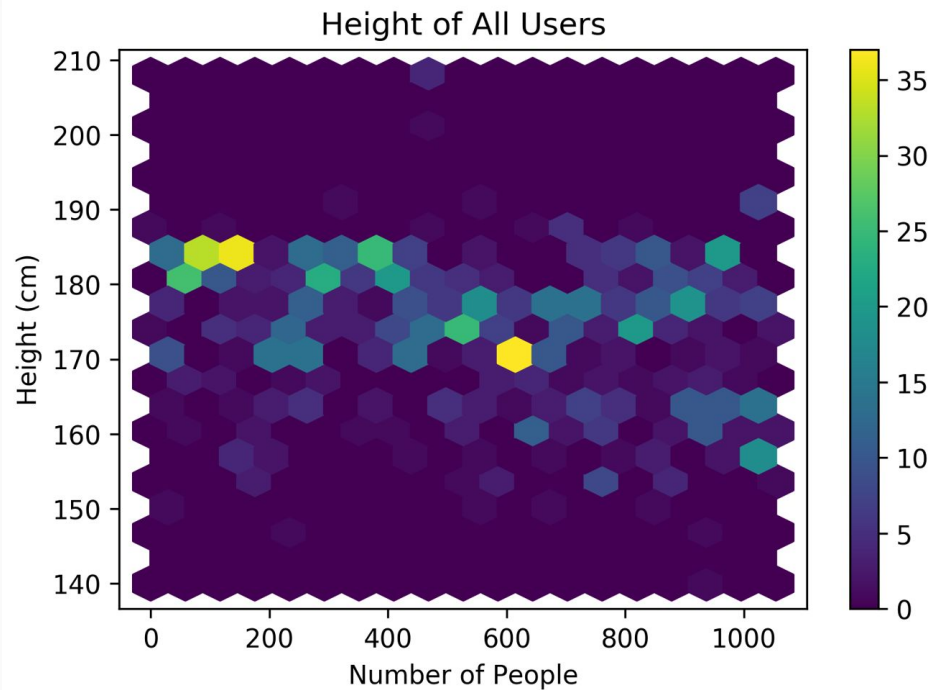
General View

30 Questions Distribution

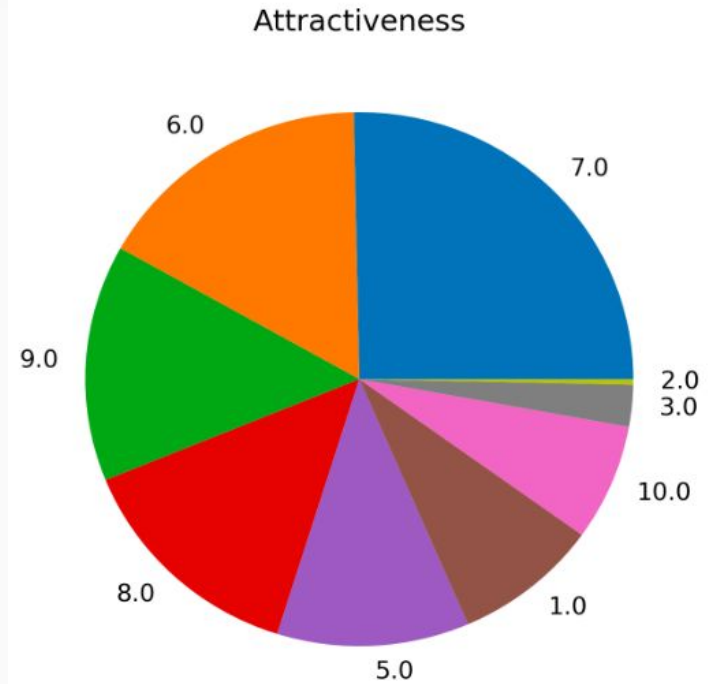
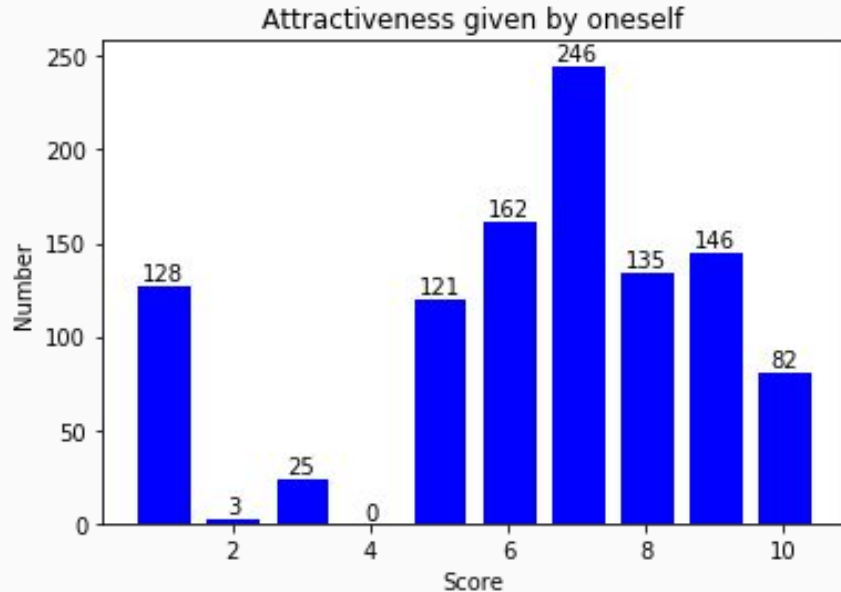


3. Interesting Points

Height

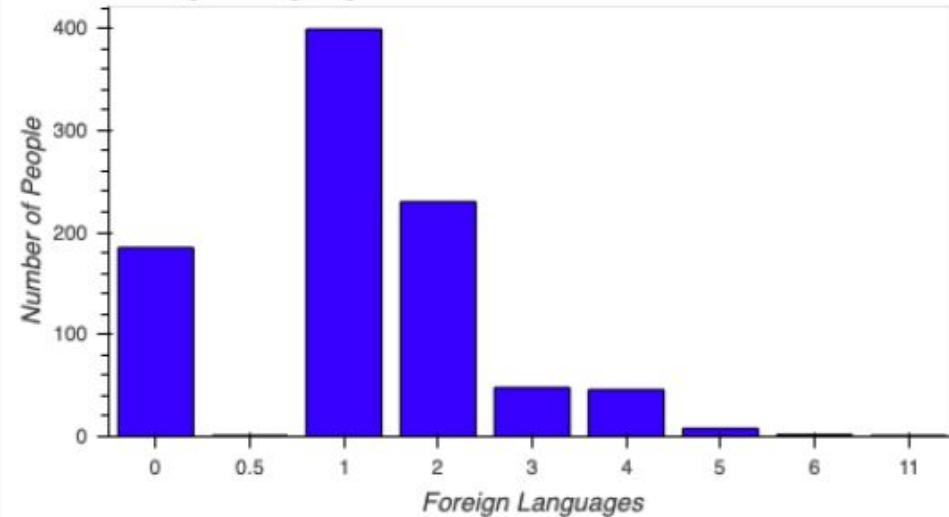


Rate own attractiveness from 1-10 realistically.

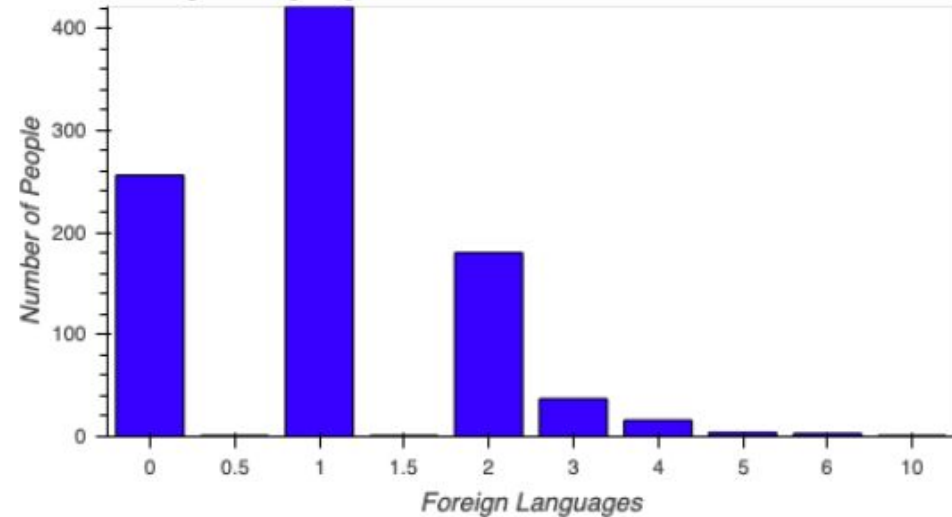


Foreign Languages Fluency

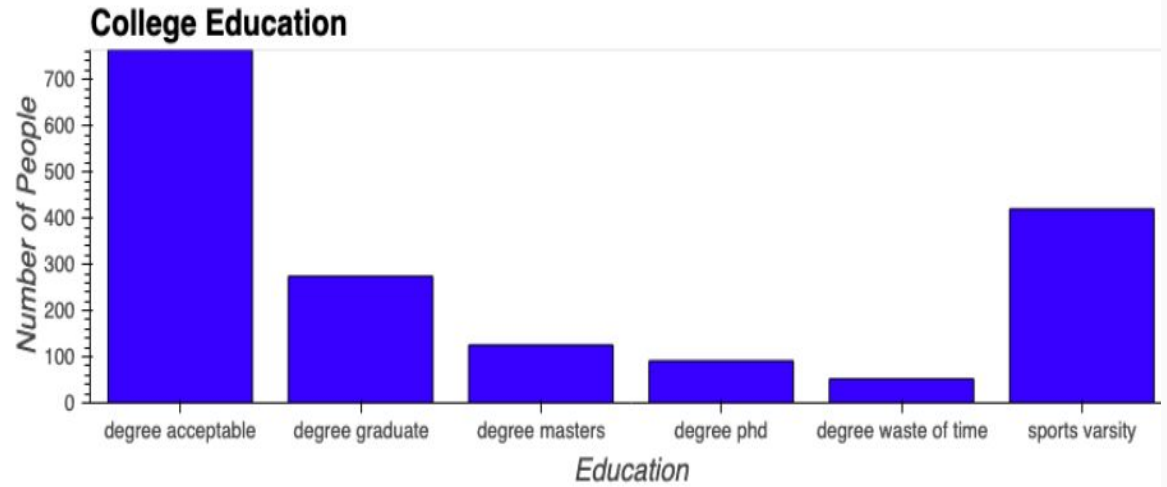
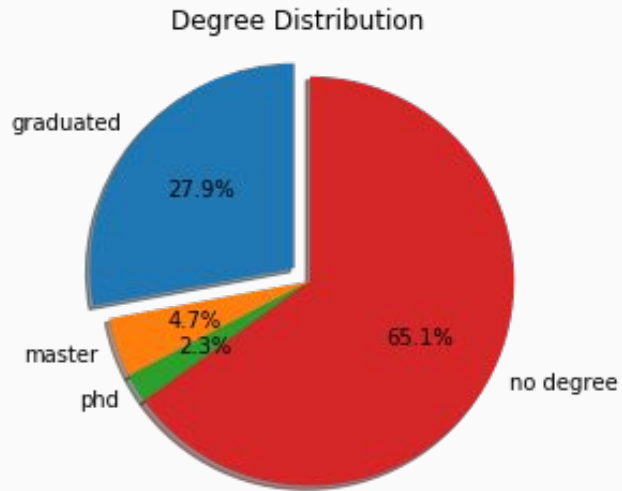
Foreign Languages Nonfluent



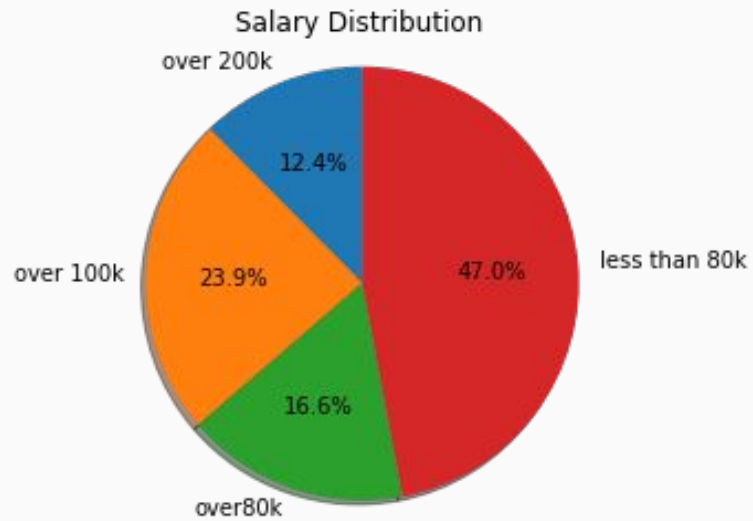
Foreign Languages Fluent



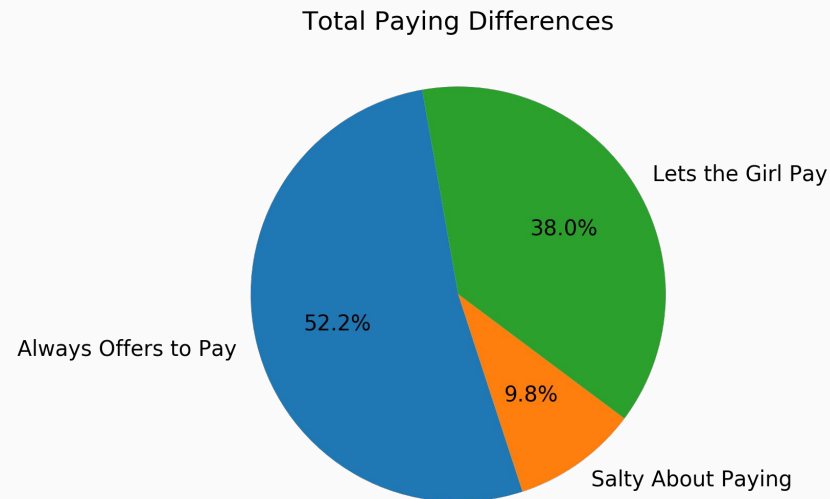
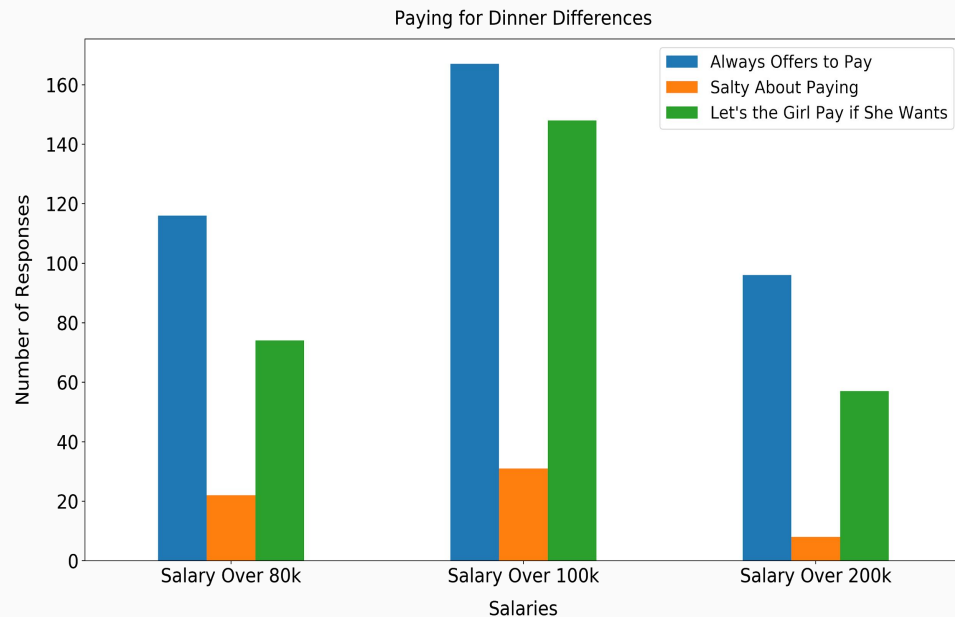
Education



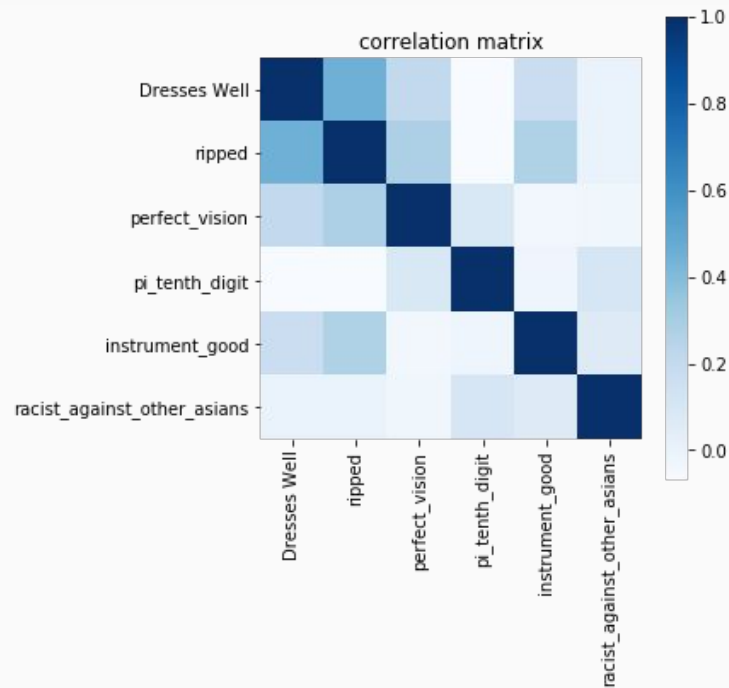
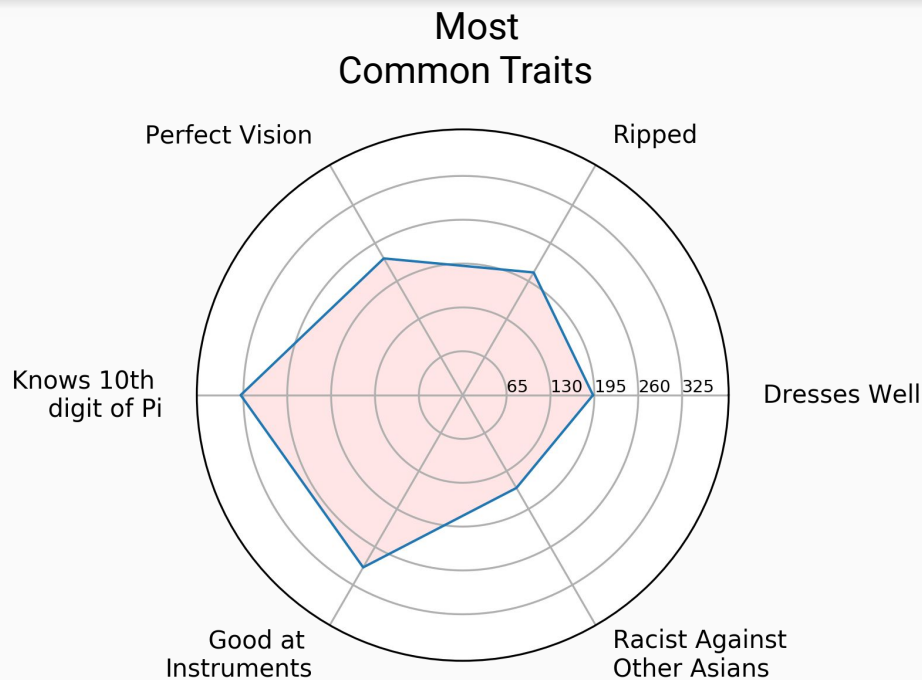
Salary



Willing to Pay?

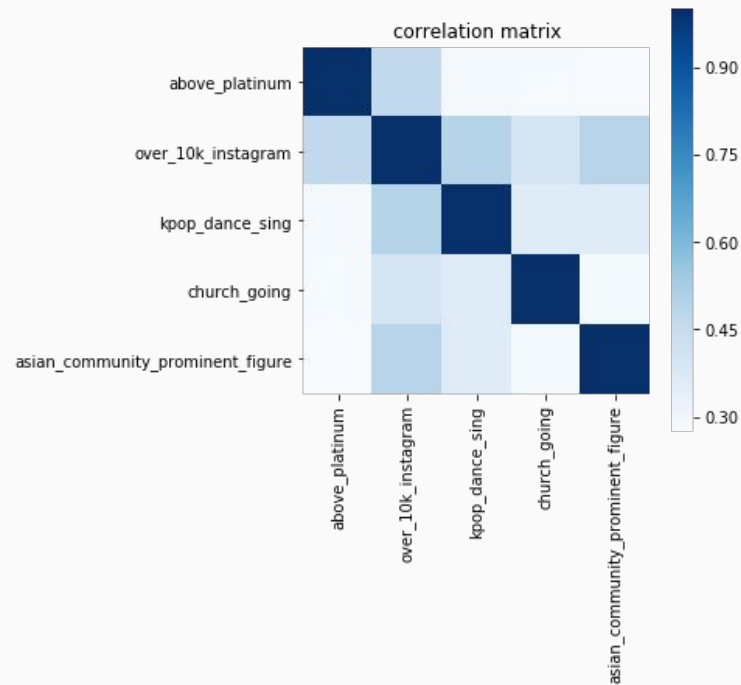
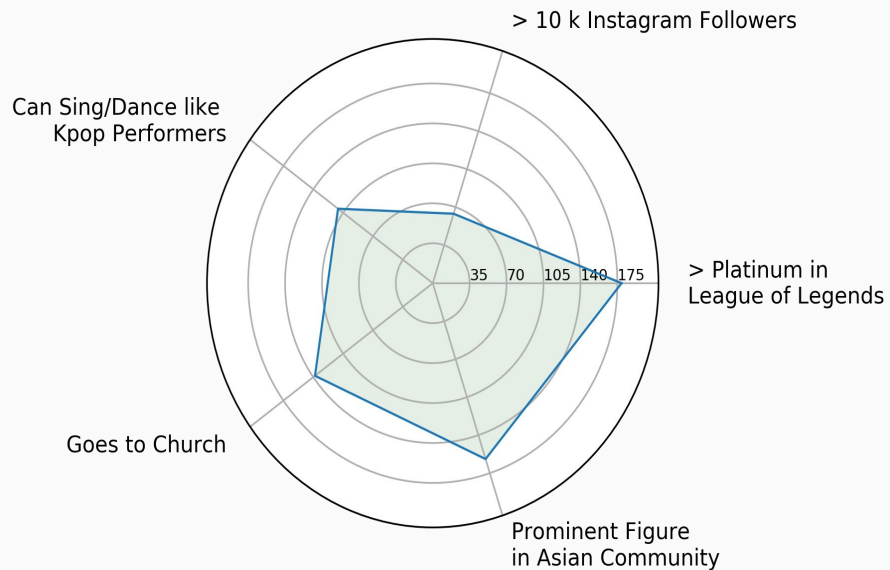


Subtle Traits



Social Life

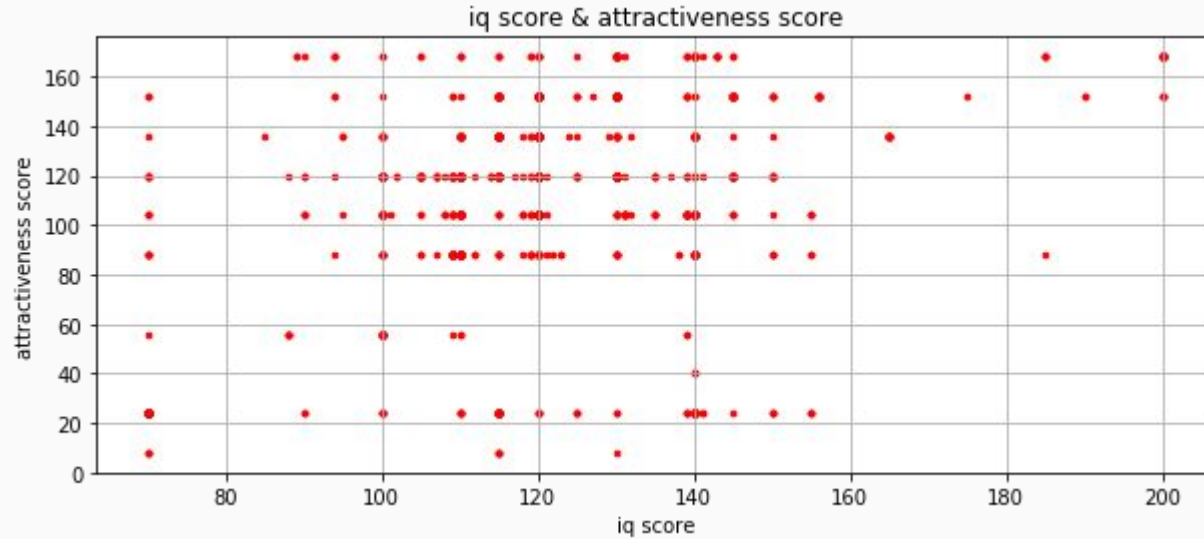
Most Common Social Life Traits



4. Deeper Analysis

More complex analysis

This left scatter shows the relationship between iq score and attractiveness score, we can see that most of the people are confident about their attractiveness, but not with their iq.

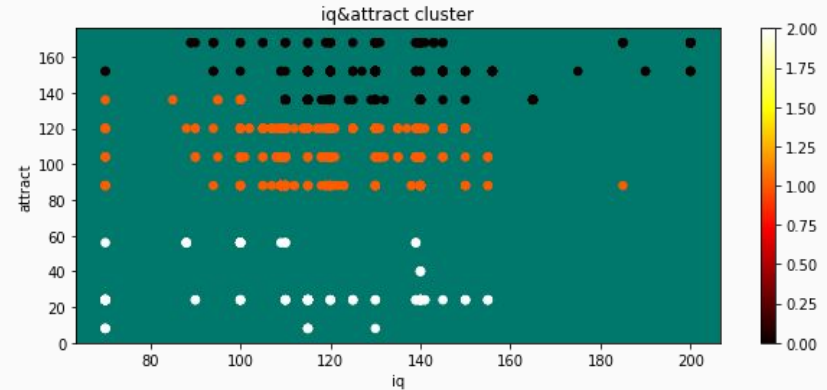
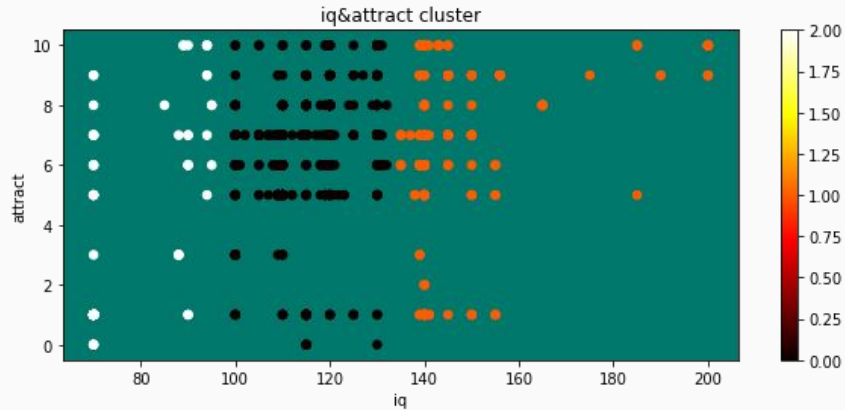


More complex analysis

If we use simple K-means Algorithm to cluster this data, we can get the result in right picture.

Here we use choose $K = 3$.

the value in x-axis and y-axis is crucial to the clustering result as you can see from the right 2 pics.

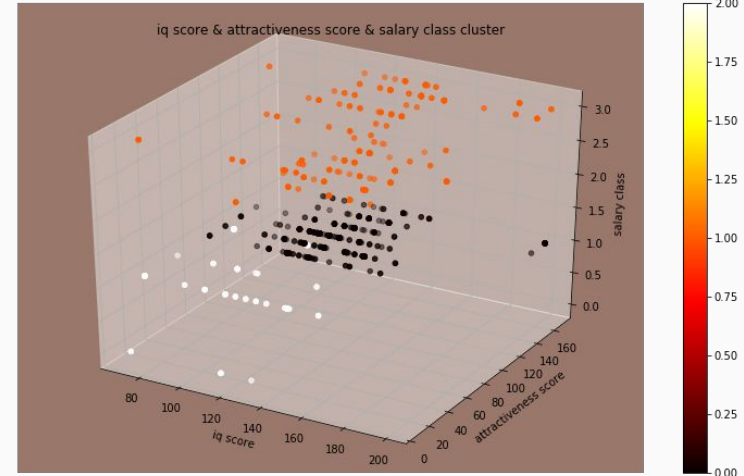
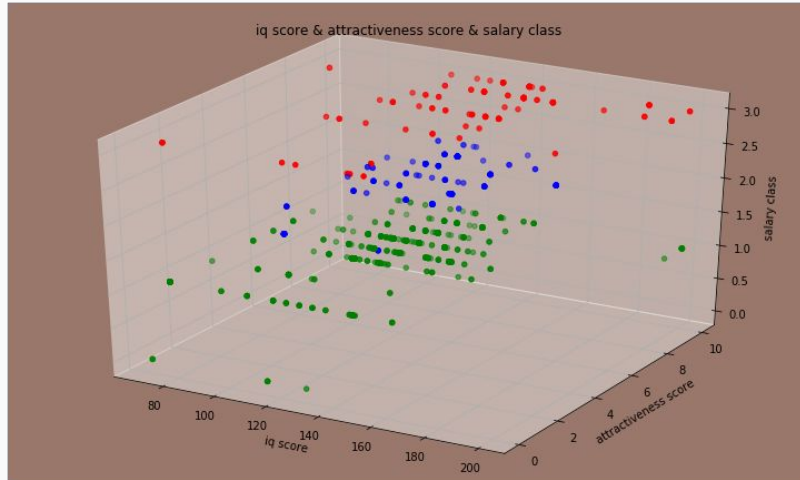


More complex analysis

What if we add another dimension of data? let's use salary

- Salary: >80k, >100k, and >200k ==>>> 1,2,3

Higher salary = Higher confidence in their IQ scores

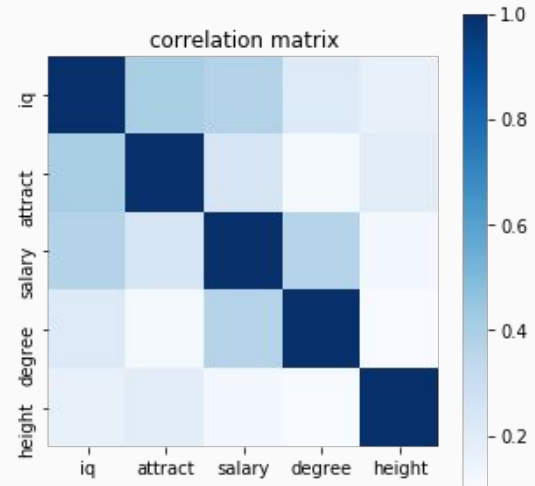
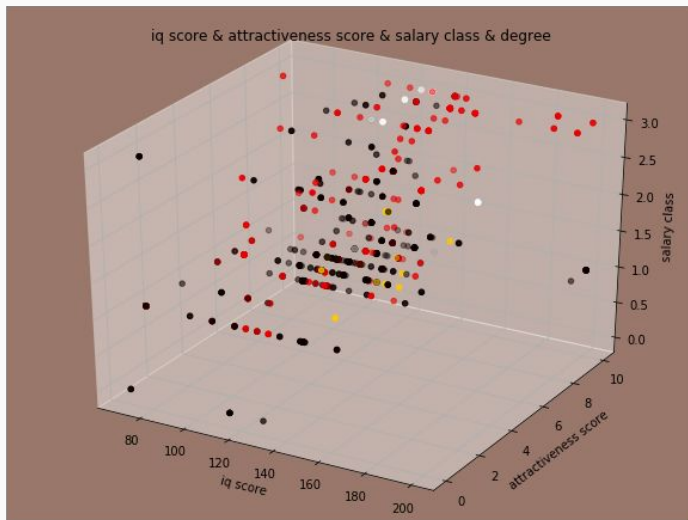


More complex analysis

Let's continue, this time we add degree.

In left picture, degree was divided into 0,1,2,3 classes corresponding to 'no degree, graduate, masters, phd. We can see that phd is very rare and is corresponding to white dot.

This time we using correlation matrix to see the relationship between these four factors



Thank You!
Happy Dating ;)