WP6

Prerequisite.

Read "Segregation, Integration, and Death: Evidence from the Korean War" by Connor Huff and Robert Schub

The original replication file is here. You are allowed to look at it, but you are not allowed to do the same operations.

Tasks

Replicate Figures 2 and 3

Use the tidyverse to replicate Figures 2 and 3. A complete replication must have the following:

1) A title clearly labeling what figure is being replicated
2) The same axis labels. It is not a requirement to have the same tick mark labels.
3) Both points and lines that are delineated to show which kind of battalions they represent.
4) The points and lines must be correct.
5) After each figure, explain how the article describes the figures.

Figure 2 uses the the Segregation_HS dataset

You will need to do some data wrangling on this dataset, which will require you to make another data frame from the original dataset. Here are some hints for this process.

n() will be necessary.

You can summarise multiple objects in the same summarise() command.

You will need to group the data by two variables.

The lines in the graph are the average fatality rate.

To get the points on the graph, you'll pass the original dataset separately to geom_point as well as a new set of aesthetics. The process looks like the following.

```
example <- ggplot(newdata, aes(x = x, y = y, group = grouping_var,
    linetype = linetype_var)) + GEOM_THAT_MAKES_A_LINE + geom_point(data = origData,
    aes(x = x, y = y, color = color_var), alpha = 0.5)
```

Figure 3 uses the Integration_HS dataset

You will need to perform a similar set of steps on the Integration_HS dataset as you did in Figure 2.

Replicate Tables 1 and 3 as coefficient plots

A complete answer will have the correct coefficient and confidence interval (use HC3 standard errors). The coefficient plot must clearly label which Table is being replicated.

In order to fully replicate the table, you will need to look up how to include fixed effects in `lm_robust()`.

Table 1 uses the Segregation_HS dataset. Table 2 uses the Integration_HS dataset

```
### Coefficient plot for Table 1
```

For Table 3, only replicate the Average Personnel models.

```
### Coefficient plot for Table 3
```