Analysis: Diversity & Innovation in Academia

Adham Rishmawi

September 13, 2022

Prompt

Does the novelty of sexes differ in different occupations?

predictors/responses

my response variable is novelty and my predictor is fields and sexes!

```r
pi <- read_csv('https://sldr.netlify.app/data/phd_innovation.csv')
```

```
## Rows: 4195 Columns: 8
## -- Column specification --------------------------------------
## Delimiter: ","
## chr (3): field, race_ethnicity, sex
## dbl (5): year, prop_same_race, prop_same_sex, novelty, uptake
##
## i Use `spec()` to retrieve the full column specification for this da
## i Specify the column types or set `show_col_types = FALSE` to quiet
```

```r
flr <- lm(novelty ~ field,
          data = pi)
summary(flr)
```

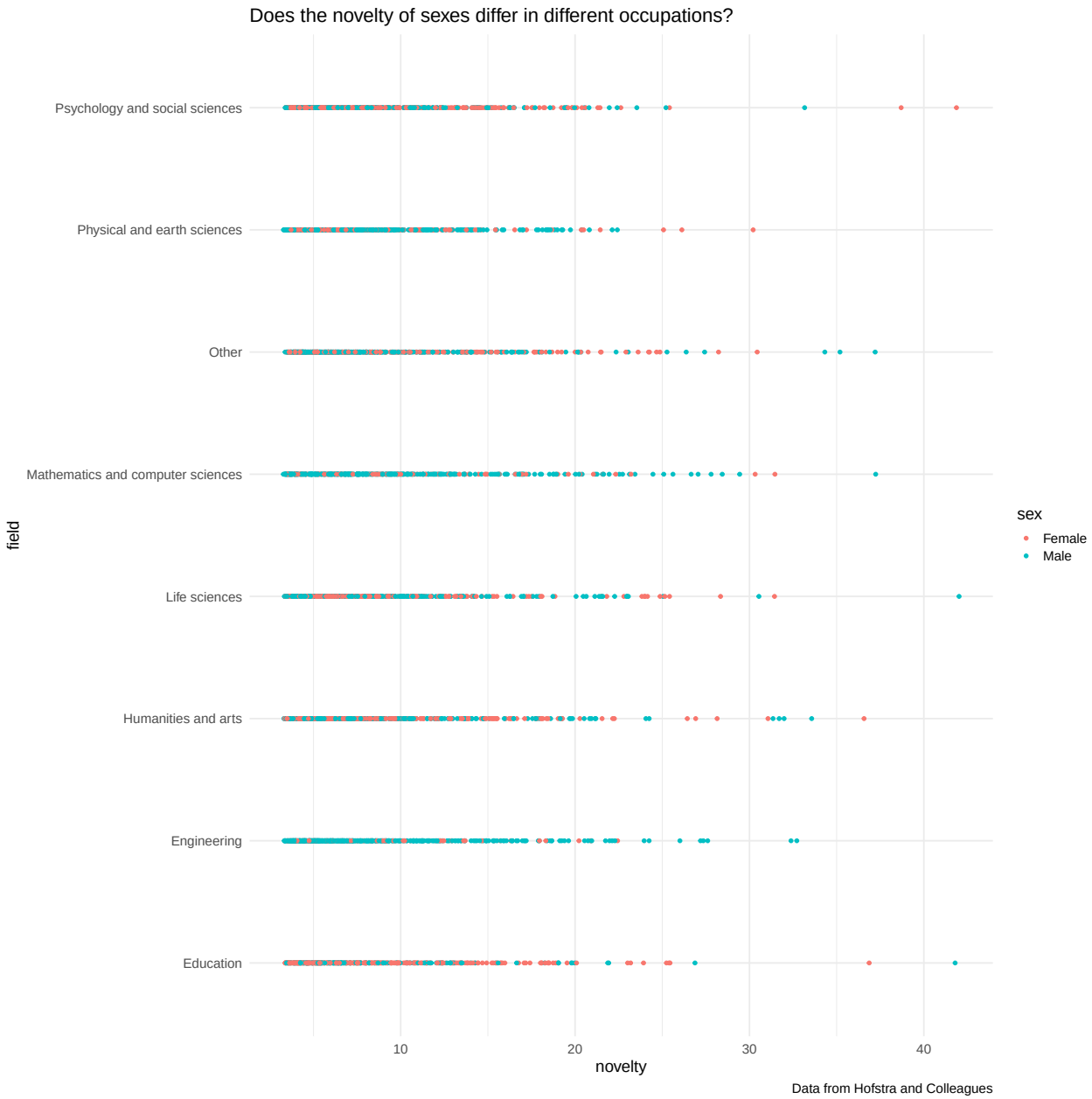
```
##
## Call:
## lm(formula = novelty ~ field, data = pi)
##
```

```
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.621 -3.540 -1.482  1.854 33.713
##
## Coefficients:
##                                            Estimate Std. Error t value
## (Intercept)                                 8.33752    0.22342  37.318
## fieldEngineering                           -0.04986    0.31086  -0.160
## fieldHumanities and arts                    0.06927    0.30694   0.226
## fieldLife sciences                         -0.02936    0.31030  -0.095
## fieldMathematics and computer sciences      0.55748    0.31072   1.794
## fieldOther                                  0.44733    0.30933   1.446
## fieldPhysical and earth sciences           -0.52660    0.31379  -1.678
## fieldPsychology and social sciences         0.46625    0.31758   1.468
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.981 on 4187 degrees of freedom
## Multiple R-squared:  0.004483,   Adjusted R-squared:  0.002819
## F-statistic: 2.694 on 7 and 4187 DF,  p-value: 0.008757
```

This tells us good information about the predictors/responses that are going to be used!

#GRAPH

```
gf_point(novelty ~ field,
         coord_flip(), color = ~ sex,
         caption = "Data from Hofstra and Colleagues",
         title = "Does the novelty of sexes differ in different occupa
         data = pi) |>
  gf_refine(coord_flip())
```


Does the novelty of sexes differ in different occupations?
Psychology and social sciences
Physical and earth sciences
Other
Mathematics and computer sciences
Life sciences
Humanities and arts
Engineering
Education
field
10
20
30
40
novelty
sex
Female
Male
Data from Hofstra and Colleagues

##voice memo is included describing my graph