

# 2024 USRA - Validation Study Analysis

Benjamin J. Zubaly

2024-05-08

## Table of contents

<b>Introduction</b>	<b>1</b>
Data Cleaning . . . . .	2
Variables in the Analysis . . . . .	3
<b>Analysis</b>	<b>9</b>
Data Exploration . . . . .	9
Descriptive Statistics . . . . .	10
Data Visualization . . . . .	15
Psychometric Properties of Religious Behavioral Intentions Scales . . . . .	39
Reliability . . . . .	39
Validity . . . . .	47
Testing Hypotheses . . . . .	51
Hypothesis 1 . . . . .	51
Hypothesis 2 . . . . .	57
Estimating the Relationship Between Disgust and Religiosity . . . . .	62
<b>Brief Discussion</b>	<b>75</b>

## Introduction

In recent decades there has been substantial interest by evolutionary scientists in the role of the behavioral immune system in regulating complex social behavior. Fincher & Thornhill (2012) proposed the hypothesis that religiosity should function to protect against parasites through (1) maintaining out-group separation and (2) maintaining in-group embeddedness. If this hypothesis is correct, religious behaviors belonging to these categories should be motivated by parasite disgust. The following analysis is an attempt to answer the question: does parasite disgust sensitivity influence religious behaviours that may function to protect against infectious

disease? This analysis will inform future studies on the basis for the relationship between disgust and religiosity. The goals of this analysis are as follows:

1. Establish the reliability and validity of our religious behavioral intentions scales (see [Variables in the Analysis](#) below).
2. Test the following hypotheses:

Hypothesis 1: Religious individuals scoring higher in pathogen disgust sensitivity will report being more likely to engage in religious behaviors that increase in-group embeddedness.

Hypothesis 2: Religious individuals scoring higher in pathogen disgust sensitivity will report being less likely to engage in religious behaviors that decrease out-group separation.

3. Estimate the relationship between disgust sensitivity and religiosity in a Christian sample.

Before describing the data set and its' variables, we will clean the data.

## Data Cleaning

I have written a script to clean the data called `data_cleaning.R` which is in the `code` directory of this repository. The following code runs the script, and this script will read in the data (which is currently in the `data` directory as `validation_data_ready.csv`) as `unclean.data`, reformats the variables into the correct data types, produces numerical item values for each scale, calculates scores for each scale, and returns the clean data as a data frame called `data`. The script uses the `groundhog` package (Simonsohn et al., 2024) to load the packages from a specific date—averting problems of package dependencies and making the code reproducible in the future—which may cause it to take longer than usual to run.

```
# Running the script to clean the data
source(file = "./code/data_cleaning.R")
```

Attached: 'Groundhog' (Version: 3.2.0)

Tips and troubleshooting: <https://groundhogR.com>

```
# Writing the new data frame to the data folder as validation_data_clean.csv
write.csv(data, file = "./data/validation_data_clean.csv", row.names = FALSE)
```

The clean data is now a data frame object in the environment called `data`.

## Variables in the Analysis

### Data Frame: Data

#### Variables:

- Demographic Variables:
  - **id**: A randomly generated ID to track participants.
  - **date\_start**: The date when participants started the survey in the format of “YYYY-MM-DD HH-MM-SS”.
  - **date\_end**: The date when participants finished the survey in the format of “YYYY-MM-DD HH-MM-SS”.
  - **time\_taken**: The amount of time it took to complete the survey (**date\_end** - **date\_start**) in minutes.
  - **rec\_meth**: The method of recruitment. This is a factor variable with four levels: “Prolific”, “Social Media”, “SONA”, and “Word of Mouth”.
  - **social\_media\_platform**: For social media recruits, the platform from which they were recruited. This is a factor variable with three levels: “Facebook”, “Instagram”, and “Other”.
  - **educ\_complete**: The highest level of education completed. This is a factor variable with five levels: “Graduate degree (MA/MSc/MPhil/other)”, “High school diploma/A-levels”, “Secondary education (e.g. GED/GCSE)”, “Technical/community college”, and “Undergraduate degree (BA/BSc/other)”.
  - **educ\_currently\_in**: The level of education currently engaged in, if applicable. This is a factor variable with four levels: “Doctorate degree (PhD/other)”, “Graduate degree (MA/MSc/MPhil/other)”, “Technical/community college”, and “Undergraduate degree (BA/BSc/other)”.
  - **age**: Age in years.
  - **gender**: Participants open-ended gender response. This is a factor variable and through investigating the reported categories, I decided to coerce responses into three categories: “Female”, “Male”, and “Non-binary”. The only response that did not match these words was one response of “man”, which was categorized as “Male”.
  - **religious\_affiliation**: Religious affiliation. This is a factor variable, and because we only selected Christian participants there is only one level: “Christianity (e.g. Baptist, Church of England, Roman Catholic, Methodist, Jehovah’s Witness, etc.)”.

- **christian\_affiliation**: Christian branch or denomination affiliation. This is a factor variable with 10 levels: “Catholic Anglican / Episcopalian”, “Protestant”, “Pentecostal / Apostolic”, “Methodist”, “Lutheran”, “Orthodox / Eastern Orthodox”, “Baptist”, “Calvinist / Reformed / Presbyterian”, “Non-denominational”, and “Other”.
- Religious Behavioral Intentions measures:
  - The items of this questionnaire were conceptualized by our group in PSYC 322, and they are written to represent the two categories of religious behaviour emphasized in Fincher & Thornhill (2012). That is, they are meant to represent religious behaviour that function to (1) increase out-group separation and (2) increase in-group embeddedness.
    - \* Because we found it difficult to come up with religious behaviors that we thought would function to increase out-group separation and would not be seen as immoral (and therefore create demand characteristics), we decided to go the opposite direction; the first subscale of the questionnaire is therefore made up of behaviors that function to *decrease* out-group separation, and we expect that this correlates negatively with parasite disgust sensitivity. For behaviors that function to increase in-group embeddedness, we tried to get at these behaviors directly. Thus, the second subscale is made up of behaviors that should function to *increase* in-group embeddedness.
    - \* Each of these behaviors had their own 7-point Likert-style response option that represents the likelihood of the person to engage in this behaviour (1=very unlikely, 2=unlikely, 3=slightly unlikely, 4=neither likely nor unlikely, 5=somewhat likely, 6=likely, 7=very likely). Therefore, this questionnaire is intended to measure our participants’ religious behavioral intentions. Item scores were summed to provide a full score (RBI.F), a score for the decrease out-group separation subscale (RBI.DOGS), and a score for the increase in-group embeddedness subscale (RBI.IIGE).
  - The full scale is displayed below:

[Instructions:] Please indicate the degree to which, as a Christian, you are likely to engage in the following behaviour by selecting the bubble of your choice.

[Decrease out-group separation:]

1. Participate in outreach or missionary action towards non-Christians in your local area if you had the chance
2. Participate in missionary action towards non-Christians in another country if you had the chance
3. Attend a different church than your “home church” one time during a year

4. Attend a different small group than your “home group” one time during a year
5. Date a non-Christian
6. Have close platonic relationships with non-Christians

*[Increase in-group embeddedness:]*

1. Regularly participate in large religious gatherings (e.g., chapel, church, etc.)
  2. Regularly participate in small-group religious activities (e.g., bible study, youth group, devotionals, etc.)
  3. Volunteer at a church (e.g., childcare, youth, worship, kitchen, etc.)
  4. Be involved in a committed relationship with another Christian
  5. Be baptized
  6. Partake in tithing/offering
  7. Be involved in a close platonic relationship with other Christians
- The Three Domains of Disgust Scale (TDDS) Tybur et al. (2009)
    - The Three Domains of Disgust Scale was developed by Tybur et al. (2009) in order to measure three relatively distinct domains of disgust based on evolutionary logic. These forms have their own subscales, sexual disgust, moral disgust (MD) and pathogen disgust (PD). It’s three factor structure has been validated in a three-factor confirmatory factor analysis, which was favored over a single factor structure (B. Olatunji et al., 2012). In order to ensure that participants are only exposed to a subminimal level of risk of discomfort—and because it was not directly relevant to our hypotheses at the time—we did not administer the sexual disgust subscale.
    - Participants were asked to rate the degree to which they feel each item is disgusting on 7-point, Likert-style response options (0=not disgusting at all, 6=extremely disgusting). The items were summed to provide a total score for each subscale. The items are as follows.

#### **Items in Order:**

1. Shoplifting a candy bar from a convenience store
2. Stepping on dog poop
3. Stealing from a neighbor
4. Sitting next to someone who has red sores on their arm
5. A student cheating to get good grades
6. Shaking hands with a stranger who has sweaty palms
7. Deceiving a friend

8. Seeing some mold on old leftovers in your refrigerator
9. Forging someone's signature on a legal document
10. Standing close to a person who has body odor
11. Cutting to the front of a line to purchase the last few tickets to a show
12. Seeing a cockroach run across the floor
13. Intentionally lying during a business transaction
14. Accidentally touching a person's bloody cut

**Items within their Subscales:**

*Moral Disgust:*

4. Stealing from a neighbor
19. Intentionally lying during a business transaction
13. Forging someone's signature on a legal document
10. Deceiving a friend
7. A student cheating to get good grades
1. Shoplifting a candy bar from a convenience store
16. Cutting to the front of a line to purchase the last few tickets to a show

*Pathogen Disgust:*

12. Seeing some mold on old leftovers in your refrigerator
15. Standing close to a person who has body odor
9. Shaking hands with a stranger who has sweaty palms
3. Stepping on dog poop
21. Accidentally touching a person's bloody cut
18. Seeing a cockroach run across the floor
6. Sitting next to someone who has red sores on their arm

- The Penn Inventory of Scrupulosity Revised (PIOS-R) (Abramowitz et al., 2002; B. O. Olatunji et al., 2007)

- The Penn Inventory of Scrupulosity (PIOS) measures religious elements of scrupulosity, including its two domains fear of Sin and fear of God, and was developed by Abramowitz et al. (2002) to assess these elements of pathological over-scrupulosity for subclinical levels of OCD symptomatology. It was further developed by B. O. Olatunji et al. (2007), who used item and factor analysis to validate a new, 15-item Penn Inventory of Scrupulosity Revised (PIOS-R).

\* In our data frame, these variables are the Penn Inventory of Scrupulosity Revised (PIOS), the Fear of God subscale (FOG), and the Fear of Sin subscale (FOS).

\* This is also measured with Likert-style response options, and it was scored by summing the items in total and within their subscales.

PIOS-R: 1. I worry that I might have dishonest thoughts 2. I fear I will act immorally 3. I feel urges to confess sins over and over again 4. I worry about heaven and hell 5. Feeling guilty interferes with my ability to enjoy things I would like to enjoy 6. Immoral thoughts come into my head and I cannot get rid of them 7. I am afraid my behaviour is unacceptable to God 8. I must try hard to avoid having certain immoral thoughts 9. I am very worried that things I did may have been dishonest 10. I am afraid I will disobey God's rules/laws 11. I am afraid of having sexual thoughts 12. I feel guilty about immoral thoughts I have had 13. I worry that God is upset with me 14. I am afraid of having immoral thoughts 15. I am afraid my thoughts are unacceptable to God

Fear of Sin: 1. I worry that I might have dishonest thoughts 2. I fear I will act immorally 3. I feel urges to confess sins over and over again 5. Feeling guilty interferes with my ability to enjoy things I would like to enjoy 6. Immoral thoughts come into my head and I cannot get rid of them 8. I must try hard to avoid having certain immoral thoughts 9. I am very worried that things I did may have been dishonest 11. I am afraid of having sexual thoughts 12. I feel guilty about immoral thoughts I have had 14. I am afraid of having immoral thoughts

Fear of God: 4. I worry about heaven and hell 7. I am afraid my behaviour is unacceptable to God 10. I am afraid I will disobey God's rules/laws 13. I worry that God is upset with me 15. I am afraid my thoughts are unacceptable to God

- Religious Social Distance Questionnaire:

- This is a slightly adapted form of the Social Distance Scale developed by Mather et al. (2017), which is a revised method of measuring social distance. Our version of the scale is intended to assess religious in-group and out-group attitudes. Mather et al. (2017) designed the measure to provide better sensitivity—particularly at the tails of the distributions—than the Bogardus Social Distance Scale (Bogardus, 1933), a long-standing method of measuring in-group and out-group attitudes in

the social psychology literature. Others have used similar methods, adapting the original Social Distance Scale to measure social distance for religious groups in North American samples (Brinkerhoff & Mackie, 1986). The items are displayed in the image below.

- This measure was administered 9 times to each participant, with blanks (i.e., “[group member]”) containing a different religious group member for each administration. Religious group members will either be (1) members of a broad religious group or (2) members of a sect of Christianity. The groups are as follows:
  - \* Broad Religious Groups (group member terminology for insertion into scale):
    - Atheism (an atheist)
    - Christianity (a Christian)
    - Sikhism (a Sikh)
    - Islam (a Muslim)
    - Buddhism (a Buddhist)
    - Hinduism (a Hindu)
    - Judaism (a Jewish person)
  - \* Christian Sects (group member terminology for insertion into scale):
    - Protestantism (a Protestant)
    - Catholicism (a Catholic)
- Instructions read: “Select the option that best describes your response to the statements below.” The order of the presentation of the different religious groups (and atheism) were randomized.
- The scale is scored by multiplying the Likert response value with the rank of the item. That is, each item is assigned a rank (1-7, in the order you see them below) representing the severity of social distance implied by the response to the item. Thus, being unwilling to accept a group member as a close relative by marriage is considered a less strong indicator of social distance than being unwilling to accept that group member in one’s country whatsoever. Once this rescaling by the rank of the item is done, the items are summed to produce a total score for the group in question.
- The variables in the data frame are denoted by `SD_religiousgroup`, where “religiousgroup” has the name of the group (e.g., `SD_protestant`). The aggregated variable of social distance towards all religious groups except Christian groups and atheists is called `SD_non_christian`, and this was calculated by taking the average total score for each of these groups.



Circle the number that best describes your response to these statements according to the following scale:

1=Strongly agree; 2=agree; 3=Neither agree or Disagree; 4=Disagree; 5=Strongly disagree

I would be willing to accept [group member] as a close relative by marriage 1 2 3 4 5

I would be willing to accept [group member] as a close personal friend 1 2 3 4 5

I would be willing to accept [group member] as a neighbor on the same street 1 2 3 4 5

I would be willing to accept [group member] as a coworker 1 2 3 4 5

I would be willing to accept [group member] as a citizen in my country 1 2 3 4 5

I would be willing to accept [group member] as a visitor in my country 1 2 3 4 5

I would not exclude [group member] from my country\* 1 2 3 4 5

## Analysis

Before moving on to the analysis, we will load the necessary packages.

```
# Defining the list of packages to install in a vector
pkg <- c("psych", "lmtest", "sandwich")

# Loading the packages using groundhog
suppressMessages(groundhog.library(pkg, date = "2024-04-22")) # Suppress message to allow for
```

## Data Exploration

To get a sense of the data, we will first generate descriptive statistics and visualize the distribution of the variables.

## Descriptive Statistics

First, we will see how long it took participants to complete the survey.

```
# Saving the descriptive stats for time_taken
time_taken_descriptives <- describe(data$time_taken)

# Displaying the results
print(time_taken_descriptives)
```

```
vars   n  mean    sd median trimmed  mad min max range skew kurtosis   se
X1     1 134 13.37 30.56      8   8.48 4.45   3 330   327  8.7      84.3 2.64
```

- The average time it took participants to complete the survey was 13.37 minutes, with a standard deviation of 30.56 minutes. Because this standard deviation is so large, the maximum value (330 minutes) is so large, and the skew is quite large (8.7), there must be an outlier influencing these statistics.
  - Indeed, if we look at the trimmed mean (under “trimmed”)—which cuts off the outer 10th percentile of scores—the mean falls to 8.48.

Next, we will describe the sample in terms of the categorical variables. The following generates frequency tables for all categorical variables.

```
# Create list of whether a variable is a factor variable or not
factor_variables <- sapply(data, is.factor)

# Create a frequency table for each factor variable
factor_tables <- lapply(data[factor_variables], table)

# Display the number of participants in the data frame
nrow(data)
```

```
[1] 134
```

```
# Display the tables
print(factor_tables)
```

```
$rec_meth
```

```
Prolific Social Media
```

```
SONA Word of Mouth
```

49	28	52	5
----	----	----	---

\$educ\_complete

Graduate degree (MA/MSc/MPhil/other)	High school diploma/A-levels
4	77
Secondary education (e.g. GED/GCSE)	Technical/community college
5	13
Undergraduate degree (BA/BSc/other)	
35	

\$educ\_currently\_in

Doctorate degree (PhD/other)	Graduate degree (MA/MSc/MPhil/other)
2	25
Technical/community college	Undergraduate degree (BA/BSc/other)
9	98

\$religious\_affiliation

Christianity (e.g. Baptist, Church of England, Roman Catholic, Methodist, Jehovah's Witness,

\$gender

Female	Male	Non-binary
81	51	1

\$christian\_affiliation

Catholic Anglican / Episcopalian	Protestant
21	19
Pentecostal / Apostolic	Methodist
12	2
Lutheran	Orthodox / Eastern Orthodox
3	7
Baptist Calvinist / Reformed / Presbyterian	
13	4
Non-denominational	Other
48	4

\$social\_media\_platform

Facebook	Instagram	Other
7	19	2

In total, there are  $N = 134$  participants in the data frame.

- Recruitment Method: The primary sources of recruitment are SONA and Prolific, followed by Social Media and Word of Mouth.
  - SONA:  $n = 52$
  - Prolific:  $n = 49$
  - Social media:  $n = 28$ 
    - \* Social Media Platform: Of those recruited through social media, most came from Instagram, followed by Facebook and other.
      - Instagram: 19
      - Facebook: 7
      - Other: 2
  - Word of Mouth:  $n = 5$
- Education: We required that participants either be in some sort of college/university/technical school or were graduated from such an institution. This is reflected in our education demographics below, providing us with a highly educated sample.
  - Completed Education Level:
    - \* Secondary education (e.g. GED/GCSE):  $n = 5$
    - \* High school diploma/A-levels:  $n = 77$
    - \* Technical/community college:  $n = 13$
    - \* Undergraduate degree (BA/BSc/other):  $n = 35$
    - \* Graduate degree (MA/MSc/MPhil/other):  $n = 4$
  - Education Level Currently In:
    - \* Technical/community college:  $n = 9$
    - \* Undergraduate degree (BA/BSc/other):  $n = 98$
    - \* Graduate degree (MA/MSc/MPhil/other):  $n = 25$
    - \* Doctorate degree (PhD/other):  $n = 2$

- Religious Affiliation: Because we selected only Christians, religious identification is all “Christianity (e.g. Baptist, Church of England, Roman Catholic, Methodist, Jehovah’s Witness, etc.)” ( $N = 134$ ).
- Christian Affiliation: Although the modal response is non-denominational for Christian affiliation, there is representation from many different Christian sects in our sample.
  - Non-denominational:  $n = 48$
  - Catholic Anglican / Episcopalian:  $n = 21$
  - Protestant:  $n = 19$
  - Baptist:  $n = 13$
  - Pentecostal / Apostolic:  $n = 12$
  - Orthodox / Eastern Orthodox:  $n = 7$
  - Calvinist / Reformed / Presbyterian:  $n = 4$
  - Lutheran:  $n = 3$
  - Methodist:  $n = 2$
  - Other:  $n = 4$
- Gender: As is typical of university samples (which our sample in large part is) gender is skewed female.
  - Female:  $n = 81$
  - Male:  $n = 51$
  - Non-binary:  $n = 1$

Now, we will investigate the numerical variables. The following generates descriptive statistics using the `psych` package for all numeric variables (except for `time_taken`).

```
# Saving a vector with the names of all numeric variables
num_vars <- data[c("RBI.DOGS", "RBI.IIGE", "RBI.F", "MD", "PD",
                  "PIOS", "FOG", "FOS", "SD_athiest", "SD_christian",
                  "SD_sikh", "SD_muslim", "SD_buddhist", "SD_hindu",
                  "SD_jewish", "SD_protestant", "SD_catholic", "age")]

# Saving the descriptive statistics for num_vars
num_vars_descriptives <- describe(num_vars)

# Displaying the results
print(num_vars_descriptives)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range
RBI.DOGS	1	134	4.82	0.84	5.00	4.85	0.74	2.33	7.00	4.67
RBI.IIGE	2	129	5.77	1.07	6.00	5.87	1.06	2.57	7.00	4.43
RBI.F	3	129	5.34	0.79	5.46	5.39	0.68	3.00	6.85	3.85
MD	4	130	31.12	6.73	32.00	31.61	5.93	0.00	42.00	42.00
PD	5	133	25.80	7.90	26.00	26.14	7.41	0.00	42.00	42.00
PIOS	6	127	28.13	11.09	27.00	28.12	11.86	1.00	56.00	55.00
FOG	7	131	9.57	4.64	9.00	9.50	4.45	0.00	20.00	20.00
FOS	8	127	18.66	7.27	18.00	18.69	5.93	0.00	40.00	40.00
SD_athiest	9	48	42.60	18.12	33.00	39.77	7.41	28.00	87.00	59.00
SD_christian	10	47	35.32	20.30	28.00	30.46	0.00	28.00	140.00	112.00
SD_sikh	11	48	44.02	21.91	33.00	40.27	7.41	28.00	132.00	104.00
SD_muslim	12	48	42.69	18.17	32.50	40.15	6.67	28.00	86.00	58.00
SD_buddhist	13	48	43.85	20.99	33.00	40.02	7.41	28.00	113.00	85.00
SD_hindu	14	48	42.48	20.04	32.00	39.02	5.93	28.00	100.00	72.00
SD_jewish	15	48	40.67	17.03	31.00	37.90	4.45	28.00	86.00	58.00
SD_protestant	16	48	38.54	21.56	28.00	33.70	0.00	28.00	140.00	112.00
SD_catholic	17	48	37.15	16.82	29.00	33.58	1.48	28.00	100.00	72.00
age	18	134	21.87	3.53	21.00	21.40	2.97	18.00	37.00	19.00
	skew	kurtosis	se							
RBI.DOGS	-0.36	0.07	0.07							
RBI.IIGE	-0.72	-0.41	0.09							
RBI.F	-0.71	0.12	0.07							
MD	-1.14	2.82	0.59							
PD	-0.47	0.27	0.69							
PIOS	0.06	-0.17	0.98							
FOG	0.16	-0.57	0.41							
FOS	0.09	0.41	0.65							
SD_athiest	1.21	0.23	2.62							
SD_christian	3.49	13.35	2.96							
SD_sikh	1.83	3.65	3.16							
SD_muslim	1.08	-0.29	2.62							
SD_buddhist	1.59	1.92	3.03							
SD_hindu	1.39	0.82	2.89							
SD_jewish	1.25	0.38	2.46							
SD_protestant	2.73	8.47	3.11							
SD_catholic	2.07	3.65	2.43							
age	1.48	2.95	0.30							

- I will not review the descriptive statistics of all of the variables here, except for a few notable cases.

– Age: The average age reflects the fact that we excluded those under 18 and partially

recruited those in university ( $M = 21.87$ ,  $SD = 3.53$ ).

- Social Distance: As should also be expected, social distance scores are lowest for Christians ( $M = 35$ ,  $SD = 20.30$ ), followed by Catholics ( $M = 37.15$ ,  $SD = 16.82$ ) and Protestants ( $M = 38.54$ ). The rest of the religious groups have higher social distance scores (44.02 –  $M = 40.67$ ), with Jewish as the next lowest social distance and Sikh with the highest social distance.

## Data Visualization

Now we will produce some bar charts and histograms to assess the distributions of these variables visually.

First, we will produce bar charts for each categorical variable.

```
# Creating a data frame with only categorical variables based on the logical vector generated
categorical_data <- data[, factor_variables]

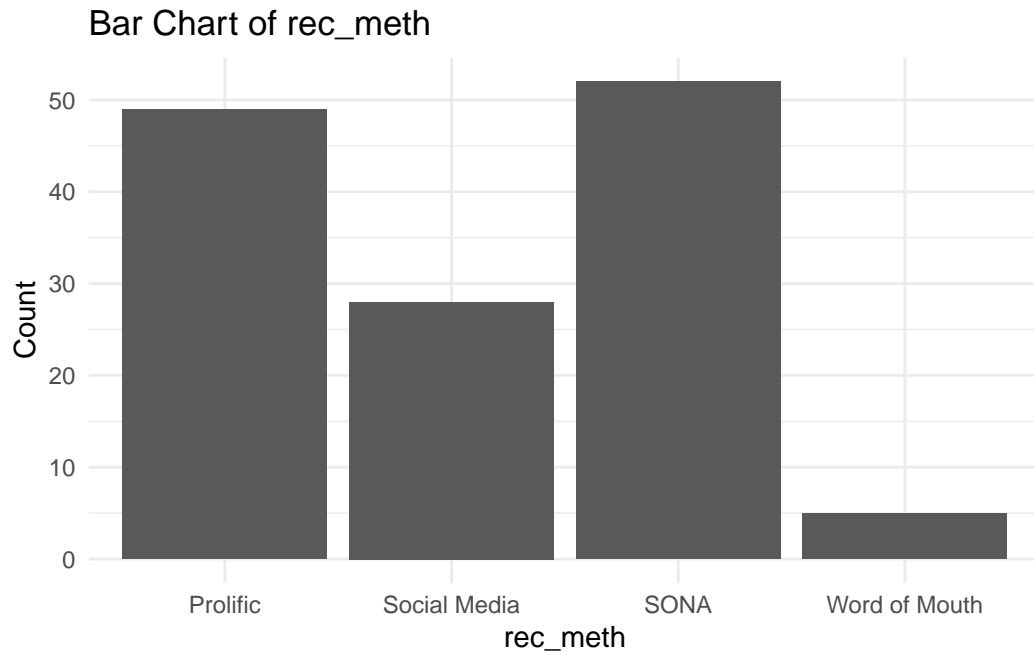
# Initialize a list to store the plots in
bar_chart_list <- list()

# Loop through each categorical variable in the data frame and create a bar chart
for (var in names(categorical_data)) {
  p <- ggplot(categorical_data, aes(x = !!sym(var))) +
    geom_bar() +
    theme_minimal() +
    labs(title = paste("Bar Chart of", var), x = var, y = "Count")

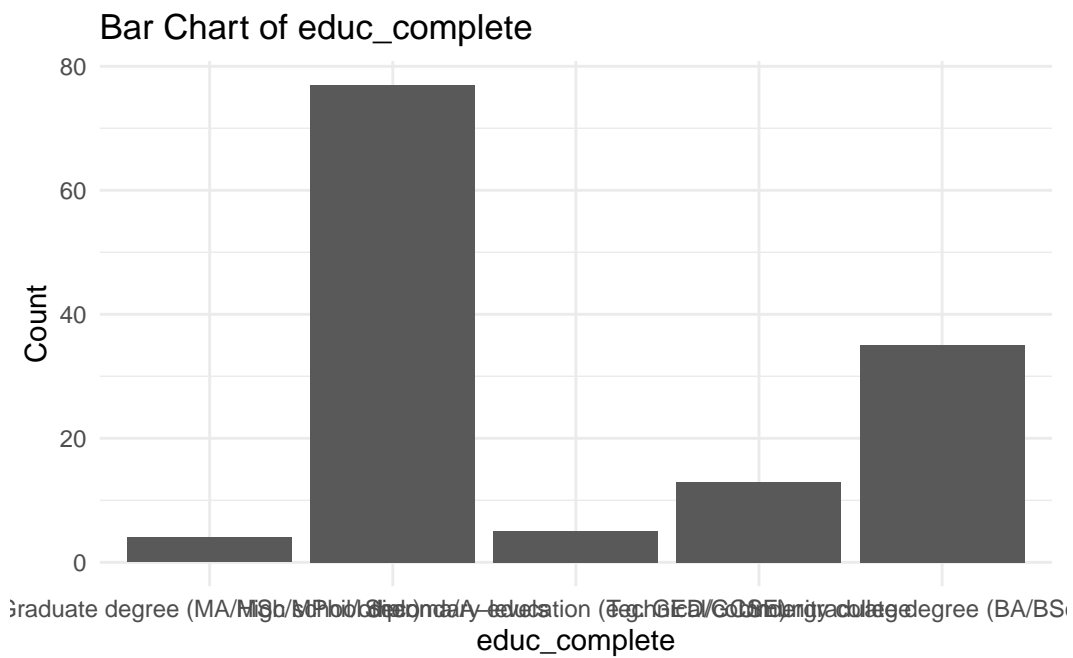
  # Save each chart in the list with the variable name as the identifier
  bar_chart_list[[var]] <- p
}

# Display each chart
print(bar_chart_list)
```

\$rec\_meth

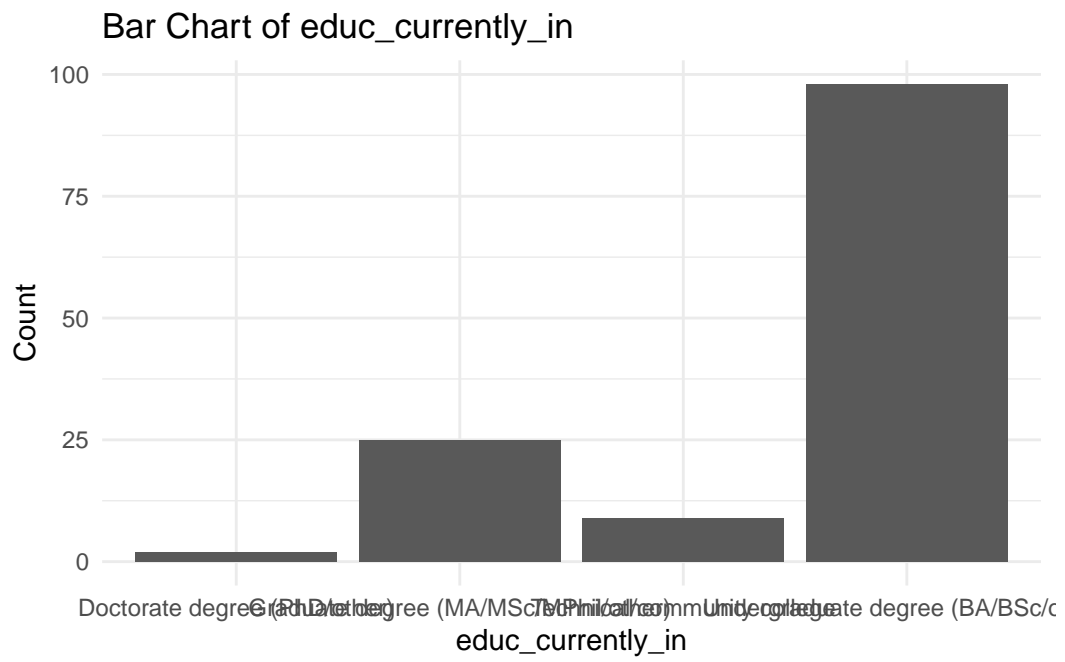


\$educ\_complete



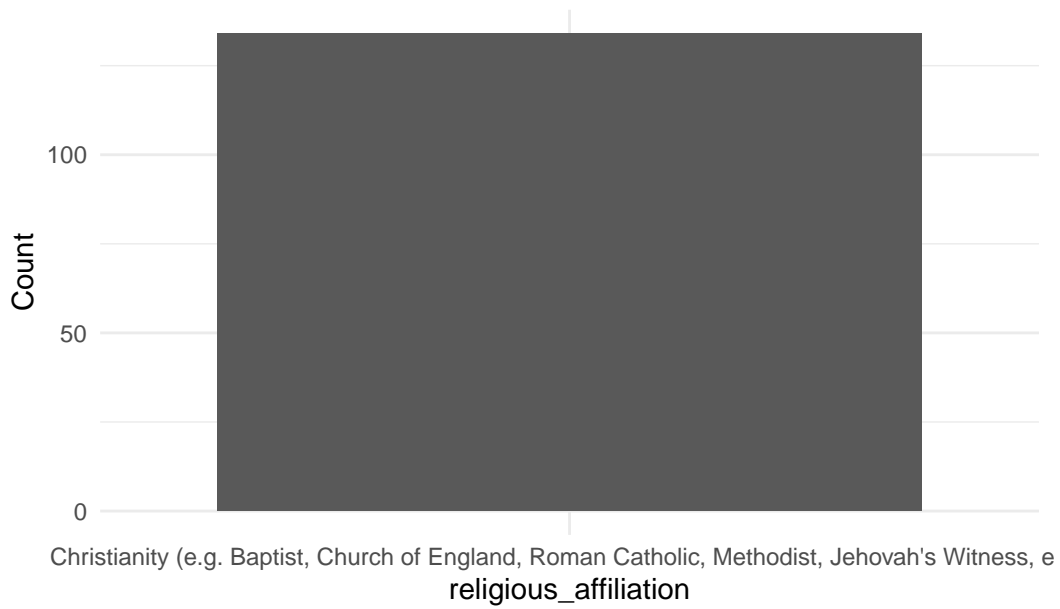


\$educ\_currently\_in



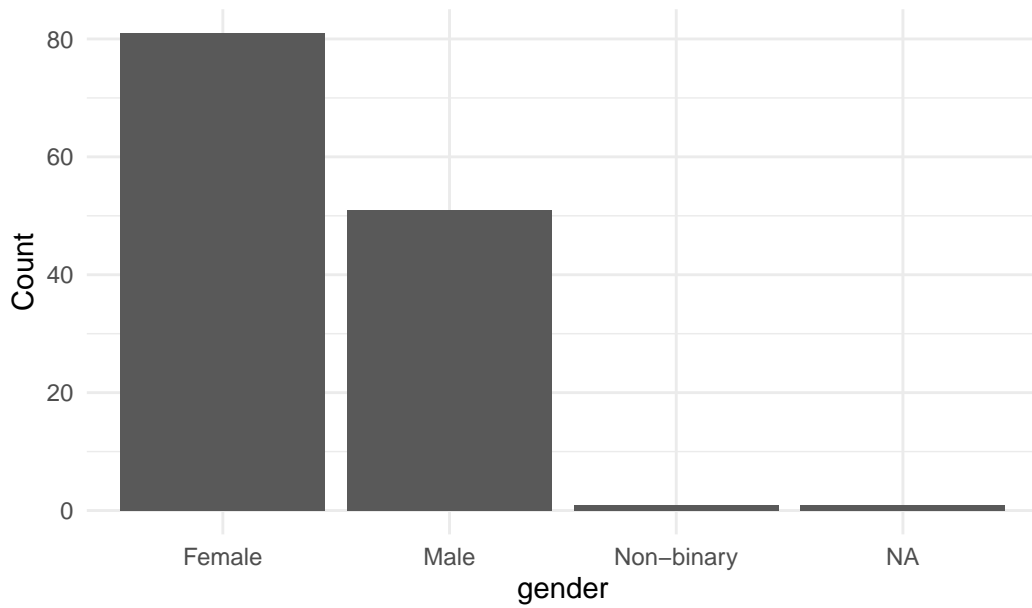
\$religious\_affiliation

Bar Chart of religious\_affiliation

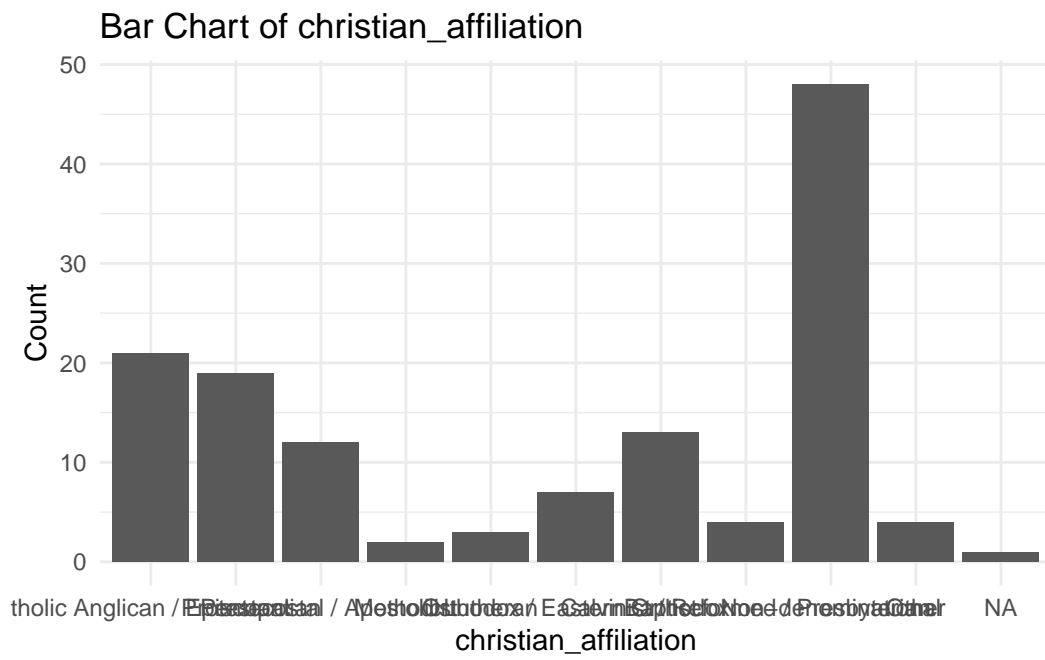


\$gender

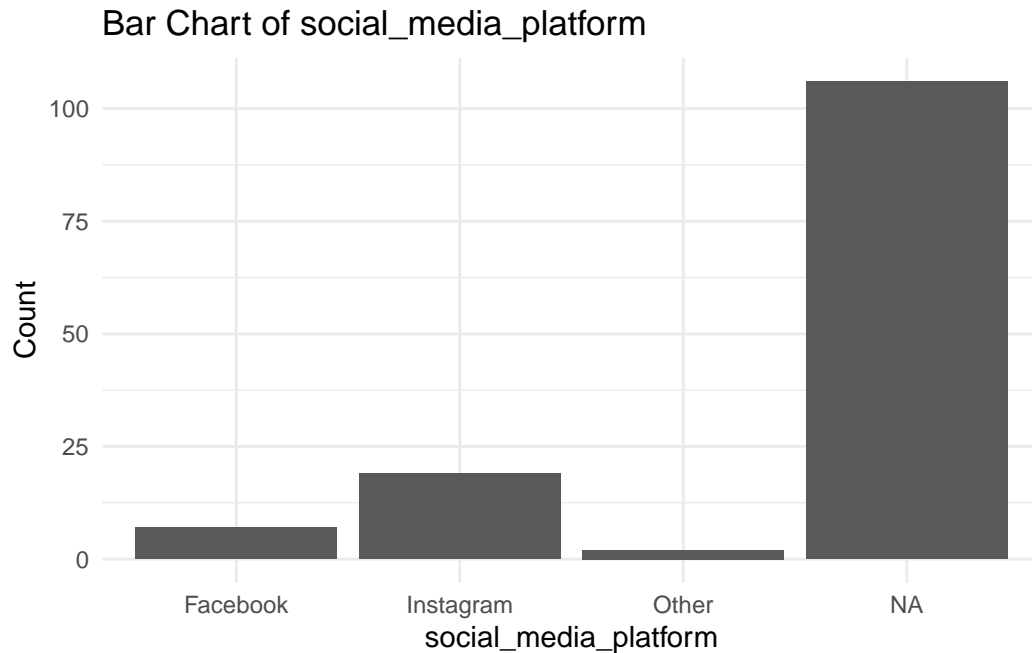
Bar Chart of gender



\$christian\_affiliation



\$social\_media\_platform



Next, we will produce a histogram for `time_taken`.

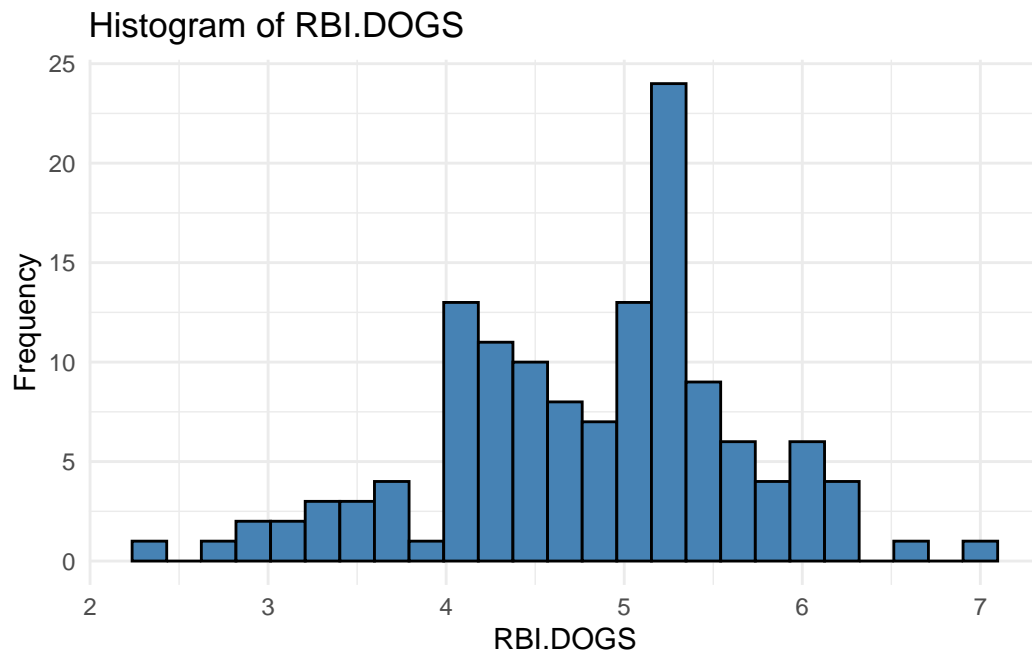
```
# Initialize a list to store the histograms in
histogram_list <- list()

# Loop through each variable in num_vars to produce a histogram
for (var in names(num_vars)) {
  p <- ggplot(num_vars, aes(x = .data[[var]])) +
    geom_histogram(bins = 25, fill = "steelblue", color = "black") +
    theme_minimal() +
    labs(title = paste("Histogram of", var), x = var, y = "Frequency")

  # Save each histogram in the list with the variable name as the identifier
  histogram_list[[var]] <- p
}

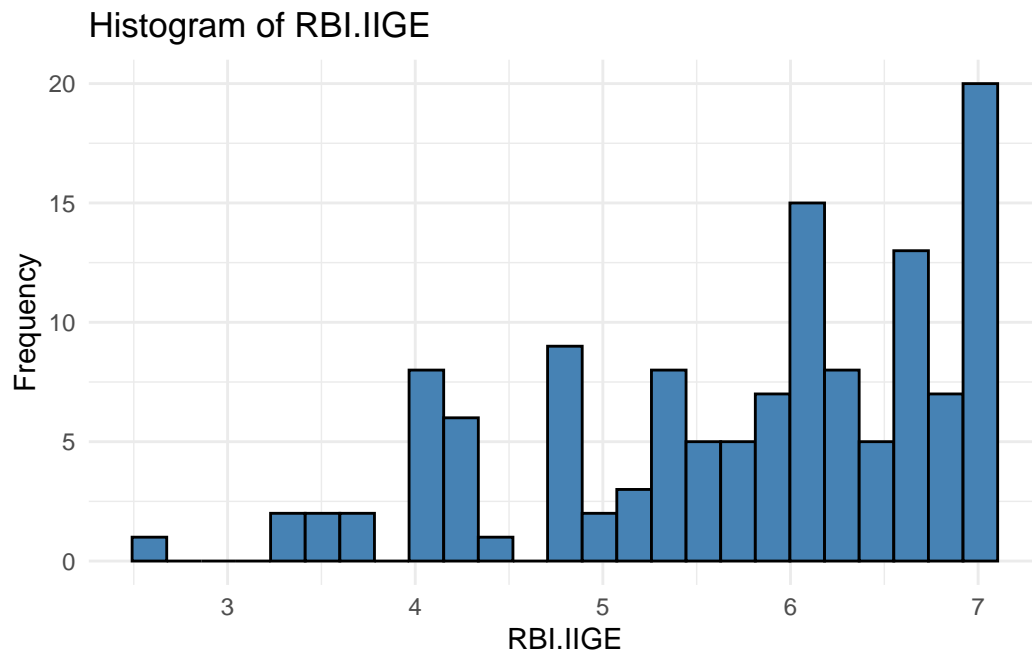
# Display the histograms
print(histogram_list)
```

`$RBI.DOGS`



`$RBI.IIGE`

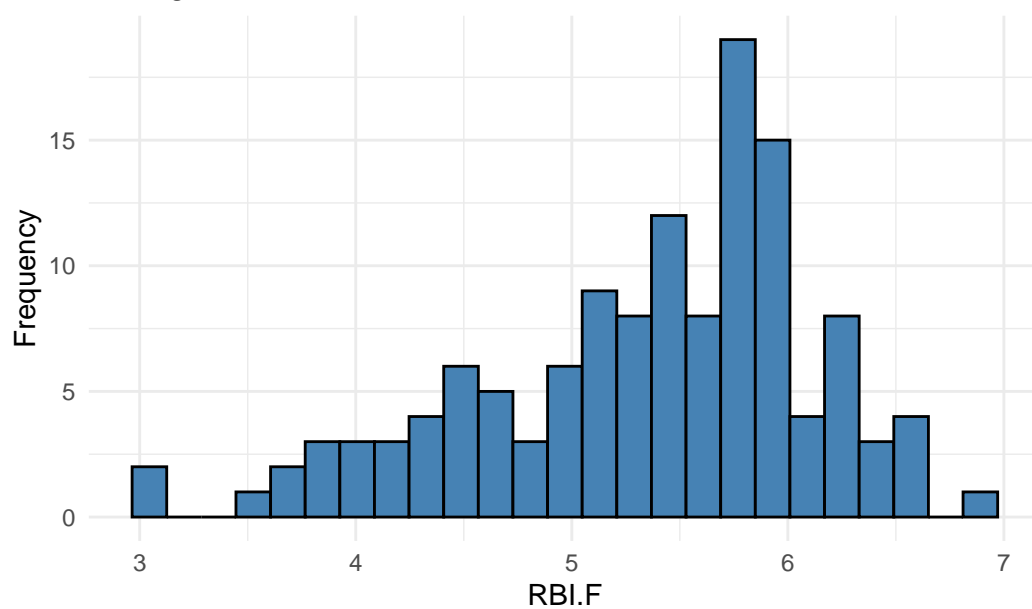
Warning: Removed 5 rows containing non-finite outside the scale range  
(`stat\_bin()`).



\$RBI.F

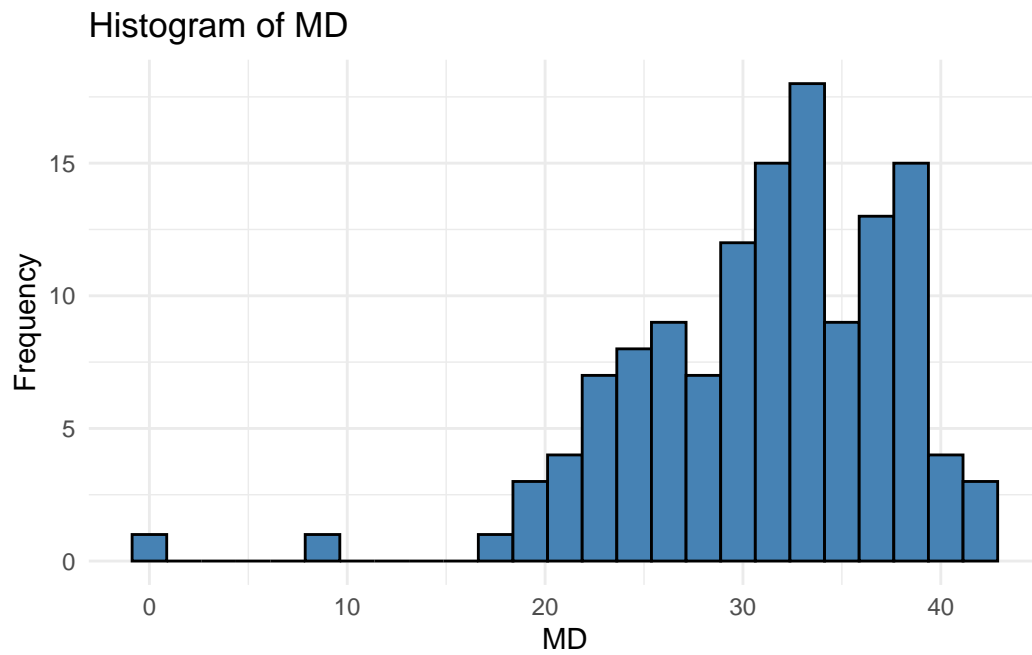
Warning: Removed 5 rows containing non-finite outside the scale range  
(`stat\_bin()`).

Histogram of RBI.F



\$MD

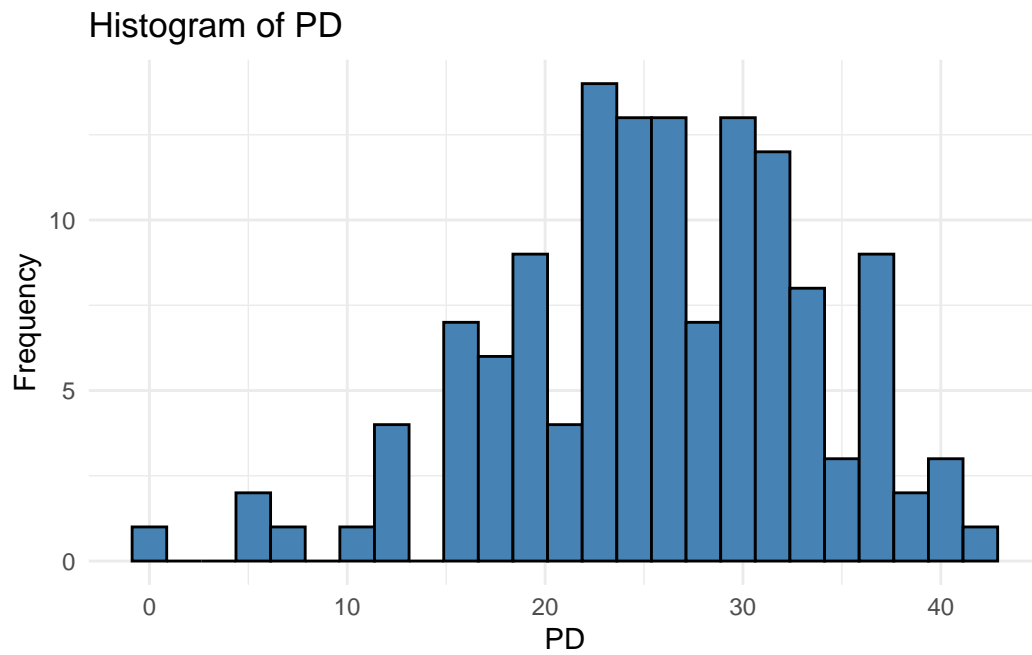
Warning: Removed 4 rows containing non-finite outside the scale range  
(`stat\_bin()`).



\$PD

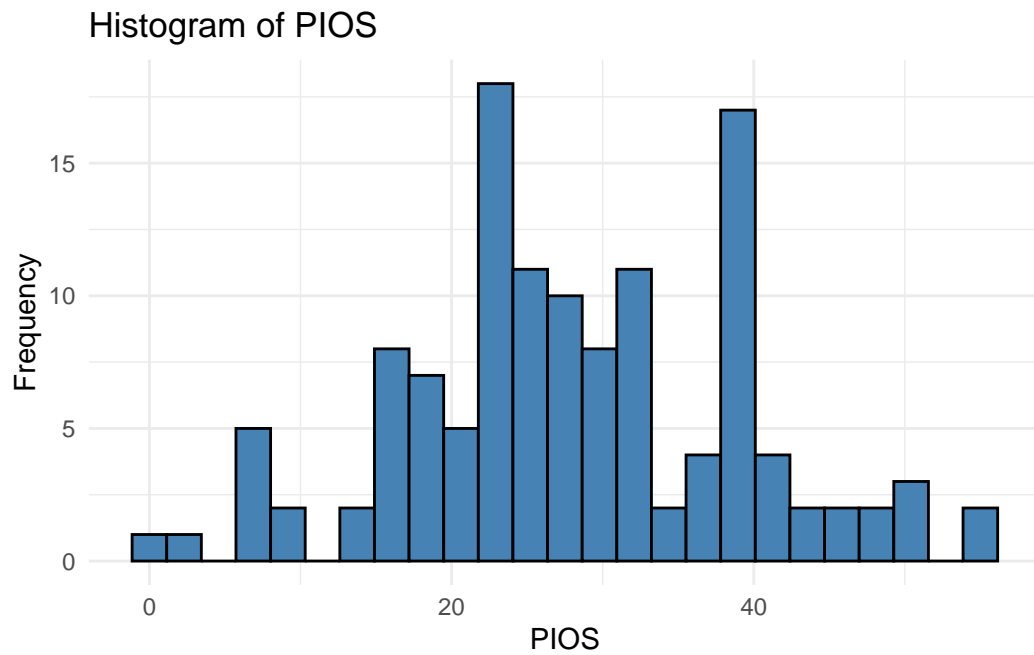
Warning: Removed 1 row containing non-finite outside the scale range  
(`stat\_bin()`).





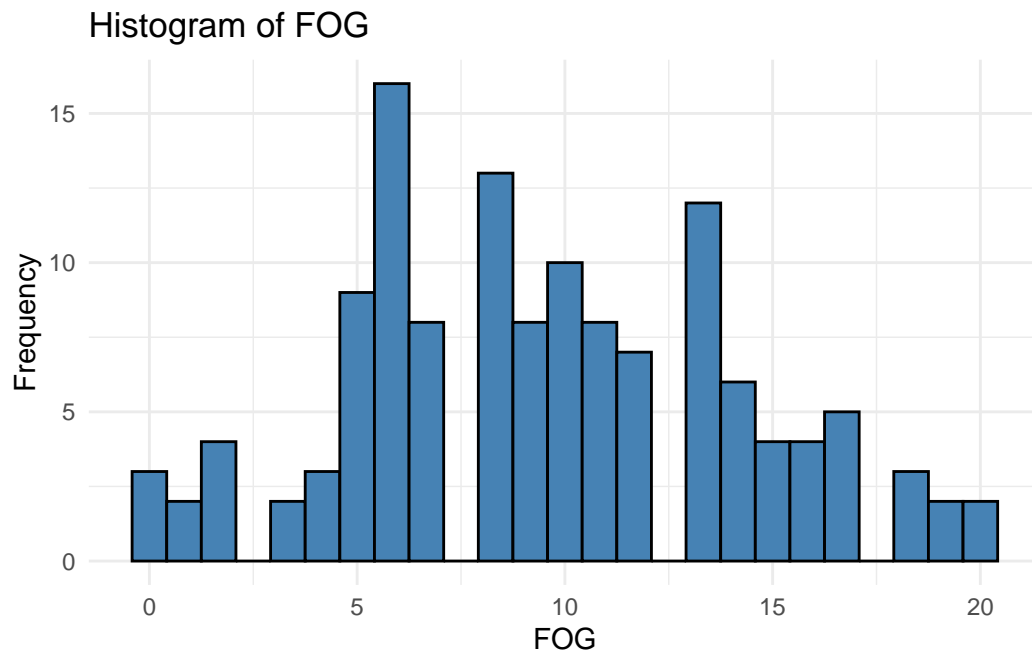
\$PIOS

Warning: Removed 7 rows containing non-finite outside the scale range  
(`stat\_bin()`).



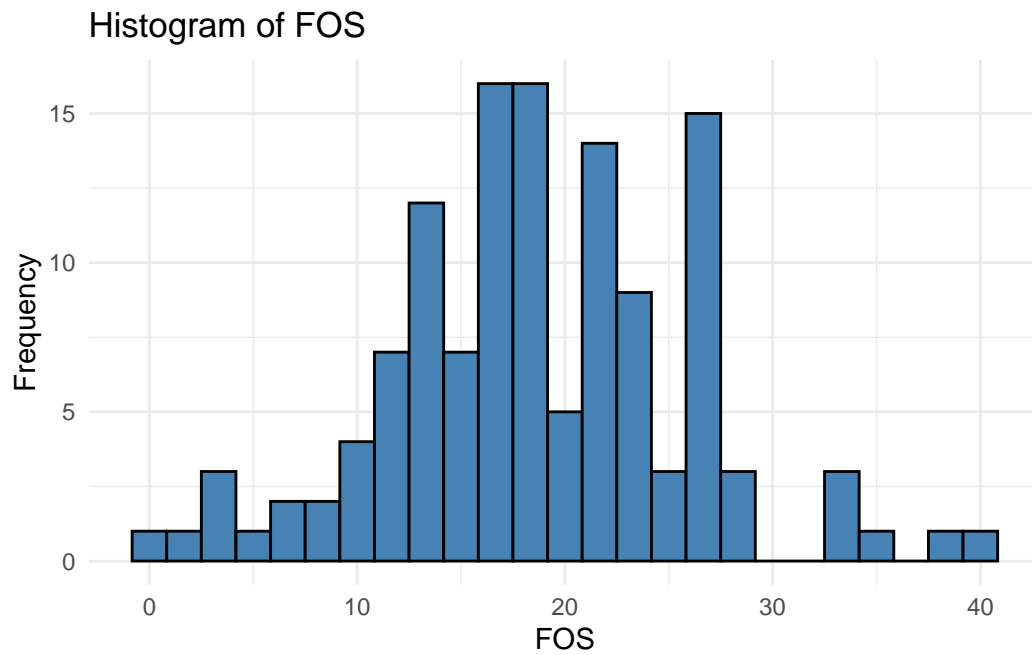
\$FOG

Warning: Removed 3 rows containing non-finite outside the scale range  
(`stat\_bin()`).



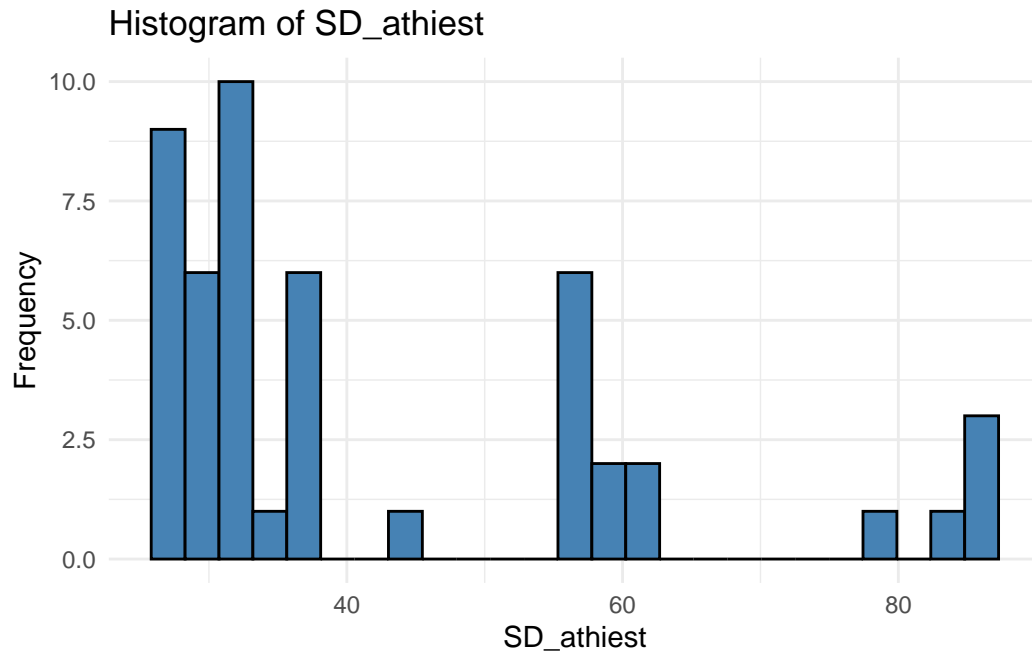
\$FOS

Warning: Removed 7 rows containing non-finite outside the scale range  
(`stat\_bin()`).



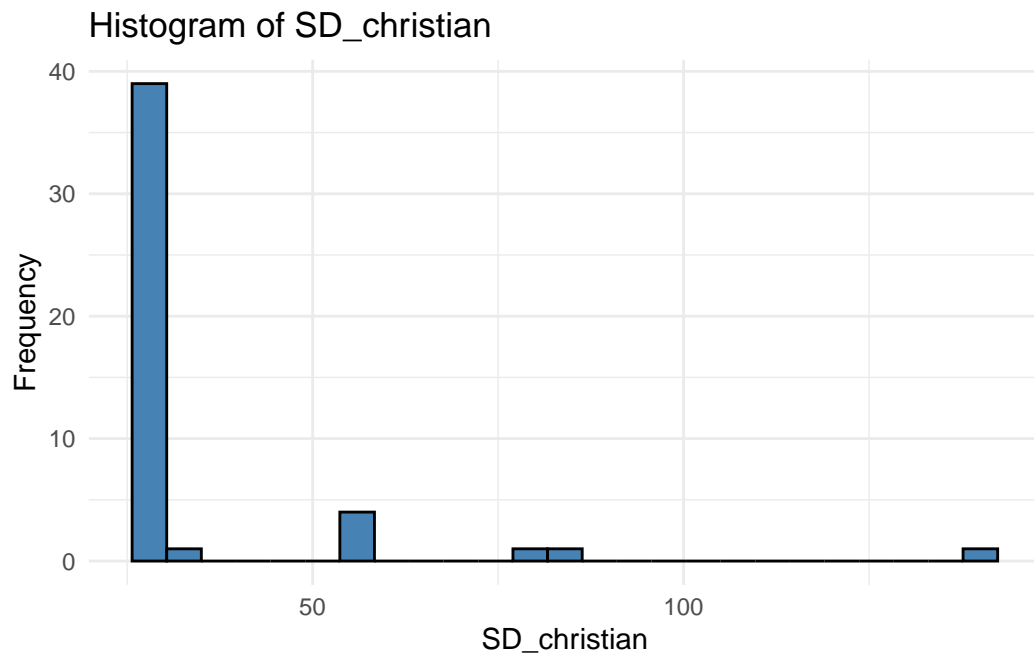
`$SD_athiest`

Warning: Removed 86 rows containing non-finite outside the scale range  
(`stat\_bin()`).



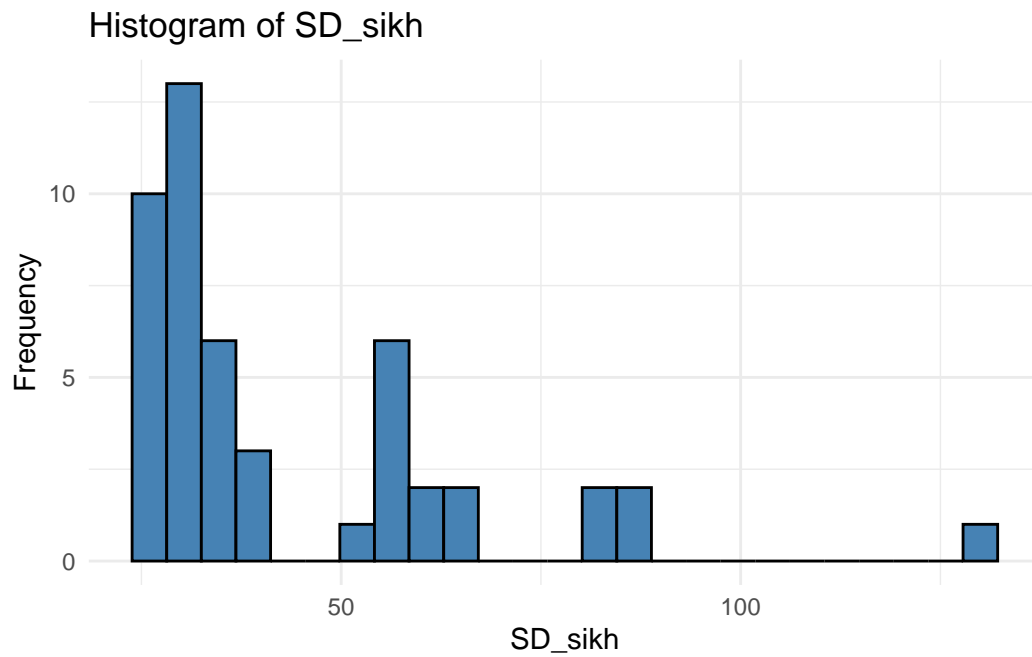
`$SD_christian`

Warning: Removed 87 rows containing non-finite outside the scale range (``stat_bin()``).



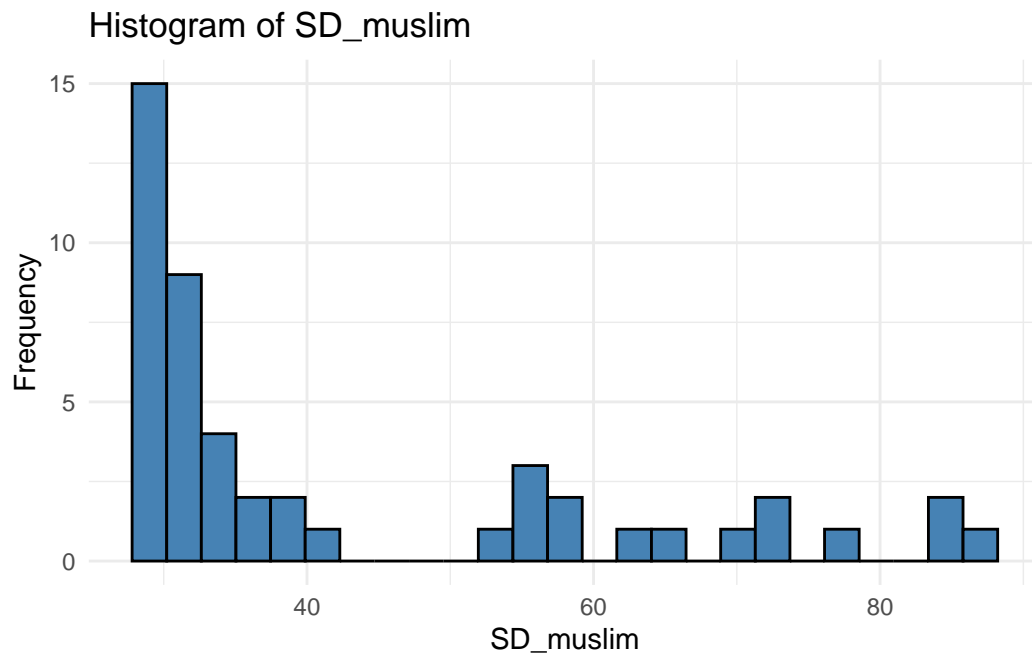
`$SD_sikh`

Warning: Removed 86 rows containing non-finite outside the scale range (``stat_bin()``).



`$SD_muslim`

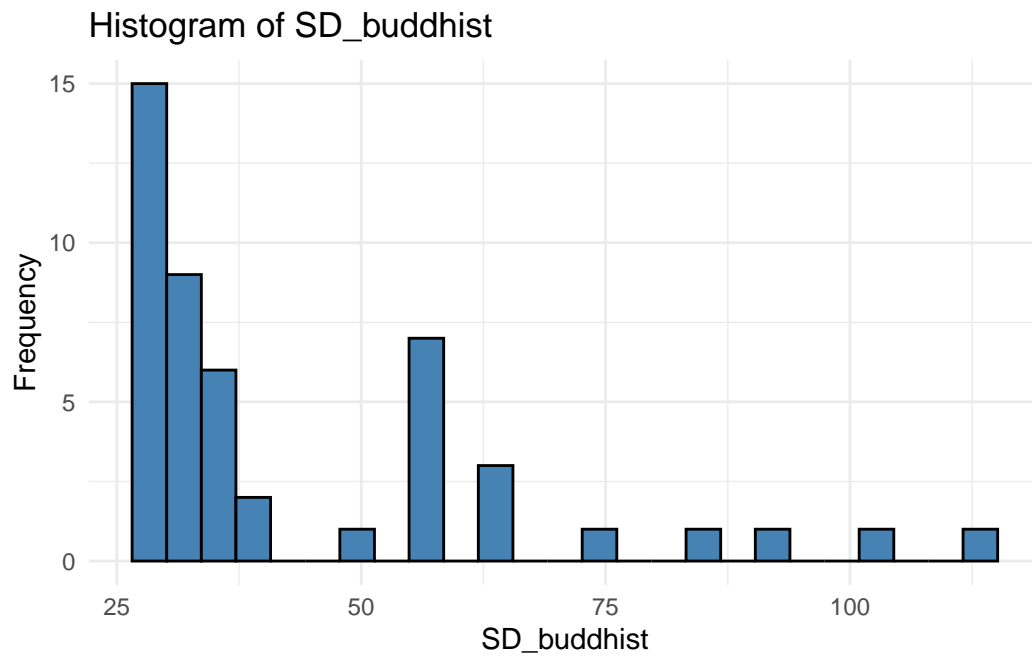
Warning: Removed 86 rows containing non-finite outside the scale range  
(`stat\_bin()`).



`$SD_buddhist`

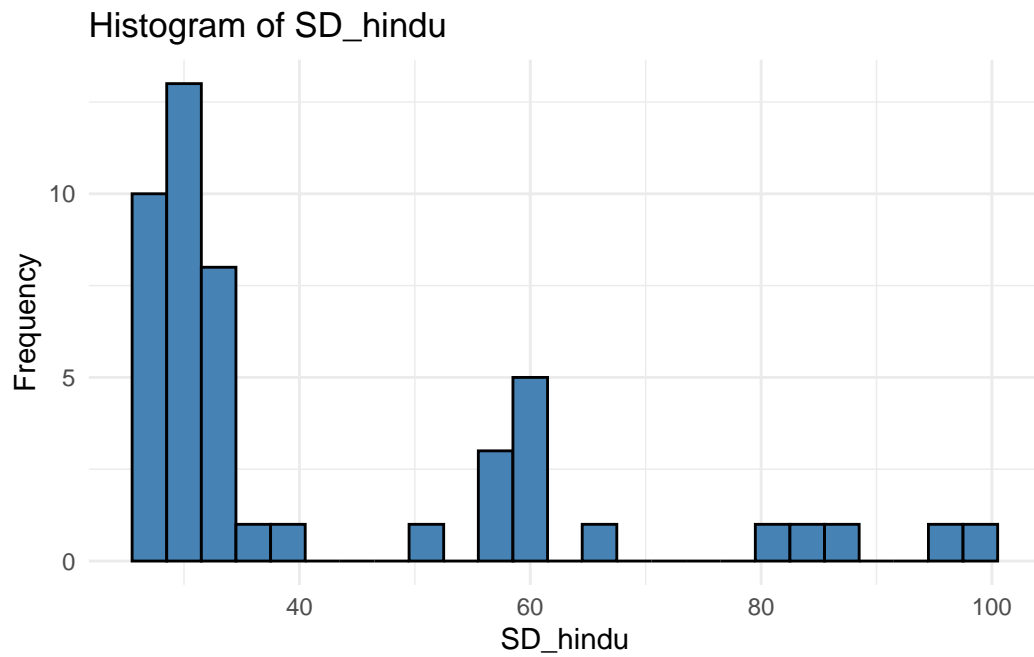
Warning: Removed 86 rows containing non-finite outside the scale range (``stat_bin()``).





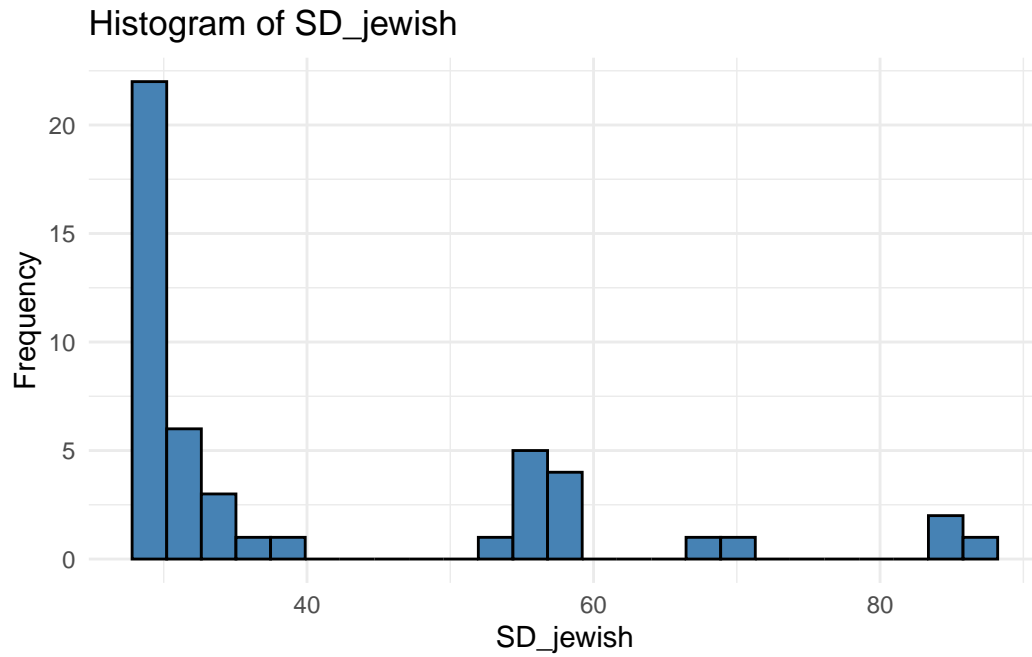
`$SD_hindu`

Warning: Removed 86 rows containing non-finite outside the scale range (``stat_bin()``).



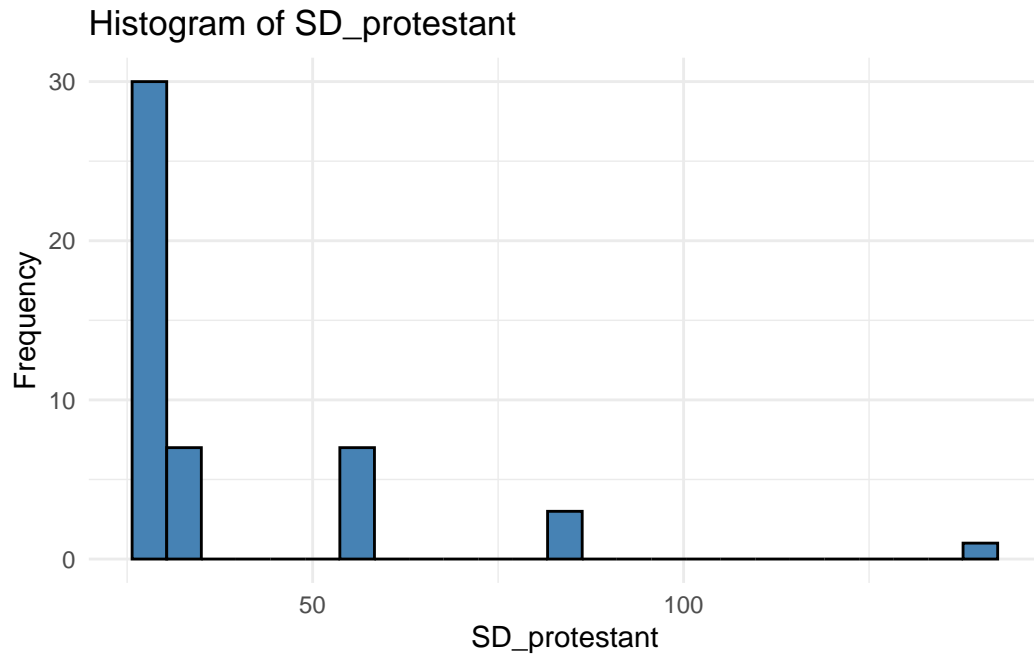
`$SD_jewish`

Warning: Removed 86 rows containing non-finite outside the scale range (``stat_bin()``).



`$SD_protestant`

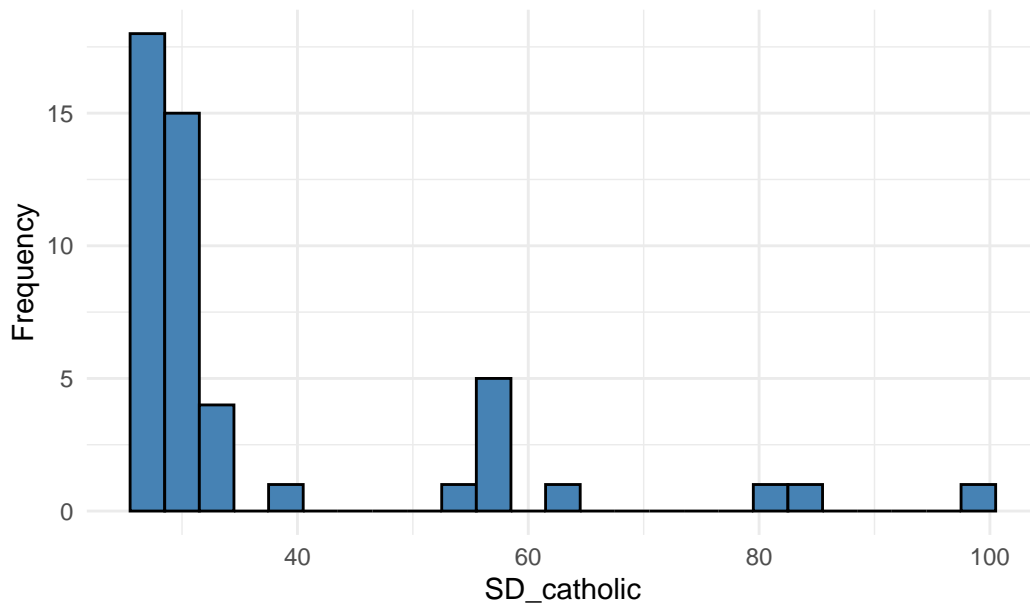
Warning: Removed 86 rows containing non-finite outside the scale range (``stat_bin()``).



`$SD_catholic`

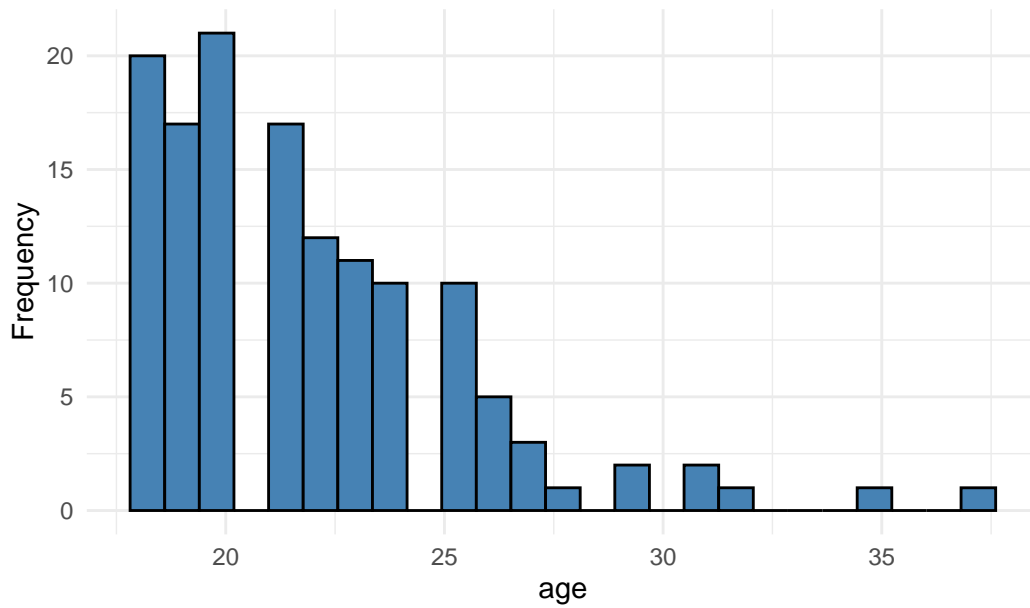
Warning: Removed 86 rows containing non-finite outside the scale range (``stat_bin()``).

Histogram of SD\_catholic



\$age

Histogram of age



- Many of the variables here are not normally distributed.
  - Age:
    - \* Age seems to be negatively skewed due to our selection of only adults.
  - Religious Behavioral Intentions:
    - \* The full religious behavioral intentions scale seems to be approximately normally distributed.
      - However, the decrease out-group separation scale has some strange perturbations, such that scores jump in frequency around 4 and decrease towards five, before spiking just above five and then continuing to decrease.
      - Then, the increase in-group embeddedness scores seem to be cut off by the range of the scale. The frequency of scores tend to increase relatively steadily until their highest, which is 7. This may reflect that we simply did not provide enough range for genuine variability on the questionnaire, or it may reflect socially desirable responding.
  - Disgust:
    - \* Moral disgust is approximately normally distributed, but there are some outliers at the low end of the scale. I originally noticed the 0 response in this scale while cleaning the raw data, and that participant filled out the rest of the survey without straightlining zeros, so I think their responses are genuine.
    - \* Pathogen disgust shows a pretty nice normal distribution, although there may be some positive skew.
  - PIOS:
    - \* The Penn Inventory of Scrupulosity almost looks to be distributed normally, but there is a spike in scores just before 40.
      - The Fear of God subscale may be negatively skewed.
      - The Fear of Sin subscale is pretty nicely distributed.
  - Social Distance:
    - \* The Social Distance scores are not normally distributed. Most scores for all religions are clumped towards the beginning of the scale, particularly for those with lower social distance scores in general (e.g., Christians). For the rest, they are still aggregated mostly at the lower range of the scale, but there are spurts of scores higher that pull the mean higher.

## Psychometric Properties of Religious Behavioral Intentions Scales

To establish the psychometric properties of our religious behavioral intentions subscales, we will first do an internal consistency analysis to establish reliability, then we will look for correlations with the other theoretically relevant variables to try to establish validity.

### Reliability

To establish internal consistency, we will analyze Chronbach's alpha scores for the subscales. We will also calculate the alpha scores if items are dropped, to assess whether items are not contributing to the internal consistency of the scales and remove them as appropriate.

1. Religious Behavioral Intentions to Decrease Out-Group Separation:

```
# Creating a data frame with only the DOGS items
DOGS_items <- select(data, DOGS.1:DOGS.6)

# Calculating Chronbach's alpha
Chronbach_DOGS <- alpha(DOGS_items)
```

Warning in alpha(DOGS\_items): Some items were negatively correlated with the first principal should be reversed.

To do this, run the function again with the 'check.keys=TRUE' option

Some items ( DOGS.5 DOGS.6 ) were negatively correlated with the first principal component and probably should be reversed.

To do this, run the function again with the 'check.keys=TRUE' option

```
# Displaying the results
print(Chronbach_DOGS$total)
```

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd
0.2212176	0.299247	0.5379125	0.06644374	0.4270364	0.1070369	4.81592	0.8420022
median_r							
-0.07060084							

```
print(Chronbach_DOGS$alpha.drop)
```

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha se
DOGS.1	0.013423552	0.11578085	0.3254350	0.02551995	0.1309413	0.14155291
DOGS.2	-0.027272388	0.08254659	0.3125957	0.01767663	0.0899736	0.14736635
DOGS.3	0.036528166	0.10684085	0.4002217	0.02336526	0.1196213	0.13862595
DOGS.4	0.006949558	0.10261772	0.4115529	0.02235910	0.1143523	0.14249058
DOGS.5	0.579261207	0.54758998	0.6410213	0.19489684	1.2103843	0.05371597
DOGS.6	0.267895534	0.39347115	0.5973212	0.11484468	0.6487262	0.09993180

	var.r	med.r
DOGS.1	0.09887512	-0.1029329
DOGS.2	0.09712237	-0.1005350
DOGS.3	0.14352826	-0.1328671
DOGS.4	0.15580372	-0.1005350
DOGS.5	0.09708139	0.2716175
DOGS.6	0.17051487	0.2716175

- The raw ( = .22) and standardized ( = .30) alpha scores are quite low for all of the items together, and items 5 and 6 are both negatively correlated with the first principal component of the items.
  - Item 5: Date a non-Christian
  - Item 6: Have close platonic relationships with non-Christians
- Looking at the alpha scores if items are dropped, when item 5 (DOGS.5) is excluded the raw ( = .58) and standardized ( = .55) alpha scores substantially increase. So we will remove that item and re-run the analysis.

```
# Re-running the analysis with only items 1-4 and 6
Chronbach_DOGS_2 <- alpha(DOGS_items[, c(1:4, 6)])
```

Warning in alpha(DOGS\_items[, c(1:4, 6)]): Some items were negatively correlated with the first principal component and should be reversed.

To do this, run the function again with the 'check.keys=TRUE' option

Some items ( DOGS.6 ) were negatively correlated with the first principal component and probably should be reversed.

To do this, run the function again with the 'check.keys=TRUE' option

```
# Displaying the results
print(Chronbach_DOGS_2$total)
```



```
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd
0.5792612 0.54759 0.6410213 0.1948968 1.210384 0.05371597 5.049254 1.071977
median_r
0.2716175
```

```
print(Chronbach_DOGS_2$alpha.drop)
```

```
raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
DOGS.1 0.3895247 0.3678987 0.4128979 0.1270235 0.5820249 0.08207802 0.08039444
DOGS.2 0.3721707 0.3517303 0.3932405 0.1194408 0.5425678 0.08510275 0.07315388
DOGS.3 0.4663998 0.4085507 0.5333289 0.1472601 0.6907620 0.06911097 0.13251741
DOGS.4 0.5058663 0.4563245 0.5610495 0.1734397 0.8393324 0.06442211 0.12981592
DOGS.6 0.7362309 0.7332626 0.7479650 0.4073201 2.7490061 0.03788809 0.03857336
med.r
DOGS.1 0.10478473
DOGS.2 0.09623195
DOGS.3 0.06629775
DOGS.4 0.10947699
DOGS.6 0.31484057
```

- After re-running the analysis, the sixth item is still negatively correlated with the first principal component of the items, and the `alpha.drop` statistics indicate that removing this item from the scale improves the internal consistency of the scale to  $= .74$  (standardized  $= .73$ ).
- Therefore, we will remove this item and run the analysis again.

```
# Re-running the analysis with only items 1-4
Chronbach_DOGS_3 <- alpha(DOGS_items[, c(1:4)])

# Displaying the results
print(Chronbach_DOGS_3$total)
```

```
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd
0.7362309 0.7332626 0.747965 0.4073201 2.749006 0.03788809 4.912313 1.349693
median_r
0.3148406
```

```
print(Chronbach_DOGS_3$alpha.drop)
```

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha se	var.r
DOGS.1	0.6319904	0.6433872	0.5621625	0.3755414	1.804162	0.05582547	0.01372745
DOGS.2	0.6125178	0.6207321	0.5419762	0.3529824	1.636659	0.05886973	0.01781150
DOGS.3	0.7080266	0.6986929	0.6906940	0.4359707	2.318873	0.04383124	0.08110905
DOGS.4	0.7328995	0.7226256	0.7069476	0.4647861	2.605235	0.03939772	0.06809032

	med.r
DOGS.1	0.3401263
DOGS.2	0.2895548
DOGS.3	0.2801703
DOGS.4	0.3401263

- The new alpha statistic for the scale is now .74 (standardized  $\alpha = .73$ ), and there are no items for which their removal increases the reliability of the scale.

To update the total scores in the clean data, we will first assign the old `RBI.DOGS` variable to a new variable called `DOGS_original` and recalculate the `RBI.DOGS` scores for participants with only the first four items.

```
# Creating the DOGS_original variable
data$DOGS_original <- data$RBI.DOGS

# Recalculating the RBI.DOGS
data <- data %>%
  rowwise() %>%
  mutate(
    RBI.DOGS = mean(c_across(DOGS.1:DOGS.4), na.rm = FALSE), # Mean for DOGS items
  ) %>%
  ungroup()
```

We will now recalculate descriptive statistics and display the histogram for the `RBI.DOGS` variable.

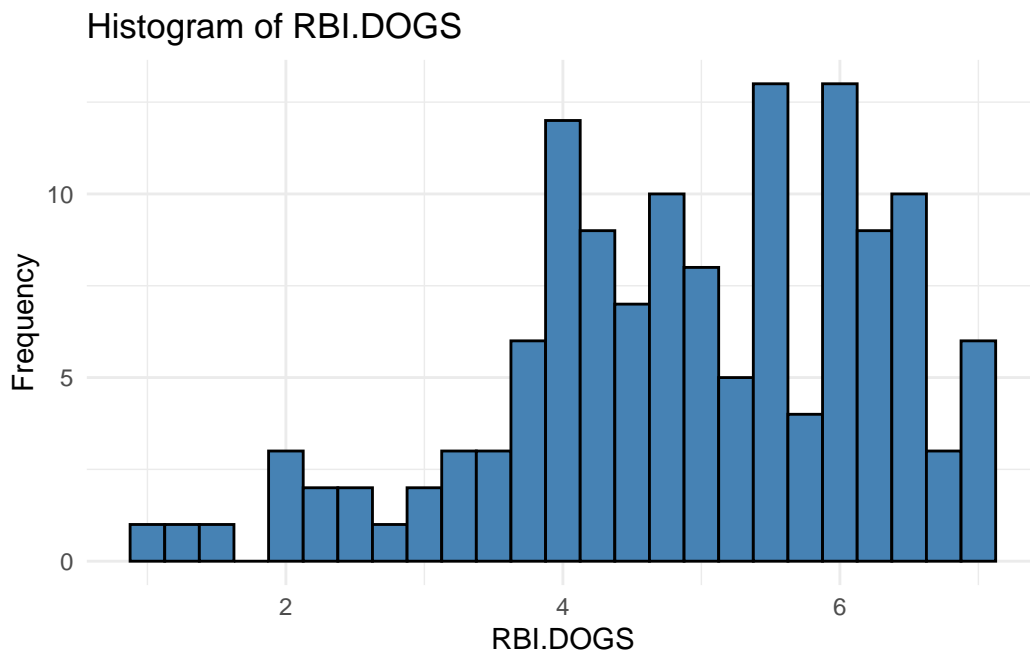
```
# Generating descriptive statistics for RBI.DOGS
RBI.DOGS_descriptives <- describe(data$RBI.DOGS)

# Display the results
print(RBI.DOGS_descriptives)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	134	4.91	1.35	5	5.02	1.48	1	7	6	-0.61	-0.07	0.12

```
# Generating the histogram for RBI.DOGS
DOGS_new_histogram <- ggplot(data, aes(x = RBI.DOGS)) +
  geom_histogram(bins = 25, fill = "steelblue", color = "black") +
  theme_minimal() +
  labs(title = "Histogram of RBI.DOGS", x = "RBI.DOGS", y = "Frequency")

print(DOGS_new_histogram)
```



## 2. Religious Behavioral Intentions to Increase In-Group Embeddedness

```
# Creating a data frame with only the IIGE items
IIGE_items <- select(data, IIGE.1:IIGE.7)

# Calculating Chronbach's alpha
Chronbach_IIGE <- alpha(IIGE_items)

# Displaying the results
print(Chronbach_IIGE$total)
```

```
raw_alpha std.alpha  G6(smc) average_r    S/N      ase    mean    sd
0.8265813 0.8117167 0.8365008 0.3811414 4.311146 0.02008378 5.761549 1.052421
median_r
0.3586863
```

```
print(Chronbach_IIGE$alpha.drop)
```

```

      raw_alpha std.alpha  G6(smc) average_r      S/N  alpha se      var.r
IIGE.1 0.7704346 0.7574861 0.7828682 0.3423558 3.123475 0.02740881 0.03610733
IIGE.2 0.7675366 0.7566002 0.7723621 0.3412722 3.108466 0.02834086 0.03112137
IIGE.3 0.7818291 0.7712560 0.7883693 0.3597746 3.371700 0.02617456 0.03069958
IIGE.4 0.8378156 0.8226340 0.8371414 0.4359874 4.638060 0.01881773 0.04313573
IIGE.5 0.8145936 0.7966374 0.8196301 0.3949982 3.917326 0.02148068 0.04980220
IIGE.6 0.7876438 0.7685943 0.7918860 0.3563209 3.321415 0.02454099 0.04394642
IIGE.7 0.8368444 0.8233999 0.8348810 0.4372808 4.662512 0.01935906 0.04139336

      med.r
IIGE.1 0.2408381
IIGE.2 0.2835710
IIGE.3 0.2835710
IIGE.4 0.4055097
IIGE.5 0.2835710
IIGE.6 0.2835710
IIGE.7 0.4055097

```

- The original alpha value ( = .83) and standardized alpha value ( = .81) are already relatively strong for the IIGE scale.
- Two items, however, are contributing less well to the internal consistency than the others, and removing them from the scale indicates that the internal consistency will not decrease substantially. They are item 4 and item 7.
  - Item 4: Be involved in a committed relationship with another Christian
  - Item 7: Be involved in a close platonic relationship with other Christians
- These items are thematically very similar to the two items removed from our Decrease Out-Group Separation scale. That is, they represent behaviors that have to do with personal relationships, whereas the rest of the items do not. This indicates to me that the items are affected by factors to do with personal relationships rather than the individual's relationship to the group, which is what we are trying to get at most particularly here. Removing these items from the scale would allow it to mirror the Decrease Out-Group Separation scale better. In addition, these items are very similar to some of the items in the Social Distance questionnaire, which we will use to validate the RBI subscales later. If this is the case, correlations between the RBI subscales and the Social Distance scores may not accurately represent concurrent validity, because some of the items are essentially the same. Therefore, I am going to remove items 4 and 7 from the scale and re-run the analysis.

```
# Re-running the analysis with only items 1-3 and 5-6
Chronbach_IIGE_2 <- alpha(IIGE_items[, c(1:3, 5:6)])

# Displaying the results
print(Chronbach_IIGE_2$total)
```

```
raw_alpha std.alpha  G6(smc) average_r      S/N      ase      mean      sd
0.8637611 0.8604012 0.8597012 0.5521071 6.163383 0.01754837 5.538433 1.326861
  median_r
0.5476191
```

```
print(Chronbach_IIGE_2$alpha.drop)
```

```
      raw_alpha std.alpha  G6(smc) average_r      S/N  alpha se      var.r
IIGE.1 0.8182228 0.8180813 0.8059262 0.5292434 4.496960 0.02431203 0.018802524
IIGE.2 0.8093727 0.8087785 0.7839428 0.5139460 4.229539 0.02579800 0.014259074
IIGE.3 0.8196335 0.8190583 0.8064363 0.5308822 4.526643 0.02422576 0.018484938
IIGE.5 0.8765977 0.8753811 0.8542974 0.6371706 7.024466 0.01710577 0.009626451
IIGE.6 0.8389000 0.8297858 0.8147013 0.5492932 4.874950 0.02041079 0.035574869
      med.r
IIGE.1 0.5359734
IIGE.2 0.5476191
IIGE.3 0.5436869
IIGE.5 0.6256166
IIGE.6 0.5293553
```

- Removal of items 4 and 7 increased the alpha value for the scale to  $\alpha = .86$  (standardized  $\alpha = .86$ ).
- While the `alpha.drop` statistics indicate that the internal consistency can be slightly increased by removing item 5, I believe that this item has strong face validity as a religious behavior that could function to increase in-group embeddedness. Therefore, I will leave it in.

– Item 5: Be baptized

To update the total scores in the clean data, we will first assign the old `RBI.IIGE` variable to a new variable called `IIGE_original` and recalculate the `RBI.IIGE` scores for participants with only items 1, 2, 3, 5, and 6. We will also update the `RBI.F` variable.

```
# Creating the IIGE_original and RBI.F_original variable
data$IIGE_original <- data$RBI.IIGE
RBI.F_original <- data$RBI.F

# Calculating new total variables
data <- data %>%
  rowwise() %>%
  mutate(
    RBI.IIGE = mean(c_across(c(IIGE.1, IIGE.2, IIGE.3, IIGE.5, IIGE.6)), na.rm = FALSE),
    RBI.F = mean(c_across(c(DOGS.1, DOGS.2, DOGS.3, DOGS.4, IIGE.1, IIGE.2, IIGE.3, IIGE.5, IIGE.6)), na.rm = FALSE)
  ) %>%
  ungroup()
```

Now we will recalculate descriptive statistics and the histogram for the RBI.IIGE variable.

```
# Generating descriptive statistics for RBI.IIGE
RBI.IIGE_descriptives <- describe(data$RBI.IIGE)

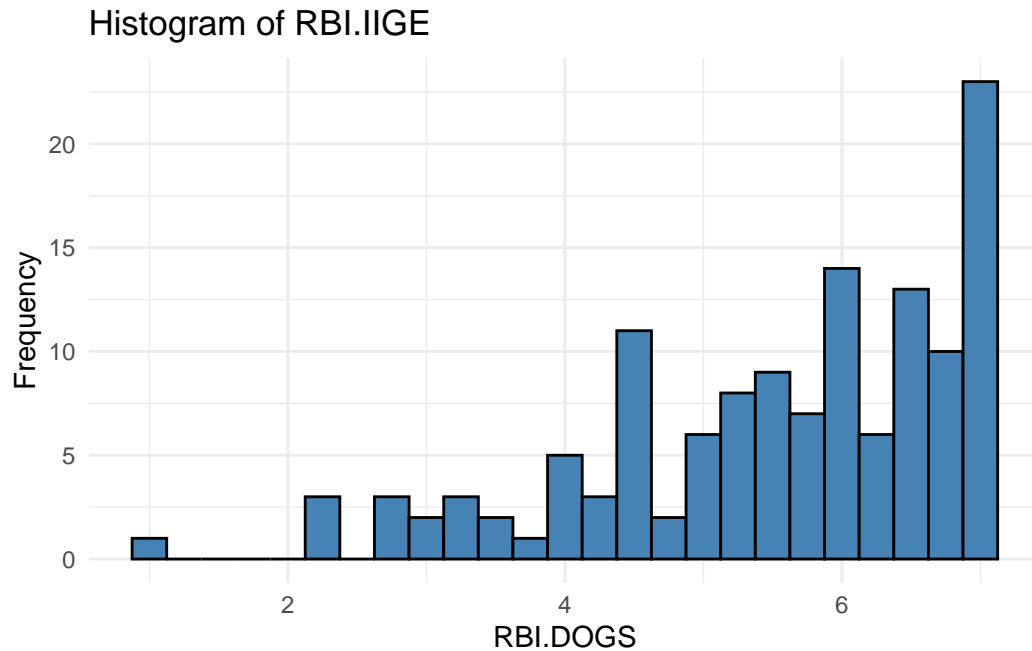
# Display the results
print(RBI.IIGE_descriptives)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	132	5.54	1.34	5.9	5.71	1.33	1	7	6	-0.93	0.28	0.12

```
# Generating the histogram for RBI.DOGS
IIGE_new_histogram <- ggplot(data, aes(x = RBI.IIGE)) +
  geom_histogram(bins = 25, fill = "steelblue", color = "black") +
  theme_minimal() +
  labs(title = "Histogram of RBI.IIGE", x = "RBI.DOGS", y = "Frequency")

print(IIGE_new_histogram)
```

Warning: Removed 2 rows containing non-finite outside the scale range (`stat\_bin()`).



## Validity

To establish the concurrent validity of the RBI scales, we will run bivariate correlations with theoretically relevant variables.

### 1. Religiosity aspects of the RBI subscales

Our scales should both be positively related to a measure that is sensitive to one's degree of religiosity. Thus, to establish the validity of the religiosity aspect of our `RBI.DOGS` and `RBI.IIGE` scales, we will run bivariate correlations with the `PIOS` and its' subscales. This will also allow us to assess the relationship between the two RBI subscales.

```
# Creating a data frame with only the RBI subscales and the PIOS and its' subscales
rel.val.data <- select(data, RBI.DOGS, RBI.IIGE, PIOS:FOS)

# Creating the correlation matrix
rel.val.cor <- corr.test(rel.val.data)

# Display the results
print(rel.val.cor)
```

```
Call:corr.test(x = rel.val.data)
```

# Correlation matrix

	RBI.DOGS	RBI.IIGE	PIOS	FOG	FOS
RBI.DOGS	1.00	0.62	0.19	0.09	0.21
RBI.IIGE	0.62	1.00	0.39	0.32	0.40
PIOS	0.19	0.39	1.00	0.89	0.96
FOG	0.09	0.32	0.89	1.00	0.73
FOS	0.21	0.40	0.96	0.73	1.00

# Sample Size

	RBI.DOGS	RBI.IIGE	PIOS	FOG	FOS
RBI.DOGS	134	132	127	131	127
RBI.IIGE	132	132	127	130	127
PIOS	127	127	127	127	127
FOG	131	130	127	131	127
FOS	127	127	127	127	127

Probability values (Entries above the diagonal are adjusted for multiple tests.)

	RBI.DOGS	RBI.IIGE	PIOS	FOG	FOS
RBI.DOGS	0.00	0	0.06	0.29	0.06
RBI.IIGE	0.00	0	0.00	0.00	0.00
PIOS	0.03	0	0.00	0.00	0.00
FOG	0.29	0	0.00	0.00	0.00
FOS	0.02	0	0.00	0.00	0.00

To see confidence intervals of the correlations, print with the short=FALSE option

- RBI Scales Correlation with Eachother:
  - The correlation between the DOGS and IIGE is a strong positive correlation ( $r = .62$ ,  $p < .01$ ).
- The RBI Scales and the PIOS:
  - The DOGS has a small, positive correlation with the PIOS ( $r = .19$ ,  $p = .03$ ).
    - \* The relationship between DOGS and the PIOS is mostly due to its relationship with the FOS subscale ( $r = .21$ ,  $p = .02$ ), and the DOGS is not significantly correlated with the FOG subscale ( $r = .09$ ,  $p = .29$ ).
  - The IIGE has a moderate, positive correlation with the PIOS ( $r = .39$ ,  $p < .01$ ).
    - \* The relationship between the IIGE and the FOS subscale ( $r = .40$ ,  $p < .01$ ) is slightly stronger than the relationship between the IIGE and the FOG subscale ( $r = .32$ ,  $p < .01$ ).

The relationships here provide partial support for the validity of the RBI subscales. First, the strong positive relationship between the subscales permits the interpretation that their



shared religious nature may underly this covariance. Second, the RBI subscales are both significantly correlated with the full PIOS scale. Third, however, the relationships between the RBI subscales—particularly for the DOGS—and the full PIOS scale are small to medium sized relationships, and stronger relationships would provide greater evidence for the validity of the religious aspects of the RBI subscales.

As an additional note, the relationships are stronger between the RBI subscales and the FOS subscale of the PIOS, and the relationship between the DOGS and the FOS subscale is very small and not statistically reliable. It is not clear to me why these relationships should be stronger for fear of sin.

## 2. In-group and out-group attitudes aspects of the RBI subscales

We can also make predictions about how the RBI subscales should relate to measures of in-group and out-group attitudes towards religious groups, as measured by our social distance questionnaire. Higher social distance is intended to indicate more negative (less positive) attitudes towards the target group. Lower social distance is intended to indicate more positive (less negative) attitudes towards the target group. The predictions are as follows.

- For the IIGE:
  1. The IIGE should correlate *negatively* with the measured social distance of our participants' religious in-group (i.e., `SD_christian`).
  2. The IIGE should correlate *positively* with the measured social distance of our participants' religious out-groups (i.e., `SD_non_christian` and `SD_athiest`).
- For the DOGS:
  1. The DOGS should correlate *negatively* with the measured social distance of our participants' religious in-group (i.e., `SD_christian`), but this relationship should not be as strong as for the IIGE.
  2. The DOGS should also correlate *negatively* with the measured social distance of our participants' religious out-groups (i.e., `SD_non_christian` and `SD_athiest`).

I should note that we originally measured social distance towards Christian sects, and this was intended to enable use to measure the relationship between the RBI subscales and these target sects within the sects of our sample. However, I do not believe that we have enough participants to allow for reliable sub-group analyses, because we only have data on social distance for the last round of data collection on SONA. Therefore, I will skip this.

```
# Creating a data frame with only the RBI subscales and the Social Distance Scores of interest
att.val.data <- select(data, RBI.DOGS, RBI.IIGE, SD_christian, SD_non_christian, SD_athiest)

# Creating the correlation matrix
att.val.cor <- corr.test(att.val.data)
```

```
# Display the results
print(att.val.cor)
```

```
Call:corr.test(x = att.val.data)
```

Correlation matrix

	RBI.DOGS	RBI.IIGE	SD_christian	SD_non_christian	SD_athiest
RBI.DOGS	1.00	0.62	-0.19	-0.14	-0.09
RBI.IIGE	0.62	1.00	-0.23	-0.14	-0.17
SD_christian	-0.19	-0.23	1.00	0.71	0.27
SD_non_christian	-0.14	-0.14	0.71	1.00	0.71
SD_athiest	-0.09	-0.17	0.27	0.71	1.00

Sample Size

	RBI.DOGS	RBI.IIGE	SD_christian	SD_non_christian	SD_athiest
RBI.DOGS	134	132	47	48	48
RBI.IIGE	132	132	46	46	46
SD_christian	47	46	47	47	47
SD_non_christian	48	46	47	48	48
SD_athiest	48	46	47	48	48

Probability values (Entries above the diagonal are adjusted for multiple tests.)

	RBI.DOGS	RBI.IIGE	SD_christian	SD_non_christian	SD_athiest
RBI.DOGS	0.00	0.00	0.97	1	1.00
RBI.IIGE	0.00	0.00	0.70	1	1.00
SD_christian	0.19	0.12	0.00	0	0.44
SD_non_christian	0.35	0.35	0.00	0	0.00
SD_athiest	0.53	0.26	0.06	0	0.00

To see confidence intervals of the correlations, print with the short=FALSE option

- For the IIGE:
  1. In line with our prediction, the relationship between the IIGE subscale and social distance towards Christians is negative, but this relationship is not significant ( $r = -.23$ ,  $p = .12$ ,  $n = 46$ ).
  2. In contrast to our predictions, the relationship between the IIGE subscale and social distance towards non-Christian religious groups is negative ( $r = -.14$ ,  $p = .35$ ,  $n = 46$ ), and the relationship between the IIGE subscale and social distance towards athiests is also negative ( $r = -.17$ ,  $p = .26$ ,  $n = 46$ ). Neither of these relationships are significant, however.
- For the DOGS:

1. In line with our predictions, the relationship between the DOGS subscale and social distance towards Christians is negative ( $r = -.19$ ,  $p = .19$ ,  $n = 47$ ) and not as strong as for the IIGE subscale. However, this relationship is not significant.
2. In line with our predictions, the relationship between the DOGS subscale and social distance towards non-Christian religious groups is negative ( $r = -.14$ ,  $p = .35$ ,  $n = 48$ ) and the relationship between the DOGS subscale and social distance towards athiests is also negative ( $r = -.09$ ,  $p = .53$ ,  $n = 48$ ). Neither of these relationships are significant, however.

This analysis does not provide evidence for the validity of the in-group and out-group attitudes aspects of the RBI subscales. Although there is low power to be able to detect small effects, none of the relationships that we predicted are supported with significant correlations.

In sum, this analysis provides very little evidence that the RBI subscales are measuring what we intend them to measure.

## Testing Hypotheses

Although we do not have strong evidence for validity of the RBI subscales, I will still test our hypotheses with these measures and interpret the results cautiously.

The hypotheses are as follows:

Hypothesis 1: Religious individuals scoring higher in pathogen disgust sensitivity will report being more likely to engage in religious behaviors that increase in-group embeddedness.

Hypothesis 2: Religious individuals scoring higher in pathogen disgust sensitivity will report being less likely to engage in religious behaviors that decrease out-group separation.

### Hypothesis 1

To test Hypothesis 1, we will conduct a simple linear regression with pathogen disgust (PD) predicting religious behavioral intention to increase in-group embeddedness (RBI.IIGE). We should find a positive relationship between them.

First, we will standardize both variables to allow for easier interpretation of the effect size.

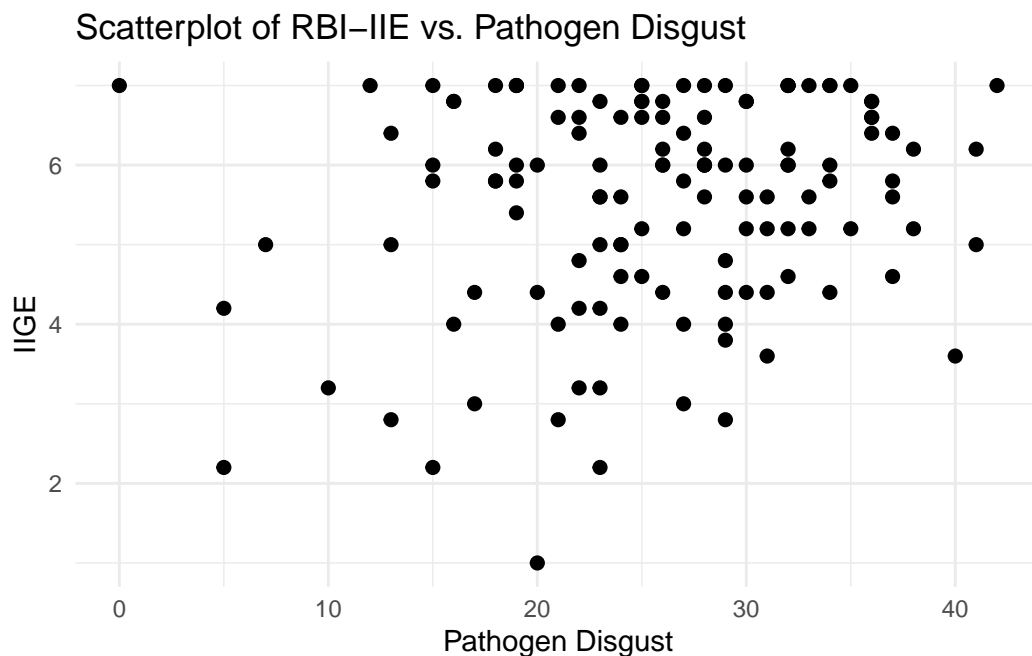
```
# Standardizing the variables
z_IIGE <- scale(data$RBI.IIGE)
z_PD <- scale(data$PD)
```

### Assumptions:

1. Dependent variable interval-ratio: Satisfied
2. Relationship between the variables is linear: Satisfied via scatterplot below—the relationship does not appear to be non-linear.

```
# Creating a scatterplot between PD and RBI.IIGE
ggplot(data, aes(x = PD, y = RBI.IIGE)) +
  geom_point(color = "black", size = 2) +
  ggtitle("Scatterplot of RBI-IIE vs. Pathogen Disgust") +
  xlab("Pathogen Disgust") +
  ylab("IIGE") +
  theme_minimal()
```

Warning: Removed 3 rows containing missing values or values outside the scale range (`geom\_point()`).



3. Independence of Residuals: Violated via the Durban-Watson test below.

```
# Fitting the linear model
hyp_1_model <- lm(data = data, z_IIGE ~ z_PD)

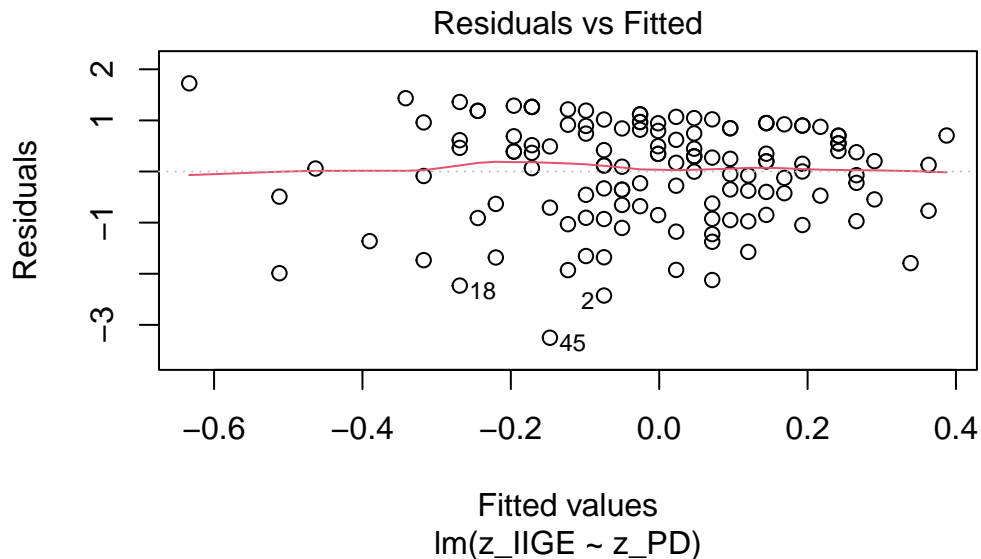
# Running the Durban-Watson test below
dwtest(hyp_1_model)
```

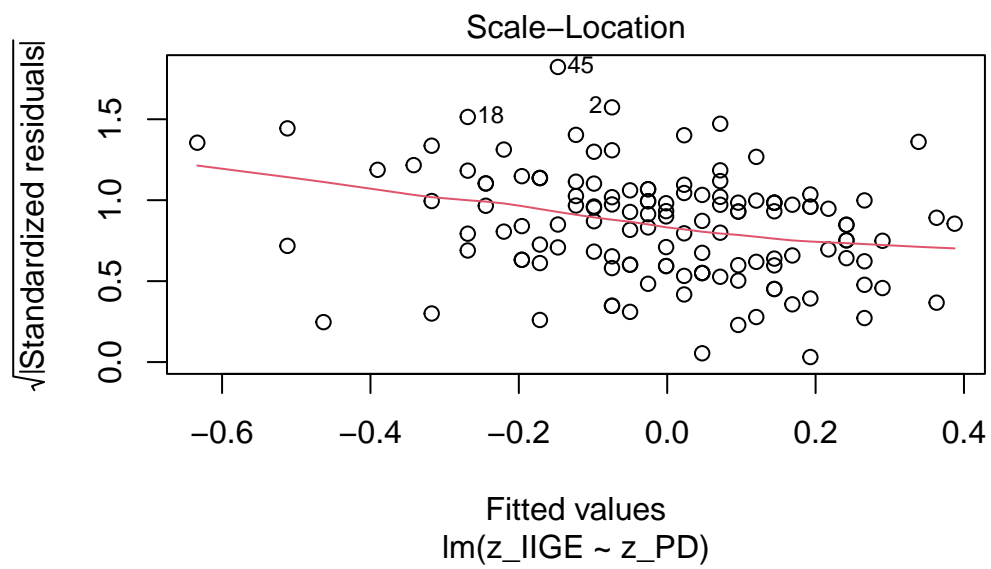
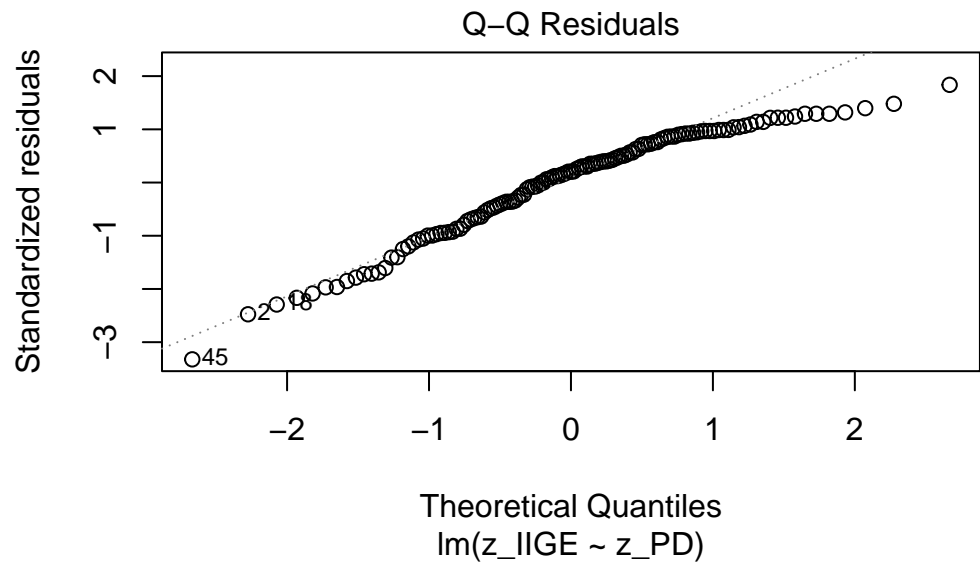
#### Durban-Watson test

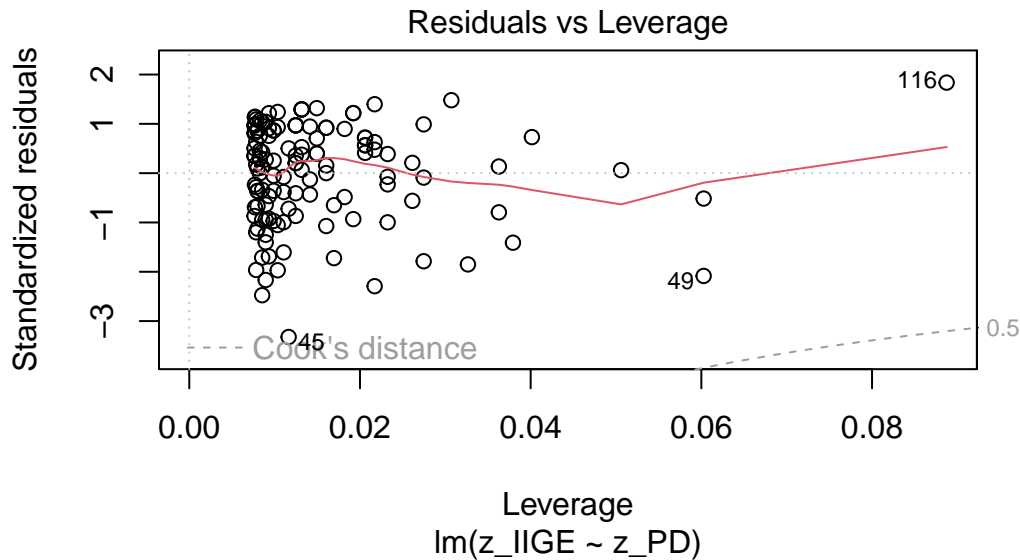
```
data: hyp_1_model
DW = 1.6643, p-value = 0.02648
alternative hypothesis: true autocorrelation is greater than 0
```

- The Durban-Watson test indicates that the autocorrelation for the model is significantly greater than zero ( $DW = 1.66$ ,  $p = .026$ ), indicating that the independence of residuals assumption is violated. However, the statistic is not extremely different from 2 in absolute terms. Because of this, we will test the model using with robust standard errors.
4. Equality of Variance: Violated via the plot of residual vs. fitted values and the Breusch-Pagan test below.

```
# Plotting residual vs. fitted values
plot(hyp_1_model)
```







```
# Breusch-Pagan test
bptest(hyp_1_model)
```

studentized Breusch-Pagan test

```
data: hyp_1_model
BP = 9.9406, df = 1, p-value = 0.001617
```

- The assumption of equality of variance is also violated here due to a significant result for the Breusch-Pagan test ( $BP(1) = 9.94$ ,  $p < .01$ ). This should be taken care of by estimating robust standard errors.
5. Residuals Normally Distributed: Violated by the Q-Q plot of residuals above.
- The residuals deviate from normality at the end of the distribution, which I suspect is due to the presence of many responses on the IIGE that are at the top end of the scale (see the histogram in [Reliability](#)). Because we have so many observations of these variables, this should not be a huge problem, but we still must be careful interpreting the results of hypothesis testing.

### Testing the Model:

To test the model, I will use robust standard errors from the `sandwich` package.

```
# Summarizing the model
summary(hyp_1_model)
```

Call:

```
lm(formula = z_IIGE ~ z_PD, data = data)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.2516	-0.6440	0.1993	0.8307	1.7256

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.006406	0.086003	-0.074	0.9407
z_PD	0.192058	0.086128	2.230	0.0275 *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9843 on 129 degrees of freedom

(3 observations deleted due to missingness)

Multiple R-squared: 0.03712, Adjusted R-squared: 0.02965

F-statistic: 4.973 on 1 and 129 DF, p-value: 0.02748

```
# Testing the model with robust standard errors
coeftest(hyp_1_model, vcov = vcovHC(hyp_1_model, type = "HC1"))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0064062	0.0856613	-0.0748	0.94050
z_PD	0.1920582	0.0969121	1.9818	0.04963 *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

- The significance test of the predictor with robust standard errors is significant ( $b = .192$ ,  $t(129) = 1.98$ ,  $p = .0496$ ), indicating that one standard deviation increase in pathogen disgust is associated with a .19 standard deviation increase in religious behavioral intention to increase in-group embeddedness. It is also worth noting here that this relationship may be larger than is represented here, because we have a restricted range at the top end of the scale, as per the scatterplot above.



- This is consistent with Hypothesis 1, although given our violations of assumptions of the regression model it may be wiser to stay agnostic.

## Hypothesis 2

To test Hypothesis 2, we will conduct a simple linear regression with pathogen disgust (PD) predicting religious behavioral intention to decrease out-group separation (RBI.DOGS). We should find a negative relationship between the variables.

First, we will standardize RBI.DOGS to allow for easier interpretation of the effect size.

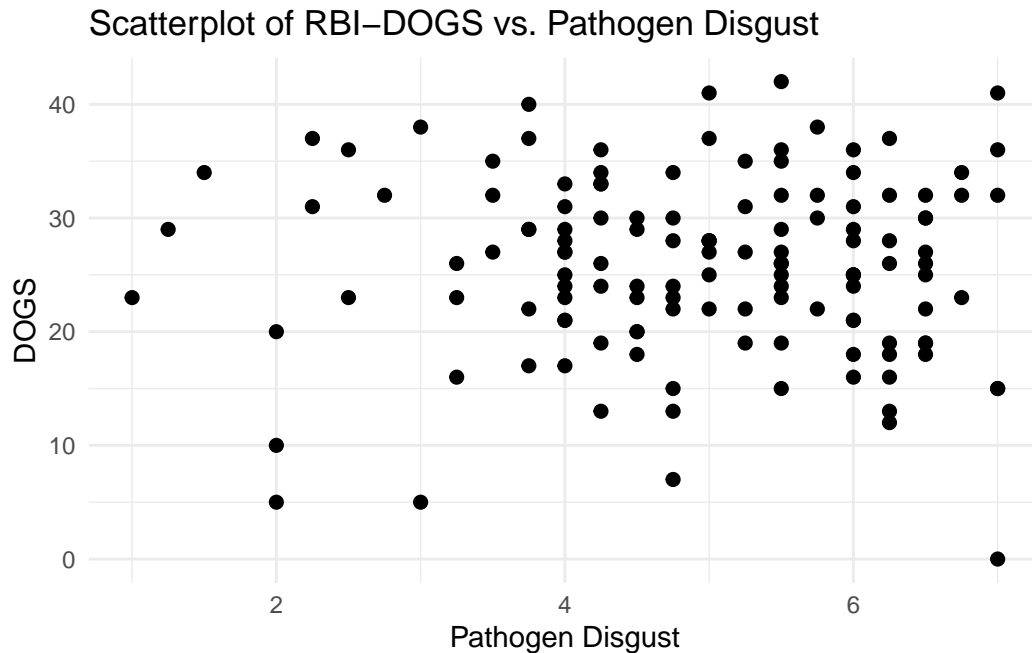
```
# Standardizing the variable
z_DOGS <- scale(data$RBI.DOGS)
```

### Assumptions:

1. Dependent variable interval-ratio: Satisfied
2. Relationship between the variables is linear: Satisfied via scatterplot below—the relationship does not appear to be non-linear.

```
# Creating a scatterplot between PD and RBI.DOGS
ggplot(data, aes(x = RBI.DOGS, y = PD)) +
  geom_point(color = "black", size = 2) +
  ggtitle("Scatterplot of RBI-DOGS vs. Pathogen Disgust") +
  xlab("Pathogen Disgust") +
  ylab("DOGS") +
  theme_minimal()
```

Warning: Removed 1 row containing missing values or values outside the scale range (`geom\_point()`).



3. Independence of Residuals: Satisfied via the Durban-Watson test below.

```
# Fitting the linear model
hyp_2_model <- lm(data = data, z_DOGS ~ z_PD)

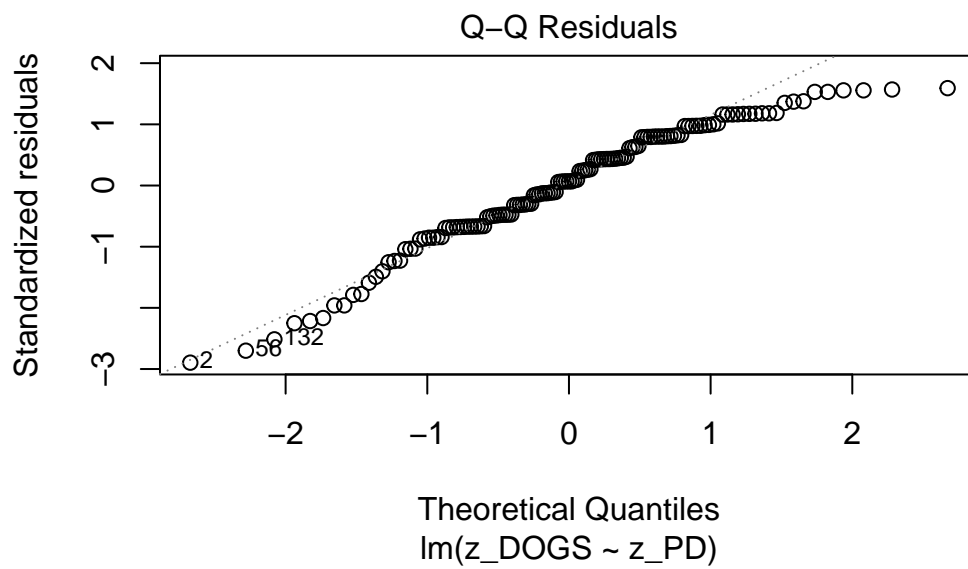
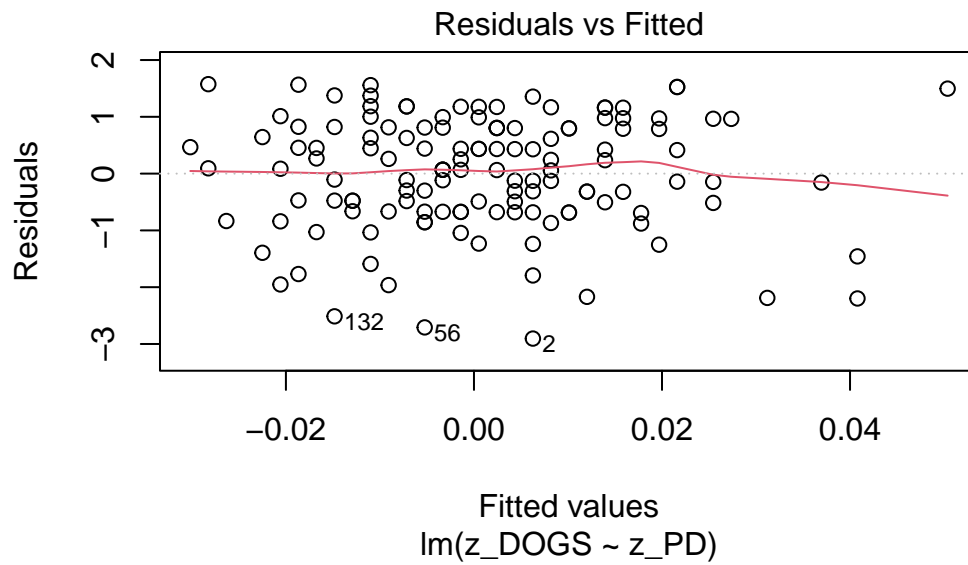
# Running the Durban-Watson test below
dwtest(hyp_2_model)
```

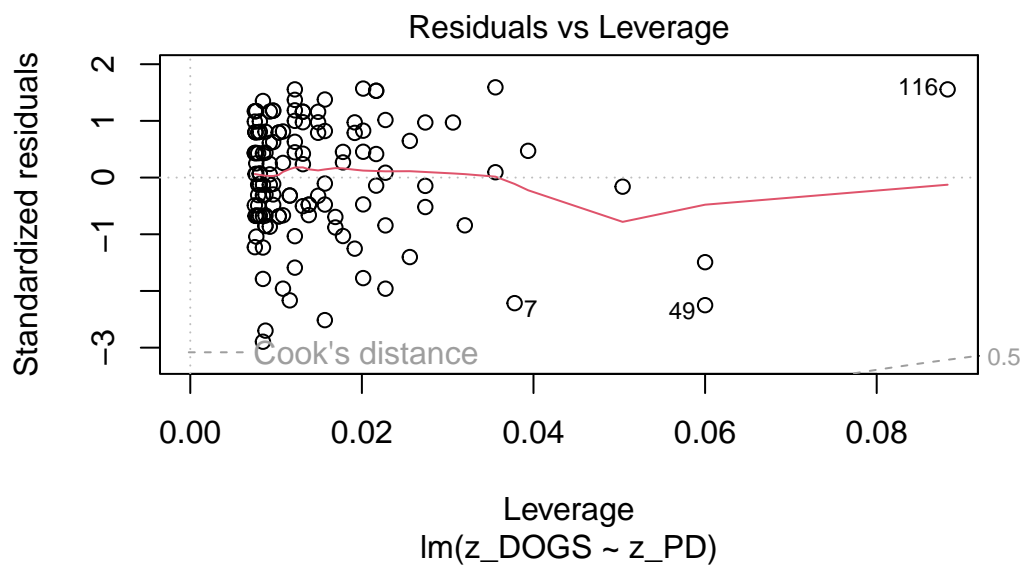
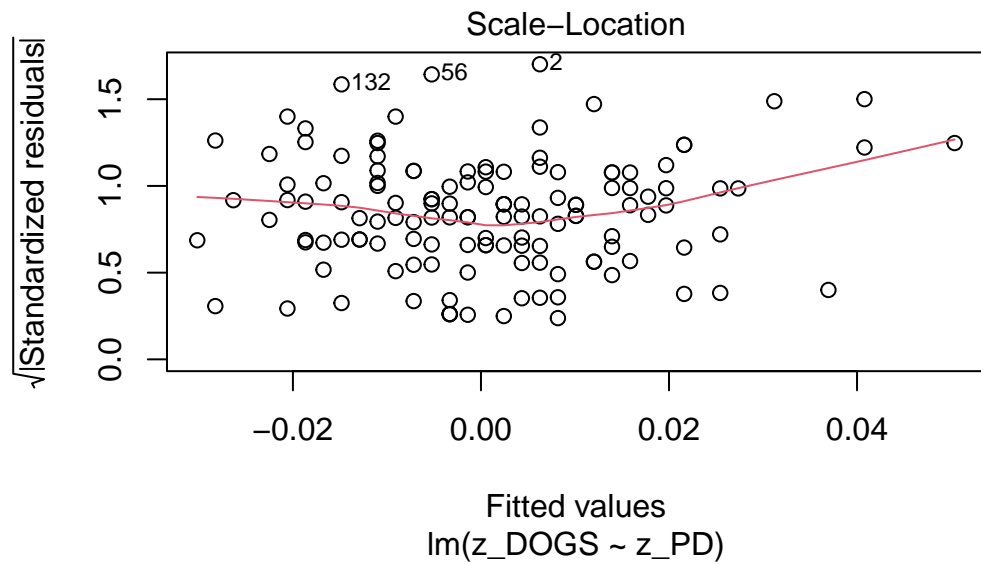
Durbin-Watson test

```
data: hyp_2_model
DW = 1.7627, p-value = 0.08388
alternative hypothesis: true autocorrelation is greater than 0
```

- The Durban-Watson test indicates that the autocorrelation for the model is not significantly greater than zero ( $DW = 1.76$ ,  $p = .08$ ).
4. Equality of Variance: Satisfied via the plot of residual vs. fitted values and the Breusch-Pagan test below.

```
# Plotting residual vs. fitted values
plot(hyp_2_model)
```





```
# Breusch-Pagan test
bptest(hyp_2_model)
```

studentized Breusch-Pagan test

```
data: hyp_2_model  
BP = 0.57777, df = 1, p-value = 0.4472
```

- The assumption of equality of variance is also not violated here due to an insignificant result for the Breusch-Pagan test ( $BP(1) = .58, p = .45$ ).
5. Residuals Normally Distributed: Violated by the Q-Q plot of residuals above.
- The residuals deviate from normality at the end of the distribution. However, we have enough observations that this should not be a huge problem.

### Testing the Model:

To test the model, I will use robust standard errors from the **sandwich** package.

```
# Summarizing the model  
summary(hyp_2_model)
```

Call:

```
lm(formula = z_DOGS ~ z_PD, data = data)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-2.90494	-0.67262	0.06829	0.80344	1.57504

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.0009042	0.0873557	0.010	0.992
z_PD	-0.0151551	0.0876859	-0.173	0.863

Residual standard error: 1.007 on 131 degrees of freedom

(1 observation deleted due to missingness)

Multiple R-squared: 0.000228, Adjusted R-squared: -0.007404

F-statistic: 0.02987 on 1 and 131 DF, p-value: 0.863

- The t test for the predictor is not significant ( $b = -.015, t(131) = -.173, p = .86$ ). Overall, this does not provide strong support for Hypothesis 2.

Overall, there is not good evidence for either hypothesis here, especially considering that we have not provided good evidence of the validity of the RBI subscales.

## Estimating the Relationship Between Disgust and Religiosity

Since we cannot rely on the RBI subscales to give us a strong indicator of religiosity, I would like to estimate the strength of the relationship between disgust and religiosity to inform future study. This is a Christian sample, so we are dealing with a restricted range and cannot estimate the relationship between disgust and religious affiliation. However, I still believe it will be informative to look at the relationship between pathogen disgust and scores on the PIOS (as well as its' subacales FOG and FOS), as indicators of the degree of one's religiosity.

First, we will standardize the PIOS variables to allow for easier interpretation of the effect size.

```
# Standardizing the variables
z_PIOS <- scale(data$PIOS)
z_FOG <- scale(data$FOG)
z_FOS <- scale(data$FOS)
```

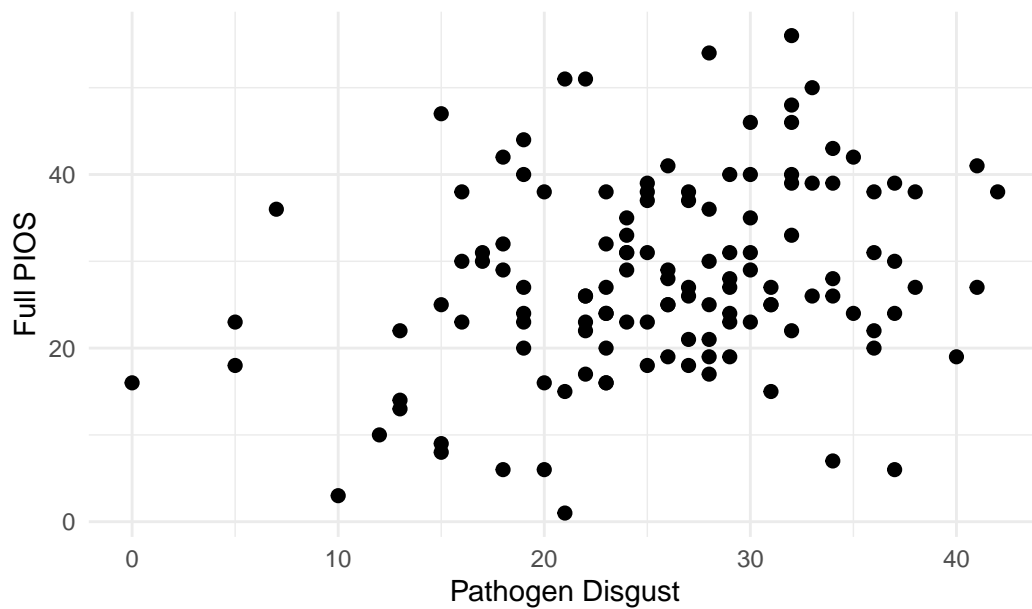
### Assumptions:

1. Dependent variables are interval-ratio: Satisfied
2. Relationships between the variables is linear: Satisfied via the scatterplots below—the relationships do not appear to be non-linear.

```
# Creating a scatterplot between PD and the PIOS
ggplot(data, aes(x = PD, y = PIOS)) +
  geom_point(color = "black", size = 2) +
  ggtitle("Scatterplot of Pathogen Disgust and the PIOS") +
  xlab("Pathogen Disgust") +
  ylab("Full PIOS") +
  theme_minimal()
```

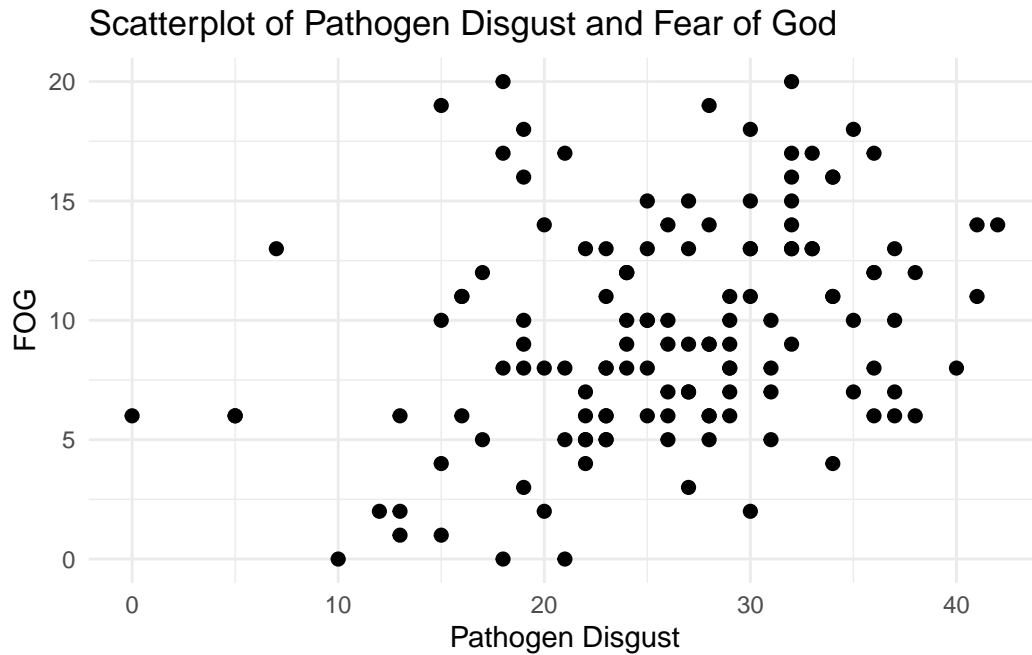
Warning: Removed 7 rows containing missing values or values outside the scale range (``geom_point()``).

Scatterplot of Pathogen Disgust and the PIOS



```
# Creating a scatterplot between PD and FOG
ggplot(data, aes(x = PD, y = FOG)) +
  geom_point(color = "black", size = 2) +
  ggtitle("Scatterplot of Pathogen Disgust and Fear of God") +
  xlab("Pathogen Disgust") +
  ylab("FOG") +
  theme_minimal()
```

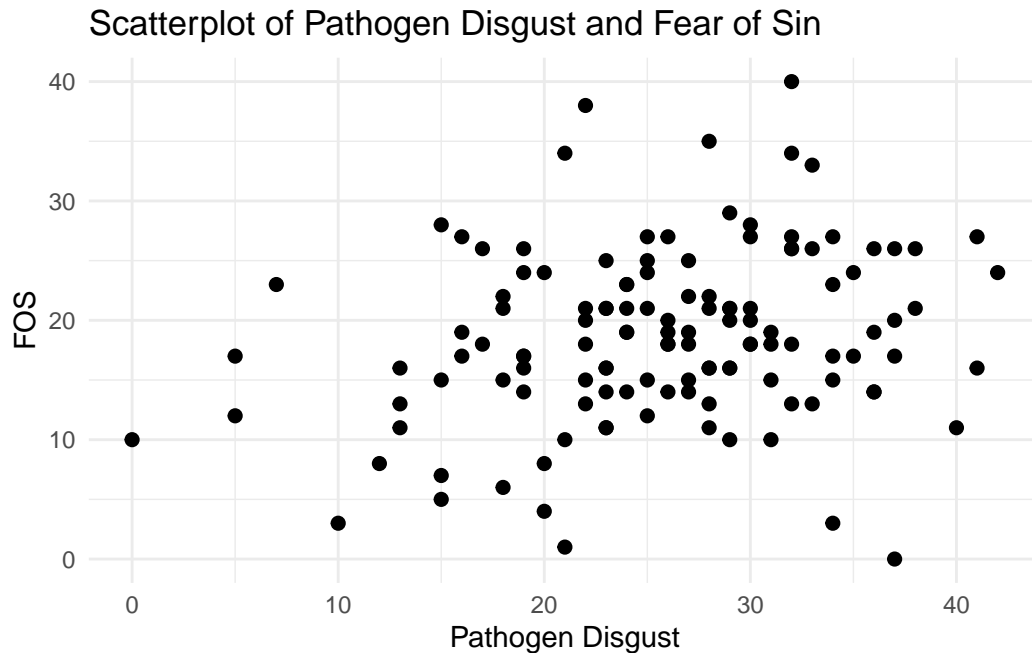
Warning: Removed 3 rows containing missing values or values outside the scale range (`geom\_point()`).



```
# Creating a scatterplot between PD and FOS
ggplot(data, aes(x = PD, y = FOS)) +
  geom_point(color = "black", size = 2) +
  ggtitle("Scatterplot of Pathogen Disgust and Fear of Sin") +
  xlab("Pathogen Disgust") +
  ylab("FOS") +
  theme_minimal()
```

Warning: Removed 7 rows containing missing values or values outside the scale range (`geom\_point()`).





3. Independence of Residuals: Satisfied via the Durban-Watson tests below.

```
# Fitting the linear models
PIOS_model <- lm(data = data, z_PIOS ~ z_PD)
FOG_model <- lm(data = data, z_FOG ~ z_PD)
FOS_model <- lm(data = data, z_FOS ~ z_PD)

# Running the Durban-Watson test below
dwtest(PIOS_model)
```

Durbin-Watson test

```
data: PIOS_model
DW = 1.9595, p-value = 0.4078
alternative hypothesis: true autocorrelation is greater than 0
```

```
dwtest(FOG_model)
```

Durbin-Watson test

```
data: FOG_model
DW = 2.0273, p-value = 0.5624
alternative hypothesis: true autocorrelation is greater than 0
```

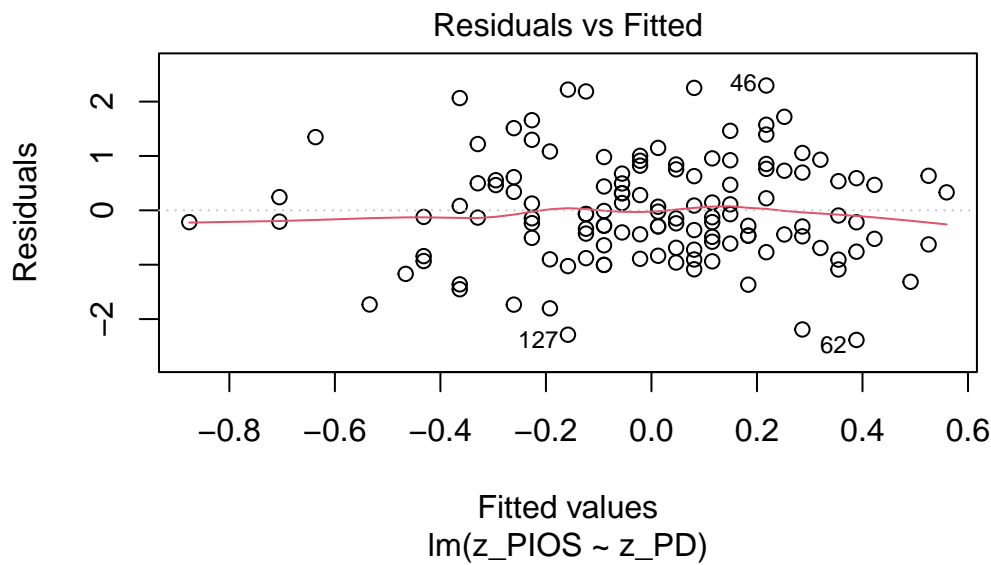
```
dwtest(FOG_model)
```

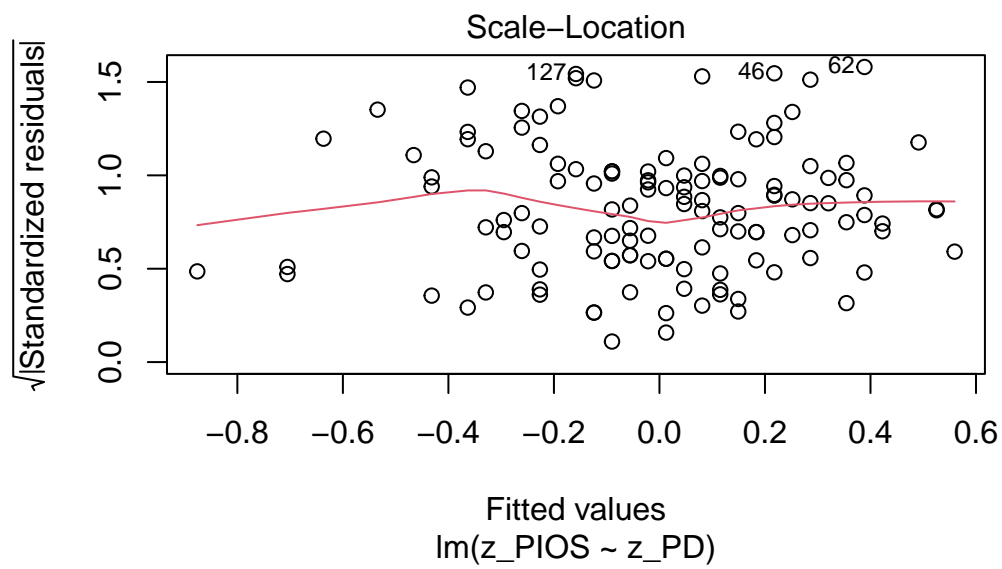
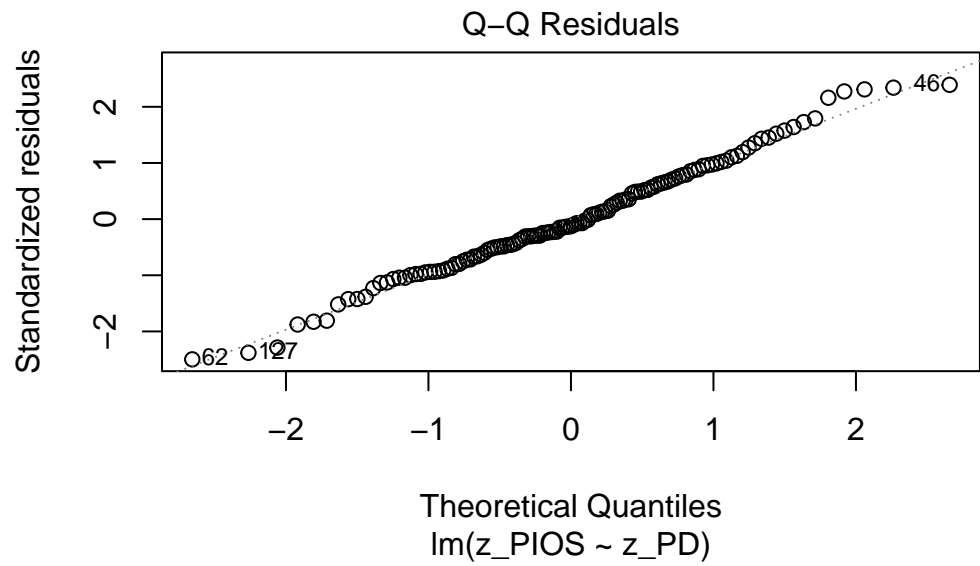
Durbin-Watson test

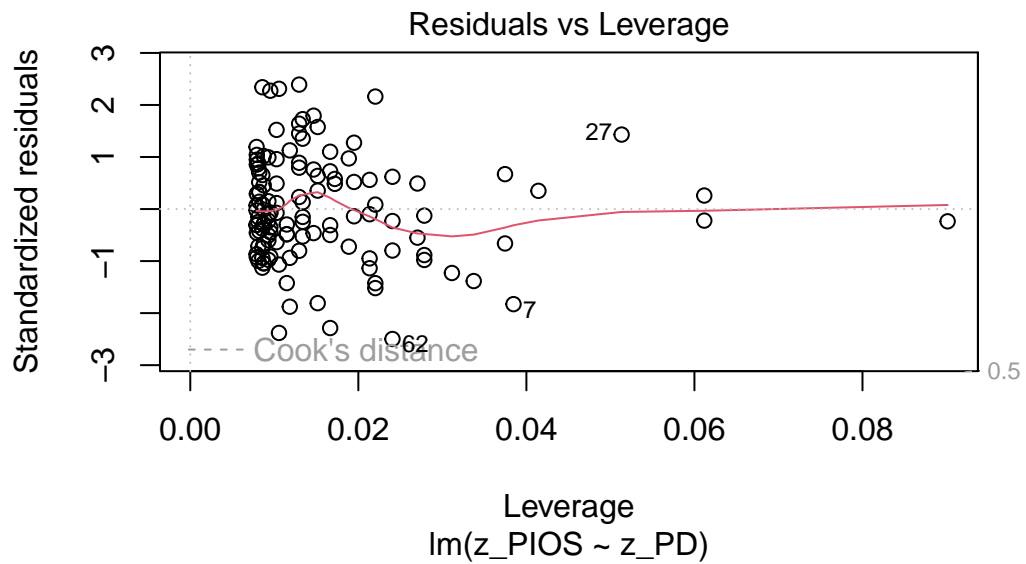
```
data: FOS_model
DW = 2.0401, p-value = 0.5891
alternative hypothesis: true autocorrelation is greater than 0
```

- The Durbin-Watson tests indicates that the autocorrelation for the models is not significantly greater than zero.
4. Equality of Variance: Satisfied via the plot of residual vs. fitted values and the Breusch-Pagan test below.

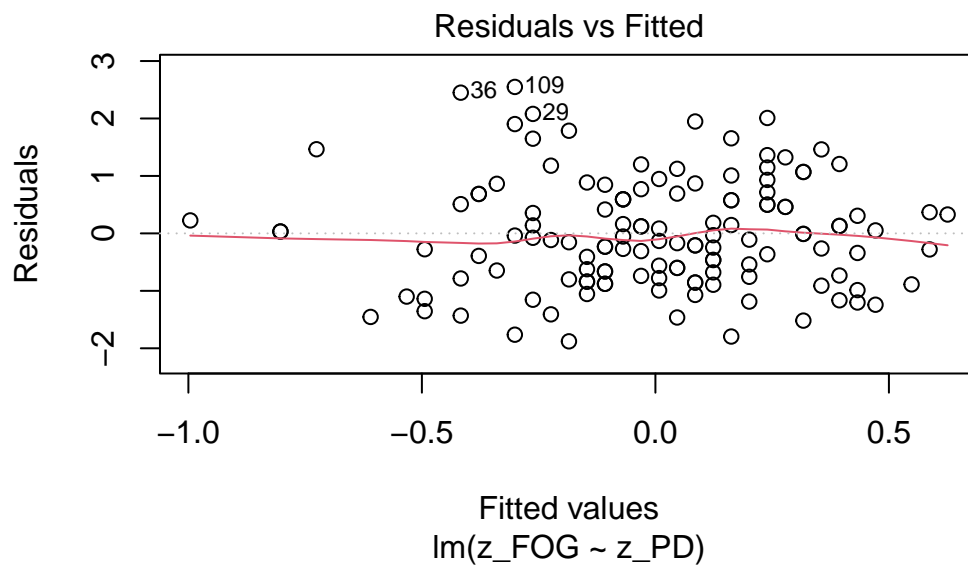
```
# Plotting residual vs. fitted values
plot(PIOS_model)
```

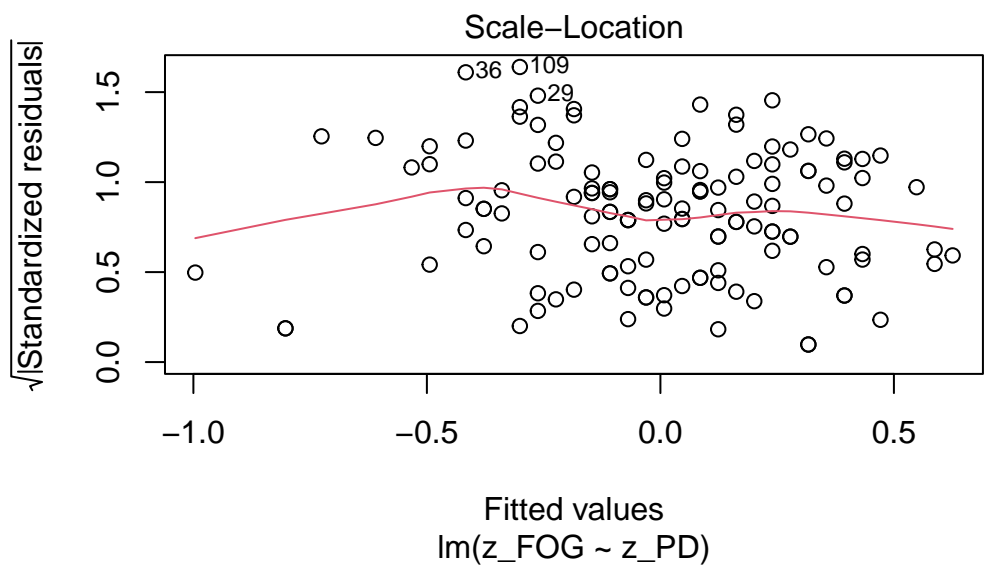
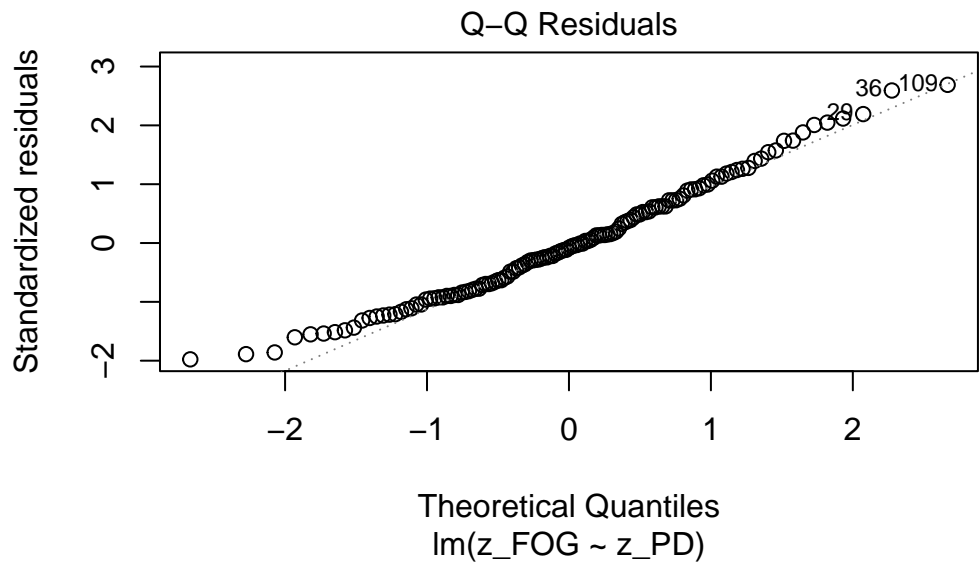


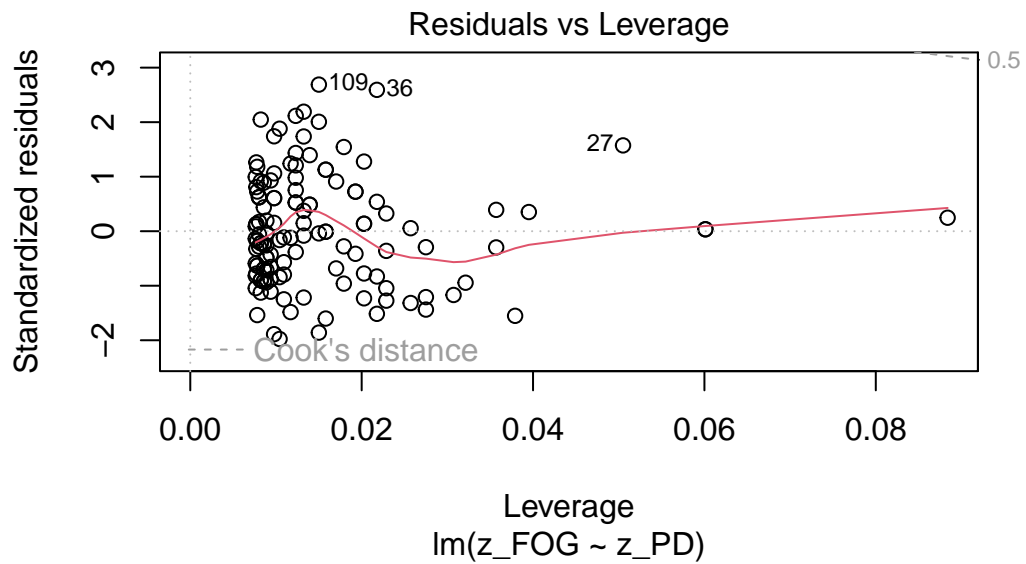




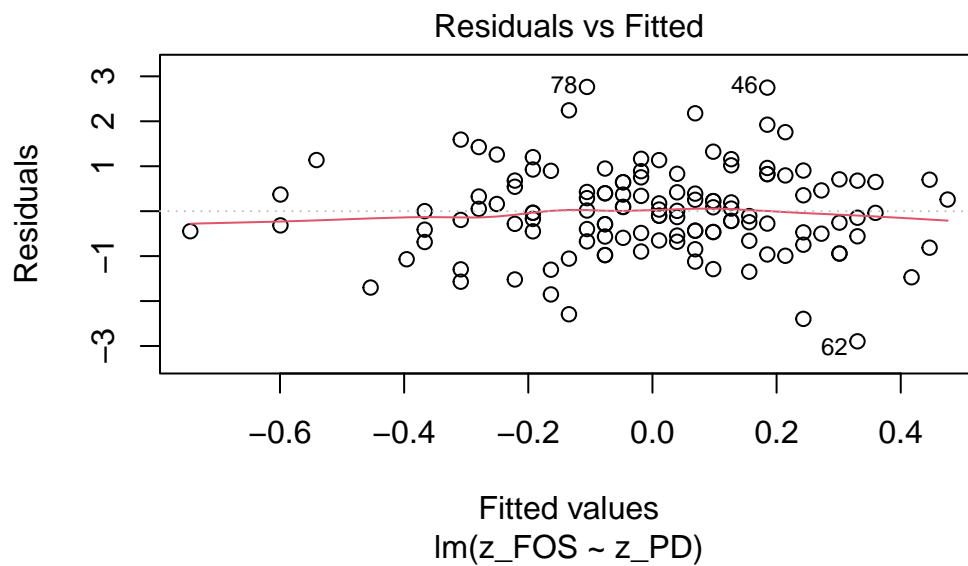
```
plot(FOG_model)
```

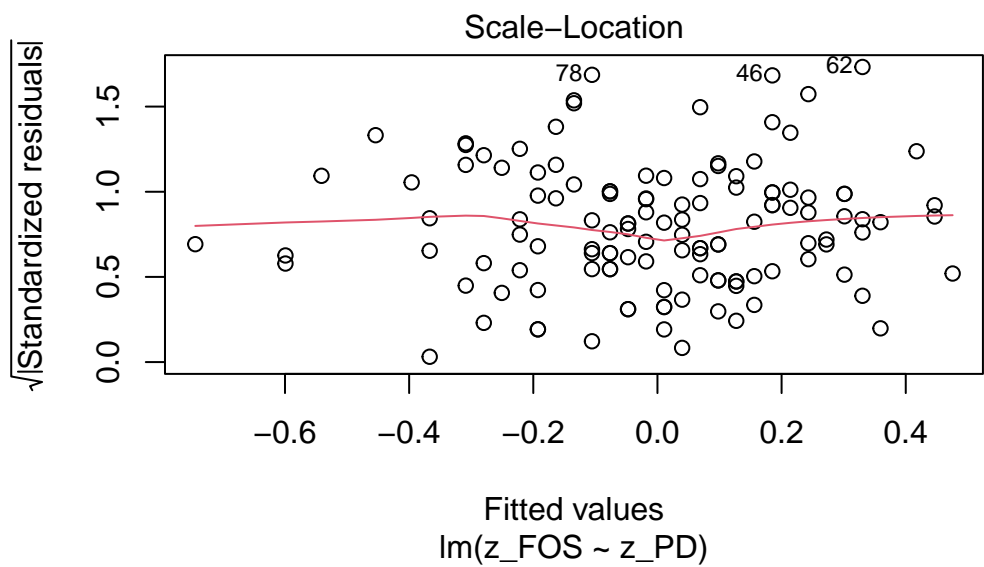
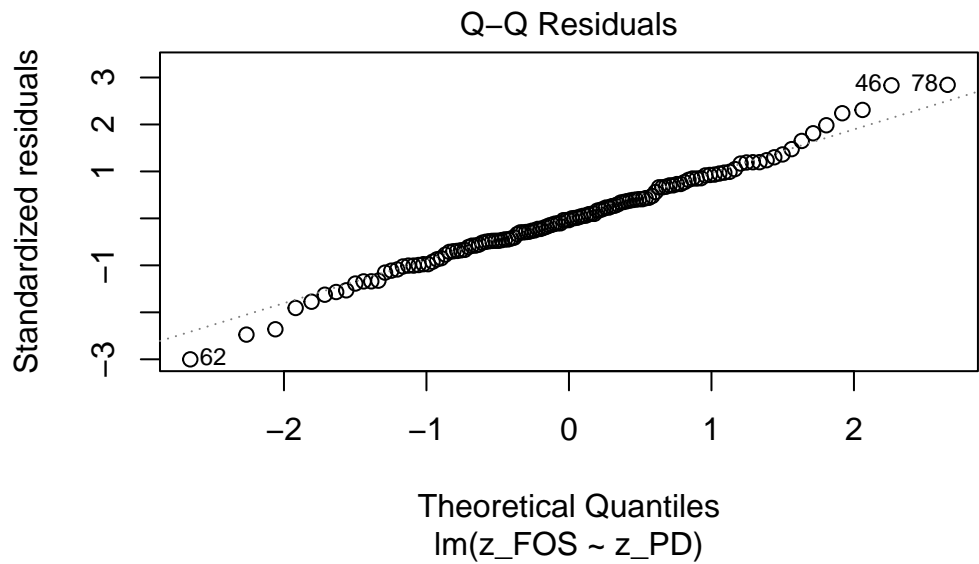


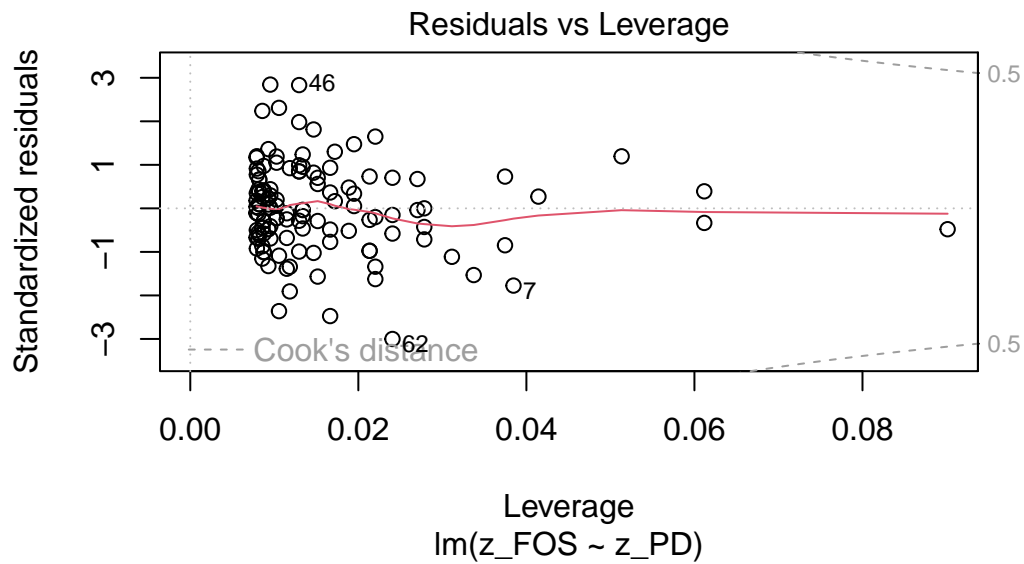




```
plot(FOS_model)
```







```
# Breusch-Pagan test
bptest(PIOS_model)
```

studentized Breusch-Pagan test

```
data: PIOS_model
BP = 0.17742, df = 1, p-value = 0.6736
```

```
bptest(FOG_model)
```

studentized Breusch-Pagan test

```
data: FOG_model
BP = 3.0277, df = 1, p-value = 0.08185
```

```
bptest(FOS_model)
```



studentized Breusch-Pagan test

data: FOS\_model

BP = 0.15368, df = 1, p-value = 0.695

- The assumption of equality of variance is also not violated for any of the models here, given the results of the Breusch-Pagan tests.

5. Residuals Normally Distributed: Supported by the Q-Q plots of residuals above.

### Testing the Models:

To test the model, I will use robust standard errors from the `sandwich` package.

```
# Summarizing the model
summary(PIOS_model)
```

Call:

```
lm(formula = z_PIOS ~ z_PD, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.3836	-0.6348	-0.1207	0.6322	2.2956

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.00571	0.08575	0.067	0.94701
z_PD	0.27006	0.08543	3.161	0.00197 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9661 on 125 degrees of freedom

(7 observations deleted due to missingness)

Multiple R-squared: 0.07403, Adjusted R-squared: 0.06662

F-statistic: 9.994 on 1 and 125 DF, p-value: 0.001971

```
summary(FOG_model)
```

Call:

```
lm(formula = z_FOG ~ z_PD, data = data)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-1.87897	-0.74769	-0.07674	0.59269	2.54936

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.0001197	0.0834616	0.001	0.998858
z_PD	0.3051016	0.0831553	3.669	0.000355 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9553 on 129 degrees of freedom

(3 observations deleted due to missingness)

Multiple R-squared: 0.0945, Adjusted R-squared: 0.08748

F-statistic: 13.46 on 1 and 129 DF, p-value: 0.0003549

```
summary(FOS_model)
```

Call:

```
lm(formula = z_FOS ~ z_PD, data = data)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-2.89624	-0.56164	-0.03585	0.64681	2.76451

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.004853	0.086694	0.056	0.9554
z_PD	0.229535	0.086369	2.658	0.0089 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9768 on 125 degrees of freedom

(7 observations deleted due to missingness)

Multiple R-squared: 0.05348, Adjusted R-squared: 0.04591

F-statistic: 7.063 on 1 and 125 DF, p-value: 0.008897

- Pathogen disgust is a significant predictor in each of the models. The standardized beta value is highest for the model predicting Fear of God ( $b = .31$ ), with fear of sin having the lowest beta value ( $b = .23$ ). The beta value for the model with PIOS as the dependent variable is  $b = .27$ .

Given that our sample is selective of already religious individuals, we might expect that the relationship between disgust and religiosity is higher for the general population.

## Brief Discussion

When we originally did this analysis for class, we concluded that there is partial support for Hypothesis 1. Give our problems with validation of the RBI subscales, I do not believe that we have provided evidence for the hypotheses here.

Although after removing a few items we ended up with scales with respectable internal consistency, there is not strong evidence of validity of the subscales either by correlations with the PIOS—which were quite small—or by correlations with in-group and out-group attitudes towards other religious groups—which were the wrong direction, very small, and not statistically reliable, although the small sample for this analysis made a significant relationship unlikely. Aside from validity problems, the IIGE showed problems with range restriction, which also could prevent it from adequately capturing the construct we are interested in. There were many responses with only the highest response option, causing an accumulation of scores at the top end of the final scale. Ultimately, I believe this indicates that we should abandon the RBI scales.

Given these problems with the measures and with regression NHST assumptions, the relationships of the relationship between parasite disgust and the RBI subscales is difficult to interpret. However, even if we assumed strong psychometric properties and robust NHSTs, the relationships that we found were either small (for the IIGE) or null (for the DOGS). This does not provide support for our hypotheses.

It will hopefully be easier to measure religiosity in a more general sample, because we can use measures of religious affiliation, as well as other measures that work better for a more general sample. Collecting a more general sample will also allow us to generalize our conclusions more confidently.

Lastly, our estimation of the relationship between disgust and religiosity suffers from the fact that the sample here is selected on the basis of religiosity, but to get a sense of how disgust sensitivity relates to the degree of religiosity the PIOS serves this purpose well enough. As a measure of scrupulosity originally for clinical contexts with OCD patients, it can provide us with a sense of the relationship between disgust and obsessive religiosity. The moderate relationships here suggest that the relationship between disgust and religiosity may be moderate or perhaps large in more general samples, given that there will be more religious variability to account for in the general sample.

Abramowitz, J. S., Huppert, J. D., Cohen, A. B., Tolin, D. F., & Cahill, S. P. (2002). Religious obsessions and compulsions in a non-clinical sample: the Penn Inventory of Scrupulosity (PIOS). *Behaviour Research and Therapy*, 40(7), 825–838. [https://doi.org/10.1016/s0005-7967\(01\)00070-5](https://doi.org/10.1016/s0005-7967(01)00070-5)

- Bogardus, E. S. (1933). A social distance scale. *Sociology & Social Research*, 17, 265–271.
- Brinkerhoff, M. B., & Mackie, M. M. (1986). The Applicability of Social Distance for Religious Research: An Exploration. *Review of Religious Research*, 28(2), 151–167. <https://doi.org/10.2307/3511469>
- Fincher, C. L., & Thornhill, R. (2012). Parasite-stress promotes in-group assortative sociality: the cases of strong family ties and heightened religiosity. *The Behavioral and Brain Sciences*, 35(2), 61–79. <https://doi.org/10.1017/S0140525X11000021>
- Mather, D. M., Jones, S. W., & Moats, S. (2017). Improving upon Bogardus: Creating a More Sensitive and Dynamic Social Distance Scale. *Survey Practice*, 10(4). <https://doi.org/10.29115/SP-2017-0026>
- Olatunji, B. O., Abramowitz, J. S., Williams, N. L., Connolly, K. M., & Lohr, J. M. (2007). Scrupulosity and obsessive-compulsive symptoms: confirmatory factor analysis and validity of the Penn Inventory of Scrupulosity. *Journal of Anxiety Disorders*, 21(6), 771–787. <https://doi.org/10.1016/j.janxdis.2006.12.002>
- Olatunji, B., Adams, T., Ciesielski, B., David, B., Sarawgi, S., & Broman-Fulks, J. (2012). The Three Domains of Disgust Scale. *Assessment*, 19, 205–225. <https://doi.org/10.1177/1073191111432881>
- Simonsohn, U., Gruson [ctb, H., & aut. (2024). *groundhog: Version-Control for CRAN, GitHub, and GitLab Packages*. <https://cran.r-project.org/web/packages/groundhog/index.html>
- Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality: individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology*, 97(1), 103–122. <https://doi.org/10.1037/a0015474>