Exercise_4_811

Marika Olijar

10/15/2021

Download the food_coded.csv file2. Load the CSV file into your R environment. Open the codebook_food.docx file for guidance.

3. Extract the first 95 rows.
4. Look at the following variables using both name and column index/number: GPA calories_chicken drink fav_cuisine father_profession mother_profession

```
colnames(food_coded)
```

```
##  [1] "GPA"                          "Gender"
##  [3] "breakfast"                    "calories_chicken"
##  [5] "calories_day"                 "calories_scone"
##  [7] "coffee"                       "comfort_food"
##  [9] "comfort_food_reasons"         "comfort_food_reasons_coded"
## [11] "cook"                         "comfort_food_reasons_coded.1"
## [13] "cuisine"                      "diet_current"
## [15] "diet_current_coded"           "drink"
## [17] "eating_changes"               "eating_changes_coded"
## [19] "eating_changes_coded1"        "eating_out"
## [21] "employment"                   "ethnic_food"
## [23] "exercise"                     "father_education"
## [25] "father_profession"            "fav_cuisine"
## [27] "fav_cuisine_coded"            "fav_food"
## [29] "food_childhood"               "fries"
## [31] "fruit_day"                    "grade_level"
## [33] "greek_food"                   "healthy_feeling"
## [35] "healthy_meal"                 "ideal_diet"
## [37] "ideal_diet_coded"             "income"
## [39] "indian_food"                  "italian_food"
## [41] "life_rewarding"               "marital_status"
## [43] "meals_dinner_friend"          "mother_education"
## [45] "mother_profession"            "nutritional_check"
## [47] "on_off_campus"                "parents_cook"
## [49] "pay_meal_out"                 "persian_food"
## [51] "self_perception_weight"       "soup"
## [53] "sports"                       "thai_food"
## [55] "tortilla_calories"            "turkey_calories"
## [57] "type_sports"                  "veggies_day"
## [59] "vitamins"                     "waffle_calories"
## [61] "weight"
```

```
food_index <- food_coded[,c(1,4,16,26,25,45)]
head(food_index)
```

```
##     GPA calories_chicken drink    fav_cuisine father_profession
## 1   2.4              430     1 Arabic cuisine         profesor 
## 2 3.654              610     2        Italian     Self employed
## 3   3.3              720     1        italian     owns business
## 4   3.2              430     2       Turkish          Mechanic 
## 5   3.5              720     2       Italian                 IT
## 6  2.25              610     2        African       Taxi Driver
##           mother_profession
## 1                unemployed
## 2                 Nurse RN 
## 3             owns business
## 4 Special Education Teacher
## 5  Substance Abuse Conselor
## 6              Hair Braider
```

New variable for how healthy each person feels but convert the scale from 1 to 10 to 1 to 100

```
food_coded$healthy_feeling
```

```
##   [1]  2  5  6  7  6  4  4  3  7  3  9  1  9  8  2  6  7  8  6  4  5  8  2  4  5
##  [26]  8  9  9  4  9  7  5  5  7  1  2  7  4  6  3 10  6  6  6  8  3  4  8  2  9
##  [51]  8  8  1  5 10  8  1  9  4  7  3  2  2  8  3  3  3  2  8  3  3  5  3  1  8
##  [76]  6  4  4  8  1  4  2  8  4  9  7  3  5  7  7  7  5  8  6  7 10  2  1  8  3
## [101]  2  3  7  4  9  2  7  5  6  5  8  9 10  9  7 10  5  9  5  7  5  5  6  1  3
```

```
food_coded.new <- food_coded
food_coded.new$health_feel_scaled <- (food_coded$healthy_feeling*10)
food_coded.new$health_feel_scaled
```

```
##   [1]  20  50  60  70  60  40  40  30  70  30  90  10  90  80  20  60  70  80
##  [19]  60  40  50  80  20  40  50  80  90  90  40  90  70  50  50  70  10  20
##  [37]  70  40  60  30 100  60  60  60  80  30  40  80  20  90  80  80  10  50
##  [55] 100  80  10  90  40  70  30  20  20  80  30  30  30  20  80  30  30  50
##  [73]  30  10  80  60  40  40  80  10  40  20  80  40  90  70  30  50  70  70
##  [91]  70  50  80  60  70 100  20  10  80  30  20  30  70  40  90  20  70  50
## [109]  60  50  80  90 100  90  70 100  50  90  50  70  50  50  60  10  30
```

6 Filter to students who are female and have GPAs that are above 3

```
gpa_female <- food_coded[food_coded$Gender=="1" & food_coded$GPA > 3.0,]
head(gpa_female)
```

```
##     GPA Gender breakfast calories_chicken calories_day calories_scone coffee
## 2 3.654      1         1              610            3            420      2
## 3   3.3      1         1              720            4            420      2
## 4   3.2      1         1              430            3            420      2
## 5   3.5      1         1              720            2            420      2
## 8   3.3      1         1              720            3            420      1
## 9   3.3      1         1              430          NaN            420      1
##                        comfort_food
## 2       chocolate, chips, ice cream
## 3   frozen yogurt, pizza, fast food
## 4  Pizza, Mac and cheese, ice cream
## 5      Ice cream, chocolate, chips 
## 8  Ice cream, cheeseburgers, chips.
## 9          Donuts, ice cream, chips
##
```

```
## 2
## 3
## 4
## 5
## 8 I eat comfort food when im stressed out from school(finals week), when I`m sad, or when i am deali
## 9
##   comfort_food_reasons_coded cook comfort_food_reasons_coded.1 cuisine
## 2                          1    3                            1       1
## 3                          1    1                            1       3
## 4                          2    2                            2       2
## 5                          1    1                            1       2
## 8                          1    3                            1       1
## 9                          2    3                            2       1
##
## 2          I eat about three times a day with some snacks. I try to eat healthy but it doesn't alway
## 3                                                        toast and fruit for breakfast, salad for lu
## 4                                                                      College diet, cheap and easy
## 5 I try to eat healthy but often struggle because of living on campus. I still try to keep the choic
## 8                                                                                           I eat a
## 9
##   diet_current_coded drink
## 2                  2     2
## 3                  3     1
## 4                  2     2
## 5                  2     2
## 8                  1     2
## 9                  1     1
##
## 2
## 3                                                                        sometimes choosing to eat fa
## 4
## 5 I have eaten generally the same foods but I do find myself eating the same food frequently due to
## 8                                                            Freshmen year i ate very unhealthy, but
## 9
##   eating_changes_coded eating_changes_coded1 eating_out employment ethnic_food
## 2                    1                     2          2          2           4
## 3                    1                     3          2          3           5
## 4                    1                     3          2          3           5
## 5                    3                     4          2          2           4
## 8                    2                     5          2          2           2
## 9                    2                     8          5          2           5
##   exercise father_education     father_profession             fav_cuisine
## 2        1                2        Self employed                   Italian
## 3        2                2         owns business                  italian
## 4        3                2             Mechanic                  Turkish
## 5        1                4                    IT                 Italian
## 8        2                3          Business guy Anything american style.
## 9      NaN                5 High School Principal                  Seafood
##   fav_cuisine_coded fav_food                               food_childhood fries
## 2                 1        1 chicken and biscuits, beef soup, baked beans     1
## 3                 1        3                 mac and cheese, pizza, tacos     1
## 4                 3        1                Beef stroganoff, tacos, pizza     2
## 5                 1        3                Pasta, chicken tender, pizza      1
## 8                 5        1      chicken, cheesey potatoes, and hot dogs     1
## 9                 1        3                            Shrimp, spaghetti     1
##   fruit_day grade_level greek_food healthy_feeling
## 2         4           4          4               5
## 3         5           3          5               6
## 4         4           4          5               7
## 5         4           4          4               6
## 8         5           2          3               3
## 9         4           1          5               7
##                                                                                      healthy_meal
## 2             Grains, Veggies, (more of grains and veggies), small protein and fruit with dairy 
## 3                                        usually includes natural ingredients; nonprocessed food
## 4                                                       Fresh fruits& vegetables, organic meats 
## 5 A lean protein such as grilled chicken, green vegetables and  brown rice or other whole grain 
## 8                      A healthy meal has a piece of meat followed by a lot of fruit and veggies
## 9                                                                                       Colorful
##
## 2 Try to eat 5-6 small meals a day. While trying to properly distribute carbs, protein, fruits, veggi
## 3                                                                        i would say my ideal diet is
## 4                                                                      Healthy, fresh veggies/fruits
## 5                                   Ideally I would like to be able to eat healthier foods in order t
## 8       My ideal diet is filled with a lot of fruit and chicken. I also really enjoy eggs any type of
## 9                                                                                              The sa
##   ideal_diet_coded income indian_food italian_food life_rewarding
## 2                3      4           4            4              1
## 3                6      6           5            5              7
## 4                2      6           5            5              2
## 5                2      6           2            5              1
## 8                2      5           1            3              3
## 9                6      5           5            5              8
##   marital_status
## 2              2
## 3              2
## 4              2
## 5              1
## 8              1
## 9              2
##                                                 meals_dinner_friend
## 2                                            Pasta, steak, chicken 
## 3 chicken and rice with veggies, pasta, some kind of healthy recipe
## 4                  Grilled chicken \nStuffed Shells\nHomemade Chili
## 5           Chicken Parmesan, Pulled Pork, Spaghetti and meatballs 
## 8                                            chicken, steak, pasta 
## 9                                                Pasta, Fish, Steak
##   mother_education         mother_profession nutritional_check on_off_campus
## 2                4                 Nurse RN                  4             1
## 3                2             owns business                 4             2
## 4                4 Special Education Teacher                 2             1
## 5                5  Substance Abuse Conselor                 3             1
## 8                2                      cook                 4             1
## 9                5 Elementary School Teacher                 2             1
##   parents_cook pay_meal_out persian_food self_perception_weight soup sports
## 2            1            4            4                      3    1      1
## 3            1            3            5                      6    1      2
## 4            1            2            5                      5    1      2
## 5            1            4            2                      4    1      1
## 8            1            5            1                      3    1      2
## 9            2            3            5                      4    2      2
##   thai_food tortilla_calories turkey_calories type_sports veggies_day vitamins
## 2         2               725             690 Basketball            4        2
## 3         5              1165             500        none           5        1
## 4         5               725             690         nan           3        1
## 5         4               940             500    Softball           4        2
## 8         1               725             500        none           4        2
## 9         5               725             345        none           3        2
##   waffle_calories                   weight
## 2             900                      155
## 3             900 I'm not answering this.
## 4            1315            Not sure, 240
## 5             760                      190
## 8            1315                      137
## 9             760                      180
```

7 Find the mean and standard deviation for the following variables, and summarize them in a data frame: chicken_calories • tortilla_calories • turkey_calories • waffle_calories

```r
food_calories <- food_coded[,c(4, 54, 55, 60)]
?sapply
m <- sapply(food_calories, mean, na.rm = T)
sd <- sapply(food_calories, sd, na.rm = T)
calories_m_sd <- rbind(m, sd)
head(calories_m_sd)
```

```
##    calories_chicken thai_food tortilla_calories waffle_calories
## m          577.3200  3.336000          947.5806       1073.4000
## sd         131.2142  1.436528          202.0902        248.6671
```

8. Summarize GPA and weight within the gender and cuisine variables.

```r
class(food_coded$weight)
```

```
## [1] "character"
```

```r
food_coded$weight <- as.numeric(food_coded$weight)
```

```
## Warning: NAs introduced by coercion
```

```r
food_men <- food_coded[food_coded$Gender == 1,]
food_women <- food_coded[food_coded$Gender == 2,]

gpa_mean_men <- tapply(food_men$GPA, food_men$cuisine, mean, na.rm = T)
```

```
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA
```

```r
gpa_sd_men <- tapply(food_men$GPA, food_men$cuisine, sd, na.rm = T)
```

```
## Warning in var(if (is.vector(x) || is.factor(x)) x else as.double(x), na.rm =
## na.rm): NAs introduced by coercion

## Warning in var(if (is.vector(x) || is.factor(x)) x else as.double(x), na.rm =
## na.rm): NAs introduced by coercion
```

```r
gpa_mean_women <- tapply(food_women$GPA, food_women$cuisine, mean, na.rm = T)
```

```
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:
## returning NA
```

```r
gpa_sd_women <- tapply(food_women$GPA, food_women$cuisine, sd, na.rm = T)
```

```
## Warning in var(if (is.vector(x) || is.factor(x)) x else as.double(x), na.rm =
## na.rm): NAs introduced by coercion
```

```r
weight_mean_men <- tapply(food_men$weight, food_men$cuisine, mean, na.rm = T)
weight_sd_men <- tapply(food_men$weight, food_men$cuisine, sd, na.rm = T)

weight_mean_women <- tapply(food_women$weight, food_women$cuisine, mean, na.rm = T)
weight_sd_women <- tapply(food_women$weight, food_women$cuisine, sd, na.rm = T)
```

1. Download the facebook-fact-check.csv
2. Load the CSV file into your R environment.

```r
fb_fact <- read.csv("facebook-fact-check.csv")
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
```

```
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

3. Extract the last 500 rows.

Hint: Check out the top_n() page to figure out how to extract the last 500 rows instead of the first 500 rows.

```
fb_extract <- fb_fact %>% slice_tail(n = 500)
head(fb_extract)
```

```
##    account_id      post_id   Category     Page
## 1 62317591679 1.015386e+16 mainstream Politico
## 2 62317591679 1.015386e+16 mainstream Politico
## 3 62317591679 1.015386e+16 mainstream Politico
## 4 62317591679 1.015386e+16 mainstream Politico
## 5 62317591679 1.015386e+16 mainstream Politico
## 6 62317591679 1.015386e+16 mainstream Politico
##                                                    Post.URL Date.Published
## 1 https://www.facebook.com/politico/posts/10153861466546680     2016-09-26
## 2 https://www.facebook.com/politico/posts/10153861478296680     2016-09-26
## 3 https://www.facebook.com/politico/posts/10153861481676680     2016-09-26
## 4 https://www.facebook.com/politico/posts/10153861491796680     2016-09-26
## 5 https://www.facebook.com/politico/posts/10153861497961680     2016-09-26
## 6 https://www.facebook.com/politico/posts/10153861505681680     2016-09-26
##   Post.Type      Rating Debate share_count reaction_count comment_count
## 1     video mostly true    yes        6857          28505          1636
## 2     video mostly true    yes          48            536           215
## 3      link mostly true    yes        1849           2352           181
## 4     video mostly true    yes          91           1294           141
## 5     video mostly true    yes          24            321            73
## 6     video mostly true    yes         337           1519           473
```

- top_n() is superceded so I tried something else!

4. Look at the even-numbered column indices only. Identify them by name.

```
row_odd <- seq_len(nrow(fb_extract)) %% 2
data_row_odd <- fb_extract[row_odd == 1, ]
colnames(data_row_odd)
```

```
##  [1] "account_id"     "post_id"        "Category"       "Page"
##  [5] "Post.URL"       "Date.Published" "Post.Type"      "Rating"
##  [9] "Debate"         "share_count"    "reaction_count" "comment_count"
```

5. Using mutate, create a new variable called post_type_coded that renames each post type to the following: • link=1 • photo = 2 • text=3 • video = 4 Hint: look up case_when within tidyverse. You can also use if_else

```
fb_extract <- fb_extract %>%
  mutate(post_type_coded = ifelse(Post.Type == "link", '1',
                                  ifelse(Post.Type == 'photo', '2',
                                         ifelse(Post.Type == 'text', '3','4'))))
```

6. Arrange page names in reverse order.

```
fb_extract <- fb_extract %>% arrange(desc(Page))
head(fb_extract)
```

```
##     account_id      post_id Category          Page
## 1 1.145179e+14 1.462399e+15     left The Other 98%
## 2 1.145179e+14 1.462468e+15     left The Other 98%
## 3 1.145179e+14 1.462507e+15     left The Other 98%
## 4 1.145179e+14 1.462536e+15     left The Other 98%
## 5 1.145179e+14 1.462680e+15     left The Other 98%
## 6 1.145179e+14 1.462684e+15     left The Other 98%
##                                                      Post.URL Date.Published
## 1 https://www.facebook.com/TheOther98/posts/1462399387104368     2016-09-19
## 2 https://www.facebook.com/TheOther98/posts/1462468047097502     2016-09-19
## 3 https://www.facebook.com/TheOther98/posts/1462507497093557     2016-09-19
## 4 https://www.facebook.com/TheOther98/posts/1462535517090755     2016-09-19
## 5 https://www.facebook.com/TheOther98/posts/1462679840409656     2016-09-19
## 6 https://www.facebook.com/TheOther98/posts/1462684057075901     2016-09-19
##   Post.Type             Rating Debate share_count reaction_count comment_count
## 1     photo no factual content                  2          26990           590
## 2     video        mostly true                 NA           7580           364
## 3      link        mostly true              26726          20354          1471
## 4      link        mostly true               3899          12225           378
## 5      link        mostly true                483           5317           101
## 6     video        mostly true                688           3329            37
##   post_type_coded
## 1               2
## 2               4
## 3               1
## 4               1
## 5               1
## 6               4
```

7. Find the mean and standard deviation for the following variables, and summarize them. • share_count • reaction_count • comment_count

```
fb_fact %>%
  summarise(share_count = mean(share_count, na.rm=T),
            reaction_count = mean(reaction_count, na.rm=T),
            comment_count = mean(comment_count, na.rm=T))
```

```
##   share_count reaction_count comment_count
## 1    4044.816       5364.285      516.1022
```

8. Summarize the mean and standard deviations in Question 7 with the “mainstream” values in the category variable.

```
fb_main <- fb_fact %>%
  filter(Category == "mainstream") %>%
  summarize(share_count_m = mean(share_count, na.rm=T), #creating 6 variables
            reaction_count_m = mean(reaction_count, na.rm=T),
            comment_count_m = mean(comment_count, na.rm=T),
            share_count_sd = sd(share_count, na.rm=T),
            reaction_count_sd = sd(reaction_count, na.rm=T),
            comment_count_sd = sd(comment_count, na.rm=T))
```

Submit Email me (laaker@wisc.edu) the link to your ps811-exercises repository when you are done.