Birthwt

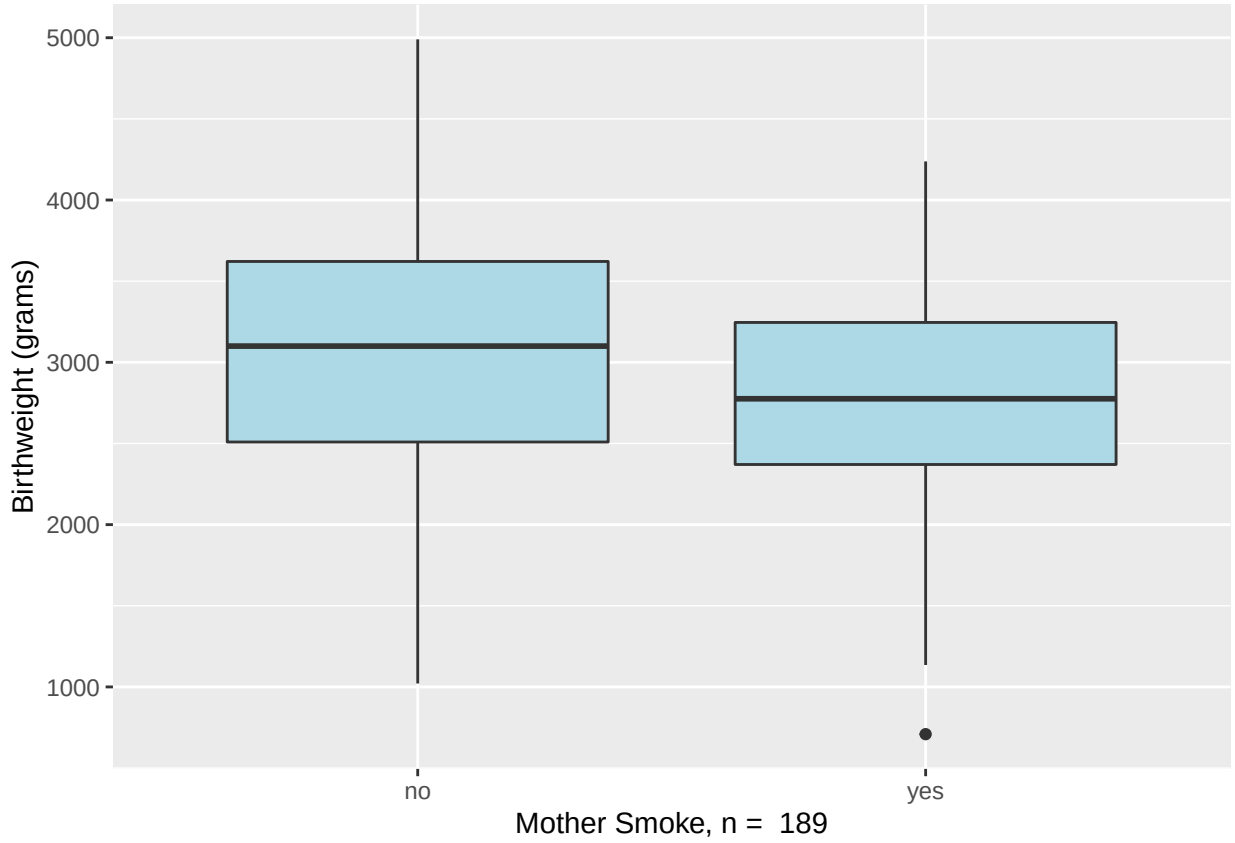
Sachin Shubham - 1524147

2/6/2021

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.5     v dplyr   1.0.3
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
birthwt <- as_tibble(MASS::birthwt)
colnames(birthwt) <- c("birthwt.below.2500",
                       "mother.age",
                       "mother.weight",
                       "race",
                       "mother.smokes",
                       "previous.prem.labor",
                       "hypertension",
                       "uterine.irr",
                       "physician.visits",
                       "birthwt.grams")

#(1) compare infants birth weight between smoking and non-smoking mothers

birthwt <- birthwt %>%
  mutate_at(c("mother.smokes", "birthwt.below.2500"),
            ~ recode_factor(.x, `0` = "no", `1` = "yes"))

qplot(x = mother.smokes, y = birthwt.grams,
      geom = "boxplot", data = birthwt,
      xlab = paste("Mother Smoke, n = ", nrow(birthwt)),
      ylab = "Birthweight (grams)",
      fill = I("lightblue"))
```




```
birthwt %>%
  group_by(mother.smokes) %>%
  summarize(num.obs = n(),
            mean.birthwt = round(mean(birthwt.grams), 0),
            sd.birthwt = round(sd(birthwt.grams), 0),
            se.birthwt = round(sd(birthwt.grams) / sqrt(num.obs), 0))
```

```
## # A tibble: 2 x 5
##   mother.smokes num.obs mean.birthwt sd.birthwt se.birthwt
## * <fct>           <int>        <dbl>      <dbl>      <dbl>
## 1 no                115         3056        753         70
## 2 yes                74         2772        660         77
```

```
birthwt.t.test <- t.test(birthwt.grams ~ mother.smokes, data = birthwt)
birthwt.t.test
```

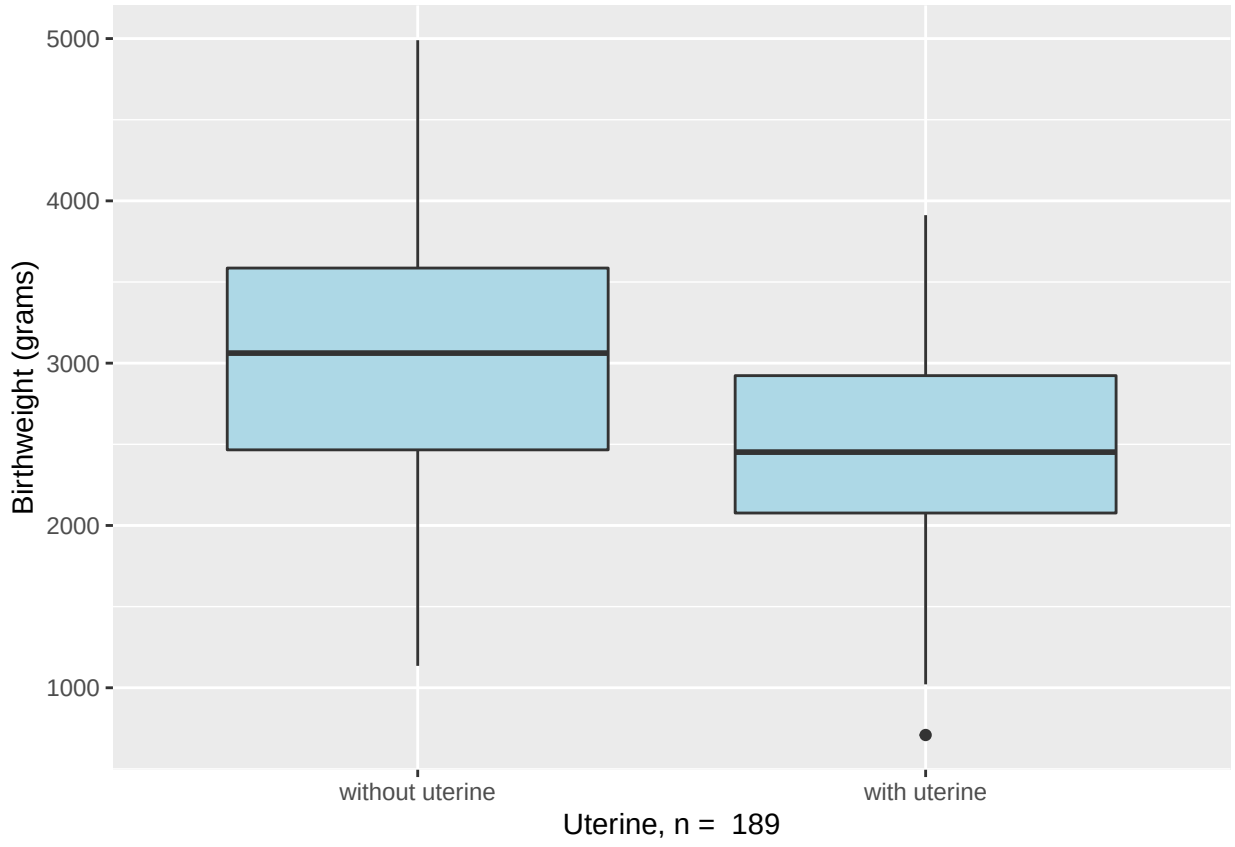
```
##
##  Welch Two Sample t-test
##
## data:  birthwt.grams by mother.smokes
## t = 2.7299, df = 170.1, p-value = 0.007003
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   78.57486 488.97860
## sample estimates:
```

```
##  mean in group no mean in group yes
##          3055.696           2771.919
```

```
birthwt.t.test$p.value
```

```
## [1] 0.007002548
```

```
#Our study finds that birth weights are on average 3055.696 in the non-smoking
#group compared to the average 2771.919 in smoking group (t-statistic 2.7299,
#p=0.007003, 95% CI [78.57486, 488.97860]g). A small p(<=0.05), reject the null
#hypothesis. Therefore, this is strong evidence that the null hypothesis is
#invalid

#(2)compare infants birth weight between a group of mothers with a presence of
#uterine irritability and a group of mothers without it

birthwt <- birthwt %>%
  mutate_at(c("uterine.irr", "birthwt.below.2500"),
            ~ recode_factor(.x, `0` = "without uterine", `1` = "with uterine"))
qplot(x = uterine.irr, y = birthwt.grams,
      geom = "boxplot", data = birthwt,
      xlab = paste("Uterine, n = ", nrow(birthwt)),
      ylab = "Birthweight (grams)",
      fill = I("lightblue"))
```



```r
birthwt %>%
  group_by(uterine.irr) %>%
  summarize(num.obs = n(),
            mean.birthwt = round(mean(birthwt.grams), 0),
            sd.birthwt = round(sd(birthwt.grams), 0),
            se.birthwt = round(sd(birthwt.grams) / sqrt(num.obs), 0))
```

```
## # A tibble: 2 x 5
##   uterine.irr     num.obs mean.birthwt sd.birthwt se.birthwt
## * <fct>             <int>        <dbl>      <dbl>      <dbl>
## 1 without uterine     161         3031        694         55
## 2 with uterine         28         2449        742        140
```

```r
birthwt.t.test <- t.test(birthwt.grams ~ uterine.irr, data = birthwt)
birthwt.t.test
```

```
##
##  Welch Two Sample t-test
##
## data:  birthwt.grams by uterine.irr
## t = 3.8615, df = 35.696, p-value = 0.000455
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  275.8913 886.6553
## sample estimates:
## mean in group without uterine    mean in group with uterine
##                      3030.702                      2449.429
```

```r
birthwt.t.test$p.value
```

```
## [1] 0.0004550226
```

```r
#Our study finds that birth weights are on average 3030.702 in the without
#uterine group compared to the average 2449.429 in with uterine group
#(t-statistic 3.8615, p=0.000455, 95% CI [275.8913, 886.6553]g). A small
#p(<=0.05), reject the null hypothesis. Therefore, this is strong evidence that
#the null hypothesis is invalid
```