bootcamp-survey

Rick Gilmore

2017-08-15 09:27:23

Contents

Goals

- Download and clean data from 2017 R Bootcamp Survey
- Visualize data
- Prepare reports in `ioslides_presentation`, `pdf_document`, and `word_document` formats

Preliminaries

Load required packages.

```
library(tidyverse)
library(googlesheets)
```

Load data and examine

The survey data are stored in a Google Sheet. We'll use the `googlesheets` package to open it and create a data frame. Documentation about the package can be found here.

There are some idiosyncrasies in using the `googlesheets` package in an R Markdown document, so I created a separate R script, `get-bootcamp-googlesheet.R` to extract the survey data and save it to a CSV under `data/survey.csv`. We can then just load this file.

```
# Created test data set for testing.
# survey <- read_csv("../data/survey.csv")
survey <- read_csv("../data/survey-test.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   X1 = col_integer(),
##   Timestamp = col_datetime(format = ""),
##   R_exp = col_character(),
##   GoT = col_integer(),
##   Age_yrs = col_integer(),
##   Sleep_hrs = col_double(),
##   Fav_date = col_date(format = ""),
##   Tidy_data = col_character()
```

```
## )
survey
```

```
## # A tibble: 50 × 8
##       X1           Timestamp   R_exp   GoT Age_yrs Sleep_hrs   Fav_date
##    <int>              <dttm>   <chr> <int>   <int>     <dbl>     <date>
## 1      1 2017-08-15 09:22:20 limited     2      47  7.309440 2017-08-15
## 2      2 2017-08-15 09:22:20    some     4      39  7.723569 2017-08-15
## 3      3 2017-08-15 09:22:20    lots     7      29  6.890581 2017-08-15
## 4      4 2017-08-15 09:22:20    none     3      46  8.133869 2017-08-15
## 5      5 2017-08-15 09:22:20    none     2      47  9.785339 2017-08-15
## 6      6 2017-08-15 09:22:20 limited     2      45 10.422163 2017-08-15
## 7      7 2017-08-15 09:22:20     pro     8      28  6.923171 2017-08-15
## 8      8 2017-08-15 09:22:20    none     6      32  8.485941 2017-08-15
## 9      9 2017-08-15 09:22:20    lots     6      33  9.388522 2017-08-15
## 10    10 2017-08-15 09:22:20    lots     4      41  7.804343 2017-08-15
## # ... with 40 more rows, and 1 more variables: Tidy_data <chr>
```

Notice that the `get-bootcamp-googlesheet.R` script changed the names of the variables a bit. We may also want to modify the levels of the `R_exp` variable to make it an ordered factor.

```
(survey_responses <- unique(survey$R_exp))
```

```
## [1] "limited" "some"    "lots"    "none"    "pro"
```

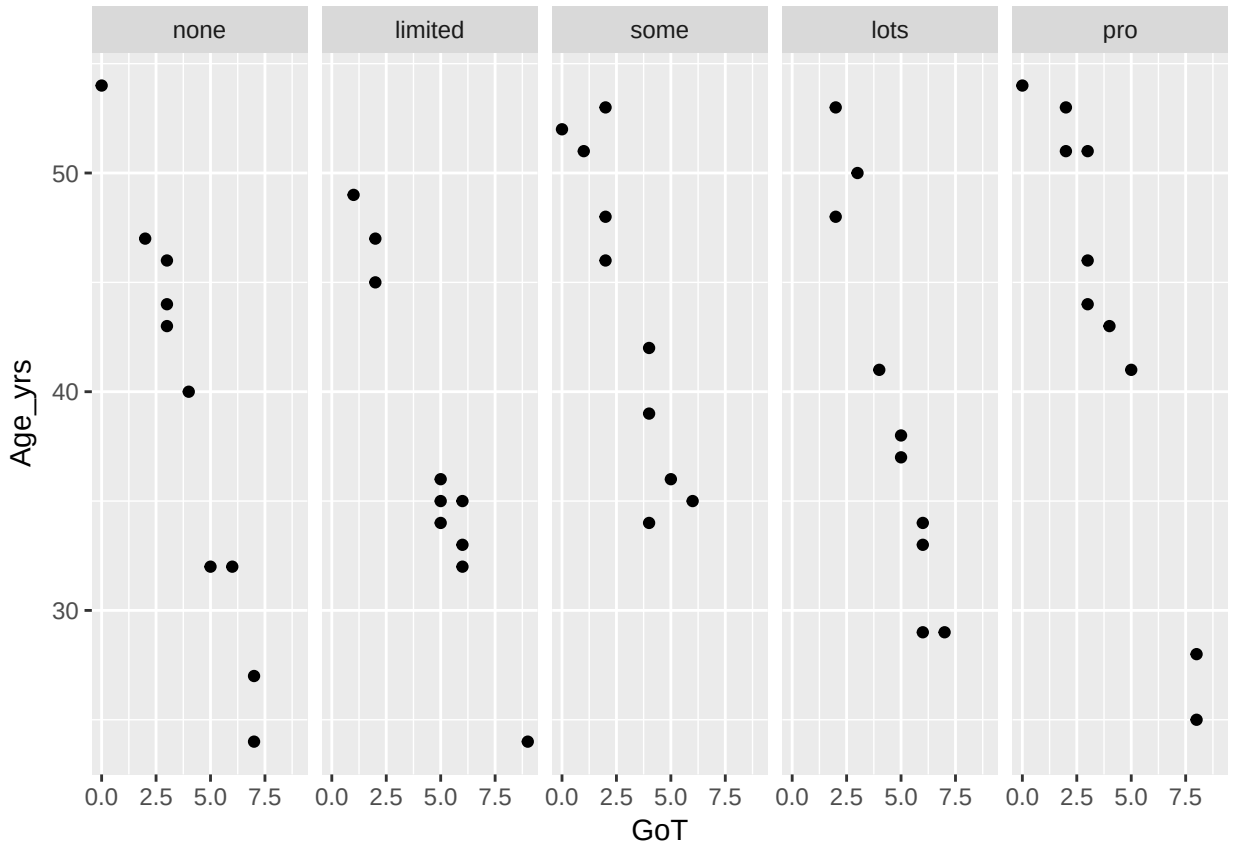
This shows us the different survey response values.

```
survey$R_exp <- ordered(survey$R_exp, levels=c("none",
                                               "limited",
                                               "some",
                                               "lots",
                                               "pro"))
```

Visualization and analysis

Now, we can ask important questions.

```
got_vs_r_exp <- survey %>%
  ggplot() +
  aes(x=GoT, y=Age_yrs) +
  facet_grid(. ~ R_exp) +
  geom_point()
got_vs_r_exp
```


none
limited
some
lots
pro
Age_yrs
50
40
30
0.0
2.5
5.0
7.5
GoT