Alexandre Zajic (Data Analysis Sample Portfolio with R)

Increase in Likelihood of Agreeing to a Second Date (Based on Ratings of the Opposite Sex)

	Male	Female
Intercept	-35.8%	-37.3%
Attractive	11.6%	6.4%
Sincere	-2.1%	*-1.2%*
Intelligent	*-1.1%*	*+1.4%*
Fun	3.9%	3.5%
Ambitious	-2.5%	-1.9%
Interesting	4.6%	4.8%
Overall Likelihood	49.5%	38.2%

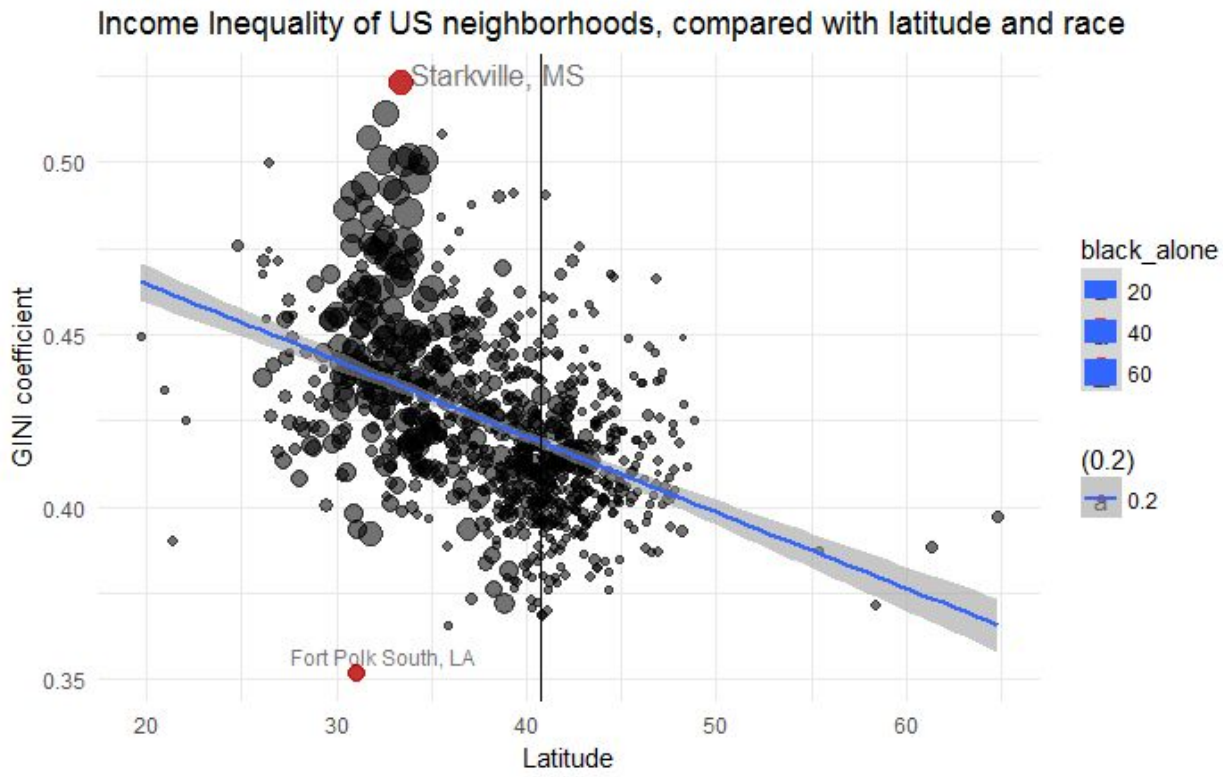
Here, I built a model in RStudio to analyze a speed dating dataset of approximately 5,200 heterosexual speed dates. After each date, the respondent would answer a questionnaire, which included their ratings of their date partner on a scale of 1-10 on the above six attributes, and whether they wanted to go on another date with them. I wanted to find what factors made a person more likely to say “yes” to a second date.

Each percentage listed in the above boxes is the change in likeliness for the respondent to agree to a second date, if they perceive their partner to be 1 point higher in a given attribute. For example, a man rating a woman 1 point higher in attractiveness would be 11.6% more likely to want a second date with the woman, while for a woman to think the same of her male partner would increase her likelihood of wanting a second date by 6.4%. Percentages in red indicate a negative correlation, while percentages in grey show a lack of a statistically significant relationship within a 95% confidence interval.

The conclusion I drew from this data was that men and women are surprisingly similar in what they do (and don’t) value in a date. Perceived ambition and sincerity correlate slightly negatively with odds of a second date for both men *and* women, while ratings of intelligence were statistically insignificant for both males and females. The data also confirms and contradicts certain pop-culture perceptions of men and women. While the data indicates that attractiveness is by far the most significant factor needed to get a man to agree to a second date, we see that it does not “crowd out” the other factors when compared to women’s

preferences. In other words, a man is just as likely to agree to a second date with a woman based on her being interesting and fun, as a woman is to agree to a second date with a man based on him being interesting and fun. The data suggests that, compared to women's preferences, men do not consider attractiveness to be important at the expense of other personality traits, but they will instead agree to more second dates on the whole.

Comparing latitude with inequality and race



This next dataset contained demographic data measuring 67 different variables on the level of Census Tracts (roughly corresponding to "neighborhoods"). We were tasked to explore the data and find an interesting conclusion that could be represented in one graph.

I explored the relationship between income inequality (GINI coefficient) and latitude, showing that the South's neighborhoods are more unequal than the North's, and inserted a vertical line at New York City's latitude for reference. My intuition was that racial disparities would also coincide with income inequality, as is shown by the graph. The larger the size of a dot, the larger the self-identified black population is as a percentage of the neighborhood population.

I also chose to analyze the two outliers at roughly the same latitude, to see why they were so far apart. Fort Polk South is a military-dominated neighborhood (probably standardizing

wages), and Starkville has a large university population (probably creating a large disparity between the poorer students, and the wealthier surroundings).