Project 2 – US births in 2014

MTH 161 – Fall 2024

Michelle Hernandez

2024-11-14

Part 1: choose a dataset and propose a question

Due Nov. 15th

For this project, we will investigate a dataset containing a sample of observations from 2014 of US births sourced from (https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/nata

In this investigation, we will determine if there is an association between the fathers age and the mothers age and whether it affects the babies overall weight when it is born.

For this model, we will use `births14.csv` data set.

```
library(tidyverse)
births14 <- read_csv("births14.csv")
```

Data

```
print(births14)
```

```
# A tibble: 1,000 x 13
    fage  mage mature     weeks premie visits gained weight lowbirthweight sex
   <dbl> <dbl> <chr>      <dbl> <chr>   <dbl>  <dbl>  <dbl> <chr>          <chr>
 1    34    34 younger m~    37 full ~     14     28   6.96 not low        male
 2    36    31 younger m~    41 full ~     12     41   8.86 not low        fema~
 3    37    36 mature mom    37 full ~     10     28   7.51 not low        fema~
 4    NA    16 younger m~    38 full ~     NA     29   6.19 not low        male
 5    32    31 younger m~    36 premie     12     48   6.75 not low        fema~
 6    32    26 younger m~    39 full ~     14     45   6.69 not low        fema~
```

```
 7    37    36 mature mom    36 premie     10     20   6.13 not low        fema~
 8    29    24 younger m~    40 full ~     13     65   6.74 not low        male
 9    30    32 younger m~    39 full ~     15     25   8.94 not low        fema~
10    29    26 younger m~    39 full ~     11     22   9.12 not low        male
# i 990 more rows
# i 3 more variables: habit <chr>, marital <chr>, whitemom <chr>
```

> **Note**