Check-in 8   Code

AUTHOR
Maxim Dokukin

Remember, **follow the instructions below and use R Markdown to create a pdf document with your code and answers to the following questions on Gradescope.** You may find a template file by clicking "Code" in the top right corner of this page.

1. Download and read the documentation for the [Childcare Costs data](#).

```r
#install.packages("tidytuesdayR")
library(tidytuesdayR)
tuesdata <- tidytuesdayR::tt_load('2023-05-09')
```

```
Downloading file 1 of 2: `childcare_costs.csv`
Downloading file 2 of 2: `counties.csv`
```

```r
tuesdata <- tidytuesdayR::tt_load(2023, week = 19)
```

```
Downloading file 1 of 2: `childcare_costs.csv`
Downloading file 2 of 2: `counties.csv`
```

```r
childcare_costs <- tuesdata$childcare_costs

library(dplyr)
```

2. Use visualization and regression to explore the following question: Is there a relationship between a county's average childcare costs and its median household income? There will be more than one way to answer this question–the important thing is to **explain** the choices you make in your analysis.

```r
median_costs <- childcare_costs |>
  group_by(county_fips_code) |>
  summarise(
    median_income = median(me_2018, na.rm = TRUE),
    median_mcsa = median(mcsa, na.rm = TRUE),
```

```
        median_mfccsa = median(mfccsa, na.rm = TRUE),
        median_mc_infant = median(mc_infant, na.rm = TRUE),
        median_mc_toddler = median(mc_toddler, na.rm = TRUE),
        median_mc_preschool = median(mc_preschool, na.rm = TRUE),
        median_mfcc_infant = median(mfcc_infant, na.rm = TRUE),
        median_mfcc_toddler = median(mfcc_toddler, na.rm = TRUE),
        median_mfcc_preschool = median(mfcc_preschool, na.rm = TRUE
    )
```
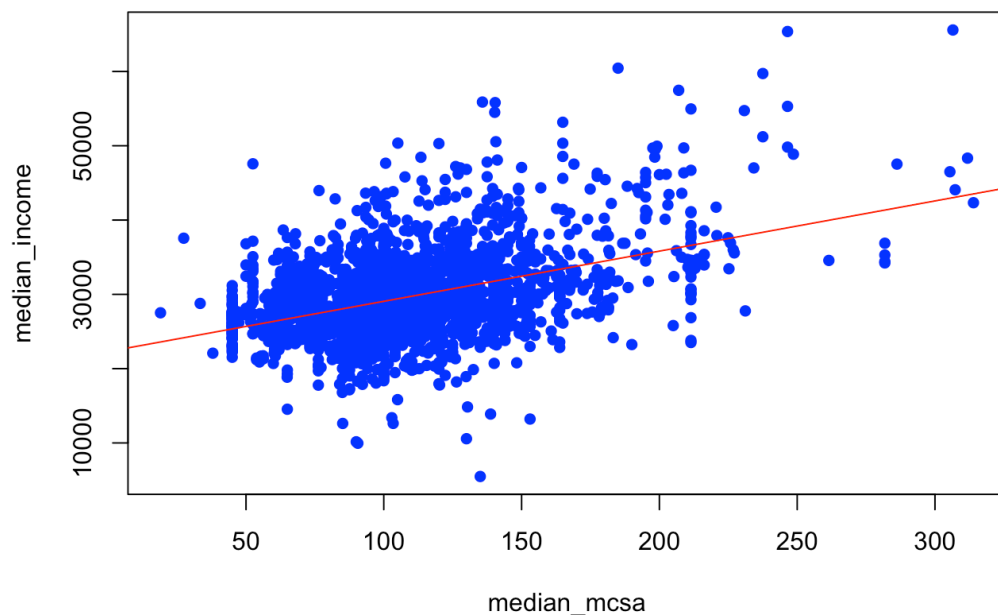
```
lm_res = lm(median_income ~ median_mcsa, data = median_costs)

plot(median_income ~ median_mcsa, data = median_costs,
     xlab = "median_mcsa",
     ylab = "median_income",
     pch  = 16,
     col = "blue",
     main = "Relationship Between Median Income and Median Child

abline(lm_res,  col = "red")
```

Relationship Between Median Income and Median Childcare Cost


```
lm_res = lm(median_income ~ median_mcsa + median_mfccsa + media
              median_mc_toddler + median_mc_preschool + median_
              median_mfcc_toddler + median_mfcc_preschool, data
```

```
summary(lm_res)
```

```
Call:
lm(formula = median_income ~ median_mcsa + median_mfccsa +
median_mc_infant +
    median_mc_toddler + median_mc_preschool +
median_mfcc_infant +
    median_mfcc_toddler + median_mfcc_preschool, data =
median_costs)

Residuals:
     Min        1Q    Median        3Q       Max
-24547.6   -2580.9   -187.2    2308.8   23520.6

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)            19003.259    351.751  54.025  < 2e-16 ***
median_mcsa              -27.089      6.273  -4.319 1.62e-05 ***
median_mfccsa             20.132      9.804   2.053  0.04012 *
median_mc_infant          -5.399      7.966  -0.678  0.49793
median_mc_toddler          4.612     12.201   0.378  0.70542
median_mc_preschool       67.596     15.630   4.325 1.58e-05 ***
median_mfcc_infant        60.161     15.196   3.959 7.71e-05 ***
median_mfcc_toddler      -89.767     31.208  -2.876  0.00405 **
median_mfcc_preschool     56.509     29.532   1.914  0.05578 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4461 on 2876 degrees of freedom
  (259 observations deleted due to missingness)
Multiple R-squared:  0.3239,    Adjusted R-squared:  0.322
F-statistic: 172.2 on 8 and 2876 DF,  p-value: < 2.2e-16
```

I decided to use median_income ~ median_mcsa correlation as the common one.

mcsa is weekly, full-time median price charged for Center-Based Care for those who are school age based on the results reported in the market rate survey report for the county or the rate zone/cluster to which the county is assigned.

Hence, this correlation should be the most representative.

However, I also investigated correlation of the income with other parameters, but did not plot it. There was significant correlation of imcome with median_mfccsa, median_mc_preschool, median_mfcc_infant, median_mfcc_toddler.

While no significant correlation exist between income and median_mc_infant, median_mc_toddler, median_mfcc_preschool.