Formula for Good First Impressions

Adam Lininger-White

4/7/2020

I INTRODUCTION

First impressions are tricky, some are strong while others are awkward but most are exceedingly average. Yet we all know somebody who makes a great first impression on everyone they meet. These blessed individuals are effortlessly charming and endearing while also making strangers feel good about themselves. In his classic book on interpersonal relations *How to Win Friends and Influence People*, Dale Carnegie argues that this skill is the greatest determinant of professional achievement and a fulfilling life. Masters of the first impression rule the world. I want to know what about these people makes us like them so much, and I found the beginnings of an answer in statistical analyses of speed dating.

Researchers at the Columbia Business School had a similar question to mine. They wanted to understand what constituted a good first impression in the world of romantic relationships. After each speed date, they measured how well the interaction went using a binary, categorical response variable: "Would you like to see this subject again, yes or no?" Participants scored each partner on several attributes and provided information about themselves in a pre-experiment survey. Using this data, researchers built linear regression models and analyzed their coefficients to determine which attributes were most important for instilling interest and leaving a good impression.

Two separate studies utilized this dataset, and both clustered subjects into groups before building models in order to analyze differences in attribute importance. One study found that women put greater weight on the intelligence and race of their partner, while men respond more to physical attractiveness [Fisman et al., 2006]. Moreover, men do not value women's intelligence or ambition when it exceeds their own. They also found that women exhibit a preference for men who grew up in affluent neighborhoods. The second study found that subjects' backgrounds, including the racial composition of the ZIP code where a subject grew up and the prevailing racial attitudes in a subject's state or country of origin, strongly influence same-race preferences [Iyengar et al., 2008].

In this paper, I perform a similar analysis with the main difference that I predict a subject's general fondness for their partner rather than their romantic interest - how much they *like* the stranger across the table as a person. To my luck, the Columbia researchers tracked this exact variable but never used it for model building or analysis. I performed linear regression to predict such fondness and found: (1) attractiveness is the most important determinant, (2) overlapping interests and perceived funness are tied for second, (3) belief the other person likes you is somewhat important, (4) intelligence and sincerity are equally important but relatively not very impactful, and (4) ambition has an insignificantly negative impact.

$$\text{Like}_{ij} = (0.30)\,\text{Attractive}_{ij} + (0.20)\,\text{Fun}_{ij} + (0.20)\,\text{Shared Interests}_{ij} + (0.12)\,\text{P(partner is interested)}_{ij}$$
$$+ (0.09)\,\text{Sincerity}_{ij} + (0.08)\,\text{Intelligence}_{ij} + (-0.02)\,\text{Ambition}_{ij} + \epsilon_{ij}$$

* subscripts indicate the score person i gave to person j on a particular attribute.

The remainder of the paper is structured as follows. Section II describes the data used for model building and how it was collected. It also briefly addresses assumptions about the data and drawbacks that limit the scope of my results. Section III outlines how the final model was selected, analyzes its appropriateness, and uses it to answer interesting questions and address pieces of conventional wisdom. Section IV concludes.

II METHODOLOGY

Data collection for the Columbia Speed Dating Experiment took place between 2002-2004. Subjects were drawn from students in graduate and professional schools at Columbia University and were recruited through a combination of mass email and fliers posted throughout the campus and handed out by research assistants. The speed dating events occurred in a popular restaurant near campus, and factors like lighting, table arrangement, and music taste and volume were constant across sessions. Upon arrival, participants received a clipboard, pen, and nametag with their unique ID number. Each clipboard included a score-card with a cover over it so that participants' responses would remain confidential. The score-card was divided into columns in which participants indicated the ID number of each person they met. Participants would then circle "yes" or "no" under the ID number to indicate whether they wanted to see the other person again. Beneath the Yes/No decision was a listing of the six attributes on which the participant had to rate his or her partner on a 1–10 Likert scale: Attractive, Sincere, Intelligent, Fun, Ambitious, Shared Interests Following these attributes were two full questions also answered with scores between 1-10: "Overall, how much do you like this person? (1=not at all, 10=a lot)" and "How probable do you think it is that this person will say 'yes' for you? (1=not probable, 10=extremely probable)" [Iyengar et al., 2008].

At the start of each session, hosts asked participants to sit at two person tables with all women on one side and men on the other. Each speed date lasted four minutes with one addional minute at the end to answer survey questions, then each man rotated until they'd met all the women.

By combining partner ratings with pre-experiment survey responses, researchers were able to create rich profiles for both participants in every speed date. All in all, the final dataset measured 195 variables across nearly 7000 dates. I found it messy and overwhelming. Other data scientists agreed and several published cleaned, simplified versions on open-source online platforms. I'm using one created by Keith McNulty, accessed via his GitHub linked here. All salient variables from the original dataset were retained, as well as some less important ones.

```
head(df)
```

```
##   gender age income goal career dec attr sinc intel fun amb shar like prob met
## 1      0  21  69487    2 lawyer   1    6    9     7   7   6    5    7    6   2
## 2      0  21  69487    2 lawyer   1    7    8     7   8   5    6    7    5   1
## 3      0  21  69487    2 lawyer   1    5    8     9   8   5    7    7   NA   1
## 4      0  21  69487    2 lawyer   1    7    6     8   7   6    8    7    6   2
## 5      0  21  69487    2 lawyer   1    5    6     7   7   6    6    6    6   2
## 6      0  21  69487    2 lawyer   0    4    9     7   4   6    4    6    5   2
```

I did some extra cleaning by removing the *income*, *goal*, and *career* variables because they mostly contained missing or inconsistently formatted data. I chose to remove *gender* when building my full model so that I could analyze the broadest, most universal ideal of a good first impression. One couldn't adopt a different demeanor for the umpteen combinations of a stranger's identity categories, so my model gives the best option if only one choice is available. Later, in my data analysis section, I reintroduce the *gender* variable and compare coefficients between genders.

```
names(df.final)
```

```
## [1] "like"  "attr"  "sinc"  "intel" "fun"   "amb"   "shar"  "prob"  "age"
```

Our response variable is the general fondness score *like*, which is continuous between 1 and 10 (inclusive) with higher values indicating greater fondness. Variables two through six are continuous over the same range and indicate how each attribute impacted the subject's overall impression (1 = awful, 10 = great). *Prob* is the survey filler's guess at the probability their partner is interested in them (1 = not likely, 10 = definitely). Almost all ratings were given in whole numbers, but some participants gave fractional scores.

Scores for each speed date were given confidentially, and participants never had the opportunity to discuss their rating systems with one another during the session. We can reasonably assume that the outcome of one speed date did not impact data for other dates, so the cases in this experiment satisfy the assumption of

independence between errors. There are, however, issues in the data we must address. Subjects voluntarily chose to participate in this study and were not randomly selected from the population. This undermines the representativeness of our sample, and makes it unclear how broadly our results apply. Looking at the demographic breakdown of participants, we see a diverse set of greographical, socioeconomic, racial, and academic backgrounds represented. This makes us more confident in the strength of the results, but recognizing that all participants are Columbia University students diminishes it again. All this is to say that the scope where my conclusions are relevant is difficult to define. They certainly apply to Columbia University students, and probably all enrollees at elite academic institutions, but I don't think they can be generalized to all young adults.

I also worried that this study only analyzed relationships between people of the opposite sex, which would be a woefully incomplete dataset for my purposes. Thankfully, the Columbia researchers were wise enough to conduct additional speed dating trials where all participants identified as either male or female. This information allows me to generalize my results to almost all interactions between two people. I lament the lack of data for folks who identify as non-binary. They are not represented, which again limits the scope of my results and analysis.

III Data Analysis.

In this section, I first outline how I arrive at my final model before using it to draw interesting conclusions and comparing different models between groups. Full code for all model building, model comparisons, graphics, and subsequent prediction can be found in this project's GitHub repository.

At the outset, I knew only that I wanted to perform regression through the origin - if a person recieves a zero for every attribute, then their overall rating should be zero. I decided to fit a model with three different techniques (maunal ascent, AIC backward selection, and LASSO), then compare their relative merits and choose one to fully implement. My job became easier when the manual ascent algorithm and automated AIC procedue returned the same model: both told me to regress on all available predictors. Using the glmnet package and its built in functions for LASSO, I found that $\lambda \approx .113$ produced the sparsest model whose standard error was still similar to the most accurate model. LASSO removed one variable *amb* from the set of utilized predictors; it's coefficient is almost zero in the full model. All three models had very similar coefficients, as shown below.

```
##            attr  sinc intel   fun    amb  shar  prob
## 1 Ascent 0.300 0.091 0.090 0.207 -0.019 0.205 0.123
## 2    AIC 0.300 0.091 0.090 0.207 -0.019 0.205 0.123
## 3  LASSO 0.277 0.078 0.065 0.202  0.000 0.190 0.096
```

I compared the residual plots for the full model and LASSO model and noticed no visually significant difference, so I elected to use the bigger model proposed by AIC and manual ascent (two votes to one majority). When building my model with the manual ascent procedure, I added predictors in the following order (from first to last): *fun, attr, prob, shar, sinc, intel, amb*.

Now that we understand where the final model comes from, we must answer the question, "Is a linear model appropriate for this prediction task?" To do so, we must prove the assumptions of linearity, constant variance, and normality of errors are true for our data. Looking at a scatter plot of residuals against fitted values, the e_{ij}'s appear evenly and randomly distributed for the middle portion of our predictions, but these assumptions get more suspect towards the tails. The same is true for the normal probability plot - the middle region closely tracks with the line but the ends slightly diverge. A reasonable explanation justifies these patterns: values of the *like* variable are restricted to the interval $[1, 10]$, so smaller $\hat{y}_i$'s are more likely to underpredict and have bigger residuals while bigger $\hat{y}_i$'s will tend to overpredict and have smaller residuals. Also, the vast majority of observations exist in the middle section of our residual plot, so it makes sense that the vertical spread over this region would be slightly larger. With these realizations in mind, it's reasonable to conclude that the response variable, fondness, is linear with constant variance and normal distribution of error terms. I considered a Box-Cox power transformation of the response variable but found it only weakly strengthened these assumptions, and the loss in interpretability outweight any gains.
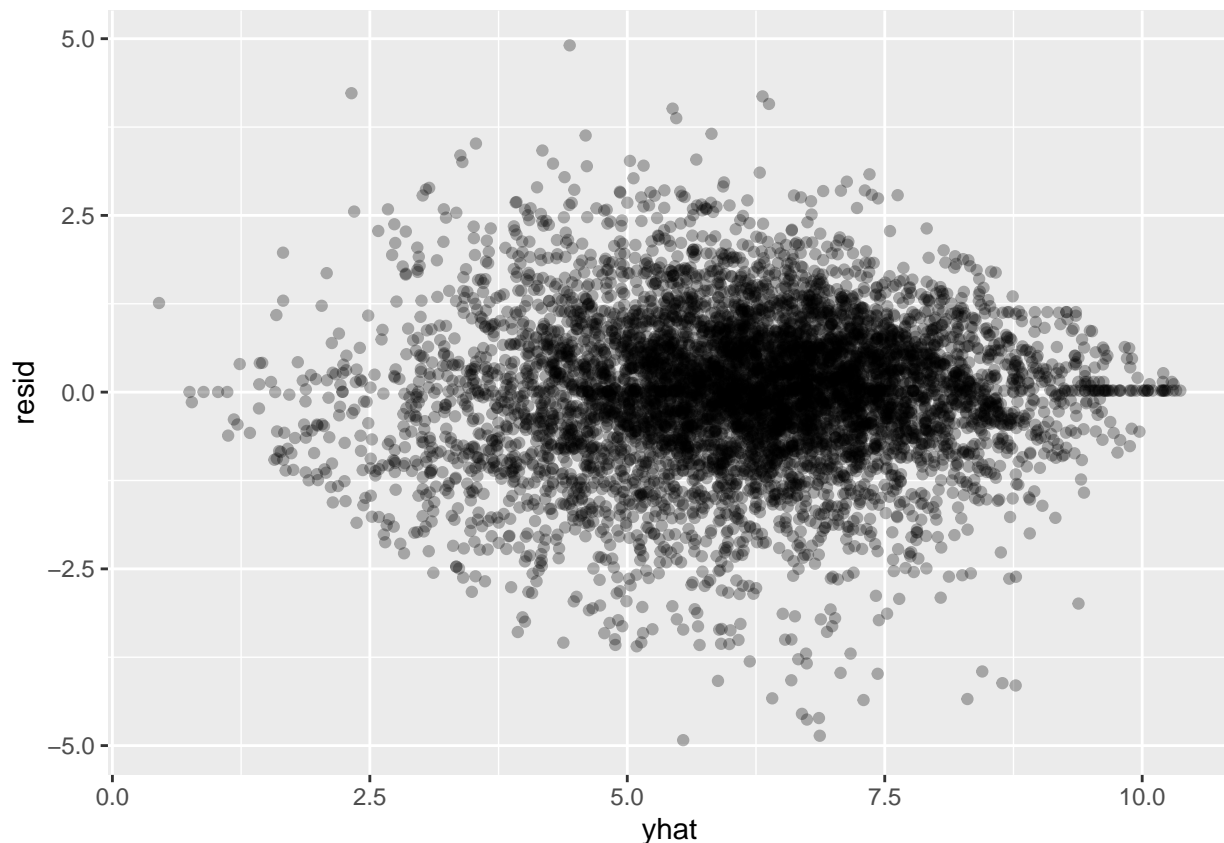
So we've seen the final model, know how it was built, and know it's appropriate for predicting how much a grad student will like someone they just met given a scoring of the stranger's attributes. Now, it's time to peek under the hood and see how these attributes determine fondness. As a refresher, here's another look at the model:

$$\text{Like}_{ij} = (0.30)\,\text{Attractive}_{ij} + (0.20)\,\text{Fun}_{ij} + (0.20)\,\text{Shared Interests}_{ij} + (0.12)\,\text{P(partner is interested)}_{ij}$$
$$+ (0.09)\,\text{Sincerity}_{ij} + (0.08)\,\text{Intelligence}_{ij} + (-0.02)\,\text{Ambition}_{ij} + \epsilon_{ij}$$

What jumps out immediately is the outsized role that three attributes play in determining Like_{ij}. Attractiveness is 1.5x more important than the second most significant predictors (funness and shared interests), which are nearly twice as powerful as the next closest variable. With scores of 10 for these three attributes and zeros in all others, one would recieve a fondness score of 7/10 - not bad at all. Thus, these are the keys for making a good impression. Sadly, this doesn't help much if one is looking for ways to improve. Maybe we can slightly improve our physical appearances by exercising regularly and dressing well, but good looks really aren't in our control. Same goes for sharing interests, assuming both partners are being genuine, one can actively look for common ground but may find none exists. So, instead of focusing on the most important predictors, let's concentrate on attributes we can control. Specifically, let's find the ones with greatest potential for improvement (on average).

First, let's consider the nonexistest, perfectly average person in our study.

```
##      like    attr     sinc    intel      fun      amb     shar     prob
## 1 6.137515 6.19762 7.170828 7.36839 6.400558 6.766965 5.46181 5.259621
```

If person j wants to improve the overall likeability score person i gives then, he or she could realistically focus on being more vivacious (*fun*), genuine and honest (*sinc*), eager to talk about complicated topics (*intel*), and outwardly friendly towards their partner (*prob*). I've chosen to ignore *amb* because its barely nonzero, and I never want to be on the side of tellig people to limit their ambitions. Our question is, "For the average

person, which of these attributes should we maximize to provide the biggest bump in likeability?" To start, we augment the perfectly average subject by replacing one of their attributes with a 10, predicting their new likeability score (CI), then setting the attribute back to its mean before flipping another and repeating. After, we compare all the newly predicted *like* scores and see which are greatest.

```
##          Variable      fit      lwr      upr
## X1      Max fun 6.889086 6.816149 6.962022
## X1.1  Max sinc 6.400735 6.338730 6.462739
## X1.2 Max intel 6.381332 6.313903 6.448760
## X1.3  Max prob 6.728387 6.660018 6.796756
```

One major cause of good first impressions is luck, getting dealt a good hand of cards. In terms of attributes one can actually work on to improve their interactions with strangers, the following are clear takeaways: (1) try hard to relate to new people and build a mutual sense of understanding, (2) bring energy and positivity to these interactions, and (3) show more friendliness and affection than normal.

In the Columbia Business School studies, researchers found differences in attribute importance for mate selection across categories like race and gender. Naturally, one might wonder whether disparities exist between these groups for predicting general fondness as well. To answer these quetions, we split subjects into groups based on some facet of their identity, build separate linear models for each group, and compare their coefficients. For simplicity, I used glmnet's LASSO to build the following models and derive their β coefficients. Stark differences exist between the attributes valued by men and women. Attractiveness remains the strongest determinant for both groups, but it also has the biggest difference in betas; its impact on men is far greater than the next closest variable, but women consider characteristics like funness and shared interests to be nearly as significant. Other noteworthy differences exist between intelligence and sincerity, both of which have stronger effects in women. Interestingly, the second biggest difference in betas occurs for *prob*, the probability the other person likes you, which men respond to more strongly than women.

```
##   Variable Beta-Male Beta-Female
## 2     attr     0.314       0.234
## 3     sinc     0.040       0.099
## 4    intel     0.044       0.098
## 5      fun     0.190       0.217
## 6     shar     0.171       0.206
## 7     prob     0.136       0.063
```

I performed a similar analysis for age, pulling the oldest and youngest quartiles of participants into separate groups before fitting models using LASSO. The printout below shows the resulting coefficients for each model. We see that younger grad students are more impacted by attractiveness and shared interests, while their older counterparts respond to funness and sense of reciprocated interest.

```
##   Variable Beta-old Beta-young
## 2     attr    0.260      0.299
## 3     sinc    0.106      0.086
## 4    intel    0.051      0.063
## 5      fun    0.219      0.167
## 6     shar    0.128      0.227
## 7     prob    0.103      0.027
```

We also might want to know more about the instances in our dataset where someobdy made a really good first impression. What combination of attributes were displayed, on average, when a participant recieved a 9 or 10 for the *like* variable? Below is a printout of the mean score for each attribute when this was was the case. Interestingly, the predictors sincerity and intelligence, which didn't have large coefficients in the final model, have the highest average scores. This indicates that people who make good first impressions in the real world, usually are very genuine and bright. Hopefully, that redeems some faith that humans aren't as superficial as our other analysis suggests. Another seemingly strange feature of this printout is that the mean scores are all < 9 even though the *like* score is ≥ 9. Again, this is because scores are bounded above by 10 so the distribution for a given attribute will be skewed down.

```
##   Center Measure     attr     sinc    intel      fun      amb     shar     prob
## 1         Mean 8.651677 8.883629 8.956607 8.763314 8.491124 8.151874 7.298817
## 2       Median 9.000000 9.000000 9.000000 9.000000 9.000000 8.000000 8.000000
```

IV CONCLUSION

If someone makes a good first impression on you, it may be difficult to put into words why you like them - maybe you'll call it a "gut feeling." In building my model, I wanted to peek into society's collective subconcious and determine the biggest factors in leaving a good first impression. I utilized a dataset compiled by a Columbia Business School study of speed dating that had datapoints for both same sex and opposite sex interactions. I did not separate subjects into groups before building my final model because I wanted to know what attributes are most important across interactions between all of people.

I find that two of the strongest predictors of a good impression, attractiveness and shared interests, are somewhat out of one's control. For the predictors one can affect, the average person can most improve their average likeability across all interactions with strangers by maximizing their energy/poisitivity and their outward friendliness towards the person. While it's impossible to create commonality where none exists, trying hard to find shared interests with new acquaintances is a great strategy for leaving good impressions. Subgroups of the population place different values on attributes when determining their fondness for an unfamiliar person. Men respond more positively to attractiveness and reciprocated interest than women, who place higher value on funness and common ground. Younger grad students react more strongly to shared interests than older grad students, but surprisingly the opposite is true for funness. While intelligence and sincerity didn't have large coefficients in our final model, these were the percieved best attributes, on average, of people who made a good first impression in the study.

These results apply to graduate school students at elite academic institutions, but everyone should be interested in the numbers behind likeability. Improving one's interpersonal skills can be a gateway to greater happiness and professional success.