

# A Complete Analysis of the Algebra I Regents Exam

John Russell

2023-01-20

Below is a sample of the data that we have collected.

| Question.Number | Standard | Percentage.Correct |
|-----------------|----------|--------------------|
| 5               | A-REI    | 0.50               |
| 24              | F-IF     | 0.11               |
| 23              | F-IF     | 0.17               |
| 3               | A-SSE    | 0.30               |
| 14              | F-IF     | 0.40               |
| 11              | F-IF     | 0.32               |

## Classification

To classify the data by difficulty levels, we will use a common technique called “Jenks Natural Breaks Classification.” This method seeks to minimize the average deviation from the class mean while maximizing the deviation from the means of the other groups. The result is the arrangement of a set of values into “natural” classes.

```
library(BAMMtools)
breaks <- getJenksBreaks(df$Percentage.Correct, 4)
breaks
```

```
## [1] 0.09 0.30 0.47 0.75
```

The above output gives us three intervals of percentage scores. We will use these to classify the exam questions as “Easy”, “Normal”, or “Hard”.

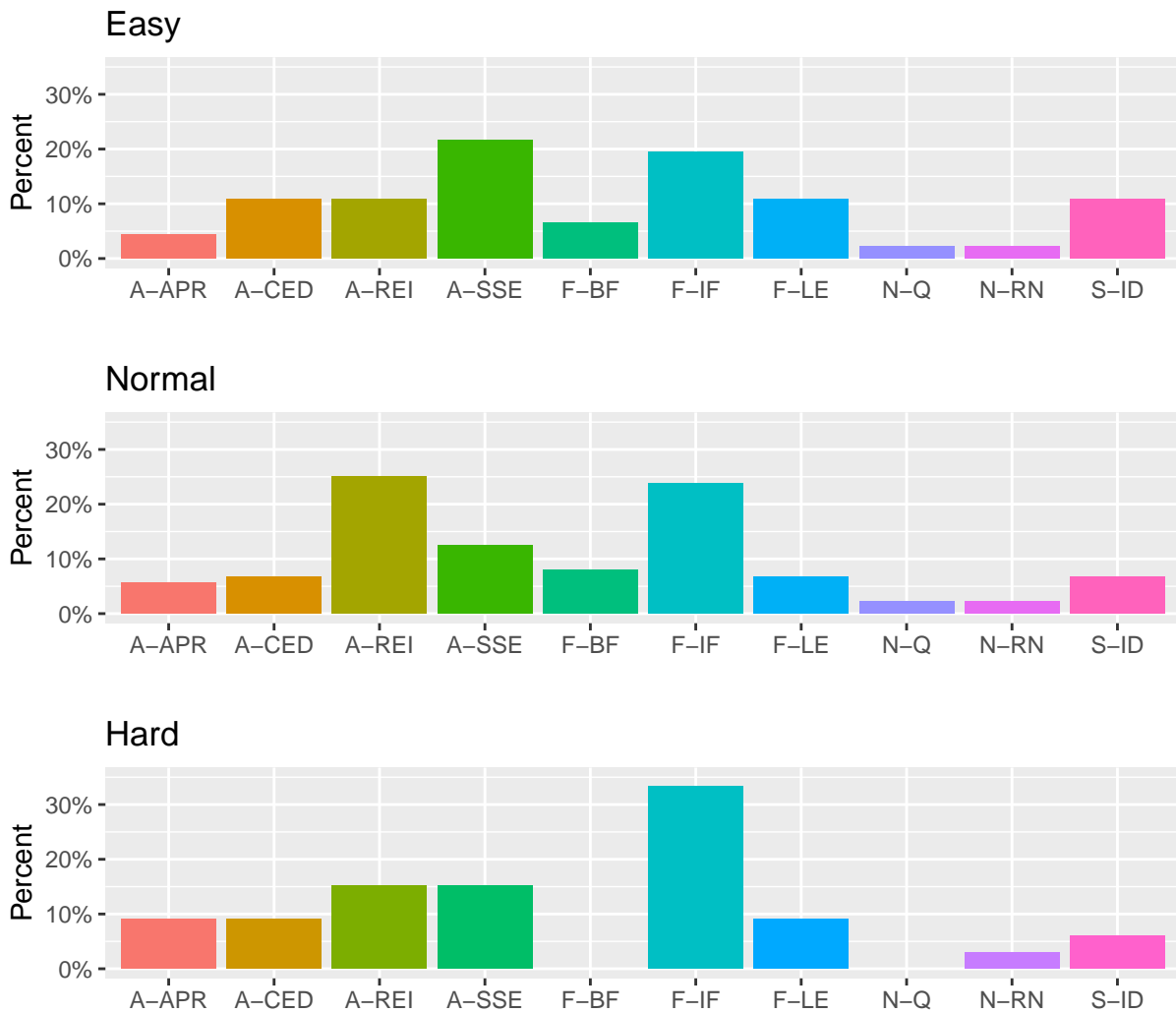
| Difficulty | Percentage Correct (Mean) |
|------------|---------------------------|
| Easy       | 56.02174                  |
| Normal     | 37.84091                  |
| Hard       | 22.57576                  |

Below is the same sample with a “Difficulty” column added.

| Question.Number | Standard | Percentage.Correct | Difficulty |
|-----------------|----------|--------------------|------------|
| 5               | A-REI    | 0.50               | Easy       |
| 24              | F-IF     | 0.11               | Hard       |
| 23              | F-IF     | 0.17               | Hard       |
| 3               | A-SSE    | 0.30               | Normal     |
| 14              | F-IF     | 0.40               | Normal     |
| 11              | F-IF     | 0.32               | Normal     |

## Plots

Distribution of Standards across Difficulty Levels



It's important to notice that F-IF, or the domain “Interpreting Functions” is found in high frequency across the three difficulty levels. Let's examine this more closely.

### Distribution of F-IF Standard across Difficulty Levels

