

I chose to analyze the [Speed Dating Experiment](#) dataset from Kaggle and to entitle our project, “Love at First Date”. The dataset gives a great deal of information about the participants. This includes age, race, field of study, school attended for undergrad, career, interests, income, and goals of participating amongst many other factors. The study collected information comparing the different pairings throughout the night. This includes how satisfied participants were with the people they met, what attributes were most important to them while rating, and how they thought they matched up compared to other people.

I am interested in this topic because unconventional dating is becoming increasingly popular in today’s society with the rise of applications such as Tinder, Coffee Meets Bagel, Bumble and the increase in online dating sites such as EHarmony, christianmingle.com, etc. Because of this, we are curious to see what qualities/characteristics are most important when meeting someone for the first time. Many people in these artificially set up dating environments have a checklist they go through when meeting people on dating apps or in speed dating situations. We are interested to see if such a “checklist” is apparent when analyzing the people that were satisfied with the overall experience.

### **Related Work**

Before deciding on visualizations, I looked at some of the work other people had done with this dataset. People viewed it from multiple perspectives, such as predicting yes or predicting overall satisfaction. I was most interested in looking at overall satisfaction.

### **Cleaning the Data**

This was a huge data set that required a great deal of work before beginning the visualizations. I first began by eliminating variables that I did not find to be of interest, whether they were redundant or I could not think of a visualization involving that variable. On a couple of the speed dating rounds, individuals were asked to use a different scoring scale. I also eliminated all of these to avoid confusion and to save time manipulating these. I then deleted variables and rows that had mostly missing data. These were not useful at all, and I did this to further reduce the size.

### **Questions I wanted to Answer:**

1. What were people’s goals for the night and how does this tie in with their interests?

People partake in such events for multiple reasons and I wanted to know and understand these reasons before diving deeper?

2. What were people looking for in a potential match?

This question gives a good idea of what characteristics people most value. I also was curious what people thought the opposite gender was looking for.

3. What were the results of this experiment?

I am curious how satisfied people were with their time spent at the speed dating event, if they have been on any dates, and how many matches were formed.



**Data Sketches**



EDA : \*\* in all visualizations including gender: 0=Female, 1=Male

1. Descriptive Statistics:

<https://public.tableau.com/profile/jordin.baldari#!/vizhome/finalstory/FinalStory>

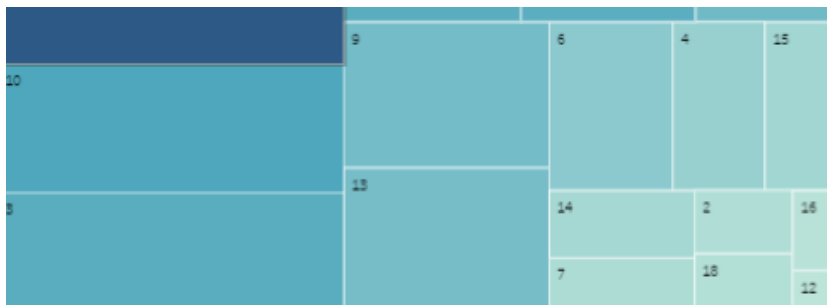
**What Are People Expecting Tab**



## 2. What Do People Want:

<https://public.tableau.com/profile/jordin.baldari#!/vizhome/finalstory/FinalStory>

### What Do People Want Tab



4= Medical Science, Pharmaceuticals, and Bio Tech  
5= Engineering  
6= English/Creative Writing/ Journalism  
7= History/Region/Philosophy  
8= Business/Econ/Finance  
9= Biological Sciences/Chemistry/Physics  
10= Education, Academia  
11= Social Work  
12= Undergrad/Undecided  
13= Political Science/International Affairs  
14= Film  
15= Fine Arts/Arts Administration  
16= Languages  
17= Architecture  
18= Other

3. Results: <https://public.tableau.com/profile/jordin.baldari#!/vizhome/finalstory/FinalStory>

How Was the Match Quality Tab



**GitHub Address:** <https://github.com/baldarja/baldariproject/blob/master/baldari.project2final.docx>

### **Evaluation:**

First looking at the descriptive statistics, the most popular interests/hobbies among both men and women were dining, music, and reading. Men were significantly more interested in gaming and sports than women, while women were significantly more interested in shopping than men. Men were slightly more optimistic about the people they would meet, coming in higher for number of people they would be happy to meet, number of people they thought would be interested in them, and the number of mutual matches they expected to get. Most people participated in the experiment because they wanted a fun night out and to meet new people. It was somewhat important to both men and women that their prospective matches was of the same race and religion as them.

I then further analyzed what people wanted in a match. The two most important attributes to men were attractiveness and intelligence. Women thought men would place a very large importance on attractiveness, much higher than the other attributes. While this was the highest ranked attribute for men, it was not as overwhelmingly so as women predicted. Perhaps men are not as shallow as women think after all! The three most important attributes to women were intelligence, attractiveness, and sincerity in decreasing order. They did more evenly distribute their points among the attributes than men. I guess it's true that women want it all and are more difficult to please. Men thought women would place the most value on attractiveness and fun. Men and women alike ranked themselves very highly in all of the different attributes, and both ranked themselves highest for sincerity and intelligence. It appears that this was a very confident group of people, but I was not surprised by this. It definitely takes confidence to participate in a speed dating experiment. I was also curious whether certain careers attracted more "yesses". An overwhelming majority of the most popular men worked in business. There was not this overwhelming majority for the most popular women, and it was pretty evenly distributed. The most common career fields of the popular women were social sciences/psychology, social work, and education. I did not find either of these figures surprising. In the society where the study was done men are expected to be the primary bread winner, explaining the strong interest in business men. Women are thought of as more of the caregivers, explaining the interest in women with less time demanding careers.

People seem to have been decently satisfied with the night ranging from the 5-6 range out of 10. However, many people have been on dates and been in contact with their matches. While many people's goal of the night was to just have fun and meet new people, I was surprised by the high number of people that have been on dates. Men have been on slightly more dates than women, meaning men are going on dates with more than one of their matches. This aligns with the fact that men have been calling their matches much more than women. I was not surprised that a great deal of the records were one sided matches. Even in more authentic dating situations, this seems to often be the case. People's perceptions of themselves did not vary much from before and after this experiment.