On the Analysis of Survey Data using R

Ivan L. Saligumba, PhD

June 22, 2022

Abstract

In this paper we are going to perform some basic statistical analysis of survey data using the free statistical tool R.

Introduction

Background of the Study

Data

```
library(tidyverse)
```

```
survey_dt <- readxl::read_xlsx("data-activity-03/Survey Data-bYHFix.xlsx", sheet = 1)
survey_dt
```

```
## # A tibble: 893 x 99
##     DT11  DT12  DT13  DT21  DT22  DT23  DT24  DT31  DT32  DT33  DT41  DT42  DT43
##    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
##  1     4     4     4     5     4     3     3     3     2     2     5     5     5
##  2     2     3     4     4     3     3     3     2     4     2     5     4     3
##  3     4     4     4     4     4     4     4     4     4     4     4     4     4
##  4     4     1     4     3     1     1     3     3     3     3     4     4     4
##  5     4     4     5     4     5     3     3     5     4     4     5     4     5
##  6     3     3     3     4     4     4     4     2     2     2     2     2     2
##  7     3     4     4     4     3     4     4     4     4     4     4     3     4
##  8     4     4     4     3     3     3     4     4     3     4     4     4     4
##  9     3     4     4     5     4     4     4     3     3     3     4     4     3
## 10     4     4     4     3     4     3     4     4     4     4     5     5     5
## # ... with 883 more rows, and 86 more variables: DT51 <dbl>, DT52 <dbl>,
## #   DT53 <dbl>, DT61 <dbl>, DT62 <dbl>, DT63 <dbl>, DT71 <dbl>, DT72 <dbl>,
## #   DT73 <dbl>, DC11 <dbl>, DC12 <dbl>, DC13 <dbl>, DC21 <dbl>, DC22 <dbl>,
## #   DC23 <dbl>, DA11 <dbl>, DA12 <dbl>, DA13 <dbl>, DA21 <dbl>, DA22 <dbl>,
## #   DA23 <dbl>, WR11 <dbl>, WR12 <dbl>, WR13 <dbl>, WR21 <dbl>, WR22 <dbl>,
## #   WR23 <dbl>, WR31 <dbl>, WR32 <dbl>, WR33 <dbl>, WR41 <dbl>, WR42 <dbl>,
## #   WR43 <dbl>, WR51 <dbl>, WR52 <dbl>, WR53 <dbl>, PU11 <dbl>, PU12 <dbl>, ...
```

```
survey_dt %>%
  select(Gender:GPA)
```

```
## # A tibble: 893 x 7
##    Gender Year  Falcuty              Uni                      SRM SRM_Mark GPA
##     <dbl> <chr> <chr>                <chr>                  <dbl> <chr>    <chr>
##  1      2 3     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 2.84
##  2      2 3     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 3.04
##  3      2 2     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 2.71
##  4      2 2     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 3
##  5      2 2     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 2.68
##  6      2 3     Literature Pedagogy  VNU University of Edu~     1 3.2      3.46
##  7      2 2     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 2.25~
##  8      2 2     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 2.7
##  9      2 2     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 3.19
## 10      2 4     Mathematics Pedagogy Thai Nguyen Universit~     1 No resu~ 2.9
## # ... with 883 more rows
```

```
starwars %>%
  glimpse
```

```
## Rows: 87
## Columns: 14
## $ name       <chr> "Luke Skywalker", "C-3PO", "R2-D2", "Darth Vader", "Leia Or~
## $ height     <int> 172, 167, 96, 202, 150, 178, 165, 97, 183, 182, 188, 180, 2~
## $ mass       <dbl> 77.0, 75.0, 32.0, 136.0, 49.0, 120.0, 75.0, 32.0, 84.0, 77.~
## $ hair_color <chr> "blond", NA, NA, "none", "brown", "brown, grey", "brown", N~
## $ skin_color <chr> "fair", "gold", "white, blue", "white", "light", "light", "~
## $ eye_color  <chr> "blue", "yellow", "red", "yellow", "brown", "blue", "blue",~
## $ birth_year <dbl> 19.0, 112.0, 33.0, 41.9, 19.0, 52.0, 47.0, NA, 24.0, 57.0, ~
## $ sex        <chr> "male", "none", "none", "male", "female", "male", "female",~
## $ gender     <chr> "masculine", "masculine", "masculine", "masculine", "femini~
## $ homeworld  <chr> "Tatooine", "Tatooine", "Naboo", "Tatooine", "Alderaan", "T~
## $ species    <chr> "Human", "Droid", "Droid", "Human", "Human", "Human", "Huma~
## $ films      <list> <"The Empire Strikes Back", "Revenge of the Sith", "Return~
## $ vehicles   <list> <"Snowspeeder", "Imperial Speeder Bike">, <>, <>, <>, "Imp~
## $ starships  <list> <"X-wing", "Imperial shuttle">, <>, <>, "TIE Advanced x1",~
```

```
filter(starwars, species == "Human")
```

```
## # A tibble: 35 x 14
##    name     height  mass hair_color skin_color eye_color birth_year sex   gender
##    <chr>     <int> <dbl> <chr>      <chr>      <chr>          <dbl> <chr> <chr>
##  1 Luke Sk~    172    77 blond      fair       blue            19   male  mascu~
##  2 Darth V~    202   136 none       white      yellow          41.9 male  mascu~
##  3 Leia Or~    150    49 brown      light      brown           19   fema~ femin~
##  4 Owen La~    178   120 brown, gr~ light      blue            52   male  mascu~
##  5 Beru Wh~    165    75 brown      light      blue            47   fema~ femin~
##  6 Biggs D~    183    84 black      light      brown           24   male  mascu~
##  7 Obi-Wan~    182    77 auburn, w~ fair       blue-gray       57   male  mascu~
##  8 Anakin ~    188    84 blond      fair       blue            41.9 male  mascu~
##  9 Wilhuff~    180    NA auburn, g~ fair       blue            64   male  mascu~
```

```
## 10 Han Solo    180    80 brown      fair       brown           29   male  mascu~
## # ... with 25 more rows, and 5 more variables: homeworld <chr>, species <chr>,
## #   films <list>, vehicles <list>, starships <list>
```

```
starwars %>%
  # filter(sex == "male" & species == "Human")
  filter(birth_year > 20 & hair_color == "brown")
```

```
## # A tibble: 7 x 14
##   name      height  mass hair_color skin_color eye_color birth_year sex   gender
##   <chr>      <int> <dbl> <chr>      <chr>      <chr>           <dbl> <chr> <chr>
## 1 Beru Whi~    165    75 brown      light      blue             47 fema~ femin~
## 2 Chewbacca    228   112 brown      unknown    blue            200 male  mascu~
## 3 Han Solo     180    80 brown      fair       brown            29 male  mascu~
## 4 Wedge An~    170    77 brown      fair       hazel            21 male  mascu~
## 5 Qui-Gon ~    193    89 brown      fair       blue             92 male  mascu~
## 6 Cliegg L~    183    NA brown      fair       blue             82 male  mascu~
## 7 Padmé Am~    165    45 brown      light      brown            46 fema~ femin~
## # ... with 5 more variables: homeworld <chr>, species <chr>, films <list>,
## #   vehicles <list>, starships <list>
```

Methodology

Results and Discussion

Conclusion

References