GMS6025C: Final Project 1 Part 3 Modeling

Emely Gazarov, Ryo Iwata, Maria Ramirez

05 May, 2024

About the data

Name: Addiction Patient Assessment Measures

Description: Data are gathered from a residential treatment facility for substance use disorders. Data include item-level responses for all questionnaires listed below. Data are available at initiation of treatment, after one month of treatment, and at treatment discharge.

Overview of variables in the data

• Variables that end in .BL, .FU, .DC are indicating that the score was collected at baseline before treatment, 30 days following up after baseline, and during discharge respectively.

Independent variables

Basic Demographics

- age
 - age
- gender
 - gender
- education
 - education
- number of days sober
 - sober days
- If a patient dropped out of the program or not
 - dropout_yn

Substance use disorders

- Flag of whether or not the patient is being treated for substance in name
 - SUD_is_Alchohol
 - SUD_is_Opioid
 - SUD is Cannabis
 - SUD_is_depressants_anxiolytic
 - SUD is Cocaine
 - SUD_is_Other_stimulant
 - SUD_is_Hallucinogen
 - SUD_is_Nicotine
 - SUD_is_Inhalant
 - SUD_is_psychoactive
- Number of substances being treated for
 - Total number of all substances that patient is being treated for

- * SUD.sum
- Total number of legal substances that patient is being treated for. (SUD_is_Alchohol, SUD is Cannabis, SUD is Nicotine, SUD is Inhalant)
 - * SUD.sum_legal
- Total number of legal substances that patient is being treated for. (SUD_is_Opioid, SUD_is_depressants_anxiolytic, SUD_is_Cocaine, SUD_is_Other_stimulant, SUD_is_psychoactive, SUD_is_Hallucinogen)
 - * SUD.sum illegal
- Flag of whether or not the patient is being treated for legal or illegals substances
 - If patient is being treated for any of the following: SUD_is_Alchohol, SUD_is_Cannabis, SUD_is_Nicotine, SUD_is_Inhalant
 - * SUD.uses_legal
 - If patient is being treated for any of the following: SUD_is_Opioid, SUD_is_depressants_anxiolytic,
 SUD_is_Cocaine, SUD_is_Other_stimulant, SUD_is_psychoactive, SUD_is_Hallucinogen
 * SUD.uses_illegal

Beliefs about AA/NA (TSPEQ)

- Categories based on number of AA/NA meetings attended throughout life
 - aana_life
- Categories based on number of AA/NA meetings attended last year
 - aana_pastyear
- Total score of questions gauging positive opinions of AA/NA
 - aana_positive
- Total score of questions gauging negative opinions of AA/NA
 - aana_negative

Childhood Experiences (ACE)

- Binary flags of whether the specified type of negative experience happened during childhood of patient
 - childhood.verbal abuse
 - childhood.physical_abuse
 - childhood.sexual_abuse
 - childhood.alone
 - childhood.neglected
 - childhood.divorced
 - childhood.parent_was_abused
 - childhood.other_was_addicted
 - childhood.other_was_stressed
 - childhood.other_was_prisoned
- Sum of number of types of negative experiences that happened during patient's childhood
 - childhood_sum

Social Support (MSPSS)

- Total score for questions that measure the perceived level of social support the respondent receives from each of the three subcategories:
 - family
 - * social.family
 - friends
 - * social.friends
 - significant other
 - * social.sig_other
 - Sum of family, friends, and significant other
 - * social

Spiritual Experiences (Brief R-COPE)

- Total score of questions asking the frequency of spiritual coping experience, ranging from 0) Not at all to 3) Nearly every day.
 - Positive coping
 - * religion_pos
 - Negative coping
 - * religion_neg
- Religion
 - Binary flag of if a patient practices religion that is in the name of the variable
 - * rel.is_Atheist
 - * rel.is_Agnostic
 - * rel.is Protestant
 - * rel.is_Catholic
 - * rel.is_Muslim
 - * rel.is_Jewish
 - * rel.is_Hindu
 - * rel.is Buddhist
 - * rel.is_Baptist
 - * rel.is_No_affiliation
 - * rel.is_Non_denominational_Christian
 - * rel.is_Other
 - Binary flag of if patient is religious or not
 - * rel.is_religious
 - * rel.is_not_religious

Stressful Life Experiences (LEC-5)

- Binary flags of whether a listed stressful life event was either 1) experienced the life event personally, 2) witnessed it happen to someone else, 3) learned about it happening from a close friend or family member, 4) experienced the event due to their job.
 - Natural disaster
 - * stress.natural_disaster.to_subj
 - * stress.natural disaster.witnessed
 - * stress.natural_disaster.learned
 - * stress.natural_disaster.exposed
 - Fire or explosion
 - * stress.fire.to_subj
 - * stress.fire.witnessed
 - * stress.fire.learned
 - * stress.fire.exposed
 - Transportation accident
 - * stress.transportation_accident.to_subj
 - * stress.transportation_accident.witnessed
 - * stress.transportation_accident.learned
 - \ast stress.transportation_accident.exposed
 - Non-transportation accident
 - $* \ \mathtt{stress.serious_accident.to_subj}$
 - * stress.serious_accident.witnessed
 - * stress.serious_accident.learned
 - * stress.serious_accident.exposed
 - Exposure to toxic substance
 - * stress.toxic.to_subj
 - * stress.toxic.witnessed
 - * stress.toxic.learned

- * stress.toxic.exposed
- Physically assaulted
 - * stress.physical_assault.to_subj
 - * stress.physical_assault.witnessed
 - * stress.physical_assault.learned
 - * stress.physical_assault.exposed
- Assaulted with a weapon
 - * stress.weapon_assault.to_subj
 - * stress.weapon_assault.witnessed
 - * stress.weapon_assault.learned
 - * stress.weapon_assault.exposed
- Sexual assault
 - * stress.sexual_assault.to_subj
 - * stress.sexual_assault.witnessed
 - * stress.sexual_assault.learned
 - * stress.sexual_assault.exposed
- Unwanted or uncomfortable sexual experience.
 - * stress.unwanted_sexual.to_subj
 - * stress.unwanted_sexual.witnessed
 - * stress.unwanted_sexual.learned
 - * stress.unwanted_sexual.exposed
- Combat or exposure to war-zone
 - * stress.combat.to_subj
 - * stress.combat.witnessed
 - * stress.combat.learned
 - * stress.combat.exposed
- Captivity
 - * stress.captivity.to_subj
 - * stress.captivity.witnessed
 - * stress.captivity.learned
 - * stress.captivity.exposed
- Life-threatening illness or injury.
 - * stress.illness.to_subj
 - * stress.illness.witnessed
 - * stress.illness.learned
 - * stress.illness.exposed
- Severe human suffering.
 - * stress.severe_suffering.to_subj
 - * stress.severe_suffering.witnessed
 - * stress.severe_suffering.learned
 - * stress.severe suffering.exposed
- Sudden violent death (for example, homicide, suicide).
 - * stress.sudden_violent_death.to_subj
 - $* \ \mathtt{stress.sudden_violent_death.witnessed}$
 - $* \ \mathtt{stress.sudden_violent_death.learned}$
 - * stress.sudden_violent_death.exposed
- Sudden accidental death.
 - * stress.sudden_accidental_death.to_subj
 - * stress.sudden_accidental_death.witnessed
 - * stress.sudden_accidental_death.learned
 - * stress.sudden_accidental_death.exposed
- Serious injury, harm, or death you caused to someone else.
 - * stress.harm_to_others.to_subj
 - * stress.harm_to_others.witnessed

```
* stress.harm_to_others.learned
```

- * stress.harm_to_others.exposed
- Any other very stressful event or experience.
 - * stress.other.to_subj
 - * stress.other.witnessed
 - * stress.other.learned
 - * stress.other.exposed
- Total score of number of catgories that happened to the patient personally and/or witnessed happened to someone else
 - * stress_to_subj
 - * stress_to_subj_and_wit
 - * stress_wit

Substance use history

- Binary flag of whether or not patient has used the following:
 - Tobacco
 - * history.tobacco
 - Alcohol
 - * history.alcohol
 - Other
 - * history.other

Dependent variables

Length of stay in treatment

- number of days in treatment
 - treatment_days

Alcohol/Drug Craving (PACS)

- Total score that assesses the alcohol/drug cravings the patient has experienced within the past week based on a variety of craving characteristics.
 - craving.BL
 - craving.FU
 - craving.DC

Quality of Life (WHOQOL-BREF)

- Total scores of assessment that measures different aspects of quality of life
 - general
 - * qol_general_health.BL
 - * qol_general_health.FU
 - * qol_general_health.DC
 - physical
 - * qol_physical.BL
 - * qol_physical.FU
 - * qol_physical.DC
 - psychological health
 - * qol_psych.BL
 - * qol_psych.FU
 - * qol_psych.DC
 - social relationships
 - * qol_social.BL

```
* qol_social.FU
  * qol_social.DC
- environment
  * qol_env.BL
  * qol_env.FU
  * qol_env.DC
- Total scores
  * qol.BL
  * qol.FU
  * qol.DC
- Recalculating quality of life scores without social component
  * no_social_qol.BL
```

Commitment to Sobriety (CSS-5)

```
• Total score of questions gauging patient's commitment to abstinence from alcohol and drug use
```

```
- commit.BL
```

- commit.DC
- commit.FU
- The difference of commitment to sobriety score between discharge and baseline
 - commit_change

Alcohol/Drug Abstinence Experiences (DAASE)

• Total score of questions that assesses patient's self-confidence in their ability to remain abstinent from alcohol in various situations.

```
    Negative Affect

           * abstain_neg.BL
           * abstain_neg.FU
           * abstain_neg.DC

    Social/Positive

           * abstain_pos.BL
           * abstain_pos.FU
           * abstain_pos.DC
       - Physical ailments
           * abstain_phy.BL
           * abstain_phy.FU
           * abstain_phy.DC
       - Situations associated with cravings
           * abstain_crv.BL
           * abstain_crv.FU
           * abstain_crv.DC
       - All previously mentioned conditions
           * abstain_total.BL
           * abstain_total.FU
           * abstain_total.DC
cleaned_addiction_data = read_rds("./addiction.rds")
```

First hypothesis:

```
+ rel.is_religious
+ religion_pos
+ aana_past_year
+ SUD.sum_illegal
+ SUD.sum_legal
,cleaned_addiction_data)
```

```
\begin{split} \hat{y} &= \beta_0 + \beta_{Age}Age \\ &+ \beta_{Male}Male? \\ &+ \beta_{EduAA}EduAA? + \beta_{EduBS}EduBS? + \beta_{EduMS}EduMS? + \beta_{EduPhD}EduPhD? \\ &+ \beta_{Religious}Religious? + \beta_{PositiveReligion}PositiveReligion \\ &+ \beta_{AANAParticipation}AANAParticipation \\ &+ \beta_{NumIllegal}NumIllegal + \beta_{NumIllegal}NumIllegal \end{split}
```

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

= 0.5 + 0.2 - 0.3
= 0.4

summary(commitment_model)

```
##
## Call:
  lm(formula = commit_change ~ age + gender + education + rel.is_religious +
##
       religion_pos + aana_past_year + SUD.sum_illegal + SUD.sum_legal,
##
       data = cleaned_addiction_data)
##
## Residuals:
                1Q Median
                                3Q
       Min
                                        Max
##
  -25.931 -1.577 -0.658
                             0.664
                                    24.143
## Coefficients:
##
                                      Estimate Std. Error t value Pr(>|t|)
                                                  0.74149
                                                             1.08
                                                                      0.282
## (Intercept)
                                       0.79804
                                                            -0.27
                                                                      0.790
## age
                                      -0.00323
                                                  0.01212
## gendermen
                                                             0.82
                                                                     0.411
                                       0.26263
                                                  0.31894
                                                            -0.91
## educationAssociate's/Some College -0.39753
                                                  0.43830
                                                                     0.365
## educationBachelor's
                                      -0.88462
                                                  0.41071
                                                            -2.15
                                                                     0.032 *
## educationMaster's
                                      -0.53454
                                                  0.63980
                                                            -0.84
                                                                     0.404
                                                            -0.92
                                                                     0.357
## educationDoctoral
                                      -0.51251
                                                  0.55629
## rel.is_religiousTRUE
                                                  0.37015
                                                             1.83
                                                                     0.067 .
                                       0.67802
## religion_pos
                                      -0.04861
                                                  0.02377
                                                            -2.04
                                                                     0.041 *
                                                            -0.44
                                                                     0.660
## aana_past_year0.25
                                      -0.14234
                                                  0.32310
## aana_past_year0.5
                                       0.26548
                                                  0.65151
                                                             0.41
                                                                     0.684
## aana_past_year0.75
                                      -0.60564
                                                  0.75880
                                                            -0.80
                                                                     0.425
## aana_past_year1
                                      -1.06685
                                                  0.80619
                                                            -1.32
                                                                     0.186
                                      0.19176
                                                                     0.174
## SUD.sum_illegal
                                                  0.14102
                                                             1.36
## SUD.sum_legal
                                       0.52078
                                                  0.20610
                                                             2.53
                                                                      0.012 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
```

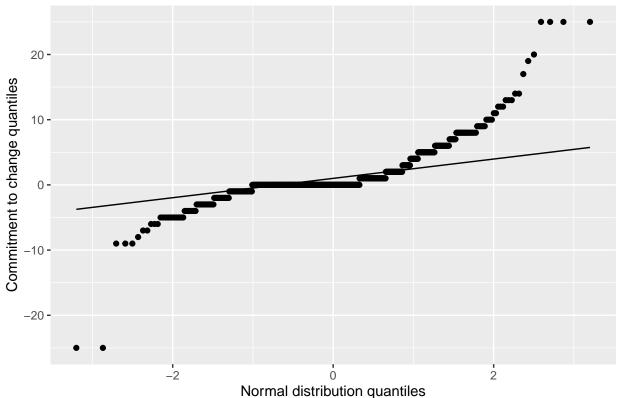
```
## Residual standard error: 3.93 on 715 degrees of freedom
## Multiple R-squared: 0.0358, Adjusted R-squared: 0.0169
## F-statistic: 1.89 on 14 and 715 DF, p-value: 0.024
```

 \bullet Hypothesis: Individuals with higher involvement in AA/NA activities will show a greater increase in their CSS-5 relative change scores.

Looking to see if difference of commitment to sobriety from baseline to discharge is normally distributed

```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(sample=commit_change)) +
    stat_qq() +
    stat_qq_line() +
    labs(
        title = "Q-Q plot of difference of commitment to sobriety at baseline to discharge",
        y = "Commitment to change quantiles",
        x = "Normal distribution quantiles")
```

Q-Q plot of difference of commitment to sobriety at baseline to discharge



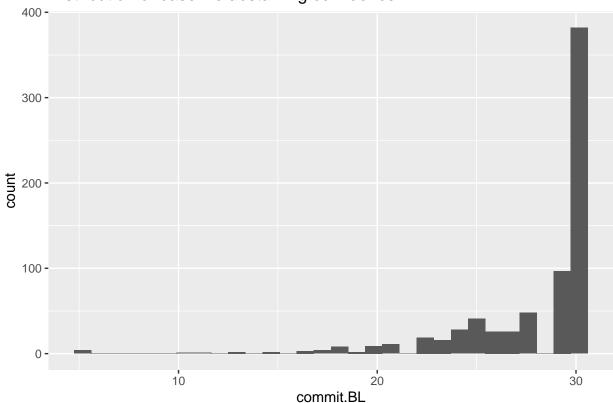
Looking to see if difference of commitment to sobriety from baseline to discharge is normally distributed

```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(commit.BL)) +
```

```
geom_histogram() +
labs(
   title = "Distribution of baseline abstaining confidence")
```

$\operatorname{stat_bin}()$ using $\operatorname{bins} = 30$. Pick better value with $\operatorname{binwidth}$.

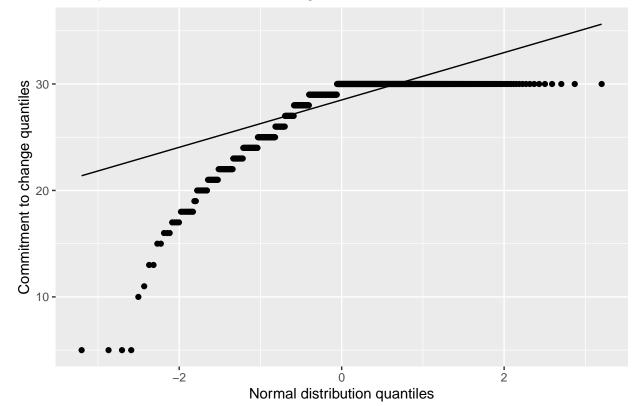
Distribution of baseline abstaining confidence



Looking to see if difference of commitment to sobriety from baseline to discharge is normally distributed

```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(sample=commit.BL)) +
    stat_qq() +
    stat_qq_line() +
    labs(
        title = "Q-Q plot of commitment to change at baseline",
        y = "Commitment to change quantiles",
        x = "Normal distribution quantiles")
```

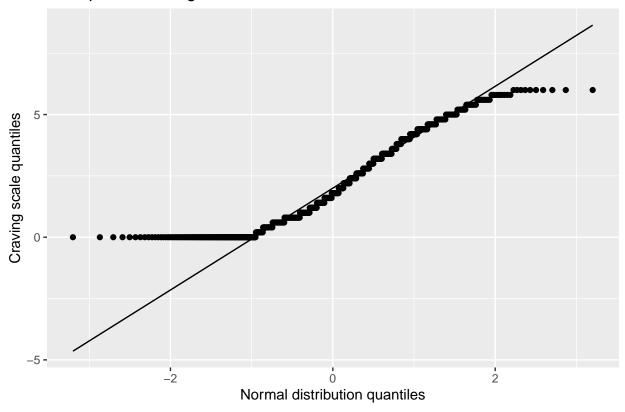
Q-Q plot of commitment to change at baseline



Looking for other normally distributed things to predict

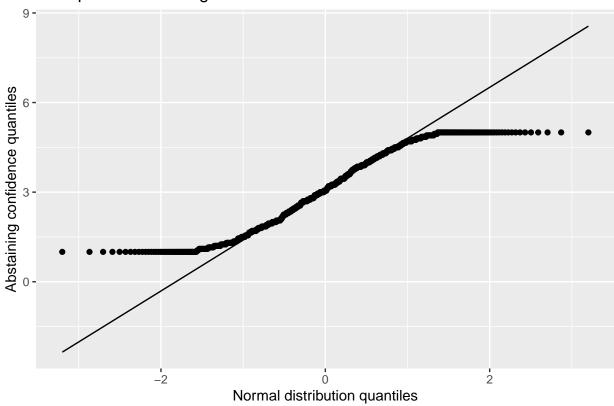
```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(sample=craving.BL)) +
    stat_qq() +
    stat_qq_line() +
    labs(
        title = "Q-Q plot of craving scale at baseline",
        y = "Craving scale quantiles",
        x = "Normal distribution quantiles")
```

Q-Q plot of craving scale at baseline



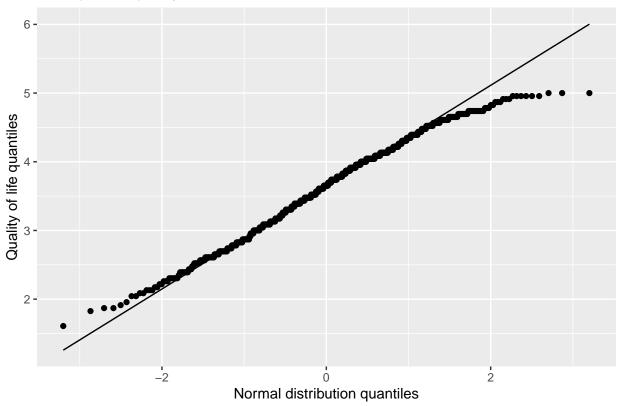
```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(sample=abstain_total.BL)) +
    stat_qq() +
    stat_qq_line() +
    labs(
        title = "Q-Q plot of abstaining confidence at baseline",
        y = "Abstaining confidence quantiles",
        x = "Normal distribution quantiles")
```

Q-Q plot of abstaining confidence at baseline



```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(sample=qol.BL)) +
    stat_qq() +
    stat_qq_line() +
    labs(
        title = "Q-Q plot of quality of life at baseline",
        y = "Quality of life quantiles",
        x = "Normal distribution quantiles")
```

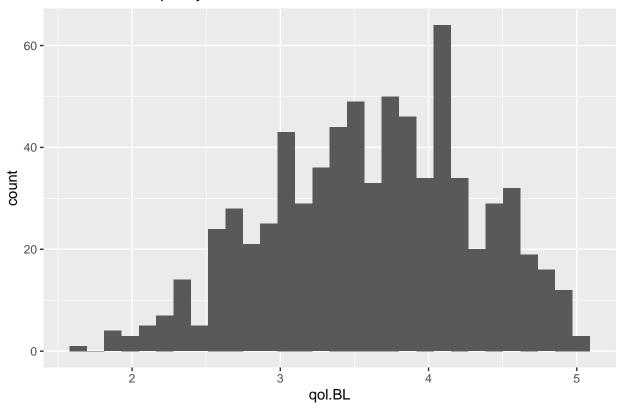
Q-Q plot of quality of life at baseline



```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(qol.BL)) +
    geom_histogram() +
    labs(
        title = "Distribution of quality of life")
```

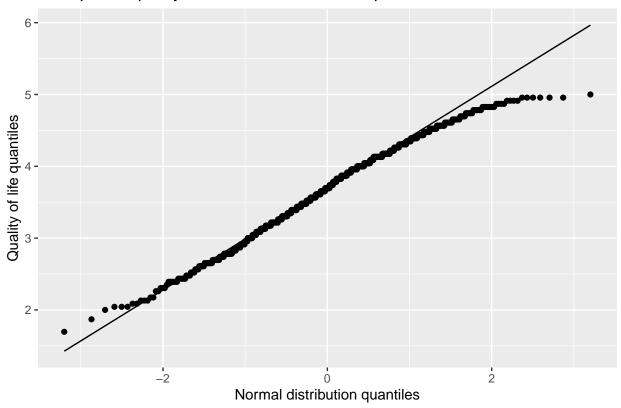
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

Distribution of quality of life



```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(sample=no_social_qol.BL)) +
    stat_qq() +
    stat_qq_line() +
    labs(
        title = "Q-Q plot of quality of life without social component at baseline",
        y = "Quality of life quantiles",
        x = "Normal distribution quantiles")
```

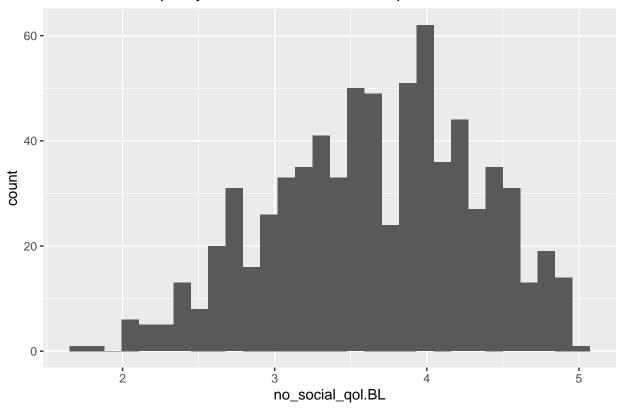
Q-Q plot of quality of life without social component at baseline



```
# and now the Q-Q plots
cleaned_addiction_data |>
    ggplot(aes(no_social_qol.BL)) +
    geom_histogram() +
    labs(
        title = "Distribution of quality of life without social component")
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

Distribution of quality of life without social component



Second hypothesis

summary(qol_model)

```
##
## Call:
## lm(formula = no_social_qol.BL ~ age + gender + education + SUD.sum_illegal +
## SUD.sum_legal + rel.is_religious + religion_pos + aana_past_year +
## aana_positive + social + stress_to_subj, data = cleaned_addiction_data)
```

```
##
## Residuals:
##
       Min
                1Q Median
                                       Max
## -1.7143 -0.3570 0.0526 0.3822
                                    1.5135
## Coefficients:
                                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)
                                                           15.15
                                                                    < 2e-16 ***
                                      2.83343
                                                 0.18707
## age
                                      0.00244
                                                 0.00175
                                                             1.40
                                                                     0.1630
                                                                     0.0781
                                                             1.76
## gendermen
                                      0.08085
                                                 0.04582
## educationAssociate's/Some College -0.07360
                                                 0.06284
                                                           -1.17
                                                                     0.2419
## educationBachelor's
                                                 0.06000
                                                             3.11
                                                                     0.0019 **
                                      0.18656
## educationMaster's
                                     -0.10187
                                                 0.09169
                                                           -1.11
                                                                     0.2669
## educationDoctoral
                                                             0.96
                                                                     0.3384
                                      0.07630
                                                 0.07965
## SUD.sum_illegal
                                                 0.02017
                                                           -2.00
                                                                     0.0462 *
                                     -0.04030
## SUD.sum_legal
                                     -0.01887
                                                 0.02949
                                                            -0.64
                                                                     0.5225
                                                 0.05299
                                                           -0.82
                                                                     0.4124
## rel.is_religiousTRUE
                                     -0.04346
## religion pos
                                      0.01671
                                                 0.00347
                                                            4.81 0.0000018 ***
## aana_past_year0.25
                                                 0.04746
                                                           -0.54
                                                                     0.5911
                                     -0.02551
## aana_past_year0.5
                                     -0.01761
                                                 0.09531
                                                           -0.18
                                                                     0.8534
## aana_past_year0.75
                                      0.12654
                                                 0.11281
                                                            1.12
                                                                     0.2623
## aana_past_year1
                                      0.04958
                                                 0.11818
                                                             0.42
                                                                     0.6750
## aana_positive
                                     -0.01472
                                                 0.00375
                                                           -3.92 0.0000971 ***
## social
                                      0.19089
                                                 0.01853
                                                            10.30
                                                                    < 2e-16 ***
                                                           -4.48 0.0000089 ***
## stress_to_subj
                                     -0.03988
                                                 0.00891
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.561 on 712 degrees of freedom
## Multiple R-squared: 0.304, Adjusted R-squared: 0.287
## F-statistic: 18.3 on 17 and 712 DF, p-value: <2e-16
```

```
Anova(qol_model, type=2, digits=4) |>
tidy() |>
kable()
```

Warning in printHypothesis(L, rhs, names(b)): one or more coefficients in the hypothesis include
arithmetic operators in their names;

the printed representation of the hypothesis will be omitted

term	sumsq	df	statistic	p.value
age	0.61383	1	1.95010	0.16301
gender	0.97971	1	3.11249	0.07812
education	7.80550	4	6.19940	0.00007
SUD.sum_illegal	1.25583	1	3.98969	0.04616
SUD.sum_legal	0.12888	1	0.40945	0.52245
rel.is_religious	0.21174	1	0.67269	0.41239
religion_pos	7.29317	1	23.16997	0.00000
aana_past_year	0.67150	4	0.53333	0.71129
aana_positive	4.83695	1	15.36669	0.00010
social	33.41702	1	106.16390	0.00000

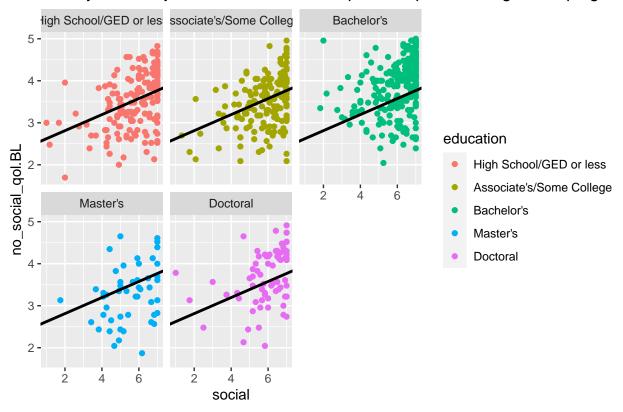
term	sumsq	df	statistic	p.value
stress_to_subj	6.30376	1	20.02669	0.00001
Residuals	224.11497	712	NA	NA

Plotting the adjusted data

```
tidy_qol_model <- tidy(qol_model)</pre>
# Getting the slope and intercept for perceived social support
coef_social <-</pre>
  pull(filter(tidy_qol_model,
              term == "social"),
       estimate)
intercept <- pull(filter(tidy gol model, term == "(Intercept)"), estimate)</pre>
# Getting the slope of the other variables
stress_to_subj_slope <- pull(filter(tidy_qol_model, term == "stress_to_subj"), estimate)</pre>
aana_positive_slope <- pull(filter(tidy_qol_model, term == "aana_positive"), estimate)</pre>
religion_pos_slope <- pull(filter(tidy_qol_model, term == "religion_pos"), estimate)
# Getting the mean of other variables
stress_to_subj_mean <- cleaned_addiction_data |>
  summarize(Mean = mean(stress_to_subj)) |>
  pull(Mean)
aana_positive_mean <- cleaned_addiction_data |>
  summarize(Mean = mean(aana_positive)) |>
  pull(Mean)
religion_pos_mean <- cleaned_addiction_data |>
  summarize(Mean = mean(religion pos)) |>
  pull(Mean)
```

- Based on the following formula:
- We are adjusting for other things.

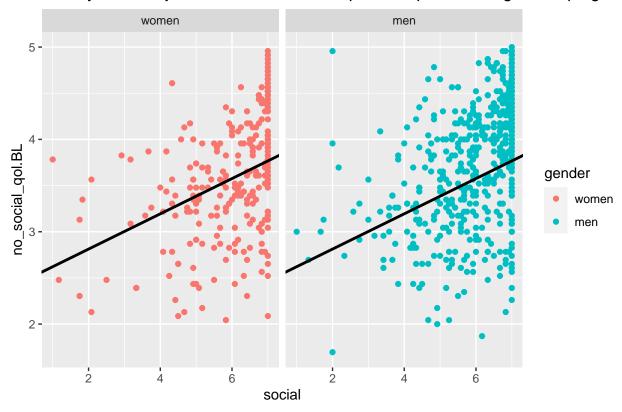
Quality of life adjusted for stress, AANA positive, positive religious coping



Quality of life adjusted for stress, AANA positive, positive religious coping



Quality of life adjusted for stress, AANA positive, positive religious coping



Looking at Model Diagnostics

```
interaction_model <- lm(no_social_qol.BL ~ age</pre>
                + gender
                + education
                + SUD.sum_illegal
                + SUD.sum_legal
                + rel.is_religious * religion_pos
                + aana_past_year
                + aana_positive
                + social
                + stress_to_subj
                ,cleaned_addiction_data)
social_only_model <- lm(no_social_qol.BL ~ age</pre>
                + gender
                + education
                + SUD.sum_illegal
                + SUD.sum_legal
                + social
                 ,cleaned_addiction_data)
with_religion_model <- lm(no_social_qol.BL ~ age</pre>
                + gender
                + education
                + SUD.sum_illegal
                + SUD.sum_legal
                + social
                + rel.is_religious * religion_pos
                ,cleaned_addiction_data)
with_aana_model <- lm(no_social_qol.BL ~ age
                + gender
                + education
                + SUD.sum_illegal
                + SUD.sum_legal
                + social
                + aana_past_year
                + aana_positive
                ,cleaned_addiction_data)
with_stress_model <- lm(no_social_qol.BL ~ age
                + gender
                + education
                + SUD.sum illegal
                + SUD.sum_legal
                + social
                + stress_to_subj
                ,cleaned_addiction_data)
```

Akaike Information Criterion (AIC) Scores of models predicting quality of life with perceived level of social support

	df	AIC
qol_model	19	1247.6
interaction_model	20	1249.5
social_only_model	11	1288.8
with_aana_model	16	1283.6
with_religion_model	14	1279.2
$with_stress_model$	12	1272.2

- Original model that takes into account perceived level of social support with degree of positive religious coping, AA/NA positive regard, and number of types of stressful life circumstances has lowest AIC score.
 - qol_model
- Other models were:
 - based on original model and included interactions terms between if a patient is religious or not with degree of positive religious coping (interaction_model)
 - only considering perceived level of social support (perceived level of social support)
 - additionally either only considered:
 - * degree of positive religious coping (with_religion_model)
 - * AA/NA positive regard (with_aana_model)
 - * number of types of stressful life circumstances (with stress model)**