RQ2_main

Quarto

The `echo: false` option disables the printing of code (only output is displayed).

```
          Q1R5                 S2_RACE_PRIME
1 Not important (3) Black or African American
2 Not important                   (1) White
3 Not important       (2) Hispanic or Latino
4     Important       (2) Hispanic or Latino
5 Not important (3) Black or African American
6 Not important           (4) Asian American
                                       S3       S3B    S5_AGE
1                                (03) Bisexual (2) Woman (2) 18-29
2 (01) Straight, that is, not gay or lesbian (2) Woman (5) 50-59
3 (01) Straight, that is, not gay or lesbian (2) Woman (6) 60-69
4 (01) Straight, that is, not gay or lesbian   (1) Man (2) 18-29
5 (01) Straight, that is, not gay or lesbian (2) Woman  (7) 70 +
6 (01) Straight, that is, not gay or lesbian   (1) Man (2) 18-29
                                 S13                                  S14
1                  (1) Grades 1-8                          (5) Rural area
2 (3) High School graduate or GED                          (5) Rural area
3        (7) Post-graduate degree                    (1) Large urban area
4    (6) Bachelors, 4-year degree             (4) Small town or small city
5   (5) Associates, 2-year degree             (4) Small town or small city
6 (3) High School graduate or GED (3) Small suburb near small town or city
                                 Q629R1                Q629R2            Q629R3
1                                  <NA>                  <NA>              <NA>
2                                  <NA>                  <NA>              <NA>
3 (1) Racial background or ethnicity (0) NO TO: Skin color (0) NO TO: Gender
4                                  <NA>                  <NA>              <NA>
5 (1) Racial background or ethnicity        (1) Skin color        (1) Gender
6                                  <NA>                  <NA>              <NA>
```

```
                                      Q629R4                          Q629R5
1                                       <NA>                            <NA>
2                                       <NA>                            <NA>
3 (0) NO TO: Sexuality or sexual orientation (0) NO TO: Immigration status
4                                       <NA>                            <NA>
5 (0) NO TO: Sexuality or sexual orientation (0) NO TO: Immigration status
6                                       <NA>                            <NA>
              Q629R6
1               <NA>
2               <NA>
3 (0) NO TO: Religion
4               <NA>
5 (0) NO TO: Religion
6               <NA>
                                                                  Q629R7    Q627
1                                                                   <NA>  (2) No
2                                                                   <NA>  (2) No
3 (0) NO TO: Accent, regardless of whether or not you have an accent (1) Yes
4                                                                   <NA>  (2) No
5 (0) NO TO: Accent, regardless of whether or not you have an accent (1) Yes
6                                                                   <NA>  (2) No
            Q633     Q619_Q626R1    Q619_Q626R2    Q619_Q626R3     Q619_Q626R4
1       (2) Some    (3) A little   (3) A little   (3) A little    (3) A little
2 (4) Not at all  (5) Don't know (5) Don't know (5) Don't know  (5) Don't know
3   (3) A little (4) None at all      (1) A lot      (1) A lot       (1) A lot
4 (5) Don't know (4) None at all      (1) A lot       (2) Some        (2) Some
5     (1) A lot    (3) A little      (1) A lot      (1) A lot       (1) A lot
6 (5) Don't know (4) None at all      (1) A lot       (2) Some (4) None at all
     Q619_Q626R5    Q619_Q626R6    Q619_Q626R7    Q619_Q626R8
1   (3) A little   (3) A little   (3) A little   (3) A little
2 (5) Don't know (5) Don't know (5) Don't know (5) Don't know
3      (1) A lot      (1) A lot      (1) A lot      (1) A lot
4      (1) A lot       (2) Some       (2) Some       (2) Some
5      (1) A lot      (1) A lot      (1) A lot      (1) A lot
6       (2) Some   (3) A little   (3) A little       (2) Some
                   educ sex_orientation          rural   race_recoded
1 No college experience          Queer          Rural          Black
2 No college experience       Straight          Rural          White
3     College graduate       Straight Urban/suburban         Latino
4     College graduate       Straight Urban/suburban         Latino
5     College graduate       Straight Urban/suburban          Black
6 No college experience       Straight Urban/suburban Asian American
  age_recoded        Religion
```

1	18-29	None
2	50-59	Protestant
3	60-69	Other Christian
4	18-29	Catholic
5	70+	Protestant
6	18-29	Agnostic

	Variable	Missing_Count
Q1R5	Q1R5	0
S2_RACE_PRIME	S2_RACE_PRIME	0
S3	S3	0
S3B	S3B	0
S5_AGE	S5_AGE	0
S13	S13	0
S14	S14	0
Q629R1	Q629R1	9180
Q629R2	Q629R2	9180
Q629R3	Q629R3	9180
Q629R4	Q629R4	9180
Q629R5	Q629R5	9180
Q629R6	Q629R6	9180
Q629R7	Q629R7	9180
Q627	Q627	0
Q633	Q633	0
Q619_Q626R1	Q619_Q626R1	0
Q619_Q626R2	Q619_Q626R2	0
Q619_Q626R3	Q619_Q626R3	0
Q619_Q626R4	Q619_Q626R4	0
Q619_Q626R5	Q619_Q626R5	0
Q619_Q626R6	Q619_Q626R6	0
Q619_Q626R7	Q619_Q626R7	0
Q619_Q626R8	Q619_Q626R8	0
educ	educ	0
sex_orientation	sex_orientation	0
rural	rural	0
race_recoded	race_recoded	0
age_recoded	age_recoded	0
Religion	Religion	0

	Variable	Missing_Proportion
Q629R1	Q629R1	0.5232
Q629R2	Q629R2	0.5232

```
Q629R3   Q629R3             0.5232
Q629R4   Q629R4             0.5232
Q629R5   Q629R5             0.5232
Q629R6   Q629R6             0.5232
Q629R7   Q629R7             0.5232
```

```r
# Remove all redundant columns
redundant_columns <- c("Q629R1", "Q629R2", "Q629R3", "Q629R4",
                       "Q629R5", "Q629R6", "Q629R7",
                       "S2_RACE_PRIME", "S13", "S3", "S5_AGE","S14")

dataset_clean <- dataset_clean[, !(names(dataset_clean) %in% redundant_columns)]

head(dataset_clean)
```

```
           Q1R5       S3B    Q627           Q633     Q619_Q626R1    Q619_Q626R2
1 Not important (2) Woman  (2) No       (2) Some    (3) A little   (3) A little
2 Not important (2) Woman  (2) No (4) Not at all  (5) Don't know (5) Don't know
3 Not important (2) Woman (1) Yes   (3) A little (4) None at all      (1) A lot
4     Important   (1) Man  (2) No (5) Don't know (4) None at all      (1) A lot
5 Not important (2) Woman (1) Yes      (1) A lot    (3) A little      (1) A lot
6 Not important   (1) Man  (2) No (5) Don't know (4) None at all      (1) A lot
     Q619_Q626R3     Q619_Q626R4    Q619_Q626R5    Q619_Q626R6    Q619_Q626R7
1   (3) A little    (3) A little   (3) A little   (3) A little   (3) A little
2 (5) Don't know  (5) Don't know (5) Don't know (5) Don't know (5) Don't know
3      (1) A lot       (1) A lot      (1) A lot      (1) A lot      (1) A lot
4       (2) Some        (2) Some      (1) A lot       (2) Some       (2) Some
5      (1) A lot       (1) A lot      (1) A lot      (1) A lot      (1) A lot
6       (2) Some (4) None at all       (2) Some   (3) A little   (3) A little
     Q619_Q626R8                  educ sex_orientation          rural
1   (3) A little No college experience           Queer          Rural
2 (5) Don't know No college experience        Straight          Rural
3      (1) A lot      College graduate        Straight Urban/suburban
4       (2) Some      College graduate        Straight Urban/suburban
5      (1) A lot      College graduate        Straight Urban/suburban
6       (2) Some No college experience        Straight Urban/suburban
    race_recoded age_recoded        Religion
1          Black       18-29            None
2          White       50-59      Protestant
3         Latino       60-69 Other Christian
4         Latino       18-29        Catholic
5          Black         70+      Protestant
```

```
6 Asian American       18-29        Agnostic
```

```r
# Rename remaining columns in dataset_clean
colnames(dataset_clean) <- c(
  "ImportanceOfStoppingDiscriminationAgainstMinorities",  # Q1R5
  "Gender",                                               # S3B
  "TreatedUnfairlyOrDiscriminatedAgainst",                # Q627
  "DiscriminationImpactOnLifeSatisfaction",              # Q633
  "PerceptionsOfDiscriminationAgainstWhites",            # Q619_Q626R1
  "PerceptionsOfDiscriminationAgainstBlacks",            # Q619_Q626R2
  "PerceptionsOfDiscriminationAgainstAsianAmericans",    # Q619_Q626R3
  "PerceptionsOfDiscriminationAgainstNativeAmericans",   # Q619_Q626R4
  "PerceptionsOfDiscriminationAgainstImmigrants",        # Q619_Q626R5
  "PerceptionsOfDiscriminationAgainstLatinos",           # Q619_Q626R6
  "PerceptionsOfDiscriminationAgainstGaysAndLesbians",   # Q619_Q626R7
  "PerceptionsOfDiscriminationAgainstMuslims",           # Q619_Q626R8
  "Education",                                           # educ
  "SexOrientation",                                      # sex_orientation
  "Rural",                                               # rural
  "Race",                                                # race_recoded
  "Age",                                                 # age_recoded
  "Religion"                                             # Religion
)

# Verify the renamed columns
colnames(dataset_clean)
```

```
 [1] "ImportanceOfStoppingDiscriminationAgainstMinorities"
 [2] "Gender"
 [3] "TreatedUnfairlyOrDiscriminatedAgainst"
 [4] "DiscriminationImpactOnLifeSatisfaction"
 [5] "PerceptionsOfDiscriminationAgainstWhites"
 [6] "PerceptionsOfDiscriminationAgainstBlacks"
 [7] "PerceptionsOfDiscriminationAgainstAsianAmericans"
 [8] "PerceptionsOfDiscriminationAgainstNativeAmericans"
 [9] "PerceptionsOfDiscriminationAgainstImmigrants"
[10] "PerceptionsOfDiscriminationAgainstLatinos"
[11] "PerceptionsOfDiscriminationAgainstGaysAndLesbians"
[12] "PerceptionsOfDiscriminationAgainstMuslims"
[13] "Education"
[14] "SexOrientation"
[15] "Rural"
```

```
[16] "Race"
[17] "Age"
[18] "Religion"
```

Model Building

```
# Regression model
model <- glm(ImportanceOfStoppingDiscriminationAgainstMinorities ~ Gender +
               TreatedUnfairlyOrDiscriminatedAgainst +
               DiscriminationImpactOnLifeSatisfaction +
               PerceptionsOfDiscriminationAgainstWhites +
               PerceptionsOfDiscriminationAgainstBlacks +
               PerceptionsOfDiscriminationAgainstAsianAmericans +
               PerceptionsOfDiscriminationAgainstNativeAmericans +
               PerceptionsOfDiscriminationAgainstImmigrants +
               PerceptionsOfDiscriminationAgainstLatinos +
               PerceptionsOfDiscriminationAgainstGaysAndLesbians +
               PerceptionsOfDiscriminationAgainstMuslims +
               Education +
                 SexOrientation +
                 Rural +
                 Race +
                 Age +
               Religion,

             data = dataset_clean, family = "binomial")

# Check multicollinearity using VIF
vif(model)
```

```
                                                GVIF Df GVIF^(1/(2*Df))
Gender                                      1.158213  3        1.024782
TreatedUnfairlyOrDiscriminatedAgainst       1.276433  1        1.129793
DiscriminationImpactOnLifeSatisfaction      1.449515  5        1.037821
PerceptionsOfDiscriminationAgainstWhites    2.249642  4        1.106660
PerceptionsOfDiscriminationAgainstBlacks    9.744467  4        1.329214
PerceptionsOfDiscriminationAgainstAsianAmericans   8.582797  4        1.308289
PerceptionsOfDiscriminationAgainstNativeAmericans  5.064934  4        1.224818
PerceptionsOfDiscriminationAgainstImmigrants      11.533688  4        1.357519
PerceptionsOfDiscriminationAgainstLatinos         12.455983  4        1.370636
```

```
PerceptionsOfDiscriminationAgainstGaysAndLesbians  6.443643  4        1.262239
PerceptionsOfDiscriminationAgainstMuslims         10.550126  4        1.342478
Education                                          1.269256  2        1.061420
SexOrientation                                     1.211796  2        1.049198
Rural                                              1.037493  1        1.018574
Race                                               2.093218  6        1.063493
Age                                                1.399241  5        1.034164
Religion                                           2.017810 16        1.022180
```

Model 2: stepwise regression

```
                                                GVIF Df GVIF^(1/(2*Df))
Gender                                      1.112836  3        1.017978
TreatedUnfairlyOrDiscriminatedAgainst       1.247497  1        1.116914
DiscriminationImpactOnLifeSatisfaction      1.377993  5        1.032582
PerceptionsOfDiscriminationAgainstWhites    2.014185  4        1.091472
PerceptionsOfDiscriminationAgainstBlacks    4.474255  4        1.205980
PerceptionsOfDiscriminationAgainstAsianAmericans  4.973272  4        1.222026
PerceptionsOfDiscriminationAgainstNativeAmericans 3.473325  4        1.168404
SexOrientation                              1.120480  2        1.028847
Race                                        1.181237  6        1.013977
```

```r
# Extract odds ratios
odds_ratios <- exp(coef(model_stepwise))

# Get confidence intervals for coefficients
conf_intervals <- exp(confint(model_stepwise))

# Extract p-values
p_values <- summary(model_stepwise)$coefficients[, 4]

# Combine everything into a single data frame
results_model_stepwise <- data.frame(
  OddsRatio = odds_ratios,
  LowerCI = conf_intervals[, 1],
  UpperCI = conf_intervals[, 2],
  PValue = p_values
)

# Display results
results_model_stepwise
```

	OddsRatio
(Intercept)	0.1612570
Gender(2) Woman	1.2541844
Gender(3) Non-binary	1.4023832
Gender(4) Something else (Specify)	1.6822476
TreatedUnfairlyOrDiscriminatedAgainst(2) No	0.8956312
DiscriminationImpactOnLifeSatisfaction(2) Some	1.0449859
DiscriminationImpactOnLifeSatisfaction(3) A little	0.9985784
DiscriminationImpactOnLifeSatisfaction(4) Not at all	0.7290754
DiscriminationImpactOnLifeSatisfaction(5) Don't know	0.8978507
DiscriminationImpactOnLifeSatisfaction(6) Refused	0.8488809
PerceptionsOfDiscriminationAgainstWhites(2) Some	1.0552392
PerceptionsOfDiscriminationAgainstWhites(3) A little	1.4658968
PerceptionsOfDiscriminationAgainstWhites(4) None at all	2.0087083
PerceptionsOfDiscriminationAgainstWhites(5) Don't know	1.5267658
PerceptionsOfDiscriminationAgainstBlacks(2) Some	0.6470676
PerceptionsOfDiscriminationAgainstBlacks(3) A little	0.5118966
PerceptionsOfDiscriminationAgainstBlacks(4) None at all	0.4806441
PerceptionsOfDiscriminationAgainstBlacks(5) Don't know	0.3885596
PerceptionsOfDiscriminationAgainstAsianAmericans(2) Some	0.8434047
PerceptionsOfDiscriminationAgainstAsianAmericans(3) A little	0.7340455
PerceptionsOfDiscriminationAgainstAsianAmericans(4) None at all	0.7180562
PerceptionsOfDiscriminationAgainstAsianAmericans(5) Don't know	0.8437969
PerceptionsOfDiscriminationAgainstNativeAmericans(2) Some	1.0871747
PerceptionsOfDiscriminationAgainstNativeAmericans(3) A little	0.9910186
PerceptionsOfDiscriminationAgainstNativeAmericans(4) None at all	0.8010391
PerceptionsOfDiscriminationAgainstNativeAmericans(5) Don't know	0.8317450
SexOrientationQueer	1.1501720
SexOrientationRefused/Don't Know	0.8955632
RaceBlack	3.5453049
RaceAsian American	2.1179304
RaceLatino	1.7842778
RaceAIAN	1.2259797
RaceNHPI	2.1738598
RaceArab/MENA	2.6687834

	LowerCI
(Intercept)	0.1294167
Gender(2) Woman	1.1636067
Gender(3) Non-binary	0.9577883
Gender(4) Something else (Specify)	0.7346264
TreatedUnfairlyOrDiscriminatedAgainst(2) No	0.8262458
DiscriminationImpactOnLifeSatisfaction(2) Some	0.8982570
DiscriminationImpactOnLifeSatisfaction(3) A little	0.8648147

```
DiscriminationImpactOnLifeSatisfaction(4) Not at all             0.6283189
DiscriminationImpactOnLifeSatisfaction(5) Don't know             0.7444418
DiscriminationImpactOnLifeSatisfaction(6) Refused                0.5760214
PerceptionsOfDiscriminationAgainstWhites(2) Some                 0.8697361
PerceptionsOfDiscriminationAgainstWhites(3) A little             1.2361494
PerceptionsOfDiscriminationAgainstWhites(4) None at all          1.7079105
PerceptionsOfDiscriminationAgainstWhites(5) Don't know           1.1695603
PerceptionsOfDiscriminationAgainstBlacks(2) Some                 0.5851886
PerceptionsOfDiscriminationAgainstBlacks(3) A little             0.4435530
PerceptionsOfDiscriminationAgainstBlacks(4) None at all          0.3821386
PerceptionsOfDiscriminationAgainstBlacks(5) Don't know           0.2770803
PerceptionsOfDiscriminationAgainstAsianAmericans(2) Some         0.7661829
PerceptionsOfDiscriminationAgainstAsianAmericans(3) A little     0.6497803
PerceptionsOfDiscriminationAgainstAsianAmericans(4) None at all  0.5854893
PerceptionsOfDiscriminationAgainstAsianAmericans(5) Don't know   0.6186419
PerceptionsOfDiscriminationAgainstNativeAmericans(2) Some        0.9827532
PerceptionsOfDiscriminationAgainstNativeAmericans(3) A little    0.8808812
PerceptionsOfDiscriminationAgainstNativeAmericans(4) None at all 0.6721337
PerceptionsOfDiscriminationAgainstNativeAmericans(5) Don't know  0.6543775
SexOrientationQueer                                              1.0217103
SexOrientationRefused/Don't Know                                 0.6695420
RaceBlack                                                        3.1422978
RaceAsian American                                               1.8661699
RaceLatino                                                       1.5707378
RaceAIAN                                                         0.9709870
RaceNHPI                                                         1.6569552
RaceArab/MENA                                                    1.9606372
                                                                   UpperCI
(Intercept)                                                      0.2002620
Gender(2) Woman                                                  1.3520497
Gender(3) Non-binary                                             2.0338100
Gender(4) Something else (Specify)                               3.5987211
TreatedUnfairlyOrDiscriminatedAgainst(2) No                      0.9708918
DiscriminationImpactOnLifeSatisfaction(2) Some                   1.2166183
DiscriminationImpactOnLifeSatisfaction(3) A little               1.1541555
DiscriminationImpactOnLifeSatisfaction(4) Not at all             0.8467606
DiscriminationImpactOnLifeSatisfaction(5) Don't know             1.0824752
DiscriminationImpactOnLifeSatisfaction(6) Refused                1.2282690
PerceptionsOfDiscriminationAgainstWhites(2) Some                 1.2828008
PerceptionsOfDiscriminationAgainstWhites(3) A little             1.7443283
PerceptionsOfDiscriminationAgainstWhites(4) None at all          2.3714874
PerceptionsOfDiscriminationAgainstWhites(5) Don't know           1.9902933
PerceptionsOfDiscriminationAgainstBlacks(2) Some                 0.7151963
```

PerceptionsOfDiscriminationAgainstBlacks(3) A little	0.5898332
PerceptionsOfDiscriminationAgainstBlacks(4) None at all	0.6009511
PerceptionsOfDiscriminationAgainstBlacks(5) Don't know	0.5405418
PerceptionsOfDiscriminationAgainstAsianAmericans(2) Some	0.9284341
PerceptionsOfDiscriminationAgainstAsianAmericans(3) A little	0.8289819
PerceptionsOfDiscriminationAgainstAsianAmericans(4) None at all	0.8780969
PerceptionsOfDiscriminationAgainstAsianAmericans(5) Don't know	1.1444465
PerceptionsOfDiscriminationAgainstNativeAmericans(2) Some	1.2028562
PerceptionsOfDiscriminationAgainstNativeAmericans(3) A little	1.1148690
PerceptionsOfDiscriminationAgainstNativeAmericans(4) None at all	0.9526623
PerceptionsOfDiscriminationAgainstNativeAmericans(5) Don't know	1.0528240
SexOrientationQueer	1.2934188
SexOrientationRefused/Don't Know	1.1835548
RaceBlack	4.0058247
RaceAsian American	2.4063192
RaceLatino	2.0288631
RaceAIAN	1.5377607
RaceNHPI	2.8313376
RaceArab/MENA	3.6086648

	PValue
(Intercept)	2.349071e-60
Gender(2) Woman	3.312281e-09
Gender(3) Non-binary	7.781804e-02
Gender(4) Something else (Specify)	1.950703e-01
TreatedUnfairlyOrDiscriminatedAgainst(2) No	7.394049e-03
DiscriminationImpactOnLifeSatisfaction(2) Some	5.695790e-01
DiscriminationImpactOnLifeSatisfaction(3) A little	9.845814e-01
DiscriminationImpactOnLifeSatisfaction(4) Not at all	3.295755e-05
DiscriminationImpactOnLifeSatisfaction(5) Don't know	2.591238e-01
DiscriminationImpactOnLifeSatisfaction(6) Refused	3.953836e-01
PerceptionsOfDiscriminationAgainstWhites(2) Some	5.874276e-01
PerceptionsOfDiscriminationAgainstWhites(3) A little	1.328535e-05
PerceptionsOfDiscriminationAgainstWhites(4) None at all	7.871687e-17
PerceptionsOfDiscriminationAgainstWhites(5) Don't know	1.800250e-03
PerceptionsOfDiscriminationAgainstBlacks(2) Some	1.797131e-17
PerceptionsOfDiscriminationAgainstBlacks(3) A little	3.214102e-20
PerceptionsOfDiscriminationAgainstBlacks(4) None at all	2.188227e-10
PerceptionsOfDiscriminationAgainstBlacks(5) Don't know	2.887783e-08
PerceptionsOfDiscriminationAgainstAsianAmericans(2) Some	5.091500e-04
PerceptionsOfDiscriminationAgainstAsianAmericans(3) A little	6.473071e-07
PerceptionsOfDiscriminationAgainstAsianAmericans(4) None at all	1.352987e-03
PerceptionsOfDiscriminationAgainstAsianAmericans(5) Don't know	2.788753e-01
PerceptionsOfDiscriminationAgainstNativeAmericans(2) Some	1.049514e-01

```
PerceptionsOfDiscriminationAgainstNativeAmericans(3) A little    8.806537e-01
PerceptionsOfDiscriminationAgainstNativeAmericans(4) None at all 1.263638e-02
PerceptionsOfDiscriminationAgainstNativeAmericans(5) Don't know  1.286757e-01
SexOrientationQueer                                               2.000785e-02
SexOrientationRefused/Don't Know                                  4.472281e-01
RaceBlack                                                         7.609542e-93
RaceAsian American                                                5.539797e-31
RaceLatino                                                        7.271025e-19
RaceAIAN                                                          8.212943e-02
RaceNHPI                                                          1.293615e-08
RaceArab/MENA                                                     2.701934e-10
```

Model Diagnostics

```r
# Set seed for reproducibility
set.seed(123)

# Split data into training (70%) and testing (30%)
library(caret)
train_index <- createDataPartition(dataset_clean$ImportanceOfStoppingDiscriminationAgainstMi
train_data <- dataset_clean[train_index, ]
test_data <- dataset_clean[-train_index, ]

# Train stepwise logistic regression model on training data

# Fit the initial model (full model)
full_model <- glm(ImportanceOfStoppingDiscriminationAgainstMinorities ~ Gender +
                TreatedUnfairlyOrDiscriminatedAgainst +
                DiscriminationImpactOnLifeSatisfaction +
                PerceptionsOfDiscriminationAgainstWhites +
                PerceptionsOfDiscriminationAgainstBlacks +
                PerceptionsOfDiscriminationAgainstAsianAmericans +
                PerceptionsOfDiscriminationAgainstNativeAmericans +
                PerceptionsOfDiscriminationAgainstImmigrants +
                PerceptionsOfDiscriminationAgainstLatinos +
                PerceptionsOfDiscriminationAgainstGaysAndLesbians +
                PerceptionsOfDiscriminationAgainstMuslims +
                Education +
                  SexOrientation +
                  Rural +
                  Race +
```

```
                  Age +
                Religion,
                   data = train_data, family = "binomial")

# Perform stepwise selection
final_model <- step(full_model,
                        direction = "both",
                        trace = 0)

# Generate predictions on the test data
predicted_probs_stepwise <- predict(final_model, newdata = test_data, type = "response")

# Convert probabilities to binary outcomes using a threshold (e.g., 0.5)
predicted_classes_stepwise <- ifelse(predicted_probs_stepwise > 0.5, "Important", "Not impor

# Convert predicted classes to factor and align levels with actual classes
predicted_classes_stepwise <- factor(predicted_classes_stepwise, levels = c("Not important",
actual_classes_stepwise <- factor(test_data$ImportanceOfStoppingDiscriminationAgainstMinoriti
                                  levels = c("Not important", "Important"))

# Create confusion matrix
conf_matrix <- confusionMatrix(predicted_classes_stepwise, actual_classes_stepwise)

# Print the confusion matrix
print(conf_matrix)
```

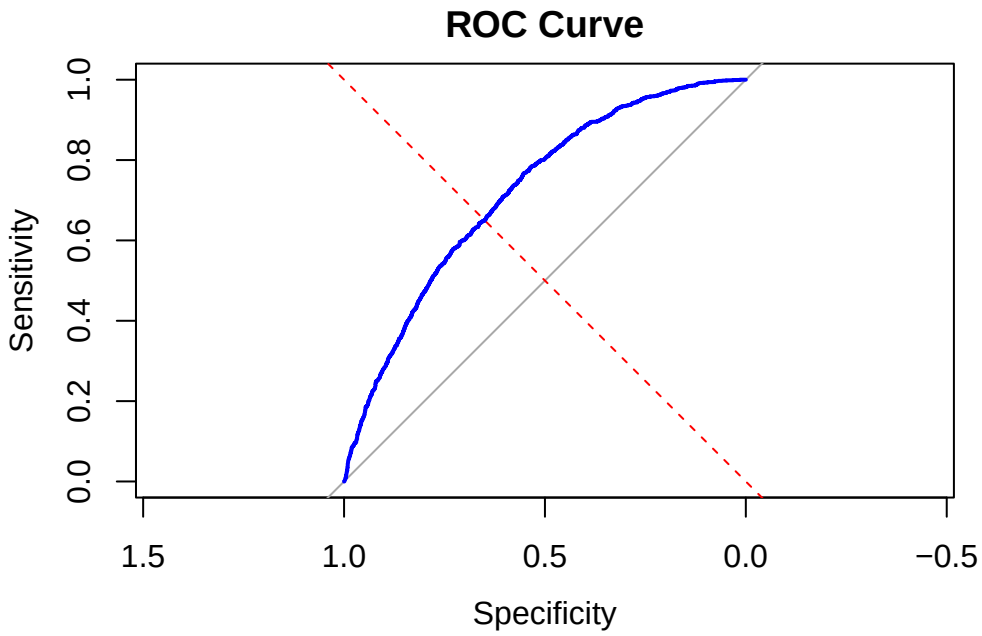
```
Confusion Matrix and Statistics

               Reference
Prediction      Not important Important
  Not important          3708      1127
  Important               200       228

               Accuracy : 0.7479
                 95% CI : (0.7359, 0.7595)
    No Information Rate : 0.7425
    P-Value [Acc > NIR] : 0.1932

                  Kappa : 0.1508

 Mcnemar's Test P-Value : <2e-16
```

```
            Sensitivity : 0.9488
            Specificity : 0.1683
         Pos Pred Value : 0.7669
         Neg Pred Value : 0.5327
             Prevalence : 0.7425
         Detection Rate : 0.7045
   Detection Prevalence : 0.9187
      Balanced Accuracy : 0.5585

       'Positive' Class : Not important
```

```r
# Calculate the ROC curve
roc_curve <- roc(test_data$ImportanceOfStoppingDiscriminationAgainstMinorities,
                 predicted_probs_stepwise,
                 levels = c("Not important", "Important"))

# Print the AUC
auc_value <- auc(roc_curve)
cat("AUC:", auc_value, "\n")
```

```
AUC: 0.7186545
```

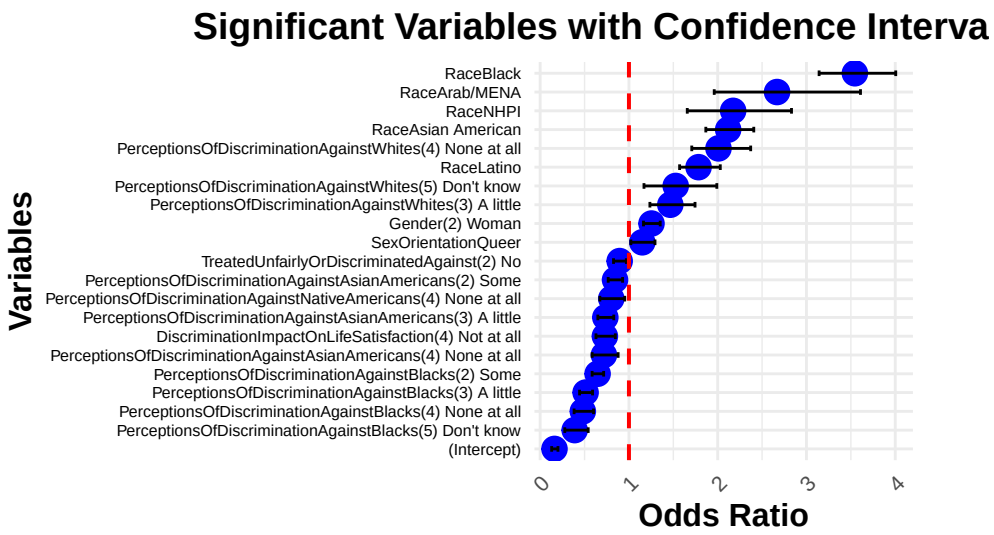
```r
# Plot the ROC curve
plot(roc_curve, col = "blue", main = "ROC Curve")
abline(a = 0, b = 1, col = "red", lty = 2)  # Add diagonal line for random guess
```




```
# Save the ROC plot
ggsave("roc_curve_plot.png", width = 8, height = 6)
```

```
## Figure 2:
# Filter for significant variables (p-value < 0.05)
significant_data <- results_model_stepwise[results_model_stepwise$PValue < 0.05, ]

# Add a column for variable names to use in the plot
significant_data <- significant_data %>%
  mutate(Variable = rownames(significant_data))

# Create the plot
ggplot(significant_data, aes(x = OddsRatio, y = reorder(Variable, OddsRatio))) +
  geom_point(size = 4, color = "blue") +  # Add points for odds ratios
  geom_errorbarh(aes(xmin = LowerCI, xmax = UpperCI), height = 0.3, color = "black") +  # Ad
  geom_vline(xintercept = 1, linetype = "dashed", color = "red", size = 0.7) +  # Add a verti
  labs(
    title = "Significant Variables with Confidence Intervals (Odds Ratio)",
    x = "Odds Ratio",
    y = "Variables"
  ) +
  theme_minimal() +  # Use a minimal theme for better aesthetics
  theme(
```

```
    axis.text.y = element_text(size = 6, color = "black"),  # Improve y-axis text readability
    axis.title.x = element_text(size = 12, face = "bold"),  # Highlight x-axis title
    axis.title.y = element_text(size = 12, face = "bold"),  # Highlight y-axis title
    plot.title = element_text(size = 14, face = "bold", hjust = 0.5),  # Center the title
    plot.margin = margin(1, 1, 1, 1, "cm"),  # Adjust margins
    axis.text.x = element_text(angle = 45, hjust = 1, size = 8),  # Reduce x-axis text size
    legend.text = element_text(size = 8),  # Reduce legend text size
    legend.title = element_text(size = 9)  # Reduce legend title size
  )
```



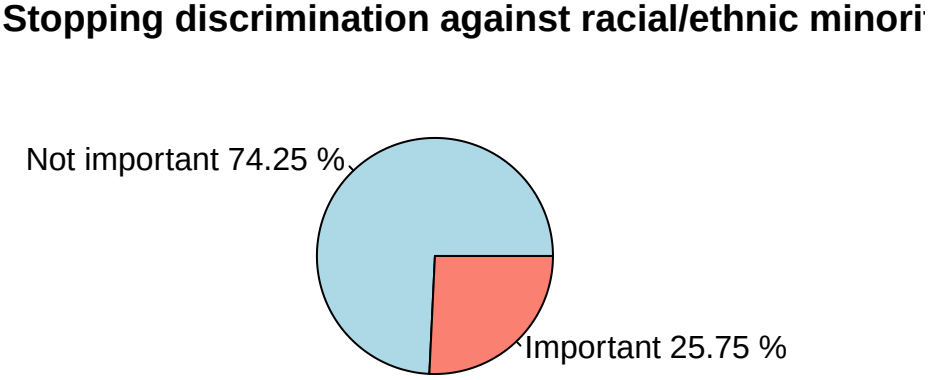

```
# Save the plot with larger dimensions
ggsave("Significant_Variables_Plot_Enhanced.png", width = 20, height = 10)
```

Figure 1: Proportion of outcome variable for RQ 1

```
# Calculate proportions
proportions <- prop.table(table(dataset$Q1R5))
proportions
```

```
(0) NO TO: Stopping discrimination against racial/ethnic minorities
                                                                0.7424907
       (1) Stopping discrimination against racial/ethnic minorities
                                                                0.2575093
```

```r
# Define custom labels
custom_labels <- c("Not important", "Important")

# Combine custom labels with percentages
pie_labels <- paste(custom_labels, round(proportions * 100, 2), "%")

# Plot the pie chart with custom labels
pie(proportions,
    labels = pie_labels,
    main = "Q1R5: Stopping discrimination against racial/ethnic minorities",
    col = c("lightblue", "salmon"))
```




```r
ggsave("Proportion_Outcome_variable_RQ2.png", width = 20, height = 10)
```