Distinguished Factors in Evaluating Speed Dating Preferences

Zihan Wan, Julie Zhang, Yue Zhang

March 2019

Abstract

We study the dating behavior using data collected from a Speed Dating experiment conducted by professors at the Columbia Business School, where random matchings of subjects were generated. The design allows us to directly observe individual decisions and preferences rather than just final matches. Previous research has focused on gender and racial differences and how this may affect the decisions an individual makes in the speed dating event. Our main questions focus on what characteristics of a partner will impact a subject's ratings of said partner in a speed date, and what characteristics of an individual may lead to them changing their standards for partner selection. Using linear regression methods with fixed effects, we find that people of African-American race are the most likely to receive the highest ratings, and that older people receive lower ratings by about 1 point for every year. We also find that after participating in speed dating, people who work in the Natural Science and Business fields change significantly differently than people who study Law. A secondary analysis using logistic regression was done on how these ratings and gender affect the decision. We find that males put significantly more weight on attractiveness, while females put more weight on sincerity. Our analysis highlights some aspects of how people choose their partners through speed dating.

1 Introduction

Constructing a family by marrying the loved one is always one of the most important things of an individual. However, people may think differently when they are seeking true love. Professors Ray Fisman and Sheena Iyengar from Columbia Business School have compiled data from a speed dating experiment and used regression to analyze how gender may affect how people choose their significant others. Their results are found in the paper "Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment", published in The Quarterly Journal of Economics (May 2006) [1]. The authors of this paper set up and conducted a speed dating experiment, where the participants voluntary attended. Each attendee had to fill out multiple surveys about themselves, along with their feedback of their speed date partners. In particular, they rated their partners on certain attributes and made a decision on whether they would like to see this person again. In this paper, they conclude females tend to focus more on a partner's intelligence and race, while males tend to focus more on physical appearance.

The purpose of this project is to analyze the same data set in different aspects than the paper. The paper mostly focused on the important factors that determine decision. We would like to investigate what the important factors are in an individuals speed dating preferences. To do so, we will attempt to answer three questions.

Our first area of focus is on how the ratings of six attributes (Intelligence, Ambition, Sincere, Fun, Shared Interest, and Physical Attractiveness) are affected by the intrinsic characteristics of the partners, such as their age and race. We also want to understand how participating in speed dating can change a person's mindset on their mate selection criteria. We will look at many characteristics of an individual, from the basic age, race, field of study, to how happy they expect to be at the speed dating event. Using these characteristics, we want to see which factors make people more likely to change their mind on what they find important in people of the opposite sex after they have completed the speed dating event. As a secondary analysis, we will use logistic regression to answer a similar question that the paper examined: how the attribute ratings contribute to the decision. Like the paper, we will also consider the effect of gender in our analysis. Namely, we want to use this study to research deeper on people's mate selection process. These three questions are connected because we study what characteristics of a partner affect the ratings, and then use the second question to look at whether speed dating would change someone's mindset on the importance of these ratings. The last question then examines how these ratings affect the outcome of a speed date, since the ultimate goal of participating in speed dating is to find a partner.

2 Description of Data

Background Information

As mentioned above, the dataset we investigate was the basis of the analysis from "Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment". Data was gathered from participants in a series of experimental speed dating events from 2002-2004. During the events, each female participant would have a four-minute conversation with every male participant. After the four-minute "first date", participants were asked to rate their partners on six attributes: attractiveness, sincerity, intelligence, fun, ambition, and shared interests. They also had to make the decision whether or not they would like to date the partner again. In addition, data of what participants find valuable in a mate was collected at different time points in the process. We will let *subject* denote the person making the decision, and *partner* denote the person being decided upon.

Variable Description

The initial dataset contains 8378 observations and 195 variables [1]. We will use the following as our independent variables:

- Gender: Female=0; Male=1
- Age: Subject's age
- Age_o: Age of partner
- Race: Black/African American=1
  European/Caucasian-American=2
  Latino/Hispanic American=3
  Asian/Pacific Islander/Asian-American=4
  Native American=5
  Other=6
- Race_o: Race of partner
- Dec: decision whether or not to date the partner again (Yes=1; No=0)
- Dec_o: decision of partner
- Match: whether both people decide to date the other (Yes=1; No=0)
- Samerace: Subject and the partner were the same race. (Yes=1; No=0)
- Exphappy: how happy participants expect to

be with the people they meet during the speed-dating event. (scale of 1-10)

- Imprace: Importance of partner being of the same race (scale of 1-10)
- Imprelig: Importance of partner being of the same religious background (scale of 1-10)
- Numyes: Number of yes's participants said on their second date during the event (new variable)
- Rating on partner's six attributes (awful=1, great=10)
  - Attr: Attractiveness
  - Sinc: Sincerity
  - Intel: Intelligence
  - Fun: Fun
  - Amb: Ambition
  - Shar: Shared interests
- What participants look for in the opposite sex before the event (Distribute a total of 100 among the following attributes)
  - Attr1_1: Attractiveness
  - Sinc1_1: Sincerity
  - Intel1_1: Intelligence
  - Fun1_1: Fun
  - Amb1_1: Ambition
  - Shar1_1: Shared interests
- What participants look for in the opposite sex the day after the event (Total of 100 points)
  - Attr1_2: Attractiveness
  - Sinc1_2: Sincerity
  - Intel1_2: Intelligence
  - Fun1_2: Fun
  - Amb1_2: Ambition
  - Shar1_2: Shared interests
- Field_cd: Field of study
  1= Law
  2= Math
  3= Social Science/Psychologist
  4= Medical Science/Pharmaceuticals/Bio Tech
  5= Engineering
  6= English/Creative Writing/ Journalism
  7= History/Religion/Philosophy
  8= Business/Econ/Finance
  9= Education, Academia
  10= Biological Sciences/Chemistry/Physics
  11= Social Work
  12= Undergrad/undecided
  13=Political Science/International Affairs
  14=Film
  15=Fine Arts/Arts Administration
  16=Languages
  17=Architecture
  18=Other

Out of the independent variables we will use, we treat gender, race (of partner), decision (of partner), same race as categorical variables, whereas the rest are treated as numerical variables. In addition, we will use iid (each person's id) as a factor to separate each individual. The dependent variables we will use are Dec, the 6 attribute ratings (Attr, Sinc, Intel, Fun, Amb, Shar) and their sum, along with Total_diff: Total absolute differences, which can be written as $\sum_{c \in C} |c1\_2 - c1\_1|$, where C is the set of 6 attributes.

Summary Statistics

The dataset contains missing values due to incomplete surveys. We will omit all observations that contain missing values of our specific variables of interest, which do not constitute a majority of the dataset. Each problem we explore will have a different dataset, because we only want to omit an observation if it has a missing value of a variable of interest. To avoid extreme values caused by a huge gap of age, we only consider participants of age 18-35. We believe our results will not be indicative of the patterns of speed dating among middle-aged individuals and thus individuals with age beyond 35 will be excluded. Table I demonstrates an overview of the 551 participants. We see a relatively even distribution of participants by gender, but not by race, with a majority of the participants being Caucasian or Asian. Each distribution of the rating is similar, with mean and median around 6 or 7. Thus if we were to look at the total sum of the 6 attributes for our analysis, we are not getting the effect to be mostly from a single attribute.

Table I. Descriptive statistics of participants

	Number of Participants	Percentage
Gender		
Male	277	50.3%
Female	274	49.7%
Race		
Black/African American	26	4.72%
European/Caucasian-American	304	55.17%
Latino/Hispanic American	42	7.62%
Asian/Pacific Islander/Asian-American	136	24.68%
Other	37	6.72%
Missing	6	1.09%%
Age		
18-25	241	43.74%
26-35	302	54.81%
Missing	8	1.45%

3 Statistical Analyses

Question 1

We will investigate how the ratings are impacted by certain characteristics of a partner. As mentioned above, there are 6 attributes that a partner is rated on by the subject from a scale of 1 to 10. As our response variable, we will consider the total sum of these 6 attributes. The predictors are the race and age of the partner, whether the subject and partner are of the same race, and whether the partner's decision was yes in the speed date. The first three predictors were chosen naturally, and we choose decision of partner as an independent variable because we suspect a speed date where the partner is impressed with a subject is more likely to have high ratings. It is more likely that both people will sense a connection or none of them will. We fit a regular linear regression model and will perform the standard diagnostics. In addition, we will include a person-based intercept because there will be correlation in the scores that a single person gives to all of their partners, but we will not analyze its significance. For instance, we looked at the observations of three individuals. For each individual, we calculate the average rating of the six attributes of the observations in which the individual's decision is no. As evident from Table II, the differences between people are very significant. The first individual rates very high overall, the second about average, and the third individual relatively low. Therefore, we decide we will consider each individual to have their own intercept (and is thus a fixed effect) in all of our models.

We perform some exploratory data analysis. Because 3 out of our 4 predictors are categorical, we look at boxplots. We have a total of 6863 observations. From Figure 1, race seems to play a small role in the total score, namely the mean score for Asian partners is slightly lower than the mean score for all other races. Decision of partner seems to play a role, where saying yes leads to a higher score. We find that whether the

Table II. Individual Differences In Rating Attributes

Individual	Attractiveness	Fun	Shared Interest	Sincere	Intelligence	Ambition
(1)	8.1	7.7	8.1	8.6	9.4	8.8
(2)	6.8	7.6	5.1	7.8	7.2	7.1
(3)	4.4	5.1	4.6	6.2	6.3	6.0

subject and the partner are of the same race does not change the ratings very much. In each case, there seem to be outliers of very low scores. If we inspect the scatterplots of age vs the total score (Figure 2), we do not see any significant patterns. We can also look at the correlation of the 6 attribute ratings. The attributes are all positively correlated, with the highest correlation of 2 attributes at 0.67, which is not very high, so we do not worry about the issue of potential collinearity. We do see that each attribute is highly correlated with the total score, which makes sense because the total score is the sum of 6 scores. Also, the total score appears to be relatively normal-shaped, it just has outliers on its lower end. For example, there are 3 total scores of 0.

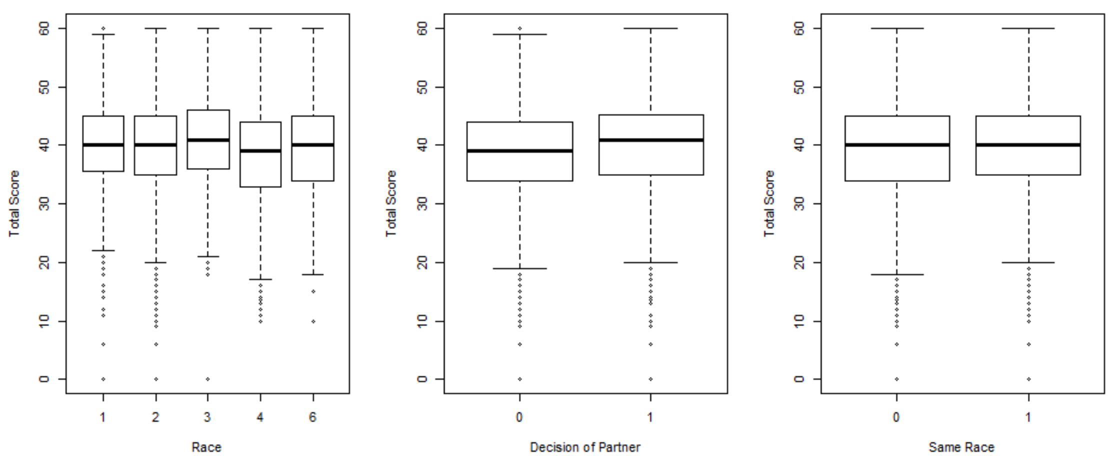

Figure 1: Boxlots of Explanatory Variables

We proceed to fit our model:

$$Total = \alpha_i + \beta_1 age\_o_i + \beta_2 race\_o_i + \beta_3 samerace + \beta_4 dec\_o_i + \epsilon_i$$

We want to check our assumptions. The linear regression model assumes normality, constant error variance, and linearity. To check normality, we look at the studentized residuals of the model. The histogram shows a bell-shaped pattern. The QQ plot of the studentized residuals against the quantiles of a normal distribution (since the studentized residuals are distributed as a t-distribution with very large degrees of freedom, and thus is approximately normal) appear to be a little curved near the smaller quantiles. This makes sense because there are a few outliers just by visual inspection with very low values of total score. However, we know that normality does not play a huge role in the analysis, and thus do not correct for it (see Figure 6 in the Appendix). The Component+Residual plots to check for nonlinearity do not show any issue. We plot the studentized residuals against the predicted values to check for nonconstant variance. We see smaller predicted values have a large range of residuals, and thus we need to perform a correction. We know that non-

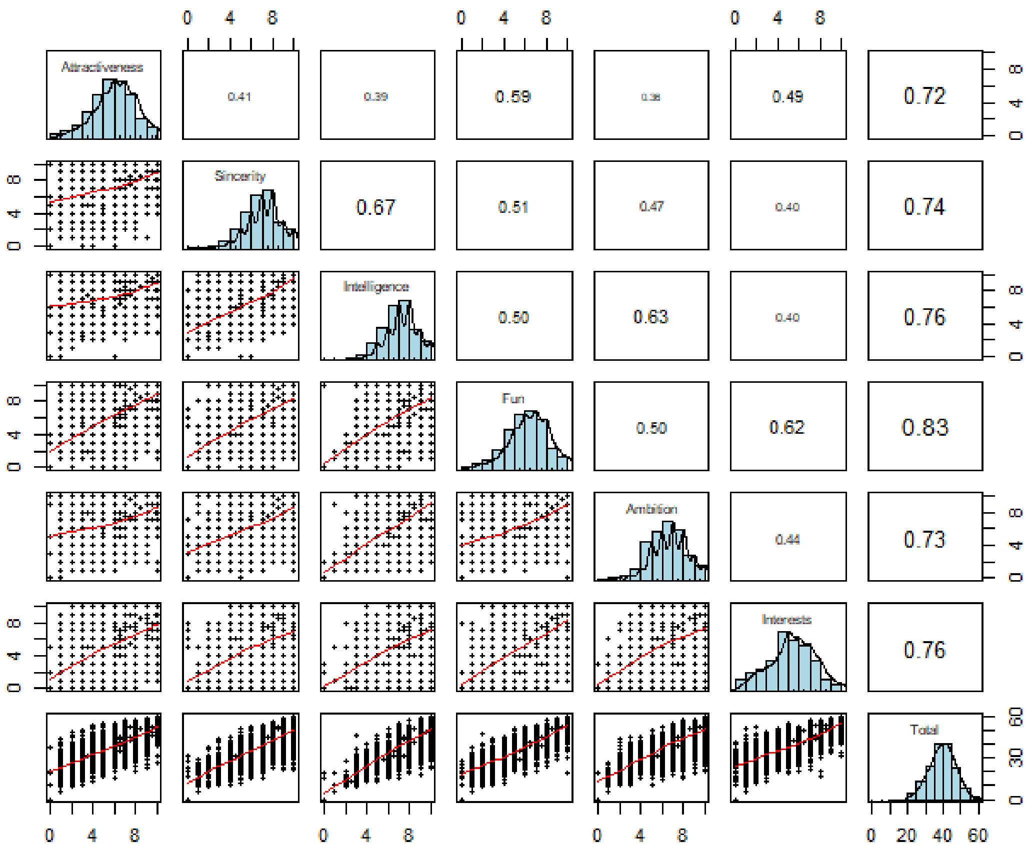

Figure 2: Correlation Plot of Ratings

constant error variance has no effect on the estimates of the coefficients, only on their standard errors. Thus we choose to use a sandwich estimator to find the standard errors that take into account heteroscedasticity. We choose our correction to be White's estimator, which estimates the variance as follows:

$$\hat{Var}(b) = (X^T X)^{-1}(X^T \text{diag}(E_1^2, ..., E_n^2)X)(X^T X)^{-1}$$

where X is the design matrix, and $E_i = Y_i - \hat{Y}_i$ is the residual for observation i [4]. There seems to a slight difference in the estimated standard errors, with the largest difference being approximately 12%. We see that all corrected standard errors except for the predictor decision of partner and same race were increased by various margins. Table VI shows the details. Thus for any future tests, we should be using these corrected standard errors.

We look at the diagnostics of the data: which data points are potential outliers and have high influence on the regression. The outlier test reveals 10 outliers, with the largest having a Bonferroni p-value of 4.3×10^{-3}. There is 1 outlier that gave all 0 points for the 6 attributes and the partner also did not say yes. For the other outliers, they either had a low total rating but their partner said yes or vice versa, so they may be regression outliers. As of the other 3 variables, it is hard to tell a definitive pattern. We look at high influence points. There are 6 high influence points: 2 of high leverage, 2 of high discrepancy, and 2 of both high leverage and discrepancy. The 2 points of high leverage have hat-value 1, and this is because it is the only observation of that individual, so if we remove this observation, the model will change because of the person-based intercept. It also does not have a studentized residual because of this. However, we can still look at its values. The other 4 points have have studentized residuals ranging from 4 to 8 in absolute value. These values are certainly extreme. We see that there are some interesting individuals who do not rate as one would expect. All of the details can be found in the Tables VII-X in the Appendix.

Question 2

In the second question, we will study what factors change a person's attitude towards mate selection in the speed dating event. Since there are two surveys conducted at different points of the event (prior to the event and the day after the event), we use total absolute difference of attendees' ratings on all 6 attributes (Attractiveness, Sincerity, Intelligence, Fun, Ambition, Shared interests) as our response variables. Participants' age, race, gender, field of study, attitude towards the importance of race and religious background, the number of yes's they made on a second date and expectation of happiness during dating are the independent variables. Since the subject we focus on is each individual, we choose a dataset that only consists of individual's unique characteristics. After removing duplicate data and NA's, our new dataset only consists of 468 observations, where each of them represents an individual. We will study the results of fitting a linear regression model by using model selection.

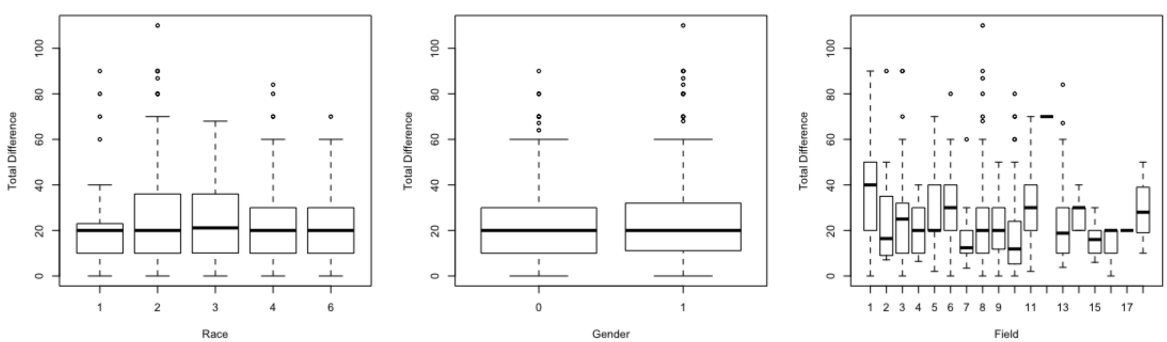

Figure 3: Boxplots of Explanatory Variables

Same as we did in the previous parts, we perform explanatory analysis before constructing our model. Since gender, race, field of study are categorical variables, we use boxplots. As shown in Figure 3, the total difference in ratings have approximately the same mean but the maximum value differs among different races. Field of study seems to have an impact on the difference due to notable fluctuations. Besides, although the distributions of total difference among each gender look very similar, we do want to study whether gender is a significant factor in this particular problem. As a result, we keep all three of them as our explanatory variables.

As for correlation, Figure 4 indicates that none our explanatory variables are significantly correlated with each other (detailed correlation table is attached in the Appendix Table XI). However, the distribution of importance of race and religion, the number of yes's are all right skewed. By checking the distribution of our dependent variable (i.e. total difference), we find a right-skewed pattern as well. Therefore, we consider

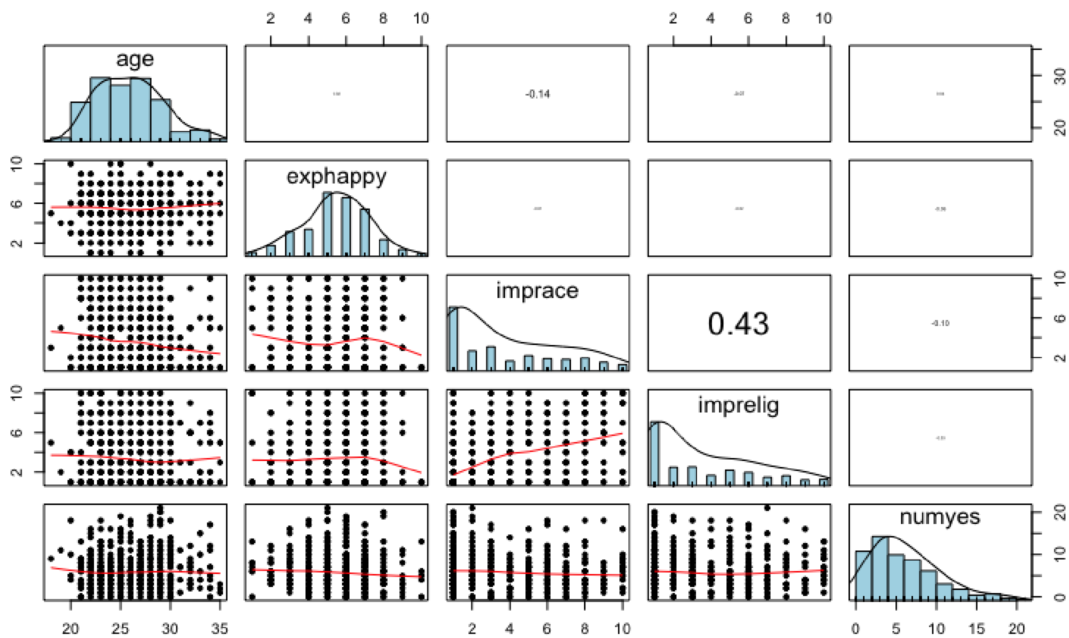

Figure 4: Correlation Plot of Explanatory Variables

transforming these variables by taking their square root. Our full model is

$$\begin{aligned} Total\_Difference_i^{\frac{1}{2}} =& \alpha_i + \beta_0 Age_i + \beta_1 Numyes_i^{\frac{1}{2}} + \beta_2 Exphappy_i + \beta_3 Imprace^{\frac{1}{2}} + \beta_4 ImpRelig^{\frac{1}{2}} \\ &+ \gamma_1 Race_i + \gamma_2 Gender_i + \gamma_3 Field_i + \epsilon_i \end{aligned}$$

Because there are many predictors, we use variable selection to find the "best" model. To carry out the model selection, we will use AIC (known as Akaike information criterion), which is a fined technique based on in-sample fit to estimate the likelihood of a model to predict/estimate the future values [2]. AIC is defined by

$$AIC = -2 \times \ln(L) + 2k$$

where $\ln(L)$ is the value of the log-likelihood, and k is the number of estimated parameters. The smaller AIC is, the better fit the model is. Regression diagnostics will be performed to check the validity of the model.

After performing a Forward-Backward hybrid model selection by AIC, we get a suggested model with AIC=660.07:

$$Total\_Difference_i^{\frac{1}{2}} = \alpha_i + \gamma_1 Field_i + \gamma_2 Gender_i + \epsilon_i$$

where i represents the individual i in both models.

When we check the regression diagnostics, the histogram and QQ plot suggest that data are normally distributed. In the residual plot, residuals are pretty symmetrically and randomly distributed, so it should be reasonable to assume the variance is constant (see Figure 7). In addition, the predictor is a categorical variable, so we may assume that the linearity assumption holds.

We also want to examine if there are any potential outliers. The outlier test reveals no significant outliers (with Bonferonni p-value smaller than $\alpha = 0.05$), which is reasonable because our predictors are categorical. However, there are 6 points that have high influence: Observation 37 and 217 have high studentized residuals; Observation 39 and 293 have the highest leverage since each of them is the only person in that specific field of study; Observation 210 and 395 have both high leverage and studentized residuals. Out of the 6 points, observation 210 has the largest Cook's distance of 0.105. Since our predictors are categorical and there is no obvious patterns about these 6 participants' field of study, we cannot really tell why these points are highly influential just based on the model we construct. Table XII-XIII showing all the influential points can be found in the Appendix.

Question 3

Our last question highly motivated by the paper. We will investigate how ratings of attributes impact the decision of the subjects and whether gender will affect decisions.

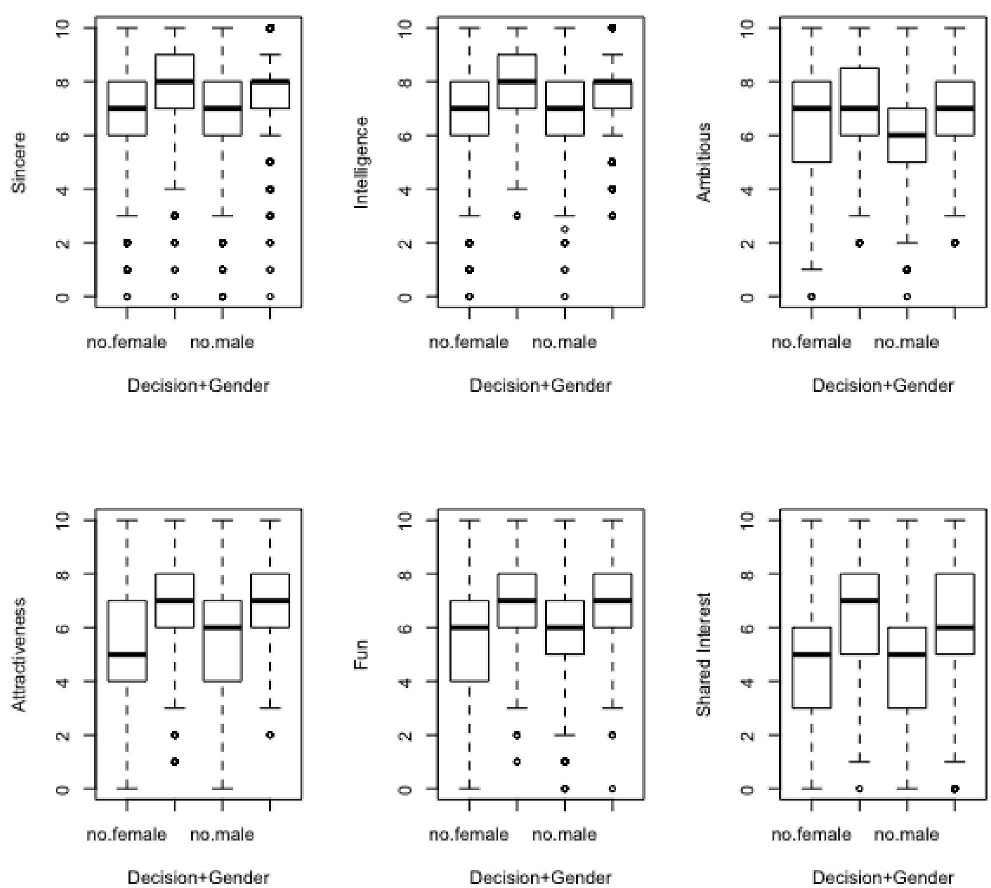

Figure 5: Boxplot of Explanatory Variables

We want to look at our data visually first. Figure 5 shows boxplots of decision and gender against each of the 6 attributes. We think the rating of attributes will tend to be highly related with the decision. Except for Ambition, the rest five box-plots all show that the observations of decision equals to one tends to have higher attribute ratings. However, gender doesn't seems act significant influence on the relationship between decision and the rating of the attributes. In order to check our hypothesizes, we will fit a logistic regression model on our data because our response variable Decision is a binary categorical variable. Therefore, we can use the coefficients of the models to estimate the weighting of the attributes on the decision. The logistic regression model is

$$P(Y = 1|X_1 = x_1, ..., X_n = x_n) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_n x_n}}{1 + e^{\beta_0+\beta_1 x+\beta_2 x_2+...+\beta_n x_n}}$$

where Y is the dependent variable and $X_1, ..., X_n$ are the independent variables [3]. We choose the ratings of the six attributes which are Attractiveness, Fun, Shared Interest, Intelligence, Ambition, and Sincere as our predictors, and Decision as our dependent variable. Again, each individual is a fixed effect.

Therefore, we will fit our first two models by:

$$\ln\left(\frac{P(Decision = 1|X)}{1 - P(Decision = 1|X)}\right) = \alpha_i + \sum_{c\in C} \beta_c \times Rating_{ijc} \quad (1)(2)$$

where $C = \{$Attractiveness, Fun, Shared Interest, Ambition, Sincere, Intelligence$\}$, and X represents all of our independent variables. Here, i means subject, j means partner. These two models will be fit to two different data sets. The first model will be fit to the data set containing only female responses and the second model will be fit to the one containing only male responses. We also include gender and its interaction term with other attributes in the third model which will be fit to the large datasets (with responses from both male and female). With the third model, we can have a better understanding on how gender affects mate selection in speed dating. The third logistic regression model is

$$\ln\left(\frac{P(Decision = 1|X)}{1 - P(Decision = 1|X)}\right) = \alpha_i + \sum_{c\in C} \beta_c \times Rating_{ijc} + \sum_{c\in C} \gamma_c \times Gender_i \times Rating_{ijc} \quad (3)$$

We can also inspect the influence points in the three models. Many of these points overlap between the 3 models, and have values we would not expect. For example, there is a male who rated his partner 10 points for all the attributes, yet said no. These results are shown in Tables XIV-XX.

4 Results

Question 1

We interpret the coefficients from our first model, which are shown in Table III. The most significant variables are the decision of the partner, with p-value of 3e-07, along with the age of the partner, with p-value 3.83e-04, and race=4 (the Asian-American category) with p-value 1.7e-09. If we use the corrected standard errors, the p-value for age will remain the same, while the p-values for decision and same race will decrease, and the others will increase. At a conventional 0.05 level of significance, the only variable that becomes non-significant is the dummy variable for race=6, or the Other category. To interpret these coefficients, we find that on average, the total score that a subject gives to a partner is 1 point higher if the partner says yes to a speed date than no. This is fairly reasonable because during a speed date, subjects will not know whether a partner will decide yes or no, but rather if the partner feels they would like to talk to the subject again, then it is more likely the partner would do the same. Also, the mean score will decrease by about 0.1 points for a partner that is 1 year older. That is, a partner that is 30 on the average will receive 1 point lower than a partner of age 20. The predictor with the largest coefficient is the indicator variable for Asian-Americans at -2.347. This means that Asian-American partners receive more than 2 points lower than the base category of African-American partners. It is interesting that all coefficients for the race indicator variables are negative, and some of the other variables have p-values less than the conventional significance level of 0.05. This seems to indicate that African-American partners receive the highest ratings on average, but we cannot overly interpret the coefficients because some have large p-values. We do not find same race to be a significant predictor, but it does have a positive coefficient, which implies people do slightly prefer to date people of the same race. We note that the R^2 of this model is 0.4815, which is not bad, considering we only have 4 predictors for such a complicated scenario.

Thus we can conclude the main factors in the rating of a partner in a speed dating experiment is their race, as African-American partners receive the highest rating on average, and the subjects in this experiment prefer younger partners.

Table III. Partner's Characteristics and Total Ratings

	Estimate	Std. Error	t-value	Pr($>\|t\|$)
(Intercept)	42.793	2.158	19.826	< 2e-16
dec_o1	0.911	0.178	5.129	3e-07
age_o	-0.094	0.026	-3.553	3.83e-04
race_o2	-0.954	0.379	-2.515	0.012
race_o3	-0.163	0.445	-0.366	0.714
race_o4	-2.347	0.389	-6.033	1.70e-09
race_o6	-0.935	0.469	-1.992	0.046
samerace1	0.301	0.203	1.484	0.138

Here, samerace1 and dec_o1 are indicator variables that take value 1 if the subject and partner are of the same race or a decision respectively. The 4 race variables are indicator variables for whether the partner is a certain race.

Question 2

The summary statistics of our second model selected by AIC is shown in Table IV. When we consider a significance level of $\alpha = 0.05$, gender (with a p-value of 0.138) is not significantly associated with the variation in total difference. The baseline field based on this coding is Law. It turns out that the ratings of people who study in fields related to Social Science, Engineering, History, Business, Education, Natural Science, Political Science, Arts and Languages differ at two time points. In particular, the p-values (4.47×10^5 and 6.75×10^7 respectively) of participants who study Business and Natural Science tend to be significantly smaller than those study in other fields. We may say that the original total absolute difference of people who study Business and Natural Sciences is significantly different than the total absolute difference of people who study Law, when the gender is the same. However, the R^2 of the model is 0.09239, which is very bad. To fix this problem, we try to add more interaction terms in our model (i.e. $Age_i * Numyes_i * Exphappy_i * Race_i * Field_i * Imprace_i * ImpRelig_i * Gender_i$). However, the subset selection process returns the same exact model as it did previously. So it would be problematic when we need precise predictions using this model which has unexplainable variability (i.e. people are fairly unpredictable).

Thus, we come to a conclusion that the main factor in the total difference of ratings in these speed dating events is field of study, with Business and Natural Science being two of the most significant fields. However, due to the low R^2 of the model, we should not expect precise predictions when we use this model.

Table IV. Participant's Field and Total Difference in Ratings

	Estimate	Std. Error	t-value	Pr($>|t|$)
(Intercept)	5.564	0.320	17.383	< 2e-16
Math	-1.132	0.645	-1.755	0.080
Social Science/Psychologist	-1.022	0.433	-2.361	0.019
Medical Science/Pharmaceuticals/Bio Tech	-1.386	0.729	-1.902	0.058
Engineering	-0.859	0.413	-2.078	0.038
English/Creative Writing/Journalism	-0.279	0.559	-0.499	0.618
History/Religion/Philosophy	-1.648	0.625	-2.635	0.009
Business/Econ/Finance	-1.464	0.355	-4.122	4.47e-05
Education/Academia	-1.177	0.467	-2.518	0.012
Biological Sciences/Chemistry/Physics	-2.012	0.399	-5.040	6.75e-07
Social Work	-0.314	0.504	-0.624	0.533
Undergrad/undecided	2.488	2.008	1.239	0.216
Political Science/International Affairs	-1.162	0.438	-2.655	0.008
Film	-0.490	0.863	-0.568	0.570
Fine Arts/Arts Administration	-1.653	0.729	-2.268	0.024
Languages	-2.688	1.184	-2.269	0.024
Architecture	-1.406	2.008	-0.700	0.484
Other	-0.599	1.183	-0.506	0.613
Gender: Male	0.314	0.212	1.484	0.138

Question 3

From Table V, we can clearly see the ratings of five out of six attributes (except Ambition) seem statistically significant to the decision with the conventional significance level $\alpha = 0.05$ for females and in general case. In male's case, Sincere also seems do not play significant role. From the coefficients, we know for each additional attractiveness point, the log(odds) of saying Yes will increase by 0.9320, for a female. In other words, each additional attractiveness point will increase the odds of saying Yes by 154% for a female subject. This way of interpretation also works for other attributes. It means a point increase in the attribute is associated with increasing the odds of saying Yes by the exponential of its coefficients. For a male, each additional attractiveness points increases the odds of saying Yes by 223%. Therefore, we think it will be a good way for us to estimate the weighting of the six attributes in mate selection. Unlike the companion paper, we think Attractiveness, Fun, and Shared Interest will be the main three attributes that affect an individual's decision based on our regression results. This is because the attributes do have the largest coefficients out of the six for both female and male subjects.

We also use the interaction term to help us find out whether gender will affect the weighting of the attributes in speed dating. The regression model we fit shows evidence that support weightings of Physical Attractiveness and Sincerity are different among different genders at significance level of 0.1. It means that each additional attractiveness point will increase 27% in the odds of a male saying Yes more than the odds of a female saying Yes. Each additional sincerity point will decrease the odds of a male saying Yes by 15% compared to the odds of a female saying Yes. Therefore, we conclude that male weighted Physical Attractiveness more than female does. Notice that the interaction terms for attractiveness and fun are positive, which means males weight these two attributes heavier than women. Contrasting this, females weight the characteristics of sincerity, ambition, intelligence, and shared interest heavier than males, albeit by a smaller factor than attractiveness.

Table V. Gender Differences In Attribute Weights

	(1)	(2)	(3)
Sincere	0.219***	0.076	0.219***
Intelligence	0.296***	0.185**	0.296***
Ambitious	0.041	-0.021	0.041
Attractiveness	0.932 ***	1.174***	0.932***
Fun	0.478***	0.589***	0.478***
Shared Interest	0.506***	0.042***	0.506***
Intelligence × Male			-0.111
Sincere × Male			-0.143 ·
Ambitious × Male			-0.062
Attractiveness × Male			0.242***
Fun × Male			0.111
Shared Interest × Male			-0.084
Subject's Gender	Female	Male	Both

Here *** means significant at 0.001 level, ** means significant at 0.01 level, · means significant at 0.1 level

5 Conclusion and Future Work

Our main results show that African-American partners tend to receive the highest ratings overall, compared to other races, and Asian-American partners receive the lower overall ratings. We also find that usually, how the partner feels about the date is likely to be reflected in the scores they receive. As for how a speed dating event would change the mindset of a person's selection process, we find that the most important variable is their field, with people in the Business and Natural Science fields being the most significant fields. Surprisingly, male and females do not have a significant difference in the average change of ratings. To determine which attributes impact decision the most, we find that males put heavier weight on physical attractiveness than women, while women weight sincerity heavier than men. This more or less matches the conclusion of a similar analysis done by our reference paper.

As we have mentioned, our paper analyzed the data in different aspects than previously done, but it is only a beginning of possible research in people's mate selection. The data set contains many variables that we haven't analyzed, and that may be important in understanding how people date. We could also try to use different statistical models to analyze our questions, as linear models may not be the most appropriate. Some limitations are that the speed dating experiment was conducted more than 10 years ago, so we do not know whether people's preferences would have changed or not. The experiment design also leads to correlation, both because the same person is rating multiple people and each person is being rated multiple times, and that a person being rated then becomes someone rating others. We do not account for all of the correlation that is present in the design, nor is it present in the paper [1]. More work can be done to account for these correlations.

References

[1] Fisman, Raymond, et al. "Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment." *The Quarterly Journal of Economics*, The President and Fellows of Harvard College and the Massachusetts Institute of Technology, 2006.

[2] Fox, John. *Applied Regression Analysis and Generalized Linear Models.* 3^{rd} ed., Sage, 2016.

[3] Weisberg, Sanford. *Applied Linear Regression.* 4^{th} ed., Wiley, 2014.

[4] White, Halbert "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity". *Econometrica.* 48 (4): 817–838, 1980.

Appendix

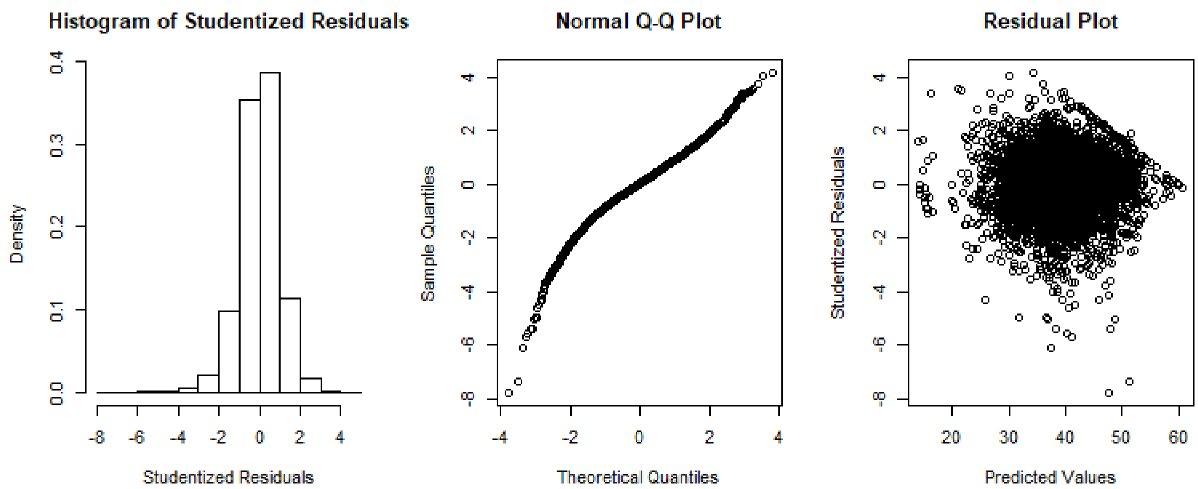

Figure 6: Checking Assumptions for Question 1

Table VI. Standard Error of Estimates

	dec_o1	age_o	race_o2	race_o3	race_o4	race_o6	samerace1
OLS Errors	0.178	0.026	0.379	0.445	0.389	0.469	0.203
Corrected Errors	0.171	0.026	0.419	0.477	0.437	0.490	0.189
Difference	−0.007	0.000	0.040	0.032	0.048	0.021	-0.014
% Difference	−3.93	0	10.55	7.19	12.34	4.48	-6.9

Table VII. Influence Test for Question 1

Observation	Studentized Residual	Hat-Value	Cook's Distance
1606	NA	1.00	NA
2335	4.01	0.33	0.015
2336	-4.98	0.33	0.022
3393	NA	1.00	NA
3501	-7.33	0.05	0.005
4787	-7.76	0.05	0.006

Table VIII. Influence Points for Question 1

Observation	iid	dec_o	age_o	race_o	samerace	attr	sinc	intel	fun	amb	shar	total
1606	146	1	27	2	1	10	9	10	10	8	8	55
2335	210	1	26	4	0	7	8	8	9	9	10	51
2336	210	1	27	2	1	1	1	1	1	1	1	6
3393	277	0	32	2	1	5	8	7	5	9	3	37
3501	283	1	26	1	0	1	1	1	1	1	1	6
4787	384	1	31	2	1	0	0	0	0	0	0	0

Table IX. Outlier Test for Question 1

Observation	Studentized Residual	Unadjusted p-value	Bonferonni p-value
4787	-7.757	1.008e-14	6.915e-11
3501	-7.333	2.528e-13	1.734e-09
6797	-6.065	1.397e-09	9.584e-06
4749	-5.688	1.343e-08	9.215e-05
2303	-5.558	2.834e-08	1.944e-04
4597	-5.375	7.945e-08	5.451e-04
112	-5.361	8.561e-08	5.874e-04
4731	-5.048	4.587e-07	3.147e-03
4780	-5.011	5.573e-07	3.823e-03
2336	-4.983	6.418e-07	4.403e-03

Table X. Outliers for Question 1

Observation	iid	dec_o	age_o	race_o	samerace	attr	sinc	intel	fun	amb	shar	total
4787	384	1	31	2	1	0	0	0	0	0	0	0
3501	283	1	26	1	0	1	1	1	1	1	1	6
6797	549	0	30	3	0	0	0	0	0	0	0	0
4749	381	1	31	2	1	1	1	1	1	1	1	6
2303	207	1	27	2	1	1	1	1	1	1	1	6
4597	373	1	26	1	0	4	3	4	2	2	0	15
12	2	0	22	2	1	8	5	6	6	9	6	40
4731	380	1	31	2	0	1	1	1	1	1	1	6
4780	384	1	24	3	0	3	3	3	3	3	3	18
2336	210	1	27	2	1	1	1	1	1	1	1	6

Table XI. Correlation of Variables in Question 2

	age	exphappy	imprace	imprelig	numyes
age	1.000	0.019	-0.139	-0.065	0.019
exphappy	0.019	1.000	-0.014	-0.021	-0.084
imprace	-0.139	-0.014	1.000	0.442	-0.077
imprelig	-0.065	-0.021	0.442	1.000	-0.025
numyes	0.019	-0.084	-0.077	-0.025	1.000

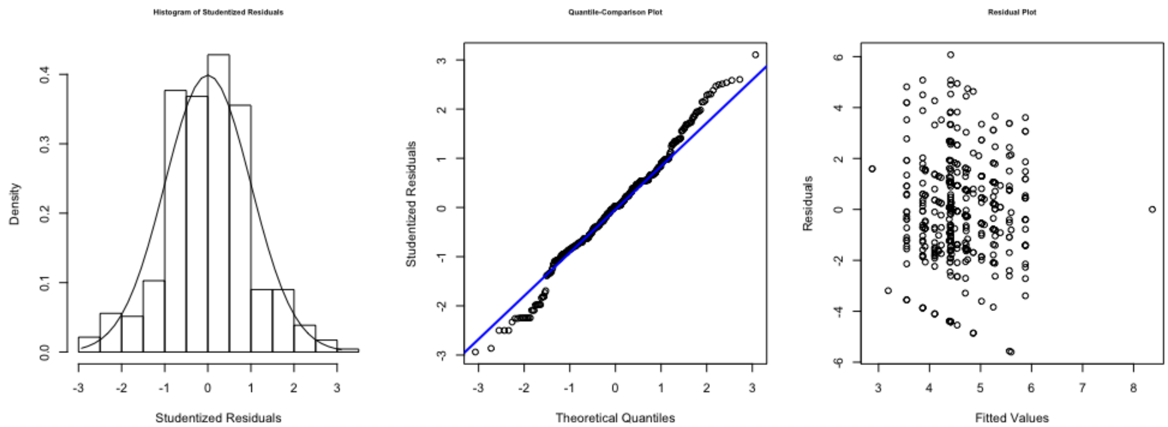

Figure 7: Model Checking in Question 2

Table XII. Influence Test for Question 2

Observation	Studentized Residual	Hat-Value	Cook's Distance
37	3.11	0.01	0.004
39	NA	1.00	NA
210	-1.98	0.34	0.105
217	-2.94	0.62	0.030
293	0.00	1.00	0.005
395	1.31	0.34	0.046

Table XIII. Influential Points for Question 2

obs	iid	Gender	Field	Total Difference
37	46	1	8	10.49
39	48	1	12	8.37
210	243	1	16	0.00
293	336	1	17	4.47
359	418	1	1	0.00
395	466	0	18	7.07

Table XIV. Standard Error of Estimates for Question 3

	(1)	(2)	(3)
Sincere	0.059	0.061	0.059
Intelligence	0.073	0.067	0.073
Ambitious	0.056	0.053	0.056
Attractiveness	0.056	0.058	0.056
Fun	0.054	0.054	0.054
Shared Interest	0.048	0.045	0.048
Intelligence × Male			0.099
Sincere × Male			0.085
Ambitious × Male			0.078
Attractiveness × Male			0.081
Fun × Male			0.076
Shared Interest × Male			0.066
Subject's Gender	Female	Male	Both

Table XV. Influence Test for Question 3 (female)

Observation	Studentized Residual	Hat-Value	Cook's Distance
140	NA	1.00	NA
426	NA	1.00	NA
779	-3.239	0.077	0.013
780	2.718	0.391	0.022
3037	-3.714	0.000	0.002
3047	-3.874	0.34	0.002

Table XVI. Influence Points for Question 3 (female)

Observation	iid	dec	attr	sinc	intel	fun	amb	shar
140	24	0	5	7	7	7	7	7
426	76	1	5	5	5	5	8	8
779	121	0	7	6	7	8	10	7
780	121	1	6	10	5	2	2	4
3037	511	0	7	6	6	7	6	5
3047	511	0	7	7	6	6	3	6

Table XVII. Influence Test for Question 3 (male)

Observation	Studentized Residual	Hat-Value	Cook's Distance
116	NA	1.00	NA
567	-4.097	0.000	0.0002
1430	NA	1.00	NA
1568	-1840.180	1.00	4.911e+18
2473	3.443	0.499	0.006

Table XVIII. Influence Points for Question 3 (male)

Observation	iid	dec	attr	sinc	intel	fun	amb	shar
116	41	1	8	8	8	8	8	8
567	104	0	10	10	10	10	10	10
1430	243	1	7	8	8	8	7	7
1568	277	0	5	8	7	5	9	3
2473	404	1	2	9	7	6	6	1

Table XIX. Influence Test for Question 3 (both)

Observation	Studentized Residual	Hat-Value	Cook's Distance
236	NA	1.00	NA
466	NA	1.00	NA
1268	-4.097	0.000	0.0009
5034	-3.419	0.468	0.025
5037	3.443	0.499	0.028
6152	-3.874	0.000	0.001

Table XX. Influence Points for Question 3 (both)

Observation	iid	gender	dec	attr	sinc	intel	fun	amb	shar
236	24	0	0	5	7	7	7	7	7
466	41	1	1	8	8	8	8	8	8
1268	104	1	0	10	10	10	10	10	10
5034	404	1	0	5	5	6	5	6	5
5037	404	1	1	2	9	7	6	6	1
6152	511	0	0	7	7	6	6	3	6