Final Project Milestone 3 - AL

Introduction

In this project, I explore data from a large-scale speed dating experiment to identify the factors that predict mutual romantic interest and understand how individuals perceive themselves and others. I aim to uncover underlying psychological constructs using exploratory factor analysis and assess their influence on speed dating outcomes.

Introduction and Data Overview

This project explores data from a large-scale speed dating experiment hosted by Columbia University from 2002 to 2004. The dataset, sourced from Kaggle, includes over 8,000 match decisions recorded during structured four-minute dates. Each participant rated their partners on six traits (attractiveness, sincerity, intelligence, fun, ambition, and shared interests) and indicated whether they would like to see them again. In addition to these evaluations, the dataset contains self-perceptions, demographic variables, and lifestyle preferences gathered from surveys before and after the events.

Link to source: https://www.kaggle.com/annavictoria/speed-dating-experiment

The data was read into R from a CSV file and cleaned to remove missing or inconsistent entries. I created tidy subsets to explore psychological constructs (e.g. self-perception, personality profiles) using exploratory factor analysis, correlation matrices, and regression models. Decisions about filtering, transforming, and visualizing the data are fully annotated throughout the document. Because each participant appears multiple times in the data (one row per date), I accounted for the repeated measures accurately by using multilevel analysis. This project investigates how individual traits and perceptions influence dating behavior while considering limitations in measurement, representation, ethics, and generalizability.

Research Question: What factors influence whether a participant says 'yes' to a speed date?

Sub-Question 1: Correlation Among Interest Variables - Heatmap

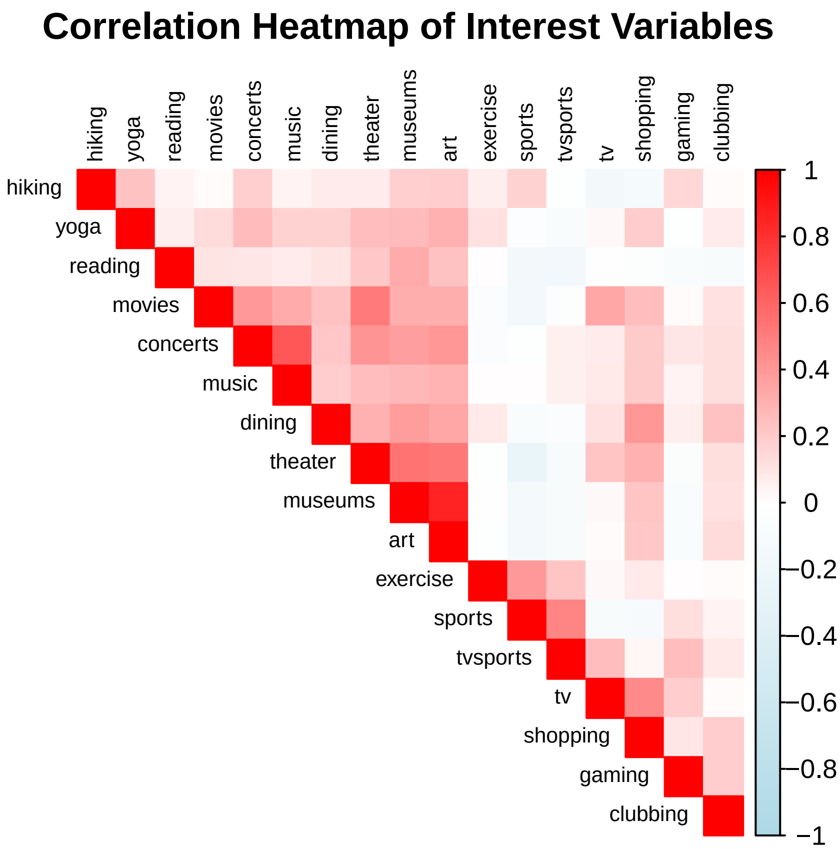
```
df <- read.csv("Speed Dating Data.csv")

interests <- df[, c("sports", "tvsports", "exercise", "dining", "museums", "art",
                    "hiking", "gaming", "clubbing", "reading", "tv", "theater",
                    "movies", "concerts", "music", "shopping", "yoga")]

# Correlation matrix
corr_matrix <- cor(interests, use = "pairwise.complete.obs")

# Heatmap
corrplot(corr_matrix, method = "color", type = "upper",
```

```
         col = colorRampPalette(c("lightblue", "white", "red"))(200),  # less blue, more red
         tl.col = "black", tl.cex = 0.7, order = "hclust",
         title = "Correlation Heatmap of Interest Variables", mar=c(0,0,2,0))
```



Explanation

1a: Stacked Bar Chart of EFA Loadings (Layperson Visual)

```
## Sub-Question 2: Exploratory Factor Analysis Model Comparison
## Visualization: Simplified Factor Loadings Stacked Bar Chart with Thresholds

library(psych)
library(dplyr)
library(tidyr)
library(ggplot2)

# Load and select interest variables
df <- read.csv("Speed Dating Data.csv")
interests <- df[, c("sports", "tvsports", "exercise", "dining", "museums", "art",
                    "hiking", "gaming", "clubbing", "reading", "tv", "theater",
                    "movies", "concerts", "music", "shopping", "yoga")]

# Run EFA with 3 factors
```

```r
efa_3 <- fa(interests, nfactors = 3, rotate = "varimax", fm = "ml")
loadings_df <- as.data.frame(unclass(efa_3$loadings))
loadings_df$interest <- rownames(loadings_df)

# Reshape data for plotting
loadings_long <- loadings_df %>%
  pivot_longer(cols = starts_with("ML"), names_to = "Factor", values_to = "Loading") %>%
  filter(abs(Loading) > 0.3) %>%
  mutate(Factor = recode(Factor,
                         "ML1" = "Cultural/Intellectual",
                         "ML2" = "Active/Sporty",
                         "ML3" = "Social/Media"))

# Order bars by descending loading
loadings_long <- loadings_long %>%
  arrange(desc(Loading)) %>%
  mutate(interest = factor(interest, levels = unique(interest[order(-Loading)])))

# Plot
ggplot(loadings_long, aes(x = interest, y = Loading, fill = Factor)) +
  geom_bar(stat = "identity", position = "stack") +
  geom_hline(yintercept = 0.5, linetype = "dotted", color = "gray40", linewidth = 0.7) +
  geom_hline(yintercept = 0.7, linetype = "dashed", color = "gray20", linewidth = 0.8) +
  coord_flip() +
  labs(title = "How Lifestyle Interests Cluster by Psychological Dimension",
       subtitle = "Dashed = very strong loading (>= 0.70), Dotted = moderate/strong (>= 0.50)",
       x = "Interest", y = "Factor Loading", fill = "Factor Meaning") +
  theme_minimal()
```

Explanation

I started by exploring how participants' lifestyle interests related to one another. Using a Pearson correlation matrix and visualizing it with a heatmap, I compared 17 continuous variables ranging from sports and yoga to reading and clubbing. Clear patterns emerged: interests like concerts, movies, and music were strongly correlated, while activities like hiking and reading appeared negatively related to gaming and clubbing. These clusters suggested underlying dimensions in the data, which led me to use EFA to identify broader behavioral groupings.

Sub-Question 2: Exploratory Factor Analysis Model Comparison

```
# EFA with 3 factors
efa_3 <- fa(interests, nfactors = 3, rotate = "varimax", fm = "ml")

# EFA with 5 factors
efa_5 <- fa(interests, nfactors = 5, rotate = "varimax", fm = "ml")

# Loadings
print(efa_3$loadings, cutoff = 0.3)
```

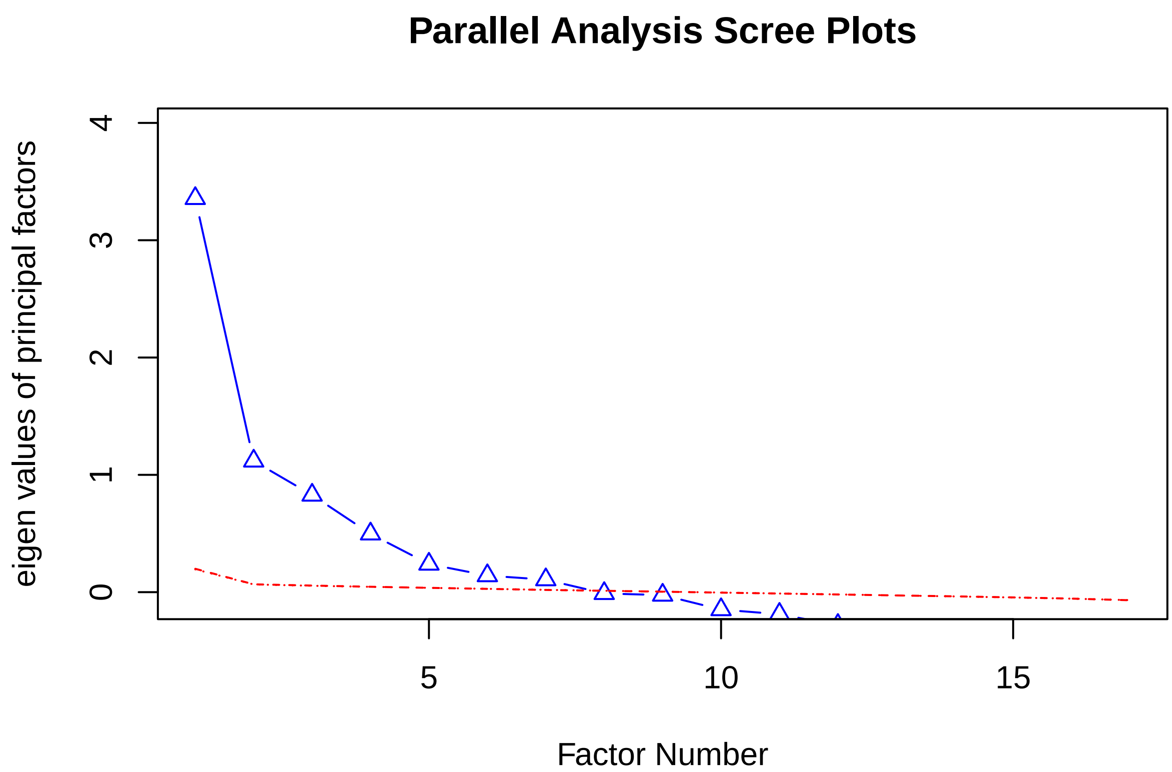
```
##
## Loadings:
##            ML1    ML3    ML2
```

```
## sports                  0.897
## tvsports                0.560
## exercise                0.431
## dining    0.315  0.362
## museums   0.898
## art       0.881
## hiking
## gaming
## clubbing
## reading   0.310
## tv               0.595
## theater   0.460  0.515
## movies           0.580
## concerts  0.342  0.454
## music            0.416
## shopping         0.591
## yoga
##
##                  ML1   ML3   ML2
## SS loadings    2.411 2.126 1.475
## Proportion Var 0.142 0.125 0.087
## Cumulative Var 0.142 0.267 0.354
```

```
print(efa_5$loadings, cutoff = 0.3)
```

```
##
## Loadings:
##          ML1    ML3    ML4    ML2    ML5
## sports                  0.851
## tvsports                0.597
## exercise                0.441
## dining                                0.576
## museums   0.918
## art       0.820
## hiking
## gaming
## clubbing                              0.341
## reading   0.349
## tv                             0.955
## theater   0.476
## movies           0.338         0.334
## concerts         0.926
## music            0.648
## shopping                       0.419  0.616
## yoga
##
##                  ML1   ML3   ML4   ML2   ML5
## SS loadings    2.149 1.673 1.422 1.399 1.223
## Proportion Var 0.126 0.098 0.084 0.082 0.072
## Cumulative Var 0.126 0.225 0.308 0.391 0.463
```

```
# Scree plot and parallel analysis to justify number of factors
fa.parallel(interests, fa = "fa", n.iter = 100, show.legend = FALSE)
```




```
## Parallel analysis suggests that the number of factors =  7  and the number of components =  NA
```

Explanation:

To dive deeper into those dimensions, I ran EFA models with both 3 and 5 factors, and compared them to the results of a parallel analysis. The 3-factor model revealed interpretable groupings: one cultural (museums, art, etc.), one athletic (sports, exercise, etc.), and one centered on media or social interests (TV, shopping, concerts, etc.). While the parallel analysis indicated up to seven factors might best fit the data, the 3 and 5-factor models still captured meaningful structure. These groupings gave me a more compact way to think about personality-related traits in the dataset.

Sub-Question 3: Does physical attractiveness correlate with the decision to match?

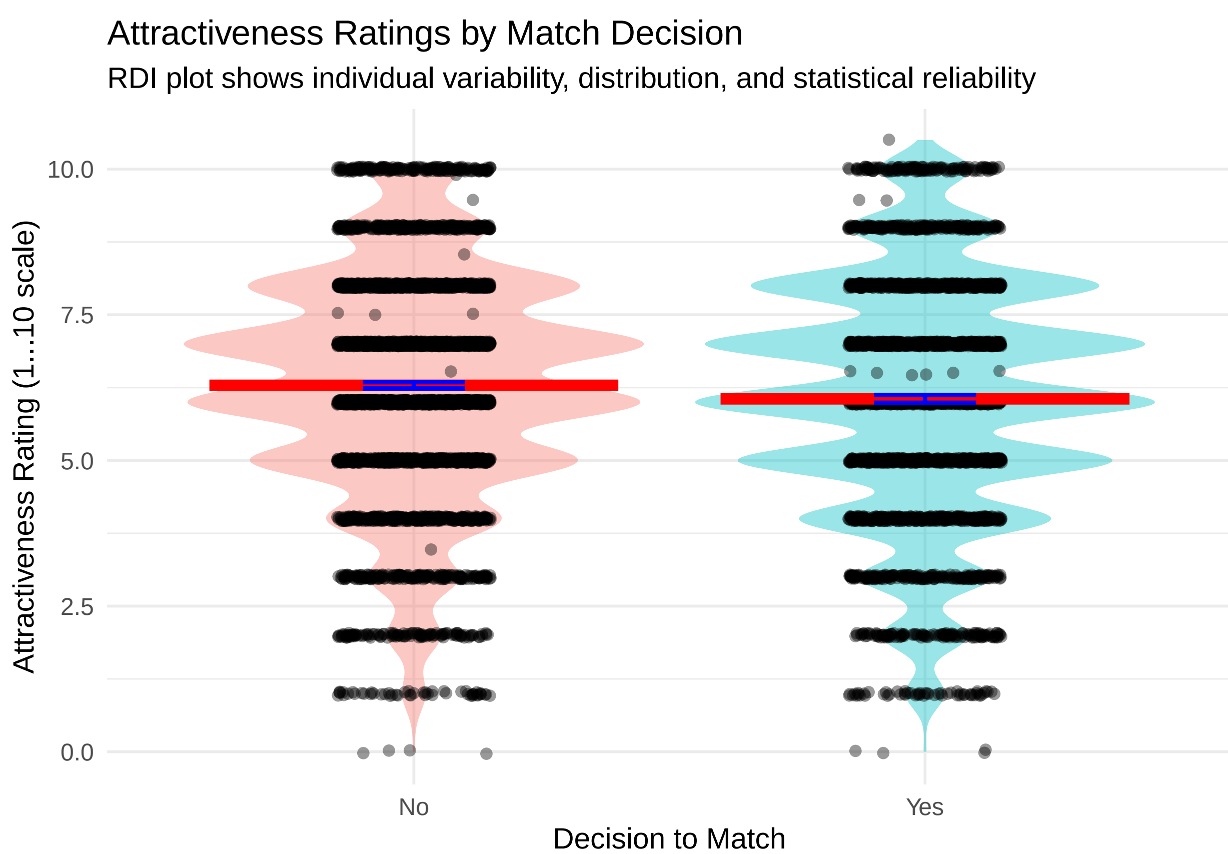
```
library(ggplot2)
library(dplyr)
library(ungeviz)
library(Hmisc)

df <- read.csv("Speed Dating Data.csv")

df_match <- df %>%
  filter(!is.na(attr_o), !is.na(dec)) %>%
  mutate(dec = factor(dec, levels = c(0, 1), labels = c("No", "Yes")))

# RDI plot instead of box plot from my presentation (for parsimony and interpretability for a lay audie
```

```r
ggplot(df_match, aes(x = dec, y = attr_o, fill = dec)) +
  geom_violin(alpha = 0.4, scale = "width", color = NA) +
  geom_jitter(width = 0.15, alpha = 0.4) +
  stat_summary(fun = "mean", geom = "hpline", width = 0.8, color = "red") +
  stat_summary(
    geom = "errorbar",
    fun.data = mean_cl_normal,
    width = 0.2,
    linewidth = 0.8,
    color = "blue"
  ) +
  labs(
    title = "Attractiveness Ratings by Match Decision",
    subtitle = "RDI plot shows individual variability, distribution, and statistical reliability",
    x = "Decision to Match",
    y = "Attractiveness Rating (1-10 scale)"
  ) +
  theme_minimal() +
  theme(legend.position = "none")
```




```r
# Probability of superiority
library(effectsize)

df_ps <- df %>%
  filter(!is.na(attr_o), !is.na(dec)) %>%
  mutate(dec = factor(dec))
```

```r
p_superiority(attr_o ~ dec, data = df_ps)
```

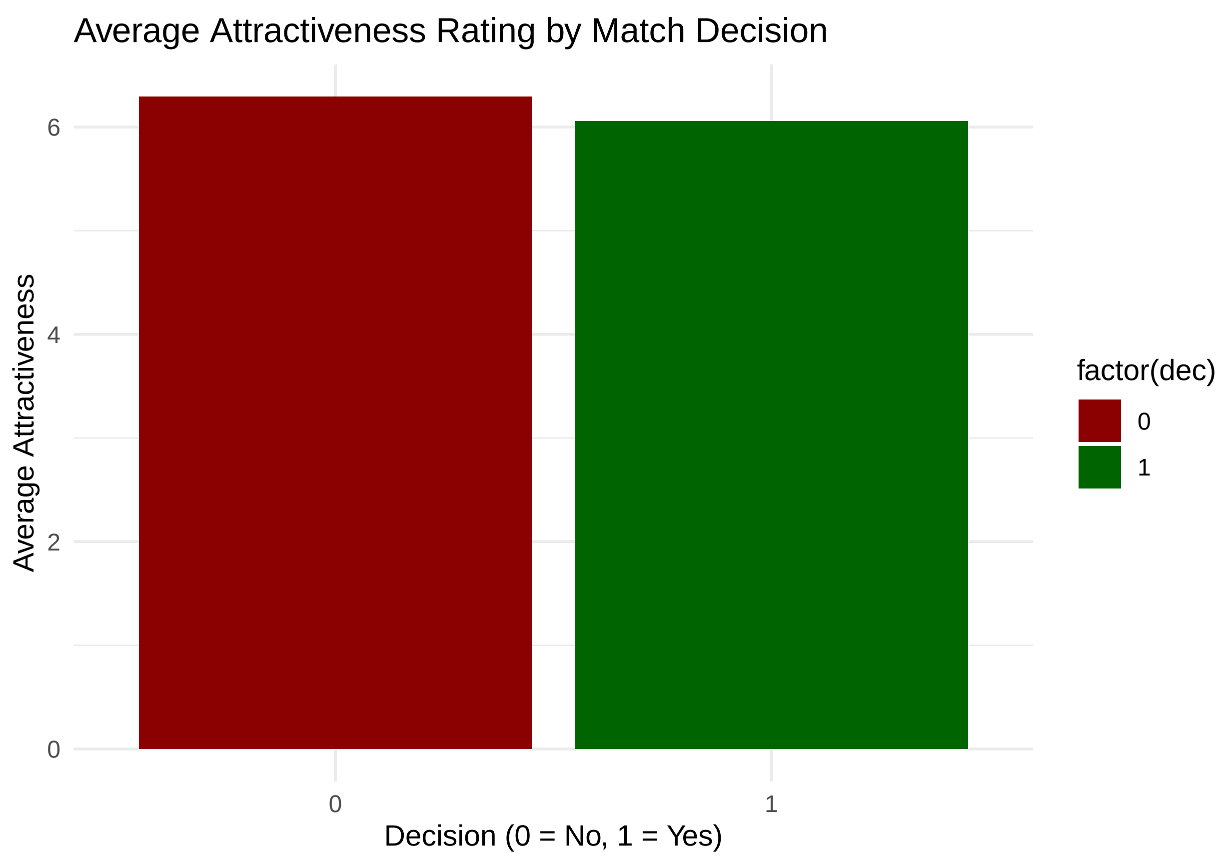
```
## Pr(superiority) |       95% CI
## ------------------------------
## 0.53            | [0.52, 0.55]
```

3a: Average Attractiveness by Match Decision (Layperson Visual)

```r
library(ggplot2)
library(dplyr)

df <- read.csv("Speed Dating Data.csv")
df_match <- df %>% filter(!is.na(attr_o))

# Plot mean attractiveness for yes/no decisions
df_match %>%
  group_by(dec) %>%
  summarise(mean_attr = mean(attr_o, na.rm = TRUE)) %>%
  ggplot(aes(x = factor(dec), y = mean_attr, fill = factor(dec))) +
  geom_col() +
  labs(title = "Average Attractiveness Rating by Match Decision",
       x = "Decision (0 = No, 1 = Yes)",
       y = "Average Attractiveness") +
  theme_minimal() + scale_fill_manual(values = c("0" = "darkred", "1" = "darkgreen"))
```



Explanation:

To examine how perceived attractiveness influences decision-making, I visualized participants' attractiveness ratings (attr_o) by match decision (dec). I used an RDI plot to provide a richer visual analysis - combining raw data points (jitter), distribution shape (violin), and statistical inferences (means and 95% CI error bars). This format reveals not only central tendencies but also the extent of overlap in ratings between "yes" and "no" decisions. Such overlap is critical in evaluating the predictive power of attractiveness. For 3a, I used a bar chart to convey the same story to a lay audience. Surprisingly, the mean rating appears slightly higher for dates participants said "no" to. This suggests that perceived physical attractiveness alone did not drive the decision to match. Other interpersonal dynamics - such as shared interests, humor, or chemistry - may have outweighed physical appearance in the final judgment. I also calculated the probability of superiority to quantify the difference in attractiveness ratings between match decisions. The result, 0.53 [95% CI: 0.52, 0.55], indicates a 53% chance that a randomly selected person someone said "no" to was rated as more attractive than one they said "yes" to. This aligns with the RDI and bar plots: attractiveness didn't strongly differentiate match decisions. The effect size is minimal, reinforcing the idea that non-physical factors likely carried more weight in these specific speed dating outcomes.

Sub-Question 4: Are men and women influenced by different attributes?

```r
library(ggplot2)
library(dplyr)
library(tidyr)
library(Hmisc)
library(ungeviz)

df <- read.csv("Speed Dating Data.csv")

df_gender <- df %>%
  select(gender, attr_o, intel_o, fun_o, sinc_o, shar_o, amb_o) %>%
  pivot_longer(cols = -gender, names_to = "trait", values_to = "rating") %>%
  filter(!is.na(rating), !is.na(gender)) %>%
  mutate(gender = factor(gender, labels = c("Female", "Male")))

# Revised from box to RDI plots once again
ggplot(df_gender, aes(x = gender, y = rating, fill = gender)) +
  geom_violin(alpha = 0.4, scale = "width") +
  geom_jitter(width = 0.15, alpha = 0.4) +
  stat_summary(fun = "mean", geom = "hpline", width = 0.8, color = "red") +
  stat_summary(geom = "errorbar", fun.data = mean_cl_normal, width = 0.2, color = "blue") +
  facet_wrap(~trait, scales = "free_y") +
  labs(
    title = "Trait Ratings by Gender",
    subtitle = "RDI plot: Individual ratings, distributions, and confidence intervals",
    x = "Gender", y = "Rating (1-10 scale)"
  ) +
  theme_minimal() +
  theme(legend.position = "none")
```

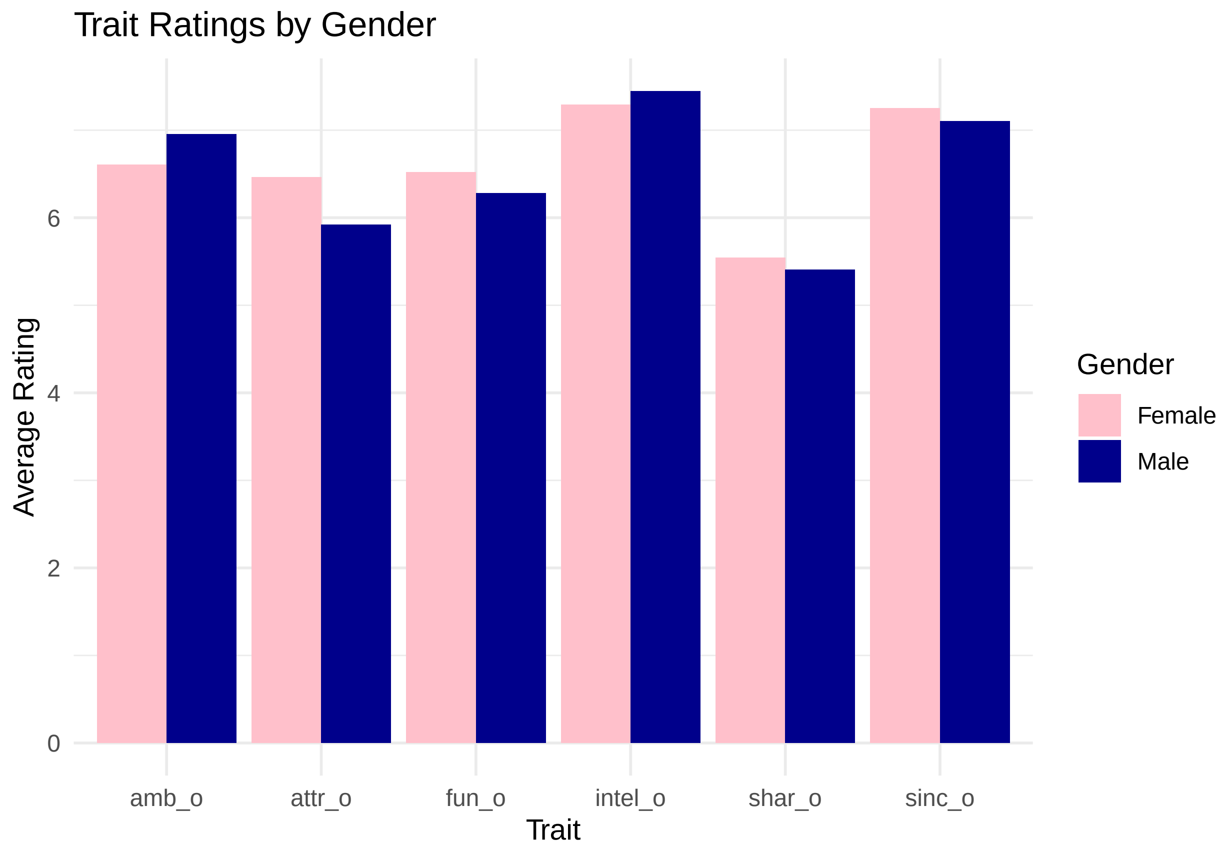
4a: Trait Ratings by Gender (Layperson Visual)

```
library(ggplot2)
library(dplyr)
library(tidyr)

df <- read.csv("Speed Dating Data.csv")

df_gender <- df %>%
  select(gender, dec, attr_o, intel_o, fun_o, sinc_o, shar_o, amb_o) %>%
  pivot_longer(cols = -c(gender, dec), names_to = "trait", values_to = "rating") %>%
  filter(!is.na(rating))

# Grouped bar chart of mean ratings by trait and gender
df_gender %>%
  group_by(gender, trait) %>%
  summarise(mean_rating = mean(rating, na.rm = TRUE)) %>%
  ggplot(aes(x = trait, y = mean_rating, fill = factor(gender))) +
  geom_col(position = "dodge") +
  labs(title = "Trait Ratings by Gender",
       x = "Trait", y = "Average Rating") +
  theme_minimal() + scale_fill_manual(values = c("0" = "pink", "1" = "darkblue"),
                    labels = c("Female", "Male"),
                    name = "Gender")
```



Explanation:

This analysis compares how men and women rated their dates across six key traits: attractiveness, intelligence, fun, sincerity, ambition, and shared interests. I first used an RDI plot to explore the data in depth for an expert audience. The violin plots show the distribution of individual ratings, while the jittered points highlight overlap and outliers. Red lines denote group means, and blue intervals show the 95% confidence range.

We see that while most differences are subtle, some trends emerge: - Men rated ambition and intelligence higher - Women rated sincerity and shared interests higher - Women gave slightly higher attractiveness and fun scores These results suggest mild gender-based preferences in how participants evaluated dates. However, the overlapping distributions indicate that evaluations were more alike than different overall, implying that first impressions in speed dating may be shaped by shared expectations, rather than polarized gender priorities.

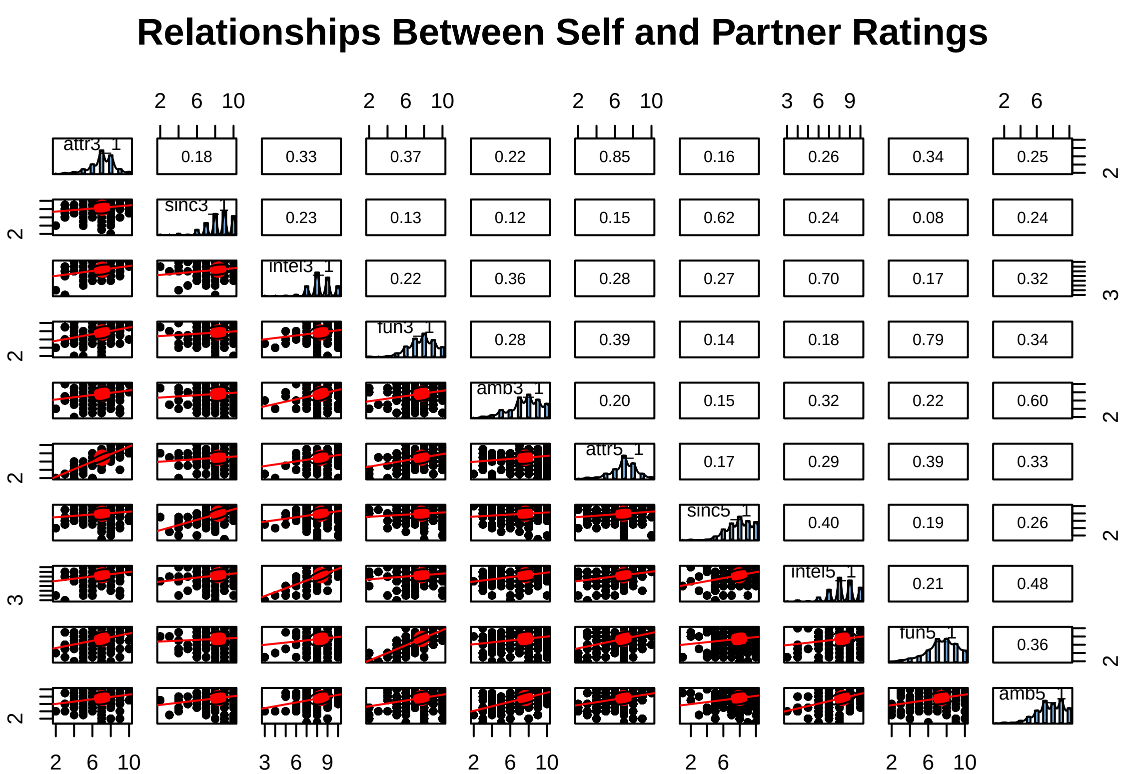
For a lay audience (4a), I simplified the visualization to a grouped bar chart that shows the mean rating per trait by gender. The direction of each difference is clear, and the relative balance helps illustrate the same takeaway: while there are some gender-based trends, no trait showed extreme divergence, suggesting that both men and women may be weighing a broader set of qualities when deciding whether to pursue a match.

Sub-Question 5: Do self-ratings align with how others perceive participants?

```
library(psych)
library(dplyr)
```

```r
# Select and clean the relevant traits
df_traits <- df %>%
  select(attr3_1, sinc3_1, intel3_1, fun3_1, amb3_1,
         attr5_1, sinc5_1, intel5_1, fun5_1, amb5_1) %>%
  na.omit()

# SPLOM (Scatterplot Matrix)
pairs.panels(df_traits,
             method = "pearson",
             hist.col = "#75AADB",
             density = TRUE,
             ellipses = TRUE,
             lm = TRUE,
             smooth = FALSE,
             jiggle = FALSE,
             factor = 2,
             cex = 0.8,
             pch = 20,
             main = "Relationships Between Self and Partner Ratings")
```




```r
# Cronbach's alpha for self-ratings
self_alpha <- alpha(df_traits[, 1:5])
print(self_alpha)
```

```
##
## Reliability analysis
## Call: alpha(x = df_traits[, 1:5])
##
```

```
##   raw_alpha std.alpha G6(smc) average_r S/N    ase mean  sd median_r
##        0.6      0.62    0.58      0.24 1.6 0.0089  7.8 0.9     0.23
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.59   0.6  0.62
## Duhachek  0.59   0.6  0.62
##
##  Reliability if an item is dropped:
##          raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## attr3_1       0.52      0.54    0.48      0.22 1.2   0.0110 0.0083  0.23
## sinc3_1       0.61      0.63    0.57      0.30 1.7   0.0089 0.0043  0.31
## intel3_1      0.52      0.53    0.47      0.22 1.1   0.0111 0.0090  0.20
## fun3_1        0.54      0.56    0.50      0.24 1.3   0.0108 0.0083  0.23
## amb3_1        0.55      0.56    0.51      0.24 1.3   0.0104 0.0081  0.23
##
##  Item statistics
##             n raw.r std.r r.cor r.drop mean  sd
## attr3_1  4906  0.66  0.67  0.55   0.42  7.1 1.4
## sinc3_1  4906  0.52  0.53  0.30   0.23  8.3 1.4
## intel3_1 4906  0.63  0.68  0.57   0.45  8.4 1.1
## fun3_1   4906  0.66  0.64  0.49   0.39  7.7 1.5
## amb3_1   4906  0.67  0.63  0.48   0.36  7.6 1.7
##
## Non missing response frequency for each item
##             2    3    4    5    6    7    8    9   10 miss
## attr3_1  0.00 0.02 0.03 0.07 0.14 0.35 0.28 0.07 0.03    0
## sinc3_1  0.00 0.00 0.02 0.01 0.06 0.15 0.25 0.28 0.23    0
## intel3_1 0.00 0.00 0.00 0.01 0.02 0.15 0.37 0.29 0.16    0
## fun3_1   0.01 0.00 0.01 0.05 0.12 0.22 0.27 0.20 0.12    0
## amb3_1   0.00 0.02 0.03 0.08 0.08 0.21 0.24 0.19 0.15    0
```

```r
# Cronbach's alpha for partner ratings
partner_alpha <- alpha(df_traits[, 6:10])
print(partner_alpha)
```

```
##
## Reliability analysis
## Call: alpha(x = df_traits[, 6:10])
##
##   raw_alpha std.alpha G6(smc) average_r S/N    ase mean  sd median_r
##       0.68      0.69    0.66      0.31 2.2 0.0072  7.6 1.1     0.31
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.67  0.68   0.7
## Duhachek  0.67  0.68   0.7
##
##  Reliability if an item is dropped:
##          raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## attr5_1       0.64      0.65    0.61      0.32 1.8   0.0085 0.0137  0.31
## sinc5_1       0.67      0.68    0.63      0.34 2.1   0.0075 0.0087  0.34
## intel5_1      0.61      0.61    0.55      0.28 1.6   0.0090 0.0085  0.29
## fun5_1        0.64      0.65    0.60      0.32 1.9   0.0083 0.0120  0.31
```

```
## amb5_1        0.59      0.60    0.56      0.27 1.5   0.0096 0.0106  0.25
##
##  Item statistics
##             n raw.r std.r r.cor r.drop mean  sd
## attr5_1  4906  0.64  0.65  0.51   0.43  6.9 1.5
## sinc5_1  4906  0.60  0.60  0.44   0.34  7.9 1.6
## intel5_1 4906  0.68  0.71  0.63   0.51  8.3 1.3
## fun5_1   4906  0.67  0.64  0.50   0.42  7.4 1.8
## amb5_1   4906  0.74  0.73  0.64   0.52  7.6 1.8
##
## Non missing response frequency for each item
##          1    2    3    4    5    6    7    8    9   10 miss
## attr5_1  0 0.01 0.03 0.03 0.09 0.15 0.35 0.24 0.08 0.03    0
## sinc5_1  0 0.01 0.00 0.01 0.04 0.11 0.18 0.25 0.20 0.19    0
## intel5_1 0 0.00 0.00 0.02 0.01 0.05 0.15 0.31 0.28 0.18    0
## fun5_1   0 0.01 0.02 0.04 0.06 0.12 0.23 0.23 0.18 0.12    0
## amb5_1   0 0.01 0.01 0.03 0.06 0.12 0.21 0.19 0.22 0.14    0
```

5a: Self vs. Partner Ratings (Layperson Visual)

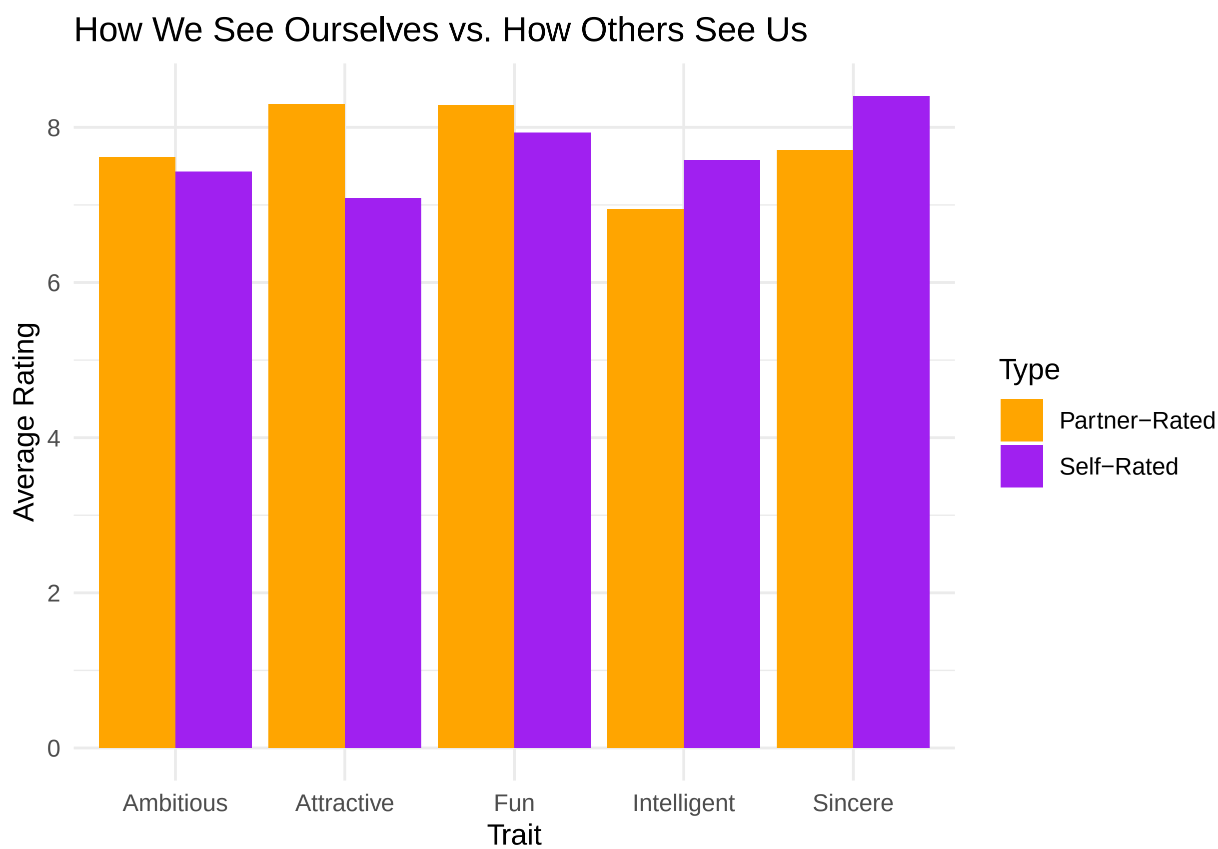
```r
library(ggplot2)

df <- read.csv("Speed Dating Data.csv")

df_self_other <- df %>%
  select(attr3_1, sinc3_1, intel3_1, fun3_1, amb3_1,
         attr5_1, sinc5_1, intel5_1, fun5_1, amb5_1)

# Calculate average scores
compare_df <- data.frame(
  Trait = rep(c("Attractive", "Sincere", "Intelligent", "Fun", "Ambitious"), each = 2),
  Type = rep(c("Self-Rated", "Partner-Rated"), times = 5),
  Score = c(
    colMeans(df_self_other[,1:5], na.rm = TRUE),
    colMeans(df_self_other[,6:10], na.rm = TRUE)
  )
)

# Plot
ggplot(compare_df, aes(x = Trait, y = Score, fill = Type)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "How We See Ourselves vs. How Others See Us",
       y = "Average Rating") +
  theme_minimal() + scale_fill_manual(values = c("Self-Rated" = "purple", "Partner-Rated" = "Orange"))
```



Explanation:

To examine how individuals view themselves versus how they believe others see them, I compared five core traits-attractiveness, sincerity, intelligence, fun, and ambition-using a scatterplot matrix and a grouped bar chart. The SPLOM revealed moderate correlations for traits like fun (r = .79) and ambition (r = .60), but weaker alignment for sincerity and attractiveness, suggesting people may be reasonably accurate in how fun or ambitious they appear, but less so in traits that are more subjective or socially sensitive. Cronbach's alpha was .60 for self-ratings and .68 for partner-ratings, indicating moderate internal consistency and justifying the use of trait profiles for comparison. The accompanying bar chart shows that individuals tended to rate themselves higher on sincerity and intelligence, while partners gave slightly higher marks for attractiveness and fun. These patterns suggest some degree of insight in self-perception, but also highlight meaningful gaps in how we see ourselves versus how others perceive us, even in brief social interactions.

Sub-Question 6: Can self-perceptions and gender predict saying "yes" to a date?

```r
library(dplyr)

# Prepare data
df_model <- df %>%
  select(dec, attr3_1, fun3_1, amb3_1, gender) %>%
  filter(!is.na(dec), !is.na(attr3_1), !is.na(fun3_1), !is.na(amb3_1), !is.na(gender))

# Convert outcome to factor for logistic regression
df_model$dec <- as.factor(df_model$dec)
```

```r
# Fit logistic regression model
log_model <- glm(dec ~ attr3_1 + fun3_1 + amb3_1 + gender, data = df_model, family = "binomial")

# Show model summary
summary(log_model)
```

```
##
## Call:
## glm(formula = dec ~ attr3_1 + fun3_1 + amb3_1 + gender, family = "binomial",
##     data = df_model)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.2566  -1.0759  -0.9277   1.2464   1.5577
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.38141    0.14742  -2.587  0.00967 **
## attr3_1     -0.07478    0.01813  -4.124 3.72e-05 ***
## fun3_1       0.02538    0.01652   1.536  0.12449
## amb3_1       0.02234    0.01367   1.633  0.10238
## gender       0.43233    0.04528   9.549  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 11254  on 8272  degrees of freedom
## Residual deviance: 11139  on 8268  degrees of freedom
## AIC: 11149
##
## Number of Fisher Scoring iterations: 4
```

Explanation:

To examine what predicts whether someone says "yes" to a date, I ran a logistic regression using gender and self-ratings of attractiveness, fun, and ambition. Gender was the strongest predictor: men were more likely to say yes than women, with a coefficient of 0.43. Self-rated attractiveness also had a significant effect; participants who viewed themselves as less attractive were more likely to say yes, with a coefficient of –0.07. Fun (0.03) and ambition (0.02) had small positive effects, but these were not statistically significant. These results suggest that match decisions may be shaped by self-image, particularly how attractive someone believes they are, and that people who rate themselves lower in attractiveness may be more open to connection.

Question 7: What Predicts a 'Yes' Decision? (A Multilevel Analysis of Self-Perception and Gender)

```r
library(lme4)
library(dplyr)
```

```r
# Prepare multilevel modeling dataset
df_ml <- df %>%
  select(dec, attr3_1, fun3_1, amb3_1, gender, iid) %>%
  filter(!is.na(dec), !is.na(attr3_1), !is.na(fun3_1), !is.na(amb3_1), !is.na(gender), !is.na(iid)) %>%
  mutate(dec = as.factor(dec),
         gender = as.factor(gender),
         iid = as.factor(iid))

# Use multilevel logistic regression (random intercept for participant)
ml_model <- glmer(dec ~ attr3_1 + fun3_1 + amb3_1 + gender + (1 | iid),
                  data = df_ml,
                  family = binomial)

# Summary
summary(ml_model)
```

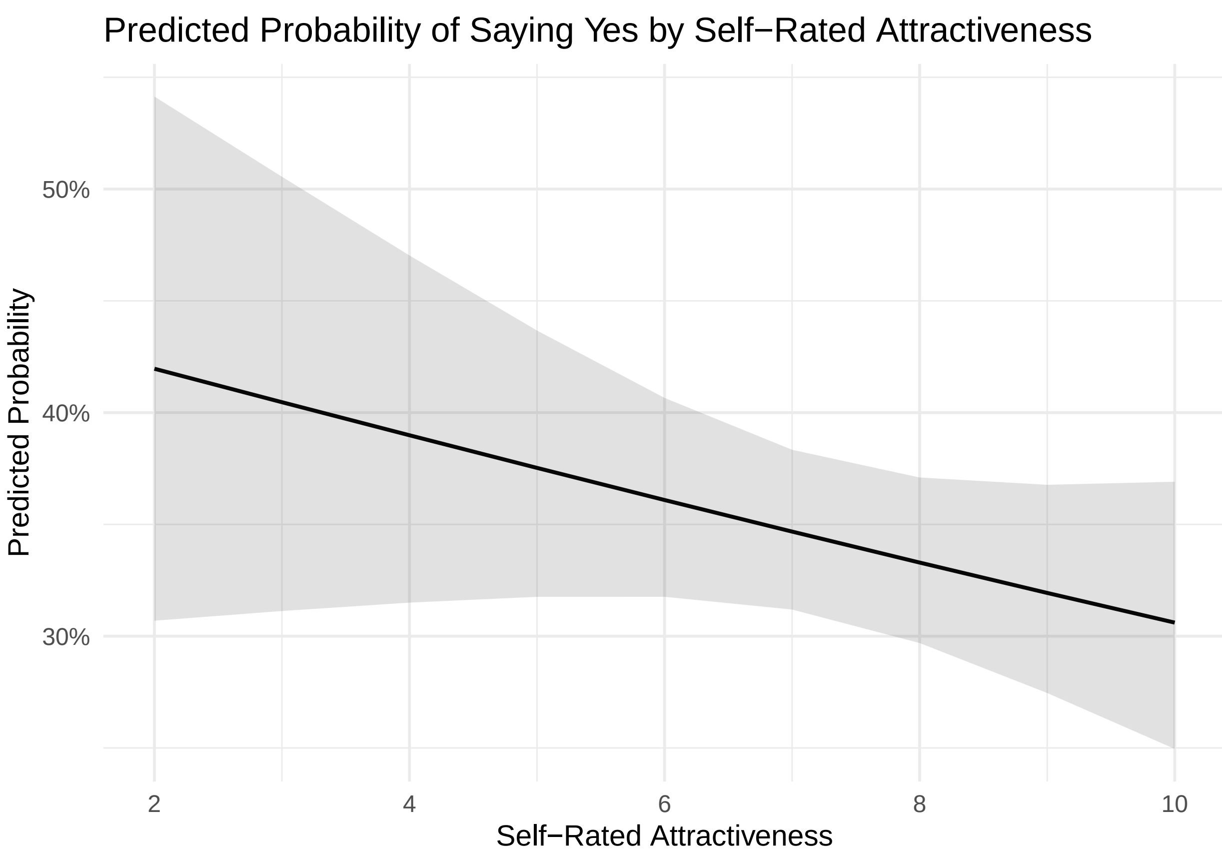
```
## Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
##  Family: binomial  ( logit )
## Formula: dec ~ attr3_1 + fun3_1 + amb3_1 + gender + (1 | iid)
##    Data: df_ml
##
##      AIC      BIC   logLik deviance df.resid
##  10215.5  10257.6  -5101.8  10203.5     8267
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6630 -0.7283 -0.4087  0.8599  3.0249
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  iid    (Intercept) 1.248    1.117
## Number of obs: 8273, groups:  iid, 542
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.46211    0.36488  -1.266    0.205
## attr3_1     -0.06178    0.04467  -1.383    0.167
## fun3_1       0.02088    0.04112   0.508    0.612
## amb3_1       0.01179    0.03325   0.354    0.723
## gender1      0.54862    0.11127   4.930  8.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) att3_1 fun3_1 amb3_1
## attr3_1 -0.442
## fun3_1  -0.378 -0.375
## amb3_1  -0.326 -0.157 -0.259
## gender1 -0.266  0.056  0.091 -0.024
```

Question 7a: Marginal Effects Plot (To Visualize Multilevel Regression Results)

```
library(ggeffects)
library(ggplot2)

# Predicted values for self-rated attractiveness
pred_attr <- ggpredict(ml_model, terms = "attr3_1 [2:10]")

# Marginal Effects Plot
plot(pred_attr) +
  labs(
    title = "Predicted Probability of Saying Yes by Self-Rated Attractiveness",
    x = "Self-Rated Attractiveness",
    y = "Predicted Probability"
  ) +
  theme_minimal()
```



Explanation

To account for the nested structure of the data-where each participant makes multiple match decisions-I fit a multilevel logistic regression with a random intercept for participant ID (iid). This model allows each person to have their own baseline log-odds of saying "yes" (log-odds being the natural logarithm of the odds, used as the linear predictor in logistic regression). By including a random intercept, the model accounts for unobserved individual-level variability that could otherwise bias fixed effect estimates. The estimated variance of the random intercept was 1.25 (standard deviation = 1.12), indicating substantial between-subject differences in overall openness to matches. Among the fixed effects, gender remained a strong and significant predictor: male participants had higher odds of saying yes (coefficient = 0.55). Self-rated attractiveness showed a negative association (–0.06), consistent with earlier results, but this effect was no longer statistically significant after accounting for individual differences. Fun and ambition showed small, positive, non-significant effects. The model's overall fit improved compared to the standard logistic regression, as indicated by a lower Akaike Information Criterion (AIC = 10215 vs. 11149). I learned that AIC is a metric used to compare model quality, balancing fit with complexity-lower AIC values indicate better relative fit with fewer unnecessary parameters. These results highlight the importance of accounting for the hierarchical structure of behavioral data, especially when individual response patterns vary substantially across repeated observations. The marginal effects plot illustrates how the predicted probability of saying "yes" varies based on self-rated attractiveness, holding other variables constant. The trend is negative: participants who rate themselves as more attractive are slightly less likely to say yes. This is consistent with the regression model, which found a negative coefficient for self-rated attractiveness. Although the effect is not statistically significant in the multilevel model, the plot suggests that participants with lower self-perceptions of attractiveness may be more open to connection.

Ethical Reflection

Checklist Reflection: Ethical Considerations in This Project

1. Does the dataset involve personal or sensitive information?
Yes. Although anonymized, the dataset captures how participants rate others-and themselves-on subjective traits like attractiveness, intelligence, and sincerity. These ratings involve implicit value judgments and socially loaded constructs, requiring careful interpretation and framing.

2. Were there systemic assumptions or biases built into the dataset?
Yes. The speed dating experiment was organized using binary gender categories and heterosexual pairings. This design excludes any other orientations, limiting the generalizability of the findings to a certain degree. It also reinforces (but in certain cases contests) gender norms about dating behavior, such as linking gender to specific preferences.

3. Were any groups disproportionately represented or underrepresented?
While the dataset is diverse in some demographic respects (e.g. race, age, and educational background), the structure and framing prioritize a narrow view of dating experiences. For example, the focus on immediate yes/no decisions may benefit extroverted or performative social styles over other forms of compatibility.

4. Did the analysis account for measurement limitations and variability?
Yes. Throughout the project, I visualized uncertainty using confidence intervals (e.g., in RDI plots and marginal effects graphs), and used Cronbach's alpha to assess internal consistency in trait constructs. I also acknowledged that constructs like self-rated attractiveness are inherently subjective and may be affected by self-esteem, social desirability bias, or cultural conditioning.

5. Were results communicated responsibly to avoid harm or misinterpretation?
Yes. In both my presentation and this document, I was careful not to claim that any one trait "causes" romantic success. Instead, I explored the relationships between self-perception, gender, and behavior, emphasizing that patterns reflect the specific context and framing of this dataset. I also highlighted overlap in trait ratings and distributions to avoid overgeneralizing based on group means.

Ethics Conclusion

This project built on the ethical frameworks I've developed across my Data Science coursework. I was taught to think critically about the absence of data, survey framing, and the manner in which data was collected. These principles carried over into this R-based analysis, where I focused on transparent data cleaning, responsible visualization, and contextualizing findings.

I chose to use multilevel modeling in part because it better reflects real-world behavioral complexity. By including a random intercept for each participant, I acknowledged that people vary in baseline willingness to say “yes” (a choice that helps prevent misleading conclusions about fixed trait effects). This decision demonstrates how ethical modeling isn’t just about data collection but also about how we quantify and represent human behavior. The requirements to compare both statistical models and visualization styles (expert vs. layperson visuals) further supports this goal of accessibility and transparency in analysis.