Quarto Set Up, Descriptive Statistics and Hypotheses Testing

Yimeng (Amy) CHENG

Department of Psychology, University of Chicago

Author Note

Yimeng (Amy) CHENG https://orcid.org/0009-0007-1238-3749

The authors declare no conflict of interest.

Correspondence concerning this article should be addressed to Yimeng (Amy) CHENG, Department of Psychology, University of Chicago, 5801 S Ellis Ave, Chicago, IL 60615, USA, Email: yimengcheng@uchicago.edu

Abstract

This assessment uses the math anxiety dataset to perform various statistical analyses and hypotheses testing.

Keywords: Math Anxiety, Cognitive Concerns, Perfectionism, Statistical Anxiety

Quarto Set Up, Descriptive Statistics and Hypotheses Testing

Table of contents

Quarto Set Up, Descriptive Statistics and Hypotheses Testing

Objective

The objective of this assignment is to practice using R to calculate and interpret descriptive statistics. Unlike in previous assignments, you will not use a built-in dataset. Instead, you will use data from a published, open-access dataset capturing relationships between math anxiety and self-perception:

Mackinnon, S. P., McCaughey, N., & Hill, T. G. (2024, July 25). The Association of Self-Efficacy, Anxiety Sensitivity, and Self-Critical Perfectionism with Statistics and Math Anxiety. https://doi.org/10.17605/OSF.IO/NZHQ6

You can access the full dataset and codebook here. The data you will need is included in the .csv file within this project's directory.

Before you begin diving into the data summary, you should take some time to familiarize yourself with the dataset, including the different categories of variables and how they can be interpreted.

Set Up

Read in the data

Run this chunk to read in the dataset and view the first few rows:

```
# A tibble: 6 x 97
       id program.type   age ethnicity.r major dal.course.code  se.1  se.2  se.3
    <dbl>        <dbl> <dbl> <chr>       <chr> <chr>           <dbl> <dbl> <dbl>
1 1.25e10            0    21 White       Psyc~ PSYO 2501, STA~     4     3     4
2 1.25e10            0    22 Middle Eas~ Biol~ STAT 1060           5     5     5
3 1.25e10            0    16 Black       Acti~ <NA>                4     5     5
4 1.25e10            0    19 White       Phar~ <NA>                5     4     4
5 1.25e10            0    18 White       Unde~ STAT 1060           4     3     4
6 1.25e10            0    19 White       neur~ stat 1060           4     4     3
```

```
# i 88 more variables: se.4 <dbl>, se.5 <dbl>, se.6 <dbl>, se.7 <dbl>,
#   se.8 <dbl>, asi.1 <dbl>, asi.2 <dbl>, asi.3 <dbl>, asi.4 <dbl>,
#   asi.5 <dbl>, asi.6 <dbl>, asi.7 <dbl>, asi.8 <dbl>, asi.9 <dbl>,
#   asi.10 <dbl>, asi.11 <dbl>, asi.12 <dbl>, asi.13 <dbl>, asi.14 <dbl>,
#   asi.15 <dbl>, asi.16 <dbl>, asi.17 <dbl>, asi.18 <dbl>, pc.1 <dbl>,
#   pc.2 <dbl>, pc.3 <dbl>, psp.1 <dbl>, psp.2 <dbl>, psp.3 <dbl>,
#   frost.ps.1 <dbl>, frost.ps.2 <dbl>, frost.ps.3 <dbl>, frost.ps.4 <dbl>, ...
```

Examine the structure of the df. Take note of data types and review factor levels.

```
spc_tbl_ [453 x 97] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ id             : num [1:453] 1.25e+10 1.25e+10 1.25e+10 1.25e+10 1.25e+10 ...
 $ program.type   : num [1:453] 0 0 0 0 0 0 0 0 0 0 ...
 $ age            : num [1:453] 21 22 16 19 18 19 21 19 18 28 ...
 $ ethnicity.r    : chr [1:453] "White" "Middle Eastern" "Black" "White" ...
 $ major          : chr [1:453] "Psychology" "Biology" "Acting" "Pharmacy" ...
 $ dal.course.code: chr [1:453] "PSYO 2501, STAT 1060" "STAT 1060" NA NA ...
 $ se.1           : num [1:453] 4 5 4 5 4 4 4 4 4 5 ...
 $ se.2           : num [1:453] 3 5 5 4 3 4 4 4 3 5 ...
 $ se.3           : num [1:453] 4 5 5 4 4 3 4 5 4 5 ...
 $ se.4           : num [1:453] 4 4 5 5 4 5 4 4 4 5 ...
 $ se.5           : num [1:453] 4 4 5 4 4 5 5 5 4 5 ...
 $ se.6           : num [1:453] 4 5 5 4 4 5 5 4 4 5 ...
 $ se.7           : num [1:453] 3 5 4 3 4 4 5 5 4 5 ...
 $ se.8           : num [1:453] 3 3 5 3 4 4 3 5 3 5 ...
 $ asi.1          : num [1:453] 4 3 5 5 4 5 3 2 4 4 ...
 $ asi.2          : num [1:453] 3 4 3 1 4 5 5 1 2 2 ...
 $ asi.3          : num [1:453] 5 4 5 3 3 5 4 2 4 2 ...
 $ asi.4          : num [1:453] 2 4 5 1 2 3 1 2 2 3 ...
```

```
$ asi.5          : num [1:453] 3 5 3 1 3 3 4 1 2 3 ...
$ asi.6          : num [1:453] 5 4 5 2 4 5 1 1 1 1 ...
$ asi.7          : num [1:453] 5 1 3 2 4 2 2 3 1 2 ...
$ asi.8          : num [1:453] 4 1 1 1 2 2 1 2 3 2 ...
$ asi.9          : num [1:453] 5 3 5 1 4 2 1 3 2 1 ...
$ asi.10         : num [1:453] 5 4 1 1 4 2 1 5 1 1 ...
$ asi.11         : num [1:453] 2 1 3 1 4 4 3 4 2 1 ...
$ asi.12         : num [1:453] 4 1 3 1 3 4 1 3 2 1 ...
$ asi.13         : num [1:453] 5 1 5 3 5 4 1 4 2 1 ...
$ asi.14         : num [1:453] 3 2 1 1 4 4 1 2 2 1 ...
$ asi.15         : num [1:453] 2 1 3 2 2 4 1 3 2 1 ...
$ asi.16         : num [1:453] 3 4 2 1 4 3 3 3 3 1 ...
$ asi.17         : num [1:453] 5 5 5 4 2 3 1 5 3 2 ...
$ asi.18         : num [1:453] 3 2 1 1 4 3 1 4 4 1 ...
$ pc.1           : num [1:453] 3 3 3 4 3 4 4 2 2 NA ...
$ pc.2           : num [1:453] 3 3 3 3 2 4 4 2 1 NA ...
$ pc.3           : num [1:453] 3 3 3 3 3 4 4 2 2 2 ...
$ psp.1          : num [1:453] 7 7 6 6 7 7 7 5 4 1 ...
$ psp.2          : num [1:453] 7 7 7 3 7 7 7 5 3 1 ...
$ psp.3          : num [1:453] 7 7 7 4 7 7 6 5 6 1 ...
$ frost.ps.1     : num [1:453] 5 4 4 4 4 5 4 2 2 1 ...
$ frost.ps.2     : num [1:453] 4 4 2 4 4 5 5 2 3 2 ...
$ frost.ps.3     : num [1:453] 3 4 3 3 4 5 5 2 2 3 ...
$ frost.ps.4     : num [1:453] 4 4 4 4 3 5 3 2 2 3 ...
$ frost.com.1    : num [1:453] 5 2 2 4 2 4 4 3 3 1 ...
$ frost.com.2    : num [1:453] 4 1 4 2 3 2 4 2 3 1 ...
$ frost.com.3    : num [1:453] 4 3 2 3 2 3 2 2 2 1 ...
```

```
$ frost.com.4    : num [1:453] 5 1 2 3 4 4 5 2 3 1 ...
$ frost.com.5    : num [1:453] 4 5 2 1 3 4 3 3 2 1 ...
$ frost.da.1     : num [1:453] 4 1 4 4 4 4 3 3 3 1 ...
$ frost.da.2     : num [1:453] 3 1 4 1 4 4 4 2 4 3 ...
$ frost.da.3     : num [1:453] 3 1 4 4 2 4 5 1 2 2 ...
$ frost.da.4     : num [1:453] 3 2 4 4 3 2 3 1 2 1 ...
$ stat.anx.1     : num [1:453] 4 3 3 2 4 3 4 NA 4 4 ...
$ stat.anx.2     : num [1:453] 5 1 1 3 4 5 5 NA 3 1 ...
$ stat.anx.3     : num [1:453] 4 1 3 1 3 3 3 NA 4 2 ...
$ stat.anx.4     : num [1:453] 5 4 5 2 5 5 4 NA 3 5 ...
$ stat.anx.5     : num [1:453] 5 4 5 2 5 5 5 NA 4 5 ...
$ stat.anx.6     : num [1:453] 3 1 3 1 3 5 2 NA 4 2 ...
$ stat.anx.7     : num [1:453] 5 1 1 2 4 2 5 3 3 1 ...
$ stat.anx.8     : num [1:453] 4 1 1 2 4 4 5 NA 2 1 ...
$ stat.anx.9     : num [1:453] 3 1 2 2 2 3 1 2 3 1 ...
$ stat.anx.10    : num [1:453] 3 1 2 2 2 2 1 NA 3 1 ...
$ stat.anx.11    : num [1:453] 2 1 2 2 1 2 1 3 4 1 ...
$ stat.anx.12    : num [1:453] 2 1 2 1 1 3 1 2 3 1 ...
$ stat.anx.13    : num [1:453] 4 1 2 3 2 5 1 3 3 1 ...
$ stat.anx.14    : num [1:453] 4 1 2 4 3 5 1 2 4 1 ...
$ stat.anx.15    : num [1:453] 4 1 2 2 2 1 1 2 3 1 ...
$ stat.anx.16    : num [1:453] 4 1 2 3 2 5 1 2 3 1 ...
$ math.anx.1     : num [1:453] 3 1 1 1 2 5 1 1 1 1 ...
$ math.anx.2     : num [1:453] 5 2 1 2 4 5 4 1 4 3 ...
$ math.anx.3     : num [1:453] 4 1 1 1 2 5 3 1 3 1 ...
$ math.anx.4     : num [1:453] 5 2 5 2 4 5 4 4 5 4 ...
$ math.anx.5     : num [1:453] 5 1 1 2 5 4 2 3 5 3 ...
```

```
 $ math.anx.6     : num [1:453] 3 1 1 1 3 2 3 1 3 1 ...
 $ math.anx.7     : num [1:453] 4 1 1 1 2 3 3 1 4 1 ...
 $ math.anx.8     : num [1:453] 5 2 5 3 5 5 5 4 5 3 ...
 $ math.anx.9     : num [1:453] 3 1 1 1 2 3 3 3 2 1 ...
 $ faculty        : num [1:453] 0 1 0 1 1 1 1 0 1 1 ...
 $ stats          : num [1:453] 0 0 0 0 1 1 0 0 1 0 ...
 $ stats.history  : num [1:453] 1 1 0 0 0 0 1 0 0 1 ...
 $ gender.category: num [1:453] 1 1 0 0 1 1 1 1 1 1 ...
 $ self.efficacy  : num [1:453] 3.62 4.5 4.75 4 3.88 ...
 $ asi.phys       : num [1:453] 3.67 2 3.33 1.67 2.67 ...
 $ asi.cog        : num [1:453] 3.33 3.5 1.83 1 3.83 ...
 $ asi.social     : num [1:453] 4.33 2.83 4.67 2.67 3.83 ...
 $ asi            : num [1:453] 3.78 2.78 3.28 1.78 3.44 ...
 $ perf.cog       : num [1:453] 3 3 3 3.33 2.67 ...
 $ perf.sp        : num [1:453] 7 7 6.67 4.33 7 ...
 $ frost.ps       : num [1:453] 4 4 3.25 3.75 3.75 5 4.25 2 2.25 2.25 ...
 $ frost.com      : num [1:453] 4.4 2.4 2.4 2.6 2.8 3.4 3.6 2.4 2.6 1 ...
 $ frost.da       : num [1:453] 3.25 1.25 4 3.25 3.25 3.5 3.75 1.75 2.75 1.75 ...
 $ stat.anx.tc    : num [1:453] 4.67 3.67 4.33 2 4.67 ...
 $ stat.anx.i     : num [1:453] 3.67 1 2.33 1.33 3.33 ...
 $ stat.anx.ah    : num [1:453] 5 1 1 2.5 4 3.5 5 3 3 1 ...
 $ stat.anx.ws    : num [1:453] 3.67 1 2 2.67 2 ...
 $ stat.anx.fst   : num [1:453] 3.67 1 2 2.67 2.33 ...
 $ stat.anx.sc    : num [1:453] 2 1 2 1.5 1 2.5 1 2.5 3.5 1 ...
 $ math.anx       : num [1:453] 4.11 1.33 1.89 1.56 3.22 ...
 - attr(*, "spec")=
  .. cols(
```

```
..   id = col_double(),
..   program.type = col_double(),
..   age = col_double(),
..   ethnicity.r = col_character(),
..   major = col_character(),
..   dal.course.code = col_character(),
..   se.1 = col_double(),
..   se.2 = col_double(),
..   se.3 = col_double(),
..   se.4 = col_double(),
..   se.5 = col_double(),
..   se.6 = col_double(),
..   se.7 = col_double(),
..   se.8 = col_double(),
..   asi.1 = col_double(),
..   asi.2 = col_double(),
..   asi.3 = col_double(),
..   asi.4 = col_double(),
..   asi.5 = col_double(),
..   asi.6 = col_double(),
..   asi.7 = col_double(),
..   asi.8 = col_double(),
..   asi.9 = col_double(),
..   asi.10 = col_double(),
..   asi.11 = col_double(),
..   asi.12 = col_double(),
..   asi.13 = col_double(),
```

```
..   asi.14 = col_double(),
..   asi.15 = col_double(),
..   asi.16 = col_double(),
..   asi.17 = col_double(),
..   asi.18 = col_double(),
..   pc.1 = col_double(),
..   pc.2 = col_double(),
..   pc.3 = col_double(),
..   psp.1 = col_double(),
..   psp.2 = col_double(),
..   psp.3 = col_double(),
..   frost.ps.1 = col_double(),
..   frost.ps.2 = col_double(),
..   frost.ps.3 = col_double(),
..   frost.ps.4 = col_double(),
..   frost.com.1 = col_double(),
..   frost.com.2 = col_double(),
..   frost.com.3 = col_double(),
..   frost.com.4 = col_double(),
..   frost.com.5 = col_double(),
..   frost.da.1 = col_double(),
..   frost.da.2 = col_double(),
..   frost.da.3 = col_double(),
..   frost.da.4 = col_double(),
..   stat.anx.1 = col_double(),
..   stat.anx.2 = col_double(),
..   stat.anx.3 = col_double(),
```

```
..   stat.anx.4 = col_double(),
..   stat.anx.5 = col_double(),
..   stat.anx.6 = col_double(),
..   stat.anx.7 = col_double(),
..   stat.anx.8 = col_double(),
..   stat.anx.9 = col_double(),
..   stat.anx.10 = col_double(),
..   stat.anx.11 = col_double(),
..   stat.anx.12 = col_double(),
..   stat.anx.13 = col_double(),
..   stat.anx.14 = col_double(),
..   stat.anx.15 = col_double(),
..   stat.anx.16 = col_double(),
..   math.anx.1 = col_double(),
..   math.anx.2 = col_double(),
..   math.anx.3 = col_double(),
..   math.anx.4 = col_double(),
..   math.anx.5 = col_double(),
..   math.anx.6 = col_double(),
..   math.anx.7 = col_double(),
..   math.anx.8 = col_double(),
..   math.anx.9 = col_double(),
..   faculty = col_double(),
..   stats = col_double(),
..   stats.history = col_double(),
..   gender.category = col_double(),
..   self.efficacy = col_double(),
```

```
  ..   asi.phys = col_double(),
  ..   asi.cog = col_double(),
  ..   asi.social = col_double(),
  ..   asi = col_double(),
  ..   perf.cog = col_double(),
  ..   perf.sp = col_double(),
  ..   frost.ps = col_double(),
  ..   frost.com = col_double(),
  ..   frost.da = col_double(),
  ..   stat.anx.tc = col_double(),
  ..   stat.anx.i = col_double(),
  ..   stat.anx.ah = col_double(),
  ..   stat.anx.ws = col_double(),
  ..   stat.anx.fst = col_double(),
  ..   stat.anx.sc = col_double(),
  ..   math.anx = col_double()
  .. )
 - attr(*, "problems")=<externalptr>

$ethnicity.r
[1] "Asian"           "Black"           "Hispanic/Latinx" "Middle Eastern"
[5] "Mixed/Other"     "White"


$major
  [1] "Acting"
  [2] "anthropology"
  [3] "Apllied Biochemistry"
  [4] "Applied Biology"
```

```
 [5] "Applied Computer Science"
 [6] "Applied Languages"
 [7] "Applied maths"
 [8] "Architecture"
 [9] "Aveiro University"
[10] "Bachelor"
[11] "Bachelor of Medical Sciences"
[12] "Bachelor of Science Medical Science"
[13] "Baltic Sea Region Studies"
[14] "Biligy"
[15] "Biochemestry"
[16] "Biochemistry"
[17] "Biochemistry & Molecular Biology"
[18] "bioinformatic"
[19] "biology"
[20] "Biology"
[21] "Biology (BSC)"
[22] "Biology and Psychology"
[23] "Biology/chemistry double major"
[24] "Biomedical imaging"
[25] "BSc Psych"
[26] "Business & IT"
[27] "bussiness"
[28] "Chemical engineering"
[29] "Chemistry"
[30] "coding"
[31] "cognitive science"
```

```
 [32] "Communication Sciences"
 [33] "comparative literature"
 [34] "computer science"
 [35] "Computer science"
 [36] "Computer Science"
 [37] "Computer Science and engineering"
 [38] "Computer Science and Engineering"
 [39] "comunication design-multimedia branch"
 [40] "Data science"
 [41] "Data Science"
 [42] "DDH and BDH"
 [43] "dentistry"
 [44] "double major in psychology and philosophy"
 [45] "Double major: Psychology and Sociology"
 [46] "Drawing"
 [47] "Economics"
 [48] "Economy administration"
 [49] "Electrical"
 [50] "Electrical Engineer"
 [51] "electrical engineering"
 [52] "Eletronics and Computer Science"
 [53] "Engineering"
 [54] "Engineering and Science of Data"
 [55] "Engineering Management"
 [56] "English and Psychology"
 [57] "Engraving and printmaking techniques"
 [58] "enineering"
```

```
 [59] "environmental health"
 [60] "European Economics Governance"
 [61] "Finance"
 [62] "Food Safety"
 [63] "Foreign Languages"
 [64] "French"
 [65] "Gap (Gestión y Administración Pública) Public administration on english"
 [66] "Geography"
 [67] "Graphic Design"
 [68] "history"
 [69] "Hull University"
 [70] "Human-Computer Interaction"
 [71] "Hydraulic engineering"
 [72] "industrial engineering"
 [73] "Industrial engineering"
 [74] "Industrial Engineering"
 [75] "Informatic"
 [76] "Information and library studies"
 [77] "Information Management"
 [78] "IT"
 [79] "kinesiology"
 [80] "Kinesiology"
 [81] "Languages and communication in international enterprises"
 [82] "Law"
 [83] "Law Justice and Society"
 [84] "Law, Justice, and Society"
 [85] "Linguistics"
```

```
 [86] "MA in European Studies"
 [87] "Majors in Psychology and Biology"
 [88] "management"
 [89] "Management"
 [90] "marine biology"
 [91] "Marine Biology"
 [92] "Marine Biolohgy"
 [93] "Marketing"
 [94] "Maste in Informatics Engineering"
 [95] "master"
 [96] "Masters Degree"
 [97] "Mathematics"
 [98] "Mechanical engineering"
 [99] "Mechanical Engineering"
[100] "medical science"
[101] "Medical Science"
[102] "Medical Science/ Mechanical Engineering"
[103] "medical sciences"
[104] "Medical sciences"
[105] "Medical Sciences"
[106] "Medicine"
[107] "Medicsl Science"
[108] "Microbiology"
[109] "Microbiology & Immunology"
[110] "Microbiology and Immunology"
[111] "Modern History"
[112] "Msc Economics"
```

```
[113] "neuro"

[114] "neuroscience"

[115] "Neuroscience"

[116] "Neuroscience and Chemistry"

[117] "Neursoscience"

[118] "nursing"

[119] "Nursing"

[120] "Nutrition and dietetics"

[121] "Open eGoverment"

[122] "oriental languages and cultures"

[123] "Pharmacy"

[124] "Philosophy"

[125] "Physicaltherapy"

[126] "Physics"

[127] "psych"

[128] "psychology"

[129] "Psychology"

[130] "PSYCHOLOGY"

[131] "Psychology (Planned)"

[132] "Psychology and Biology"

[133] "Psychology and History"

[134] "psychology and neuroscience"

[135] "psychology and sociology"

[136] "Pure Mathematics and Law, Justice, and Society"

[137] "Recreation Therapy"

[138] "Science Education"

[139] "Security and strategy studies"
```

```
[140] "Social Psychology"

[141] "Social Psychology and Psychology of Work"

[142] "Sociology"

[143] "Software and Data Engineering"

[144] "Sustainability"

[145] "Systems Engineering"

[146] "tax economics"

[147] "Textile Engineering"

[148] "Therapeutic Recreation"

[149] "Tribology"

[150] "undecided"

[151] "Undecided"

[152] "undecided between oncologyst and psychiatrist"

[153] "Undicided"

[154] "UNIVERSITY JUAN CARLOS 1"

[155] "University od Kazimeirz Wielki in Bydgoszcz, Poland"

[156] "University Polytechnic of Valencia"

[157] "Urban Planning and Policy Design"


$dal.course.code
 [1] "1060, 2501"

 [2] "1o60, 2501"

 [3] "2501"

 [4] "I have not yet, will next semester."

 [5] "MATH 1060"

 [6] "MATH1060"

 [7] "N/A"
```

```
 [8] "posy 2501 posy 3502"
 [9] "psyo 2501"
[10] "Psyo 2501"
[11] "PSYO 2501"
[12] "PSYO 2501 STAT 1060"
[13] "PSYO 2501, PSYO 3502"
[14] "PSYO 2501, STAT 1060"
[15] "PSYO 3502"
[16] "psyo2501"
[17] "PSYO2501"
[18] "PSYO2501, STAT1060"
[19] "PYSO2501"
[20] "SCIE 1506"
[21] "SCIE 1506, PSYO 2501"
[22] "SCIE 1506: PSYO 2501"
[23] "SCIE 1506; MATH 1215; PSYO 2501"
[24] "STAT 0160, PSYO 2501"
[25] "stat 1060"
[26] "Stat 1060"
[27] "STAT 1060"
[28] "STAT 1060 & PSYO 2501"
[29] "stat 1060 and psyo 2501"
[30] "Stat 1060 and Psyo 2501"
[31] "STAT 1060 and PSYO 2501"
[32] "STAT 1060 and STAT 2080"
[33] "STAT 1060 PSYO 2501"
[34] "STAT 1060 PSYO2501"
```

```
[35] "STAT 1060, 2060, 2080"

[36] "STAT 1060, PSYO 2080"

[37] "stat 1060, psyo 2501"

[38] "STAT 1060, PSYO 2501"

[39] "STAT 1060, PSYO 2501, PSYO 3502"

[40] "STAT 1060, PSYO 2501, STAT 2080,"

[41] "STAT 1060, PSYO2501, PSYO3502,"

[42] "STAT 1060, STAT 2060, PSYO 2501"

[43] "STAT 1060, STAT 2060, STAT 2080, STAT 2450, PSYO 2501"

[44] "STAT 1060, STAT 2080, and PSYO 2501"

[45] "STAT 1060, STAT 2080, PSYO 2501"

[46] "STAT 1060, STAT 2080, PSYO 2501, PSYO 3502"

[47] "STAT 1060, STAT 2080, STAT 2450"

[48] "STAT 1060, STAT 2501"

[49] "STAT 1060, STAT2060, STAT 2080, PSYO 2501"

[50] "STAT 1060; PSYO 2501"

[51] "STAT 1060; PSYO2501"

[52] "STAT 1060; STAT 2080"

[53] "STAT 1060; STAT 2501"

[54] "STAT 2060"

[55] "STAT 2080"

[56] "stat 2080, psyo 2501"

[57] "STAT 2080, STAT 1060, PSYO 2501"

[58] "STAT 2501"

[59] "Stat1060"

[60] "STAT1060"

[61] "STAT1060 & PSYO2501"
```

```
[62] "STAT1060 and PSYO2501"
[63] "Stat1060, 2060, 2080 PSYO2501, 3502"
[64] "Stat1060, and psyo2501"
[65] "Stat1060, psyo 2501"
[66] "STAT1060, PSYO 2501, PSYO 3502"
[67] "STAT1060, PSYO2501"
[68] "stat1060, stat 2080"
[69] "STAT1060, STAT2080, PSYO2501"
[70] "STAT1060, STAT2501"
[71] "STAT1060; PSYO2501"
[72] "STAT1060; PSYO2501; ENGM2032"
[73] "Stats 1060, PSYO 2501"
[74] "Stats 1060/ENGM2032"
[75] "STATS 1060; PSYO 2501"
[76] "yes"
```

Data Preparation

Not all variables are currently the data type we'd like them to be. Based on what you see when you look at the structure of the df and using the data's codebook, which columns are not currently in the correct data type? What changes do you need to make?

Answer: Program type should not be numeric but should be factor. Id should also not be numeric but should be character.

In the chunk below, perform simple data cleaning to retype the columns as needed.

In an effort to simplify things at least a little, let's drop the columns that are item responses, keeping participant info and composite survey scores. Use the codebook to identify which columns are item responses, then modify the dataframe to drop those columns. We're assigning this to a new object, `math.anx`, to keep the original data intact.

Descriptive Statistics

Center and Spread

We can see basic summary statistics from the wide data (`math.anx`):

```
     id                 age        gender.category           ethnicity.r
 Length:453         Min.   :16.00   Min.   :0.0000   Asian          : 41
 Class :character   1st Qu.:19.00   1st Qu.:0.0000   Black          :  9
 Mode  :character   Median :21.00   Median :1.0000   Hispanic/Latinx: 13
                    Mean   :21.96   Mean   :0.6954   Middle Eastern : 26
                    3rd Qu.:23.00   3rd Qu.:1.0000   Mixed/Other    : 17
                    Max.   :53.00   Max.   :2.0000   White          :347


 program.type              major                  dal.course.code
 No :314      Psychology      : 78   STAT 1060           : 76
 Yes:139      Neuroscience    : 50   STAT 1060, PSYO 2501: 25
              undecided       : 48   PSYO 2501           : 19
              psychology      : 28   STAT1060            : 17
              Medical Sciences: 20   STAT 1060; PSYO 2501: 10
              Undecided       : 18   (Other)             :110
              (Other)         :211   NA's                :196
   faculty           stats        stats.history    self.efficacy
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :1.000
 1st Qu.:1.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:3.500
 Median :1.0000   Median :0.0000   Median :1.0000   Median :4.000
 Mean   :0.9978   Mean   :0.3296   Mean   :0.6887   Mean   :3.908
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:4.375
 Max.   :2.0000   Max.   :1.0000   Max.   :1.0000   Max.   :5.000
                  NA's   :1
```

```
   asi.phys        asi.cog        asi.social         asi
Min.   :1.000   Min.   :1.000   Min.   :1.167   Min.   :1.056
1st Qu.:1.500   1st Qu.:1.333   1st Qu.:2.333   1st Qu.:1.833
Median :1.833   Median :2.167   Median :2.833   Median :2.333
Mean   :2.147   Mean   :2.276   Mean   :2.936   Mean   :2.452
3rd Qu.:2.833   3rd Qu.:3.000   3rd Qu.:3.500   3rd Qu.:3.000
Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :4.875

   perf.cog        perf.sp         frost.ps      frost.com       frost.da
Min.   :1.000   Min.   :1.000   Min.   :1.00   Min.   :1.00   Min.   :1.000
1st Qu.:2.000   1st Qu.:4.333   1st Qu.:3.25   1st Qu.:2.00   1st Qu.:2.500
Median :3.000   Median :5.333   Median :3.75   Median :2.80   Median :3.250
Mean   :2.833   Mean   :5.122   Mean   :3.66   Mean   :2.85   Mean   :3.146
3rd Qu.:3.667   3rd Qu.:6.333   3rd Qu.:4.25   3rd Qu.:3.60   3rd Qu.:4.000
Max.   :5.000   Max.   :7.000   Max.   :5.00   Max.   :5.00   Max.   :5.000
NA's   :7
 stat.anx.tc      stat.anx.i     stat.anx.ah     stat.anx.ws
Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
1st Qu.:3.333   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:1.000
Median :3.667   Median :2.667   Median :3.000   Median :1.667
Mean   :3.670   Mean   :2.738   Mean   :3.176   Mean   :1.792
3rd Qu.:4.333   3rd Qu.:3.333   3rd Qu.:4.000   3rd Qu.:2.000
Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
NA's   :3       NA's   :2       NA's   :1       NA's   :1
 stat.anx.fst    stat.anx.sc       math.anx
Min.   :1.000   Min.   :1.000   Min.   :1.000
1st Qu.:1.333   1st Qu.:1.000   1st Qu.:2.000
```

```
 Median :2.000   Median :2.000   Median :2.556
 Mean   :2.095   Mean   :2.048   Mean   :2.614
 3rd Qu.:2.667   3rd Qu.:2.625   3rd Qu.:3.222
 Max.   :5.000   Max.   :5.000   Max.   :5.000
 NA's   :3       NA's   :1       NA's   :2
```

But that's kind of a mess, and it doesn't let us use the summary statistics for any purpose beyond just looking at them. We can use the long version to calculate summary statistics for each survey score.

The dataset is not currently in a tidy format, at least not for our purposes. The survey scores are spread across multiple columns. If we want to use a survey score as a grouping variable or calculate summary statistics, we need to reshape the data so that each survey score is in a single column.

Now pivot the data so that the survey scores are in a single column. You'll want to retain the participant information in the long format as well.

```
# A tibble: 6 x 13
  id          age gender.category ethnicity.r program.type major dal.course.code
  <chr>     <dbl>           <dbl> <fct>       <fct>        <fct> <fct>
1 12490341~    21               1 White       No           Psyc~ PSYO 2501, STA~
2 12490341~    21               1 White       No           Psyc~ PSYO 2501, STA~
3 12490341~    21               1 White       No           Psyc~ PSYO 2501, STA~
4 12490341~    21               1 White       No           Psyc~ PSYO 2501, STA~
5 12490341~    21               1 White       No           Psyc~ PSYO 2501, STA~
6 12490341~    21               1 White       No           Psyc~ PSYO 2501, STA~
# i 6 more variables: faculty <dbl>, stats <dbl>, stats.history <dbl>,
#   self.efficacy <dbl>, Survey_Type <chr>, Score <dbl>
```

With the data in a tidy format, we can both calculate summary statistics and visualize the

data more easily. Let's start with summarizing mean, median, sd, min/max, and range:

```
      id                 age        gender.category          ethnicity.r
 Length:7248        Min.   :16.00   Min.   :0.0000   Asian          : 656
 Class :character   1st Qu.:19.00   1st Qu.:0.0000   Black          : 144
 Mode  :character   Median :21.00   Median :1.0000   Hispanic/Latinx: 208
                    Mean   :21.96   Mean   :0.6954   Middle Eastern : 416
                    3rd Qu.:23.00   3rd Qu.:1.0000   Mixed/Other    : 272
                    Max.   :53.00   Max.   :2.0000   White          :5552


 program.type              major                  dal.course.code
 No :5024     Psychology      :1248   STAT 1060           :1216
 Yes:2224     Neuroscience    : 800   STAT 1060, PSYO 2501: 400
              undecided       : 768   PSYO 2501           : 304
              psychology      : 448   STAT1060            : 272
              Medical Sciences: 320   STAT 1060; PSYO 2501: 160
              Undecided       : 288   (Other)             :1760
              (Other)         :3376   NA's                :3136
    faculty           stats         stats.history     self.efficacy
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :1.000
 1st Qu.:1.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:3.500
 Median :1.0000   Median :0.0000   Median :1.0000   Median :4.000
 Mean   :0.9978   Mean   :0.3296   Mean   :0.6887   Mean   :3.908
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:4.375
 Max.   :2.0000   Max.   :1.0000   Max.   :1.0000   Max.   :5.000
                  NA's   :16
 Survey_Type            Score
 Length:7248        Min.   :1.000
```

```
 Class :character   1st Qu.:1.889
 Mode  :character   Median :2.750
                    Mean   :2.847
                    3rd Qu.:3.667
                    Max.   :7.000
                    NA's   :20
```

```
# A tibble: 16 x 7
   Survey_Type   Mean Median    SD   Min   Max Range
   <chr>        <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>
 1 asi           2.45   2.33 0.794  1.06  4.88  3.82
 2 asi.cog       2.28   2.17 1.02   1     5     4
 3 asi.phys      2.15   1.83 0.955  1     5     4
 4 asi.social    2.94   2.83 0.920  1.17  5     3.83
 5 frost.com     2.85   2.8  0.980  1     5     4
 6 frost.da      3.15   3.25 1.01   1     5     4
 7 frost.ps      3.66   3.75 0.816  1     5     4
 8 math.anx      2.61   2.56 0.874  1     5     4
 9 perf.cog      2.83   3    0.928  1     5     4
10 perf.sp       5.12   5.33 1.43   1     7     6
11 stat.anx.ah   3.18   3    1.25   1     5     4
12 stat.anx.fst  2.09   2    0.894  1     5     4
13 stat.anx.i    2.74   2.67 0.999  1     5     4
14 stat.anx.sc   2.05   2    1.14   1     5     4
15 stat.anx.tc   3.67   3.67 0.966  1     5     4
16 stat.anx.ws   1.79   1.67 0.807  1     5     4
```

There's a problem with the summary statistics above. Missing values for some scores for some participants mean that the summary calculations are also missing. In some cases that may

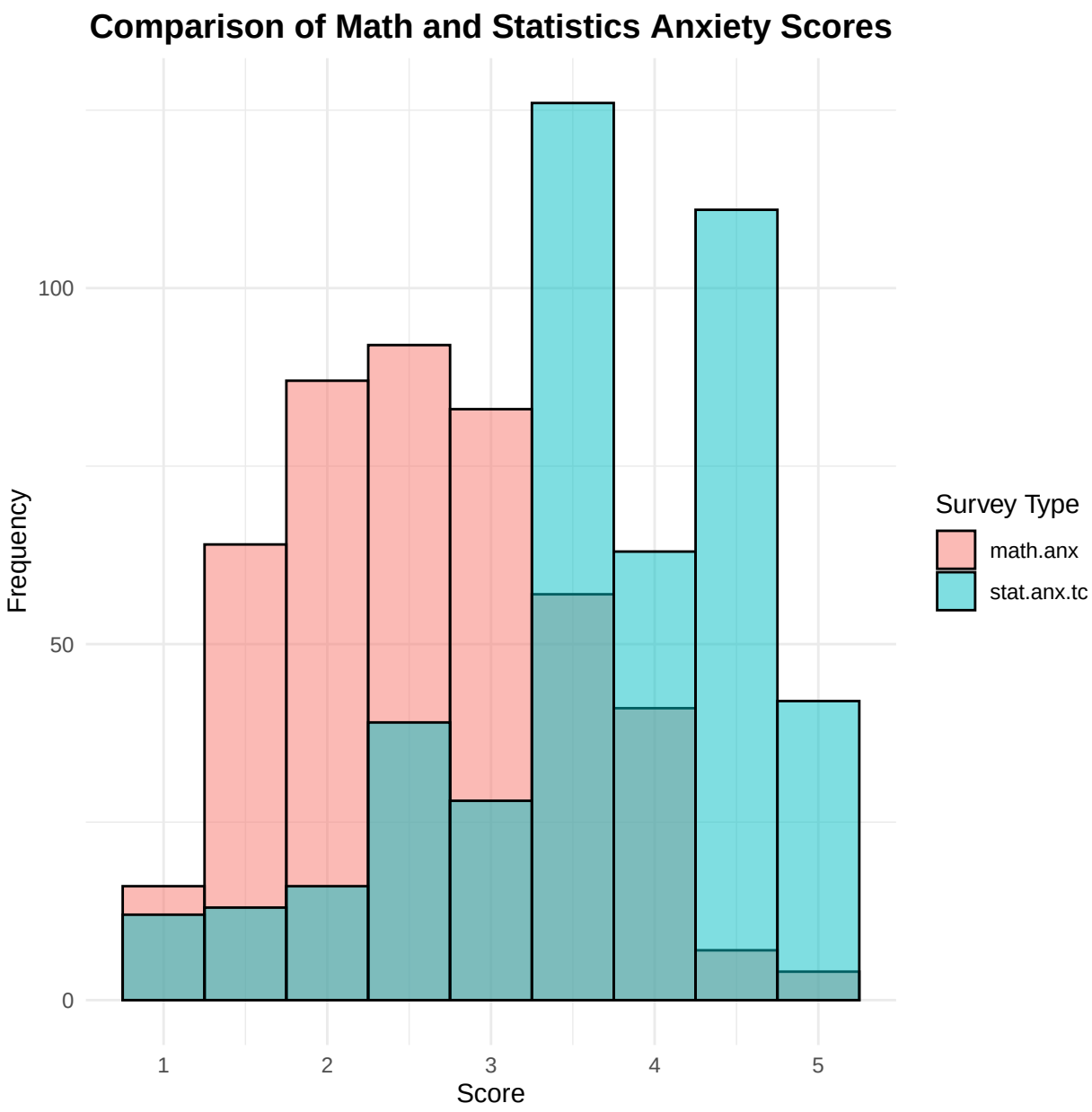
be what we want, but in this case we have enough data that we can just ignore missing values.

As for some reason I just could not load flextable, I will use kableExtra instead to create APA formatted table.

```
# A tibble: 16 x 7
   Survey_Type   Mean Median    SD   Min   Max Range
   <chr>        <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>
 1 asi           2.45   2.33  0.79  1.06  4.88  3.82
 2 asi.cog       2.28   2.17  1.02  1     5     4
 3 asi.phys      2.15   1.83  0.95  1     5     4
 4 asi.social    2.94   2.83  0.92  1.17  5     3.83
 5 frost.com     2.85   2.8   0.98  1     5     4
 6 frost.da      3.15   3.25  1.01  1     5     4
 7 frost.ps      3.66   3.75  0.82  1     5     4
 8 math.anx      2.61   2.56  0.87  1     5     4
 9 perf.cog      2.83   3     0.93  1     5     4
10 perf.sp       5.12   5.33  1.43  1     7     6
11 stat.anx.ah   3.18   3     1.25  1     5     4
12 stat.anx.fst  2.09   2     0.89  1     5     4
13 stat.anx.i    2.74   2.67  1     1     5     4
14 stat.anx.sc   2.05   2     1.14  1     5     4
15 stat.anx.tc   3.67   3.67  0.97  1     5     4
16 stat.anx.ws   1.79   1.67  0.81  1     5     4
```

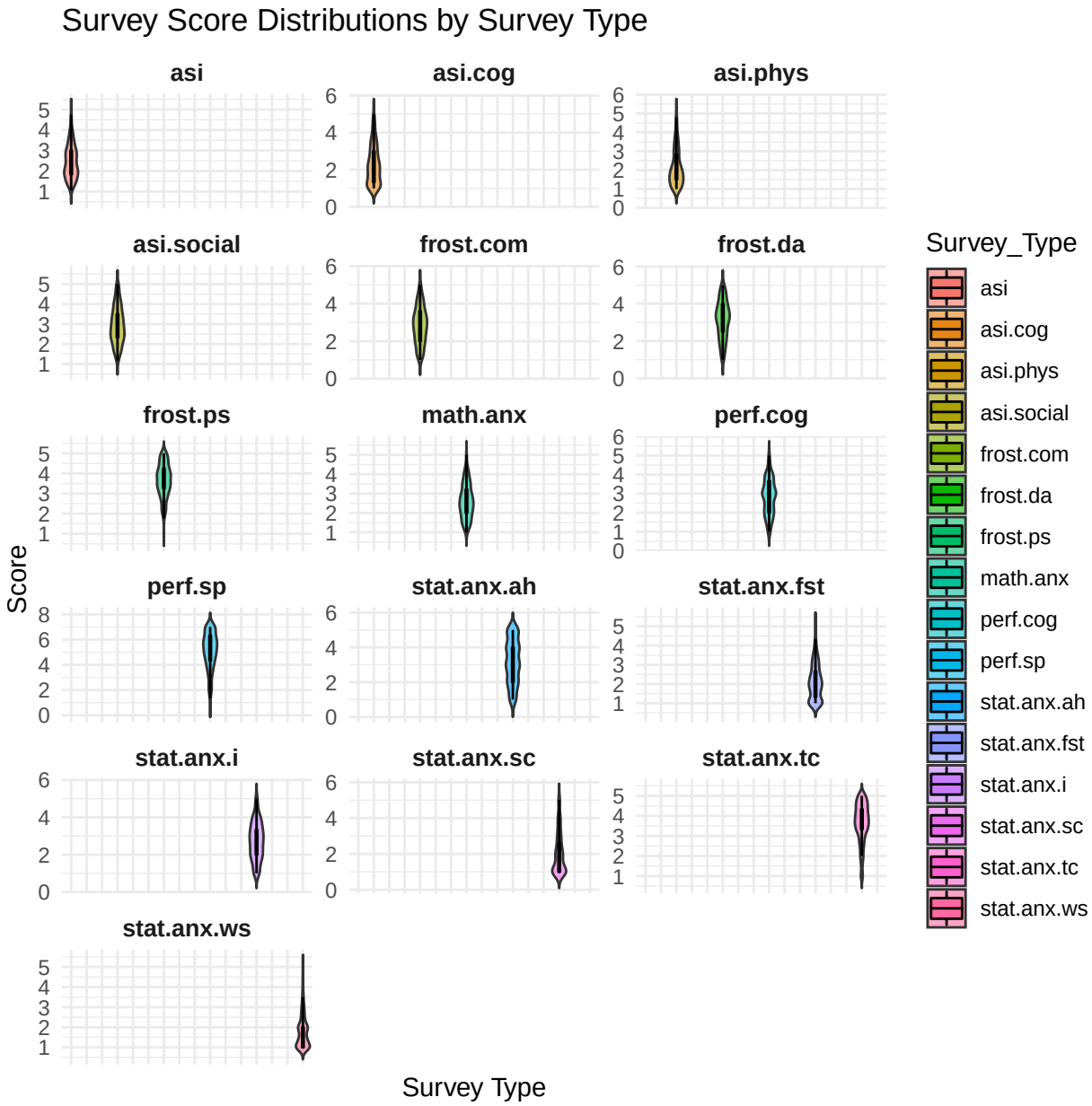
We can also visualize the data to get a better sense of the distribution of scores. Let's start with a histogram of the math anxiety scores.

Extra step: Math Anxiety vs. Statistics Anxiety Scores

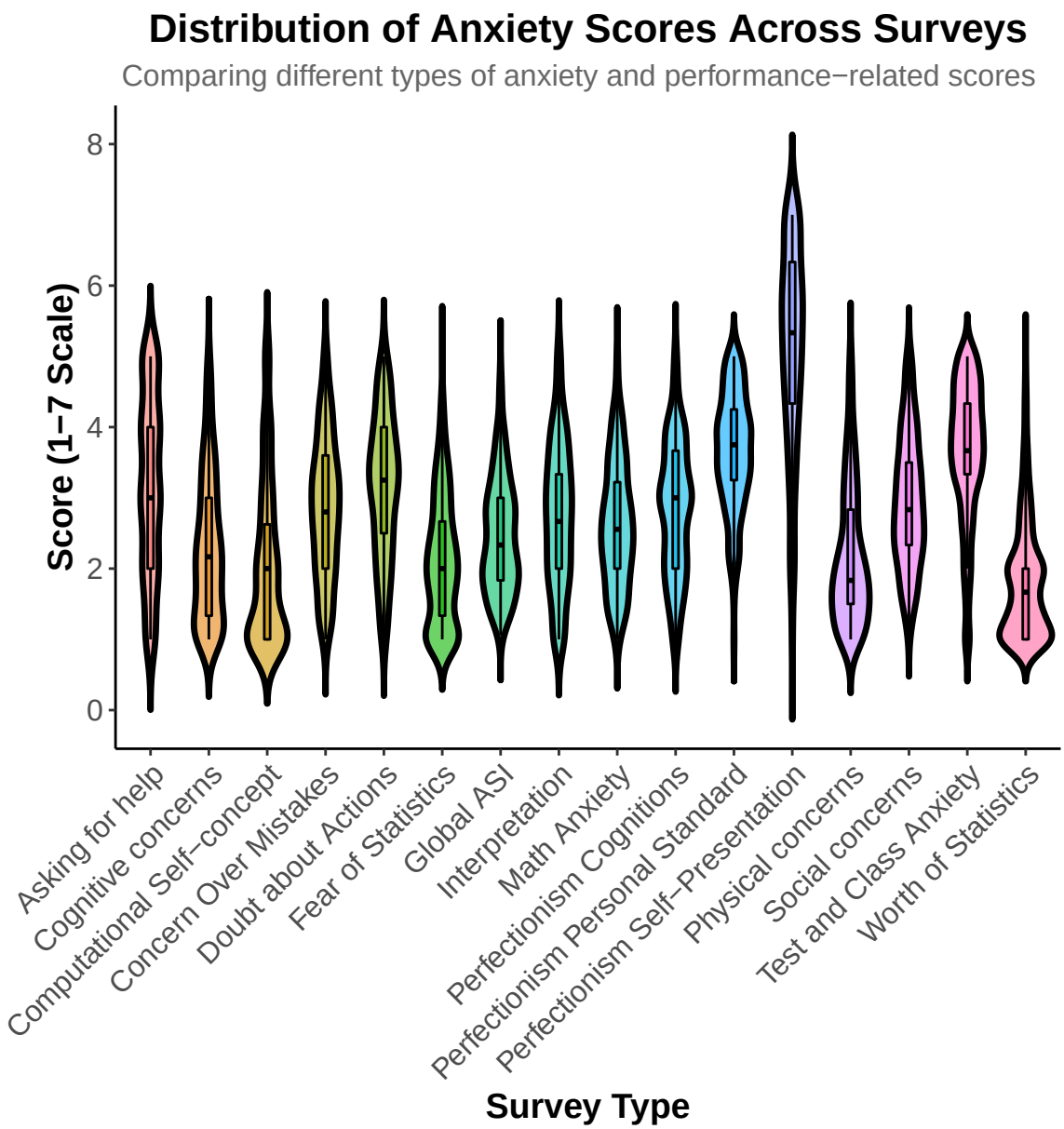


Now generate a violin plot of all the survey scores to see differences in distributions across surveys.

Trying to use facets in creating violin plots



Creating publication quality plots using theme and labs layers



I will summarize Math Anxiety Scores and include an additional measure (Interquartile Range, IQR).

1. Create a summary by program type and survey type

```
# A tibble: 32 x 9
   `Program Type` `Survey Type`        Mean Median    SD   Min   Max Range   IQR
   <fct>          <chr>               <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
 1 No             Asking for help      3.24   3     1.28  1     5       4   2.5
```

```
 2 No            Cognitive concerns   2.24   2.17  1.02  1     5      4    1.67
 3 No            Computational Self~  2.11   2     1.17  1     5      4    2
 4 No            Concern Over Mista~  2.93   3     0.99  1     5      4    1.4
 5 No            Doubt about Actions  3.13   3.25  1.04  1     5      4    1.5
 6 No            Fear of Statistics   2.1    2     0.88  1     5      4    1.33
 7 No            Global ASI           2.42   2.33  0.79  1.06  4.56   3.5  1.17
 8 No            Interpretation       2.88   3     1.01  1     5      4    1.67
 9 No            Math Anxiety         2.74   2.67  0.89  1     5      4    1.33
10 No            Perfectionism Cogn~  2.67   2.67  0.84  1     4      3    1.33
# i 22 more rows
```

2. Create a Violin Plot for Score Distribution by program type

3. Create a Box-plot Using the Summarized Data

Briefly describe what you see in the figures and tables you created. What do the summary statistics tell you about the data? What do the visualizations tell you about the data? What patterns do you see? Are there any other variables you think would be interesting to look at? Do the descriptives give you any ideas for interesting research questions to ask of these data?

In your answer below, use Quarto markdown references to the tables and figures you created. You can refer to the `d2mr-apaquarto` repository for examples of how to do this.

Descriptive Analysis of Survey Scores by Program Type

Summary Statistics Overview

The summary statistics presented in Table **?@tbl-summary-stats** provide a numerical overview of survey scores across different program types. We see that the mean values across survey types range between 2 and 4, with standard deviations around 0.8 to 1.3. The interquartile range (IQR) highlights some variation in score distributions, suggesting potential differences in how students from different programs experience these anxiety and performance-related factors.

Visualization Insights

The boxplots (Figure (**?**)) and violin plots (Figure (**?**)) offer a clearer depiction of score distributions.

Boxplots reveal that while median scores are similar between STEM and Non-STEM students, the spread of scores varies. Some survey types, such as "Asking for Help" and "Perfectionism Concern," exhibit more variability, particularly within STEM students.

Violin plots provide additional insights into the density of scores, showing that for certain survey types (e.g., "Fear of Statistics," "Math Anxiety"), there is a higher concentration of students reporting lower scores, while in other areas, scores are more evenly distributed.

Emerging Patterns

STEM vs. Non-STEM Differences

STEM students tend to have slightly lower self-reported anxiety levels in some domains (e.g., "Doubt in Abilities"), but the variability is larger in others (e.g., "Fear of Statistics").

Correlation

Now that we have a better sense of the distribution of the survey scores, let's look at the relationships between them. We can start by calculating the correlation matrix for the survey scores.

```
# Create a matrix-like object of only survey scores (use the wide data)


# Convert long format to wide format (each survey type as a column)
survey_wide <- math.anx %>%
  select(id, Survey_Type, Score) %>%
  pivot_wider(names_from = Survey_Type, values_from = Score)


# Remove ID column (only keep numeric survey scores)
survey_scores_only <- survey_wide %>% select(-id)
```

```r
# Calculate the correlation matrix for the survey scores using `cor()`

cor_matrix <- cor(survey_scores_only, use = "pairwise.complete.obs")

# Print the correlation matrix
print(cor_matrix)
```

```
                              Physical concerns Cognitive concerns
Physical concerns                    1.00000000          0.6104022
Cognitive concerns                   0.61040216          1.0000000
Social concerns                      0.42791870          0.4956596
Global ASI                           0.82894000          0.8634367
Perfectionism Cognitions             0.16369919          0.2732041
Perfectionism Self-Presentation      0.20709222          0.2685102
Perfectionism Personal Standard      0.09021988          0.1355116
Concern Over Mistakes                0.35952255          0.4502603
Doubt about Actions                  0.30020636          0.4592351
Test and Class Anxiety               0.17721312          0.1796388
Interpretation                       0.21965993          0.2484067
Asking for help                      0.22701725          0.2511848
Worth of Statistics                  0.15488534          0.0344268
Fear of Statistics                   0.16790178          0.1302702
Computational Self-concept           0.23987004          0.2037421
Math Anxiety                         0.27357472          0.3188073
                              Social concerns Global ASI
Physical concerns                  0.42791870  0.8289400
```

Cognitive concerns	0.49565959	0.8634367
Social concerns	1.00000000	0.7694673
Global ASI	0.76946731	1.0000000
Perfectionism Cognitions	0.25910804	0.2826583
Perfectionism Self-Presentation	0.49308778	0.3886977
Perfectionism Personal Standard	0.18086165	0.1638611
Concern Over Mistakes	0.43035140	0.5048039
Doubt about Actions	0.38014927	0.4662523
Test and Class Anxiety	0.23322814	0.2372117
Interpretation	0.18440695	0.2663737
Asking for help	0.31693391	0.3201459
Worth of Statistics	0.07629927	0.1069411
Fear of Statistics	0.06953729	0.1500991
Computational Self-concept	0.14855479	0.2427754
Math Anxiety	0.26011814	0.3464399

	Perfectionism Cognitions
Physical concerns	0.1636991913
Cognitive concerns	0.2732041013
Social concerns	0.2591080409
Global ASI	0.2826583229
Perfectionism Cognitions	1.0000000000
Perfectionism Self-Presentation	0.2664517224
Perfectionism Personal Standard	0.5142173496
Concern Over Mistakes	0.3717105905
Doubt about Actions	0.2591455376
Test and Class Anxiety	0.1070434544
Interpretation	0.0500130468

Asking for help	0.0868452869
Worth of Statistics	-0.1154640536
Fear of Statistics	0.0008872539
Computational Self-concept	-0.0125320593
Math Anxiety	0.1080537730
	Perfectionism Self-Presentation
Physical concerns	0.20709222
Cognitive concerns	0.26851019
Social concerns	0.49308778
Global ASI	0.38869768
Perfectionism Cognitions	0.26645172
Perfectionism Self-Presentation	1.00000000
Perfectionism Personal Standard	0.29131917
Concern Over Mistakes	0.55070268
Doubt about Actions	0.41823450
Test and Class Anxiety	0.27374624
Interpretation	0.19306364
Asking for help	0.36731936
Worth of Statistics	-0.04312240
Fear of Statistics	0.01742979
Computational Self-concept	0.07631246
Math Anxiety	0.30460046
	Perfectionism Personal Standard
Physical concerns	0.09021988
Cognitive concerns	0.13551164
Social concerns	0.18086165
Global ASI	0.16386106

Perfectionism Cognitions	0.51421735
Perfectionism Self-Presentation	0.29131917
Perfectionism Personal Standard	1.00000000
Concern Over Mistakes	0.28286000
Doubt about Actions	0.12809854
Test and Class Anxiety	0.08910384
Interpretation	0.05993551
Asking for help	0.03566139
Worth of Statistics	-0.13208476
Fear of Statistics	-0.03903485
Computational Self-concept	-0.15299556
Math Anxiety	0.07759856

	Concern Over Mistakes	Doubt about Actions
Physical concerns	0.35952255	0.30020636
Cognitive concerns	0.45026034	0.45923508
Social concerns	0.43035140	0.38014927
Global ASI	0.50480391	0.46625226
Perfectionism Cognitions	0.37171059	0.25914554
Perfectionism Self-Presentation	0.55070268	0.41823450
Perfectionism Personal Standard	0.28286000	0.12809854
Concern Over Mistakes	1.00000000	0.52249297
Doubt about Actions	0.52249297	1.00000000
Test and Class Anxiety	0.27252702	0.25306174
Interpretation	0.28159070	0.26160905
Asking for help	0.39857539	0.30934985
Worth of Statistics	0.05667524	0.02508179
Fear of Statistics	0.09497412	0.10421510

Computational Self-concept	0.19279361	0.22770277
Math Anxiety	0.38419165	0.36252468

	Test and Class Anxiety	Interpretation
Physical concerns	0.17721312	0.21965993
Cognitive concerns	0.17963885	0.24840669
Social concerns	0.23322814	0.18440695
Global ASI	0.23721168	0.26637370
Perfectionism Cognitions	0.10704345	0.05001305
Perfectionism Self-Presentation	0.27374624	0.19306364
Perfectionism Personal Standard	0.08910384	0.05993551
Concern Over Mistakes	0.27252702	0.28159070
Doubt about Actions	0.25306174	0.26160905
Test and Class Anxiety	1.00000000	0.59758255
Interpretation	0.59758255	1.00000000
Asking for help	0.40468884	0.41472269
Worth of Statistics	0.17147396	0.28669050
Fear of Statistics	0.23629023	0.34096286
Computational Self-concept	0.35498764	0.46617663
Math Anxiety	0.62564167	0.65102083

	Asking for help	Worth of Statistics
Physical concerns	0.22701725	0.15488534
Cognitive concerns	0.25118483	0.03442680
Social concerns	0.31693391	0.07629927
Global ASI	0.32014593	0.10694111
Perfectionism Cognitions	0.08684529	-0.11546405
Perfectionism Self-Presentation	0.36731936	-0.04312240
Perfectionism Personal Standard	0.03566139	-0.13208476

Concern Over Mistakes	0.39857539	0.05667524
Doubt about Actions	0.30934985	0.02508179
Test and Class Anxiety	0.40468884	0.17147396
Interpretation	0.41472269	0.28669050
Asking for help	1.00000000	0.13966498
Worth of Statistics	0.13966498	1.00000000
Fear of Statistics	0.24941419	0.52637725
Computational Self-concept	0.30035710	0.49802552
Math Anxiety	0.43278254	0.23459598

	Fear of Statistics	Computational Self-concept
Physical concerns	0.1679017761	0.23987004
Cognitive concerns	0.1302701632	0.20374211
Social concerns	0.0695372868	0.14855479
Global ASI	0.1500991180	0.24277541
Perfectionism Cognitions	0.0008872539	-0.01253206
Perfectionism Self-Presentation	0.0174297923	0.07631246
Perfectionism Personal Standard	-0.0390348469	-0.15299556
Concern Over Mistakes	0.0949741150	0.19279361
Doubt about Actions	0.1042150962	0.22770277
Test and Class Anxiety	0.2362902305	0.35498764
Interpretation	0.3409628581	0.46617663
Asking for help	0.2494141931	0.30035710
Worth of Statistics	0.5263772466	0.49802552
Fear of Statistics	1.0000000000	0.53796229
Computational Self-concept	0.5379622946	1.00000000
Math Anxiety	0.2873601409	0.48724833

	Math Anxiety

```
Physical concerns                 0.27357472
Cognitive concerns                0.31880731
Social concerns                   0.26011814
Global ASI                        0.34643989
Perfectionism Cognitions          0.10805377
Perfectionism Self-Presentation   0.30460046
Perfectionism Personal Standard   0.07759856
Concern Over Mistakes             0.38419165
Doubt about Actions               0.36252468
Test and Class Anxiety            0.62564167
Interpretation                    0.65102083
Asking for help                   0.43278254
Worth of Statistics               0.23459598
Fear of Statistics                0.28736014
Computational Self-concept        0.48724833
Math Anxiety                      1.00000000
```

What a mess! Put this into an APA-style table using flextable and `flextable::theme_apa()`.

Sorry still can't use flextable. I will use another package: kableExtra

	Survey Score	Physical concerns	Cognitive concerns
1	Physical concerns	1.00	0.61
2	Cognitive concerns	0.61	1.00
3	Social concerns	0.43	0.50
4	Global ASI	0.83	0.86
5	Perfectionism Cognitions	0.16	0.27
6	Perfectionism Self-Presentation	0.21	0.27
7	Perfectionism Personal Standard	0.09	0.14

8	Concern Over Mistakes	0.36	0.45
9	Doubt about Actions	0.30	0.46
10	Test and Class Anxiety	0.18	0.18
11	Interpretation	0.22	0.25
12	Asking for help	0.23	0.25
13	Worth of Statistics	0.15	0.03
14	Fear of Statistics	0.17	0.13
15	Computational Self-concept	0.24	0.20
16	Math Anxiety	0.27	0.32

	Social concerns	Global ASI	Perfectionism Cognitions
1	0.43	0.83	0.16
2	0.50	0.86	0.27
3	1.00	0.77	0.26
4	0.77	1.00	0.28
5	0.26	0.28	1.00
6	0.49	0.39	0.27
7	0.18	0.16	0.51
8	0.43	0.50	0.37
9	0.38	0.47	0.26
10	0.23	0.24	0.11
11	0.18	0.27	0.05
12	0.32	0.32	0.09
13	0.08	0.11	-0.12
14	0.07	0.15	0.00
15	0.15	0.24	-0.01
16	0.26	0.35	0.11

	Perfectionism Self-Presentation	Perfectionism Personal Standard

1	0.21	0.09
2	0.27	0.14
3	0.49	0.18
4	0.39	0.16
5	0.27	0.51
6	1.00	0.29
7	0.29	1.00
8	0.55	0.28
9	0.42	0.13
10	0.27	0.09
11	0.19	0.06
12	0.37	0.04
13	-0.04	-0.13
14	0.02	-0.04
15	0.08	-0.15
16	0.30	0.08

	Concern Over Mistakes	Doubt about Actions	Test and Class Anxiety
1	0.36	0.30	0.18
2	0.45	0.46	0.18
3	0.43	0.38	0.23
4	0.50	0.47	0.24
5	0.37	0.26	0.11
6	0.55	0.42	0.27
7	0.28	0.13	0.09
8	1.00	0.52	0.27
9	0.52	1.00	0.25
10	0.27	0.25	1.00

11	0.28	0.26	0.60
12	0.40	0.31	0.40
13	0.06	0.03	0.17
14	0.09	0.10	0.24
15	0.19	0.23	0.35
16	0.38	0.36	0.63

	Interpretation	Asking for help	Worth of Statistics	Fear of Statistics
1	0.22	0.23	0.15	0.17
2	0.25	0.25	0.03	0.13
3	0.18	0.32	0.08	0.07
4	0.27	0.32	0.11	0.15
5	0.05	0.09	-0.12	0.00
6	0.19	0.37	-0.04	0.02
7	0.06	0.04	-0.13	-0.04
8	0.28	0.40	0.06	0.09
9	0.26	0.31	0.03	0.10
10	0.60	0.40	0.17	0.24
11	1.00	0.41	0.29	0.34
12	0.41	1.00	0.14	0.25
13	0.29	0.14	1.00	0.53
14	0.34	0.25	0.53	1.00
15	0.47	0.30	0.50	0.54
16	0.65	0.43	0.23	0.29

	Computational	Self-concept	Math Anxiety
1	0.24	0.27	
2	0.20	0.32	
3	0.15	0.26	

4	0.24	0.35
5	-0.01	0.11
6	0.08	0.30
7	-0.15	0.08
8	0.19	0.38
9	0.23	0.36
10	0.35	0.63
11	0.47	0.65
12	0.30	0.43
13	0.50	0.23
14	0.54	0.29
15	1.00	0.49
16	0.49	1.00

I guess that's an improvement, but still not great. If you want, try out the `apaTables` package to see if it can make a more readable correlation matrix.

Visualize the correlation matrix using a heatmap using the `corrplot` package.

Briefly describe what you see in the correlation matrix. Do you see any general trends? Anything potentially interesting?

Interpretation of the Correlation Matrix Heatmap

The heatmap displays the correlation coefficients between different survey scores, with **darker blue** shades representing stronger *positive* correlations and **darker red** shades indicating stronger *negative* correlations.

General Trends Observed:. Math Anxiety Shows Distinct Patterns

Math Anxiety appears to be more strongly correlated with Fear of Statistics and Statistics Anxiety, which makes sense given that they all relate to quantitative performance-related anxieties.

Interestingly, Math Anxiety does not seem to be as strongly correlated with social or physical anxieties, suggesting a domain-specific nature.

Cognitive and Performance Anxiety Correlations

Cognitive Performance and Cognitive Anxiety exhibit moderate correlations, suggesting that individuals with high cognitive anxiety may also struggle with performance-related tasks.

Perfectionism Concern shows moderate correlations with Doubt in Abilities, indicating that individuals with perfectionistic tendencies may experience self-doubt.

Lower Correlations for Some Constructs

Some constructs, such as Spatial Performance, show weaker correlations with most other variables, which suggests that spatial skills may be less influenced by anxiety factors in this dataset.

Potentially Interesting Insights:

The clustering of certain anxiety types suggests that anxiety may not be a single unified construct but rather have distinct components that interact in specific ways.

Investigating subgroup differences (e.g., STEM vs. Non-STEM students) could be interesting, as certain anxiety types may be more relevant in different academic or career paths.

The relationship between Cognitive Anxiety and Performance-Related Scores (e.g., Math Anxiety, Perfectionism Concern) might indicate that high anxiety could predict lower confidence or performance expectations.

Select at least one demographic variable to explore in relation to the survey scores. Build on the plot you created above to include the demographic variable as a grouping variable. You can use grouping aesthetics, faceting, or both.

As before, describe what you see in the scatterplot. Does the demographic variable you chose seem to be related to one or both survey scores? To the association between them? If you see any kind of relationships, what might they mean?

Interpretation of the Scatterplot

In Figure Figure **??**, the scatterplot illustrates the relationship between Global ASI and Math Anxiety, while incorporating ethnicity as a grouping variable.

Key Observations:

Positive Association Across Groups:

The trend lines for each ethnic group indicate a general positive relationship between Global ASI and Math Anxiety. This suggests that higher Global ASI scores tend to be associated with higher Math Anxiety scores, regardless of ethnicity.

Variability by Ethnicity:

Some ethnic groups appear to have steeper regression lines, suggesting a stronger association between Global ASI and Math Anxiety.

Others have flatter slopes, implying that Global ASI may not be as predictive of Math Anxiety for those groups.

Confidence Intervals:

The gray shaded areas around the regression lines represent confidence intervals. Some groups exhibit wider confidence intervals, indicating greater uncertainty in the estimates, possibly due to smaller sample sizes or greater individual variability.

Does Ethnicity Moderate the Relationship?

The differences in slopes suggest that the strength of the relationship between Global ASI and Math Anxiety varies by ethnicity.

Some groups show a steeper increase in Math Anxiety with rising Global ASI, while others have a weaker or even nearly flat relationship.

This indicates a potential moderation effect of ethnicity, meaning that the impact of Global ASI on Math Anxiety might not be uniform across different ethnic backgrounds.

Further moderation analysis will be performed to analyse the relationship.

Research Question 1: Is there a significant difference in Math Anxiety across different ethnic groups?

Hypotheses:

Null Hypothesis (H_0): The mean Math Anxiety scores do not differ across ethnic groups. Alternative Hypothesis (H_A): At least one ethnic group has a significantly different Math Anxiety mean score.

```
             Df Sum Sq Mean Sq F value Pr(>F)
ethnicity.r   5    5.4  1.0802   1.422  0.215
Residuals   445  338.0  0.7596
2 observations deleted due to missingness
```

As the p value is insignificant, the result suggests that the mean Math Anxiety scores do not differ across ethnic groups.

Research Question 2: Does Ethnicity Moderate the Relationship?

Now I aim to test whether ethnicity moderates the relationship between Global ASI (Anxiety Sensitivity Index) and Math Anxiety. Specifically, examining whether the effect of Global ASI on Math Anxiety varies across different ethnic groups.

Hypotheses

Main Effect Hypothesis.

- **Null Hypothesis (H_0)**: There is **no significant relationship** between Global ASI and Math Anxiety.
- **Alternative Hypothesis (H_A)**: There is a **significant positive relationship** between Global ASI and Math Anxiety.

$$\beta_1 \neq 0$$

where:

- β_1 represents the effect of Global ASI on Math Anxiety.

Moderation Hypothesis.

- **Null Hypothesis (H_0)**: The relationship between Global ASI and Math Anxiety **does not depend on ethnicity** (i.e., there is no interaction).
- **Alternative Hypothesis (H_A)**: The relationship between Global ASI and Math Anxiety **is moderated by ethnicity** (i.e., there is a significant interaction effect).

$$\beta_3 \neq 0$$

where:

- β_3 represents the **interaction effect** between Global ASI and ethnicity.

Regression Model for Moderation Analysis

The linear regression model including an **interaction term** between Global ASI and ethnicity is:

$$\text{Math Anxiety} = \beta_0 + \beta_1(\text{Global ASI}) + \beta_2(\text{Ethnicity}) + \beta_3(\text{Global ASI} \times \text{Ethnicity}) + \epsilon$$

where:

- β_0 is the intercept.
- β_1 is the main effect of Global ASI.
- β_2 represents the effect of ethnicity.
- β_3 represents the interaction effect (i.e., moderation effect).
- ϵ is the residual error.

Run the Linear Regression Model with Interaction

```
Call:
lm(formula = `Math Anxiety` ~ `Global ASI` * ethnicity.r, data = survey_wide)

Residuals:
     Min       1Q   Median       3Q      Max
-1.86290 -0.56199 -0.07265  0.51327  2.29991

Coefficients:
                                       Estimate Std. Error t value Pr(>|t|)
(Intercept)                             2.31219    0.42969   5.381 1.21e-07
`Global ASI`                            0.02862    0.16804   0.170   0.8648
ethnicity.rBlack                       -0.75726    1.04302  -0.726   0.4682
ethnicity.rHispanic/Latinx             -1.09981    0.88004  -1.250   0.2121
ethnicity.rMiddle Eastern              -0.70756    0.65577  -1.079   0.2812
ethnicity.rMixed/Other                 -0.79781    0.76510  -1.043   0.2976
ethnicity.rWhite                       -0.71210    0.45280  -1.573   0.1165
`Global ASI`:ethnicity.rBlack           0.42820    0.40291   1.063   0.2885
`Global ASI`:ethnicity.rHispanic/Latinx 0.76574    0.37411   2.047   0.0413
`Global ASI`:ethnicity.rMiddle Eastern  0.46155    0.26128   1.766   0.0780
`Global ASI`:ethnicity.rMixed/Other     0.57487    0.31628   1.818   0.0698
`Global ASI`:ethnicity.rWhite           0.37546    0.17673   2.125   0.0342

(Intercept)                            ***
`Global ASI`
ethnicity.rBlack
```