Topic ideas

Outliers: Matthew Hawkins, Aston Yong, Catherine Ning

October 7, 2021

Data Set 1

Introduction and Data

(Introduce and discuss data here) Bechdel Test Data Set

The data set comes from FiveThirtyEight, and, more specifically, from an article written on FiveThirtyEight's website.

The data was collected in a community-sourced manner, scoring a movie on a Bechdel test and then linking that score with an existing IMDB link which includes an array of other variables associated with that movie.

The observations are movies produced between 1970 and 2013, and the characteristics are different descriptors of the movie, regarding their production, their revenue, and their score on the Bechdel Test

The Bechdel test is a measure of the representation of women in a film, and it is on a scale of 0-3 First: The film must have two named women to get the first point.

Second: The two women must talk to each other (or two women in the film must have a conversation with each other) to get the second point.

Third: The conversation must be about something other than a man to get the third point.

Research questions

Does the Bechdel Test score have a significant effect on box office revenue? Additionally, what factors (year, budget, plot, etc.) correlate with Bechdel test score?

```
## Rows: 8,839
## Columns: 5
## $ year    <dbl> 1888, 1892, 1895, 1895, 1896, 1896, 1896, 1896, 1897, 1898, 18~
## $ id      <dbl> 8040, 5433, 6200, 5444, 5406, 5445, 6199, 4982, 9328, 4978, 54~
## $ imdb_id <chr> "0392728", "0000003", "0132134", "0000014", "0000131", "022334~
## $ title   <chr> "Roundhay Garden Scene", "Pauvre Pierrot", "The Execution of M~
## $ rating  <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0,~
```

```
## Rows: 1,794
## Columns: 34
## $ year          <dbl> 2013, 2012, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 20~
## $ imdb          <chr> "tt1711425", "tt1343727", "tt2024544", "tt1272878", "tt0~
## $ title         <chr> "21 &amp;amp; Over", "Dredd 3D", "12 Years a Slave", "2 Guns~
## $ test          <chr> "notalk", "ok-disagree", "notalk-disagree", "notalk", "m~
## $ clean_test    <chr> "notalk", "ok", "notalk", "notalk", "men", "men", "notal~
## $ binary        <chr> "FAIL", "PASS", "FAIL", "FAIL", "FAIL", "FAIL", "FAIL", ~
## $ budget        <dbl> 1.30e+07, 4.50e+07, 2.00e+07, 6.10e+07, 4.00e+07, 2.25e+~
## $ domgross      <chr> "25682380", "13414714", "53107035", "75612460", "9502021~
```

```
## $ intgross      <chr> "42195766", "40868994", "158607035", "132493015", "95020~
## $ code          <chr> "2013FAIL", "2012PASS", "2013FAIL", "2013FAIL", "2013FAI~
## $ budget_2013   <dbl> 13000000, 45658735, 20000000, 61000000, 40000000, 225000~
## $ domgross_2013 <chr> "25682380", "13611086", "53107035", "75612460", "9502021~
## $ intgross_2013 <chr> "42195766", "41467257", "158607035", "132493015", "95020~
## $ period_code   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ decade_code   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ imdb_id       <chr> "1711425", "1343727", "2024544", "1272878", "0453562", "~
## $ plot          <chr> NA, NA, "In the antebellum United States, Solomon Northu~
## $ rated         <chr> NA, NA, "R", "R", "PG-13", "PG-13", "R", "R", "PG-13", "~
## $ response      <lgl> NA, NA, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, ~
## $ language      <chr> NA, NA, "English", "English, Spanish", "English", "Engli~
## $ country       <chr> NA, NA, "USA, UK", "USA", "USA", "USA", "USA", "UK", "US~
## $ writer        <chr> NA, NA, "John Ridley (screenplay), Solomon Northup (base~
## $ metascore     <dbl> NA, NA, 97, 55, 62, 29, 28, 55, 48, 33, 90, 58, 52, 78, ~
## $ imdb_rating   <dbl> NA, NA, 8.3, 6.8, 7.6, 6.6, 5.4, 7.8, 5.7, 5.0, 7.5, 7.4~
## $ director      <chr> NA, NA, "Steve McQueen", "Baltasar Kormákur", "Brian Hel~
## $ released      <chr> NA, NA, "08 Nov 2013", "02 Aug 2013", "12 Apr 2013", "25~
## $ actors        <chr> NA, NA, "Chiwetel Ejiofor, Dwight Henry, Dickie Gravois,~
## $ genre         <chr> NA, NA, "Biography, Drama, History", "Action, Comedy, Cr~
## $ awards        <chr> NA, NA, "Won 3 Oscars. Another 131 wins & 137 nomination~
## $ runtime       <chr> NA, NA, "134 min", "109 min", "128 min", "118 min", "98 ~
## $ type          <chr> NA, NA, "movie", "movie", "movie", "movie", "movie", "mo~
## $ poster        <chr> NA, NA, "http://ia.media-imdb.com/images/M/MV5BMjExMTEzO~
## $ imdb_votes    <dbl> NA, NA, 143446, 87301, 43608, 25735, 123837, 85871, 1897~
## $ error         <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
```

```
## Rows: 1,794
## Columns: 36
## $ year          <dbl> 2013, 2012, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 20~
## $ imdb          <chr> "tt1711425", "tt1343727", "tt2024544", "tt1272878", "tt0~
## $ title         <chr> "21 &amp;amp; Over", "Dredd 3D", "12 Years a Slave", "2 Guns~
## $ test          <chr> "notalk", "ok-disagree", "notalk-disagree", "notalk", "m~
## $ clean_test    <chr> "notalk", "ok", "notalk", "notalk", "men", "men", "notal~
## $ binary        <chr> "FAIL", "PASS", "FAIL", "FAIL", "FAIL", "FAIL", "FAIL", ~
## $ budget        <dbl> 1.30e+07, 4.50e+07, 2.00e+07, 6.10e+07, 4.00e+07, 2.25e+~
## $ domgross      <chr> "25682380", "13414714", "53107035", "75612460", "9502021~
## $ intgross      <chr> "42195766", "40868994", "158607035", "132493015", "95020~
## $ code          <chr> "2013FAIL", "2012PASS", "2013FAIL", "2013FAIL", "2013FAI~
## $ budget_2013   <dbl> 13000000, 45658735, 20000000, 61000000, 40000000, 225000~
## $ domgross_2013 <chr> "25682380", "13611086", "53107035", "75612460", "9502021~
## $ intgross_2013 <chr> "42195766", "41467257", "158607035", "132493015", "95020~
## $ period_code   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ decade_code   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ imdb_id       <chr> "1711425", "1343727", "2024544", "1272878", "0453562", "~
## $ plot          <chr> NA, NA, "In the antebellum United States, Solomon Northu~
## $ rated         <chr> NA, NA, "R", "R", "PG-13", "PG-13", "R", "R", "PG-13", "~
## $ response      <lgl> NA, NA, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, ~
## $ language      <chr> NA, NA, "English", "English, Spanish", "English", "Engli~
## $ country       <chr> NA, NA, "USA, UK", "USA", "USA", "USA", "USA", "UK", "US~
## $ writer        <chr> NA, NA, "John Ridley (screenplay), Solomon Northup (base~
## $ metascore     <dbl> NA, NA, 97, 55, 62, 29, 28, 55, 48, 33, 90, 58, 52, 78, ~
## $ imdb_rating   <dbl> NA, NA, 8.3, 6.8, 7.6, 6.6, 5.4, 7.8, 5.7, 5.0, 7.5, 7.4~
## $ director      <chr> NA, NA, "Steve McQueen", "Baltasar Kormákur", "Brian Hel~
```

```
## $ released      <chr> NA, NA, "08 Nov 2013", "02 Aug 2013", "12 Apr 2013", "25~
## $ actors        <chr> NA, NA, "Chiwetel Ejiofor, Dwight Henry, Dickie Gravois,~
## $ genre         <chr> NA, NA, "Biography, Drama, History", "Action, Comedy, Cr~
## $ awards        <chr> NA, NA, "Won 3 Oscars. Another 131 wins & 137 nomination~
## $ runtime       <chr> NA, NA, "134 min", "109 min", "128 min", "118 min", "98 ~
## $ type          <chr> NA, NA, "movie", "movie", "movie", "movie", "movie", "mo~
## $ poster        <chr> NA, NA, "http://ia.media-imdb.com/images/M/MV5BMjExMTEzO~
## $ imdb_votes    <dbl> NA, NA, 143446, 87301, 43608, 25735, 123837, 85871, 1897~
## $ error         <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ id            <dbl> 4194, 3532, 5033, 4450, 4054, 4989, 4975, 4707, 4012, 41~
## $ rating        <dbl> 1, 3, 3, 1, 2, 2, 1, 3, 3, 1, 3, 3, 3, 3, 1, 3, 3, 3, 3,~
```

We combined the two data sets above into a single data set for use in the project. Additionally, we will filter out missing values as we use the data to avoid errors and faulty conclusions.

Data Set 2

Introduction and Data

(Introduce and discuss data here) Speed Dating Data Set

The data comes from a 2006 study by Raymond Fisman, Sheena S. Iyengar, Emir Kamenica, Itamar Simonson in The Quarterly Journal of Economics. The study of Columbia University students seeks to analyze gender differences in dating preferences.

Data was collected from a Speed Dating experiment on Columbia students where Fisman and Iyengar generated random matchings and random variation in the number of potential partners. Participants had 4 minute conversations with each other and answered a survey of questions after. The meetings took place at local restaurants/bars during 2002-2004. There were 392 subjects and over 8,000 observations.

The dataset is substantial with over 8,000 observations with regards to a speed dating test conducted over the period of 2002-2004. These observations record attributes including gender, race, income, as well as perceived attractiveness in a scale of 1-10.

Research questions

(Discuss research questions here)

Do race, gender, and income impact the rating by their partner during speed dating? Is there a relationship between age/background variables and desired date chracteristics?

Glimpse of data sets

Data set 1

```
## Rows: 1,794
## Columns: 34
## $ year          <dbl> 2013, 2012, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 20~
## $ imdb          <chr> "tt1711425", "tt1343727", "tt2024544", "tt1272878", "tt0~
## $ title         <chr> "21 &amp;amp; Over", "Dredd 3D", "12 Years a Slave", "2 Guns~
## $ test          <chr> "notalk", "ok-disagree", "notalk-disagree", "notalk", "m~
## $ clean_test    <chr> "notalk", "ok", "notalk", "notalk", "men", "men", "notal~
## $ binary        <chr> "FAIL", "PASS", "FAIL", "FAIL", "FAIL", "FAIL", "FAIL", ~
## $ budget        <dbl> 1.30e+07, 4.50e+07, 2.00e+07, 6.10e+07, 4.00e+07, 2.25e+~
## $ domgross      <chr> "25682380", "13414714", "53107035", "75612460", "9502021~
## $ intgross      <chr> "42195766", "40868994", "158607035", "132493015", "95020~
```

```
## $ code          <chr> "2013FAIL", "2012PASS", "2013FAIL", "2013FAIL", "2013FAI~
## $ budget_2013   <dbl> 13000000, 45658735, 20000000, 61000000, 40000000, 225000~
## $ domgross_2013 <chr> "25682380", "13611086", "53107035", "75612460", "9502021~
## $ intgross_2013 <chr> "42195766", "41467257", "158607035", "132493015", "95020~
## $ period_code   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ decade_code   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ imdb_id       <chr> "1711425", "1343727", "2024544", "1272878", "0453562", "~
## $ plot          <chr> NA, NA, "In the antebellum United States, Solomon Northu~
## $ rated         <chr> NA, NA, "R", "R", "PG-13", "PG-13", "R", "R", "PG-13", "~
## $ response      <lgl> NA, NA, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, ~
## $ language      <chr> NA, NA, "English", "English, Spanish", "English", "Engli~
## $ country       <chr> NA, NA, "USA, UK", "USA", "USA", "USA", "USA", "UK", "US~
## $ writer        <chr> NA, NA, "John Ridley (screenplay), Solomon Northup (base~
## $ metascore     <dbl> NA, NA, 97, 55, 62, 29, 28, 55, 48, 33, 90, 58, 52, 78, ~
## $ imdb_rating   <dbl> NA, NA, 8.3, 6.8, 7.6, 6.6, 5.4, 7.8, 5.7, 5.0, 7.5, 7.4~
## $ director      <chr> NA, NA, "Steve McQueen", "Baltasar Kormákur", "Brian Hel~
## $ released      <chr> NA, NA, "08 Nov 2013", "02 Aug 2013", "12 Apr 2013", "25~
## $ actors        <chr> NA, NA, "Chiwetel Ejiofor, Dwight Henry, Dickie Gravois,~
## $ genre         <chr> NA, NA, "Biography, Drama, History", "Action, Comedy, Cr~
## $ awards        <chr> NA, NA, "Won 3 Oscars. Another 131 wins & 137 nomination~
## $ runtime       <chr> NA, NA, "134 min", "109 min", "128 min", "118 min", "98 ~
## $ type          <chr> NA, NA, "movie", "movie", "movie", "movie", "movie", "mo~
## $ poster        <chr> NA, NA, "http://ia.media-imdb.com/images/M/MV5BMjExMTEzO~
## $ imdb_votes    <dbl> NA, NA, 143446, 87301, 43608, 25735, 123837, 85871, 1897~
## $ error         <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
```

Data set 2

```
## Rows: 8,378
## Columns: 195
## $ iid      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3~
## $ id       <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3~
## $ gender   <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ idg      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 5~
## $ condtn   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ wave     <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ round    <dbl> 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 1~
## $ position <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 9~
## $ positin1 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ order    <dbl> 4, 3, 10, 5, 7, 6, 1, 2, 8, 9, 10, 9, 6, 1, 3, 2, 7, 8, 4, 5,~
## $ partner  <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10,~
## $ pid      <dbl> 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 11, 12, 13, 14, 15, 1~
## $ match    <dbl> 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0~
## $ int_corr <dbl> 0.14, 0.54, 0.16, 0.61, 0.21, 0.25, 0.34, 0.50, 0.28, -0.36, ~
## $ samerace <dbl> 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1~
## $ age_o    <dbl> 27, 22, 22, 23, 24, 25, 30, 27, 28, 24, 27, 22, 22, 23, 24, 2~
## $ race_o   <dbl> 2, 2, 4, 2, 3, 2, 2, 2, 2, 2, 2, 2, 4, 2, 3, 2, 2, 2, 2, 2, 2~
## $ pf_o_att <dbl> 35.00, 60.00, 19.00, 30.00, 30.00, 50.00, 35.00, 33.33, 50.00~
## $ pf_o_sin <dbl> 20.00, 0.00, 18.00, 5.00, 10.00, 0.00, 15.00, 11.11, 0.00, 0.~
## $ pf_o_int <dbl> 20.00, 0.00, 19.00, 15.00, 20.00, 30.00, 25.00, 11.11, 25.00,~
## $ pf_o_fun <dbl> 20.00, 40.00, 18.00, 40.00, 10.00, 10.00, 10.00, 11.11, 10.00~
## $ pf_o_amb <dbl> 0.00, 0.00, 14.00, 5.00, 10.00, 0.00, 5.00, 11.11, 0.00, 0.00~
## $ pf_o_sha <dbl> 5.00, 0.00, 12.00, 5.00, 20.00, 10.00, 10.00, 22.22, 15.00, 0~
## $ dec_o    <dbl> 0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0~
```

```
## $ attr_o   <dbl> 6, 7, 10, 7, 8, 7, 3, 6, 7, 6, 8, 7, 10, 9, 10, 7, 5, 7, 8, 6~
## $ sinc_o   <dbl> 8, 8, 10, 8, 7, 7, 6, 7, 7, 6, 7, 6, 10, 9, 10, 8, 3, 7, 6, 5~
## $ intel_o  <dbl> 8, 10, 10, 9, 9, 8, 7, 5, 8, 6, 6, 10, 10, 9, 10, 7, 4, 7, 9,~
## $ fun_o    <dbl> 8, 7, 10, 8, 6, 8, 5, 6, 8, 6, 9, 6, 10, 9, 10, 5, 3, 7, 9, 7~
## $ amb_o    <dbl> 8, 7, 10, 9, 9, 7, 8, 8, 8, 6, 7, 6, 10, 9, 7, 7, 5, 7, 8, 9,~
## $ shar_o   <dbl> 6, 5, 10, 8, 7, 7, 7, 6, 9, 6, 4, 5, 10, 9, 8, 7, 3, 5, 7, 7,~
## $ like_o   <dbl> 7.0, 8.0, 10.0, 7.0, 8.0, 7.0, 2.0, 7.0, 6.5, 6.0, 7.0, 8.0, ~
## $ prob_o   <dbl> 4, 4, 10, 7, 6, 6, 1, 5, 8, 6, 2, 4, 10, 7, 1, 5, 3, 6, 8, 9,~
## $ met_o    <dbl> 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 1~
## $ age      <dbl> 21, 21, 21, 21, 21, 21, 21, 21, 21, 21, 24, 24, 24, 24, 24, 2~
## $ field    <chr> "Law", "Law", "Law", "Law", "Law", "Law", "Law", "Law", "Law"~
## $ field_cd <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2~
## $ undergra <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ mn_sat   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ tuition  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ race     <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2~
## $ imprace  <dbl> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 8~
## $ imprelig <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 4~
## $ from     <chr> "Chicago", "Chicago", "Chicago", "Chicago", "Chicago", "Chica~
## $ zipcode  <dbl> 60521, 60521, 60521, 60521, 60521, 60521, 60521, 60521, 60521~
## $ income   <dbl> 69487, 69487, 69487, 69487, 69487, 69487, 69487, 69487, 69487~
## $ goal     <dbl> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 6~
## $ date     <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 3~
## $ go_out   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ career   <chr> "lawyer", "lawyer", "lawyer", "lawyer", "lawyer", "lawyer", "~
## $ career_c <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sports   <dbl> 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3~
## $ tvsports <dbl> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 8~
## $ exercise <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7~
## $ dining   <dbl> 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10, 10, 10, 10,~
## $ museums  <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 5~
## $ art      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5~
## $ hiking   <dbl> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 8~
## $ gaming   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 4~
## $ clubbing <dbl> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 5~
## $ reading  <dbl> 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 10, 10, 10, 10, 10, 10, 10, 10,~
## $ tv       <dbl> 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 8~
## $ theater  <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, 7~
## $ movies   <dbl> 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 8, 8, 8, 8, 8, 8, 8, ~
## $ concerts <dbl> 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 7, 7, 7, 7, 7, 7, 7, ~
## $ music    <dbl> 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 5~
## $ shopping <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 8~
## $ yoga     <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 7~
## $ exphappy <dbl> 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4~
## $ expnum   <dbl> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 2~
## $ attr1_1  <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 45, 45, 45, 45, 45, 4~
## $ sinc1_1  <dbl> 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 5, 5, 5, 5, 5, 5, 5, ~
## $ intel1_1 <dbl> 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 25, 25, 25, 25, 25, 2~
## $ fun1_1   <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 20, 20, 20, 20, 20, 2~
## $ amb1_1   <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 0, 0, 0, 0, 0, 0, 0, ~
## $ shar1_1  <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 5, 5, 5, 5, 5, 5, 5, ~
## $ attr4_1  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc4_1  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel4_1 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

```
## $ fun4_1   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb4_1   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar4_1  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr2_1  <dbl> 35, 35, 35, 35, 35, 35, 35, 35, 35, 35, 65, 65, 65, 65, 65, 6~
## $ sinc2_1  <dbl> 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 0, 0, 0, 0, 0, 0, 0, ~
## $ intel2_1 <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 10, 10, 10, 10, 10, 1~
## $ fun2_1   <dbl> 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 25, 25, 25, 25, 25, 2~
## $ amb2_1   <dbl> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ shar2_1  <dbl> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ attr3_1  <dbl> 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 7, 7, 8~
## $ sinc3_1  <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 5, 5, 5, 5, 5, 5, 5, 5, 5, 9~
## $ fun3_1   <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 10, 10, 10, 10, 10, 10, 10, 10,~
## $ intel3_1 <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 9~
## $ amb3_1   <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 3, 3, 3, 3, 3, 3, 3, 3, 3, 8~
## $ attr5_1  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc5_1  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel5_1 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun5_1   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb5_1   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ dec      <dbl> 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0~
## $ attr     <dbl> 6, 7, 5, 7, 5, 4, 7, 4, 7, 5, 5, 8, 5, 7, 6, 8, 7, 5, 7, 6, 7~
## $ sinc     <dbl> 9, 8, 8, 6, 6, 9, 6, 9, 6, 6, 7, 5, 8, 9, 8, 7, 5, 8, 6, 7, 9~
## $ intel    <dbl> 7, 7, 9, 8, 7, 7, 7, 7, 8, 6, 8, 6, 9, 7, 7, 8, 9, 7, 8, 8, 1~
## $ fun      <dbl> 7, 8, 8, 7, 7, 4, 4, 6, 9, 8, 4, 6, 6, 6, 9, 3, 6, 5, 9, 7, 7~
## $ amb      <dbl> 6, 5, 5, 6, 6, 6, 6, 5, 8, 10, 6, 9, 3, 5, 7, 6, 7, 9, 4, 9, ~
## $ shar     <dbl> 5, 6, 7, 8, 6, 4, 7, 6, 8, 8, 3, 6, 4, 7, 8, 2, 9, 5, 5, 8, 9~
## $ like     <dbl> 7, 7, 7, 7, 6, 6, 6, 6, 7, 6, 6, 7, 6, 7, 8, 6, 8, 5, 5, 8, 8~
## $ prob     <dbl> 6, 5, NA, 6, 6, 5, 5, 7, 7, 6, 4, 3, 7, 8, 6, 5, 7, 6, 6, 7, ~
## $ met      <dbl> 2, 1, 1, 2, 2, 2, 2, NA, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, ~
## $ match_es <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, N~
## $ attr1_s  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc1_s  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel1_s <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun1_s   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb1_s   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar1_s  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr3_s  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc3_s  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel3_s <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun3_s   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb3_s   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ satis_2  <dbl> 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5, 5, 5, 5, 5, 5, 5, 5, 5, N~
## $ length   <dbl> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, N~
## $ numdat_2 <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, NA, NA, NA, NA, NA, NA, NA, NA,~
## $ attr7_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc7_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel7_2 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun7_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb7_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar7_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr1_2  <dbl> 19.44, 19.44, 19.44, 19.44, 19.44, 19.44, 19.44, 19.44, 19.44~
## $ sinc1_2  <dbl> 16.67, 16.67, 16.67, 16.67, 16.67, 16.67, 16.67, 16.67, 16.67~
## $ intel1_2 <dbl> 13.89, 13.89, 13.89, 13.89, 13.89, 13.89, 13.89, 13.89, 13.89~
## $ fun1_2   <dbl> 22.22, 22.22, 22.22, 22.22, 22.22, 22.22, 22.22, 22.22, 22.22~
```

```
## $ amb1_2   <dbl> 11.11, 11.11, 11.11, 11.11, 11.11, 11.11, 11.11, 11.11, 11.11~
## $ shar1_2  <dbl> 16.67, 16.67, 16.67, 16.67, 16.67, 16.67, 16.67, 16.67, 16.67~
## $ attr4_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc4_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel4_2 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun4_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb4_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar4_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr2_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc2_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel2_2 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun2_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb2_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar2_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr3_2  <dbl> 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, N~
## $ sinc3_2  <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, N~
## $ intel3_2 <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, N~
## $ fun3_2   <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, N~
## $ amb3_2   <dbl> 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, N~
## $ attr5_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc5_2  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel5_2 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun5_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb5_2   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ you_call <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, N~
## $ them_cal <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, N~
## $ date_3   <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, N~
## $ numdat_3 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ num_in_3 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr1_3  <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 30, 30, 30, 30, 30, 3~
## $ sinc1_3  <dbl> 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 5, 5, 5, 5, 5, 5, 5, ~
## $ intel1_3 <dbl> 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 40, 40, 40, 40, 40, 4~
## $ fun1_3   <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 1~
## $ amb1_3   <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 0, 0, 0, 0, 0, 0, 0, ~
## $ shar1_3  <dbl> 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 10, 10, 10, 10, 10, 1~
## $ attr7_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc7_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel7_3 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun7_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb7_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar7_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr4_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc4_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel4_3 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun4_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb4_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar4_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr2_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc2_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel2_3 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun2_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb2_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ shar2_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ attr3_3  <dbl> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, N~
```

```
## $ sinc3_3  <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, N~
## $ intel3_3 <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, N~
## $ fun3_3   <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, N~
## $ amb3_3   <dbl> 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, N~
## $ attr5_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ sinc5_3  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ intel5_3 <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ fun5_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ amb5_3   <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```