

Project cover sheet

Project title: What is the attribute that influences the most how people decide whether they want to see someone again romantically? Explore some other interesting questions with this data.

Group members: Anthony Nsimba, Donal Conlon, Jingchen Zhang, Robert White

All group members have read and agreed to the final version of all documents.

Signatures

1. Anthony Nsimba

2. Donal Conlon

3. Jingchen Zhang

4. Robert White

(If you cannot provide a digital or scanned signature, please type your name and send an e-mail to rafael.deandrademoral@mu.ie confirming that you have read and agreed to the final version of all documents)

First Dates: Analysis of Speed Dating dataset and discovering what influences love at first sight.

By Anthony Nsimba, Donal Conlon, Jingchen Zhang, Robert White

It seems the heart wants what the heart wants, and it can figure it out quickly too.

Research from Columbia Business School professors Ray Fisman and Sheena Lyengar gathered data on speed dating for over 500 subjects, both male and female. The combination of 4-minute dates, as well surveyed data containing categorical information such as education, income, interests, and career created interesting results with which we can analyse and draw insights as to what attributes influence whether males or females want to see each other again.

Authors Anthony Nsimba, Donal Conlon, Jingchen Zhang and Robert White have cleaned, wrangled, and analysed this dataset to quantify several hypothesis and research questions, namely:

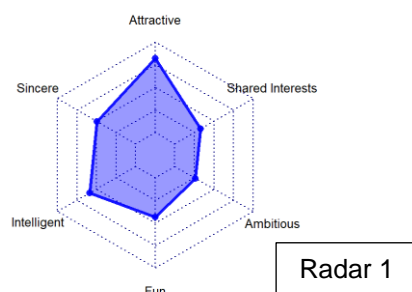
- What attributes males and females claim to look for in a potential date.
- What attributes do males and females perceive the other to look for in a partner?
- Create a linear model to investigate what attributes(s) are most influential in how people decide whether they want to see someone again romantically?
- Finally, we use the data collected to optimise a dating profile for matchmaking success.

What attributes do males and females claim to look for in a potential date?

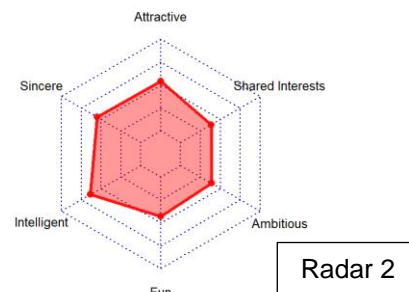
We first had to engage in a process of data standardization. As per the pre-event survey, candidates were asked to distribute their preference across six attributes in a partner; attractiveness, sincerity, intelligence, fun, ambition, and shared interests. However, scoring was not measured equally, with some subjects scoring over a 100-point scale, and many others over a 10-point scale. This discrepancy forced us to normalise data via each participants' scores for individual attributes, as a percentage of the sum of all attributes. We also removed any missing data, and split the data by gender, to highlight any inherent bias.

Radars are created by means of an empty data frame, populated by a nested loop.

Amount of influence attributes have in men seeking women



Amount of influence attributes have in women seeking men



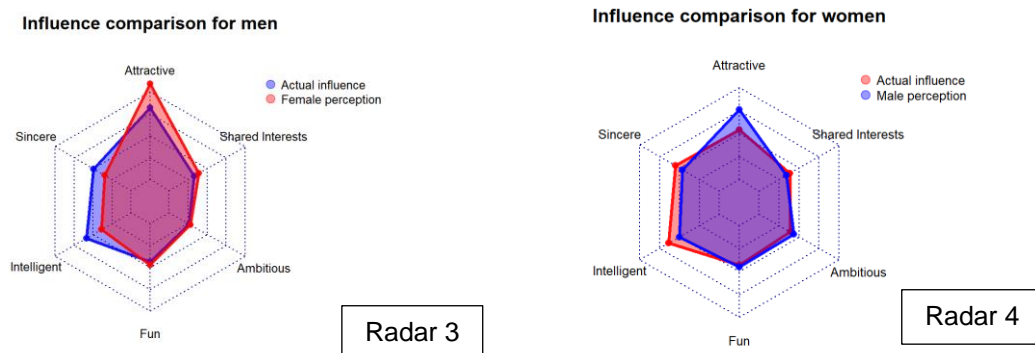
It is clear from radar 1 that men hold attractiveness as the most desirable attribute in a partner. Women on the other hand hold ambition in higher regard, as per Radar 2. Both genders held similar weights for fun, sincerity, shared interest, and intelligence.

Assuming there is a strong correlation between intelligence and wealth (those with higher degrees of intelligence tend to be more educated and have higher paying jobs), it can be concluded that men

look for more superficial qualities in their partners. Similarly, if the same assumption is made vis-à-vis intelligence == wealth, we can also conclude females are inclined towards men with status and power owing to the strong scores in ambition and intelligence.

What attributes do males and females perceive the other to look for in a partner?

No further data cleaning was needed following our previous graph. Below we hypothesize how each gender rated the attributes their partner following a match. Radar 3 explores the attributes males find most desirable against the attribute's females perceived them to find most desirable. While men do in fact place the most importance on attractiveness, it is not to the extent that women thought. Elsewhere, both sets of respondents placed equal weight on fun, shared interest, and ambition.



By the same token, Radar 4 illustrates a similarity in the weight of fun, shared interest, and ambition. Women hold greater importance on attributes of sincerity and intelligence and less importance on the attractiveness attribute than men thought.

It is interesting that both men and women overestimated the weight of importance of attractiveness and underestimated the weight of importance of sincerity and intelligence for the other gender. Perhaps this shows the skewed perspective of love and romantic attraction in today's society.

The results of the radar plots seem to align with societal norms. However, it is important to note that while the participants tended to respond in this manner, their actual response in the speed-dating experiment may not have been reflective of what they had initially said. Thus, further analysis must be performed on the actual data from the speed dating events themselves, in order to confirm their original thoughts or portray a cognitive dissonance.

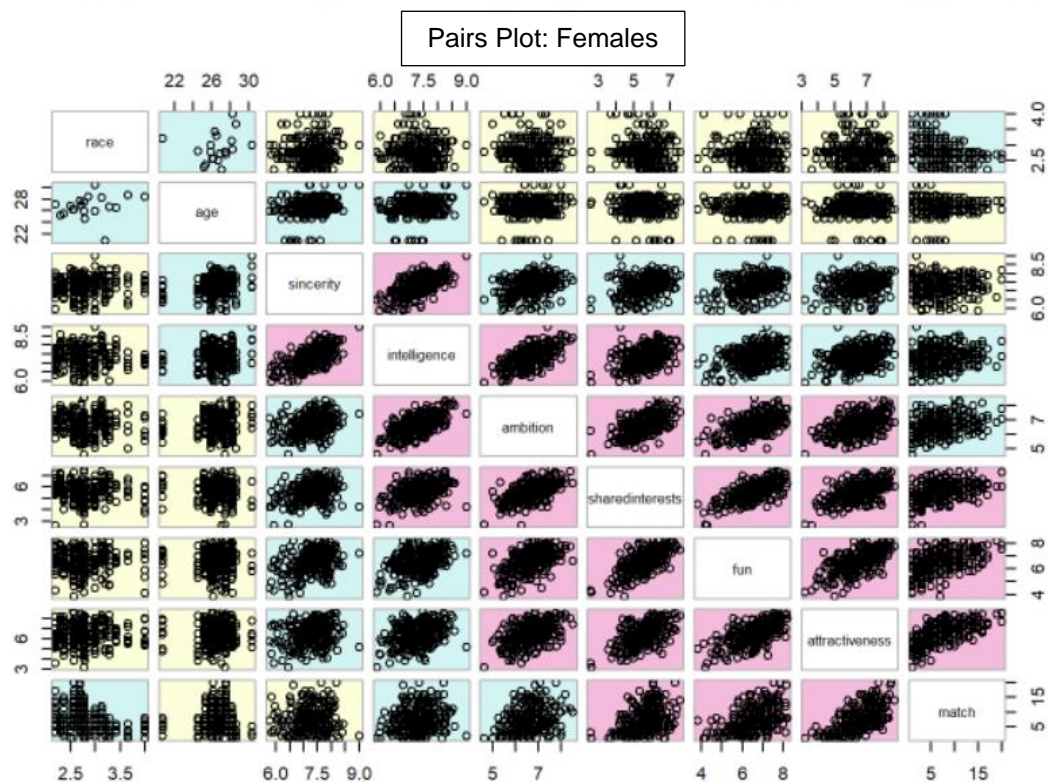
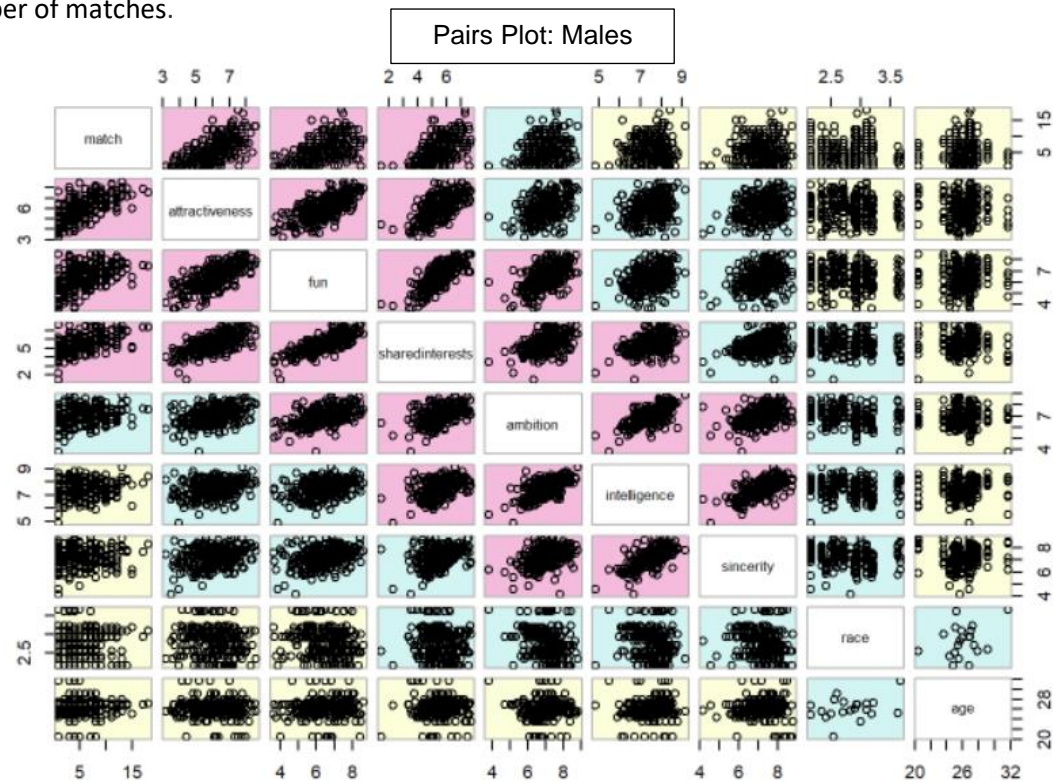
Linear Model - Discovering the statistically significant attributes that led to a match.

Here we ran a linear regression model on the ratings compiled by daters about each partner on the night, to assess what attributes were actually significant in striking lucky or striking out.

To do this, the 'lm_model_data' subset was created by calculating the mean rating of the variable's attractiveness, sincerity, intelligence, fun, ambition, shared interests, age, and race, and by calculating the number of matches per subject. The subjects with zero matches were removed from

the subset. The data was then split by gender, to investigate if men or women had different preferences.

The next step for both genders was to construct a pairs plot to examine any linear relationships. Attributes with a pink background represent attributes with a strong linear association with the number of matches.



We then used the 'allfits' function from the R package 'Leaps', to examine the best regression for each number of predictors in the model and calculate the r-squared for each number of predictors to get the best model fit. The allfits function also told us what attributes to include for a model with each number of predictors.

The model with eight predictors was chosen as it had the highest R-squared value and gave us the most information about our data. All model assumptions are met with both models below.

Regression Coefficients: Males

```
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.1973     3.7581  -0.319  0.75030
## attractiveness  2.3925     0.2316  10.330 < 2e-16 ***
## sharedinterests -0.1027     0.3315  -0.310  0.75691
## sincerity      -1.4769     0.4507  -3.277  0.00119 **
## race           -3.5352     0.4690  -7.538 7.86e-13 ***
## ambition       -0.8638     0.3911  -2.209  0.02806 *
## intelligence    0.7701     0.5178   1.487  0.13813
## age            0.2927     0.1205   2.428  0.01585 *
## fun            1.0028     0.3358   2.986  0.00309 **
```

Attractiveness, fun and shared interests had the highest correlation with the number of matches. Ambition had a moderate correlation and intelligence, sincerity, race, and age had a weak correlation.

From the model summary, attractiveness and race were again the most significant factors in determining a match. Fun and sincerity had a larger significant effect than the female counterparts and ambition and age are also significant factors.

Similar to males, attractiveness, fun and shared interests had strong positive correlation with the number of matches. Ambition, race, and intelligence had moderate correlation and age and sincerity had weak correlation. All variables were included in the model as adding this attributes explained more variation in the model.

Regression Coefficients: Females

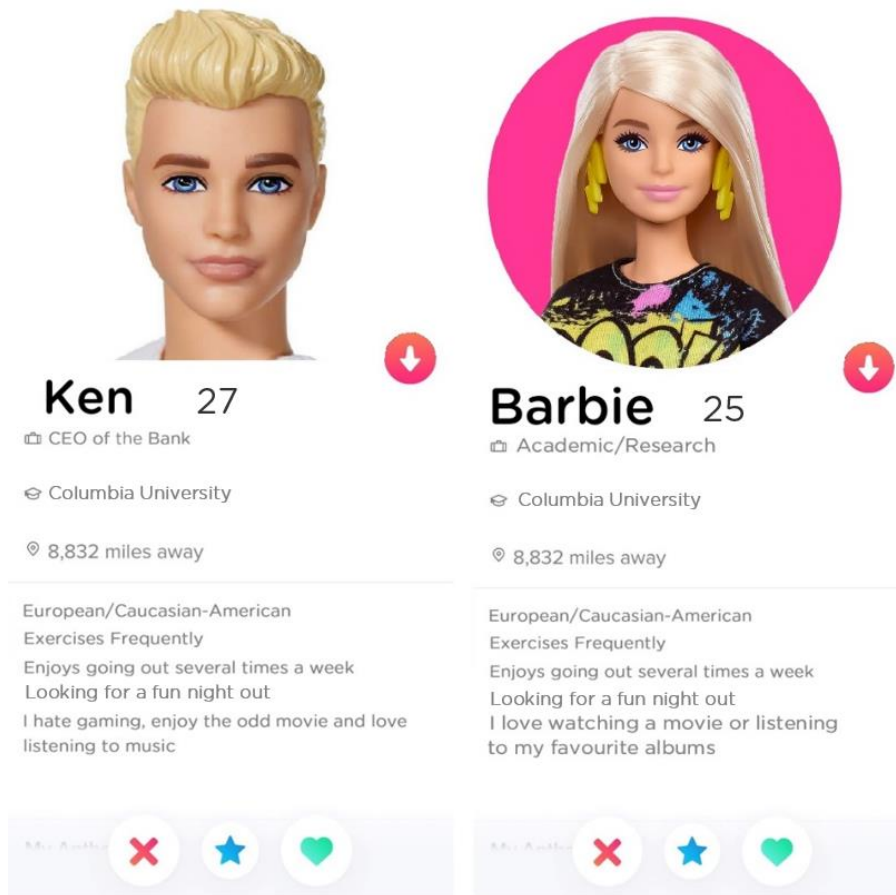
```
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.1973     3.7581  -0.319  0.75030
## attractiveness  2.3925     0.2316  10.330 < 2e-16 ***
## race           -3.5352     0.4690  -7.538 7.86e-13 ***
## sincerity      -1.4769     0.4507  -3.277  0.00119 **
## fun            1.0028     0.3358   2.986  0.00309 **
## age            0.2927     0.1205   2.428  0.01585 *
## ambition       -0.8638     0.3911  -2.209  0.02806 *
## intelligence    0.7701     0.5178   1.487  0.13813
## sharedinterests -0.1027     0.3315  -0.310  0.75691
```

Females rated attractiveness and race as the most significant factors where they got a match. A unit increase in the mean attractiveness rating would lead to a predicted increase of 2.33 matches when the other attributes remain unchanged. Sincerity, fun, ambition, and age are other significant factors for females.

Ken & Barbie: Creating the optimum dating profile

As a final avenue of interest with the dataset, we sought to optimise a dating profile based on all the relevant information gathered. The hypothesis here hinges on the fact that attractiveness is something largely hereditary and genetic, and all other attributes are deeply rooted in personality, far from malleable or flexible to each prospective date. But what if we were to engineer a profile with the most popular 'categorical' data?

In order to find the most dateable male and female, we began by omitting all the 'undatable's' – the subjects who failed to find a match. This subset was further split into males and females before finding the most popular answer in a range of variables that we deemed suitably flexible.



The results of our query have furnished us with two fictitious characters we have aptly named – ‘Ken & Barbie’.

The resulting characters certainly paint a slightly one-dimensional view of the optimum profile; however, such are the most frequent characteristics of the ‘matched’ subjects in this experiment.

Conclusion

Attractiveness is the main attribute that influences whether both males and females want to see each other again. This is the most significant factor flagged by subjects prior to the experiment, and exhibited by subjects during the experiment. Race, sincerity, fun age, and ambition are also significant factors.

Perhaps attraction to a person is heavily dependent on appearance and beauty, but people are not as honest with themselves in admitting so directly. Yet the data shows that beauty and physical attraction tends to be quite important in finding a romantic partner.