What factors affect baby weight?

Submitted by Sonia Chhay

[1. Executive summary](#)

[2. Introducing the dataset](#)
- [2.1 Variables](#)
- [2.2 Exploring the dataset](#)
  - [2.2.1 Babies' birth weight between smoking and non-smoking mothers](#)
  - [2.2.2 Technical analysis](#)

[3 What factors affect a baby's birth weight?](#)
- [3.1 What is the relationship between birth weight and maturity level?](#)
  - [3.1.1 Technical analysis](#)
- [3.2 How do babies' birth weight compare when looking at maturity level and maternal smoking status?](#)
  - [3.2.1 Technical analysis](#)

[4. Checking validity of statistical tests](#)

[5. Conclusion](#)

All code and output: [Here](#)

1. Executive summary

Smoke from a cigarette typically contains thousands of health-harming chemicals, toxins, and carcinogens which increase risk of cancer development and other diseases. Yet, despite the implementation of the WHO Framework Convention on Tobacco[1] and global efforts to control smoking, the tobacco industry continues to grow and contribute to high smoking prevalence rates.

The highly addictive nature of nicotine causes many users, including those pregnant, to become dependent on tobacco products. In fact, cigarettes are among the most frequently used drugs during pregnancy (Kataoka et al., 2018). Tobacco consumption during pregnancy can cause additional severe consequences such as compromising maternal and fetal health, reduced infant birth weight, premature delivery, and increased perinatal mortality ("Weight, fertility, and pregnancy", 2018).

Various studies have proposed that there are advantages in abstinence from smoking during pregnancy to allow the fetus to go through full development (Kataoka et al., 2018). This analysis aims to analyze the effects of maternal smoking for different gestational periods on the birth weight of newborns.

The research question is two-fold:

- What is the relationship between birth weight and maturity level?
- How do babies' birth weight compare in relation to maturity level and maternal smoking status?

---

[1] The WHO Framework Convention on Tobacco was put into effect in February 2005 in response to the globalization of the tobacco epidemic. It reaffirms the right of all people to the highest standard of health.

2. Introducing the dataset

2.1 Variables

For this report, R was used to analyze and visualize the data (code given under table of contents). The dataset consists of three variables for 409 babies -- their birth weight, gestation period, and whether their mother smokes commercial tobacco during pregnancy.

Variables in dataset:

- bwt: Baby's birth weight (oz)
- gestation: Fetal development period from the time of conception until birth (days)
- smoke: Indication of whether the mother smokes during pregnancy (Yes = 1, No = 0)

A sample of the data is shown below:

```
Observations: 409
Variables: 3
$ bwt       <int> 68, 92, 91, 69, 65, 71, 77, 77, 75, 96, 127, 86, 100, 109...
$ gestation <int> 223, 224, 225, 232, 232, 234, 238, 238, 239, 241, 242, 24...
$ smoke     <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
```

Variables created for further analysis:

- maturity: Gestational age (1 if baby was premature and spent less than 259 days in the womb, 3 if gestational age was beyond 293, and 2 otherwise)
- MatSmoke: Combination of maturity level and maternal smoking status

Knowing maturity and maternal smoking status for each maternity level (MatSmoke) is helpful in examining various factors across different groups. For example, maturity would allow for different factors to be compared between premature, mature, and postmature babies. MatSmoke would also aid in this process by taking maternal smoking status into consideration.

Thus, the maturity variable was created by converting gestational age to a factor with three levels, and the MatSmoke variable was created by combining maturity level and maternal smoking status.

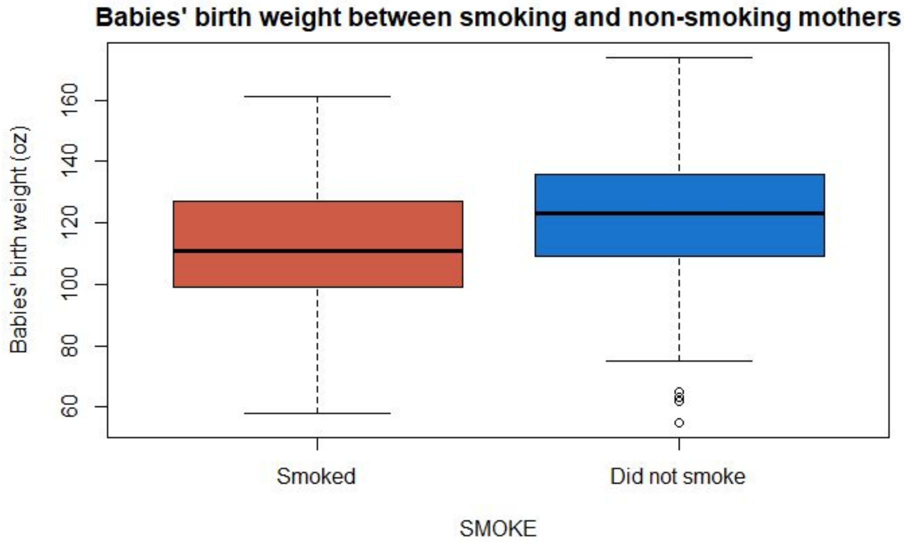
2.2 Exploring the dataset

The dataset contains babies' birth weight for 163 smoking and 246 non-smoking mothers. It is important to consider the relationship between the birth weight of a newborn and maternal smoking as tobacco consumption often leads to problems for health and physical development for both the mother and fetus.

NOTE:

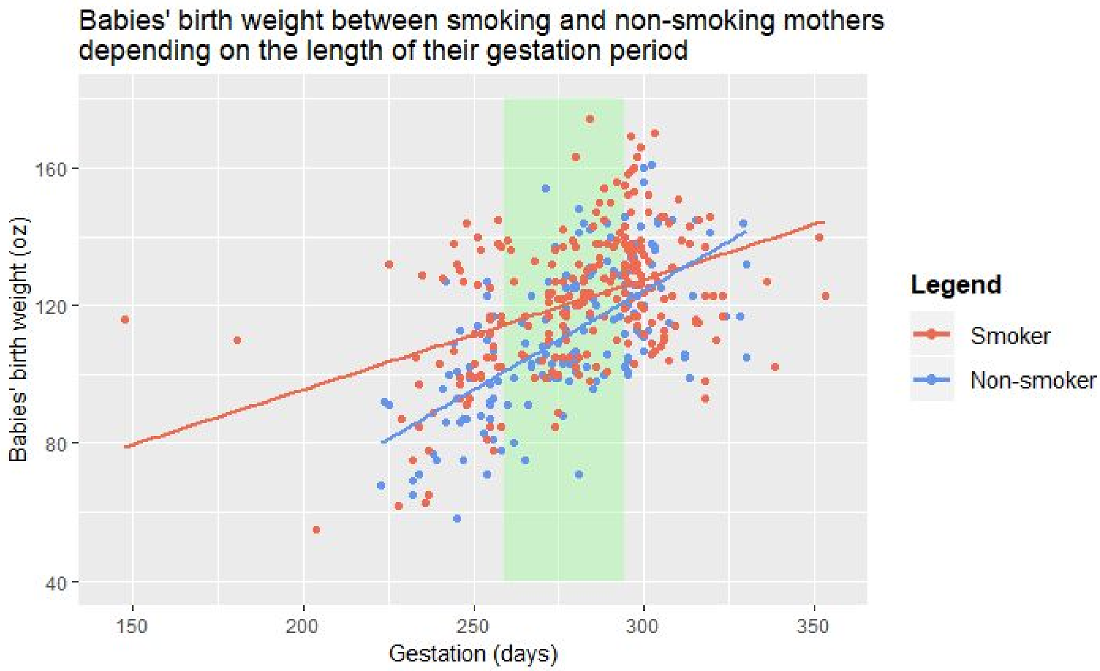
- *Healthy newborns generally weigh between 80 oz to 130 oz*
- *Normal duration of pregnancy lasts about 260 days to 293 days*

2.2.1 Babies' birth weight between smoking and non-smoking mothers



The box plot above shows that babies born to mothers who smoked are more likely to have a healthy birth weight compared to babies born to mothers who did not smoke. However, the range of birth weights for babies born to mothers who smoked is much wider compared to babies born to mothers who did not smoke, and the median (typical birth weight) is lower.

The gestation period should be taken into consideration as it can also affect birth weight. In the graph below, babies born during normal gestation (highlighted in green) tend to weigh around 90 oz to 140 oz, which is within the expected healthy range. However, for premature and postmature babies born to smoking and non-smoking mothers, their birth weight tends to be more spread out between 55 oz to 174 oz.



Thus, it is suggested that babies born during the normal gestational period are more likely to have an average birth weight. However, this raises a few questions: What other factors play a significant role in a baby's birth weight? How do premature, mature, and postmature babies compare to each other?

2.2.2 Technical analysis

The significance of the information from Part 2.2.1 can be checked through a pooled two-sample t-test. This particular test is appropriate to use because the results of a conducted F-test gave evidence of equal variance (p-value=0.43>0.05).

The p-value obtained from the pooled two-sample t-test was $3.672*10^{-6}$, suggesting there is evidence of a significant difference in mean birth weights between babies born to mothers who were smokers and babies born to mothers who were non-smokers.

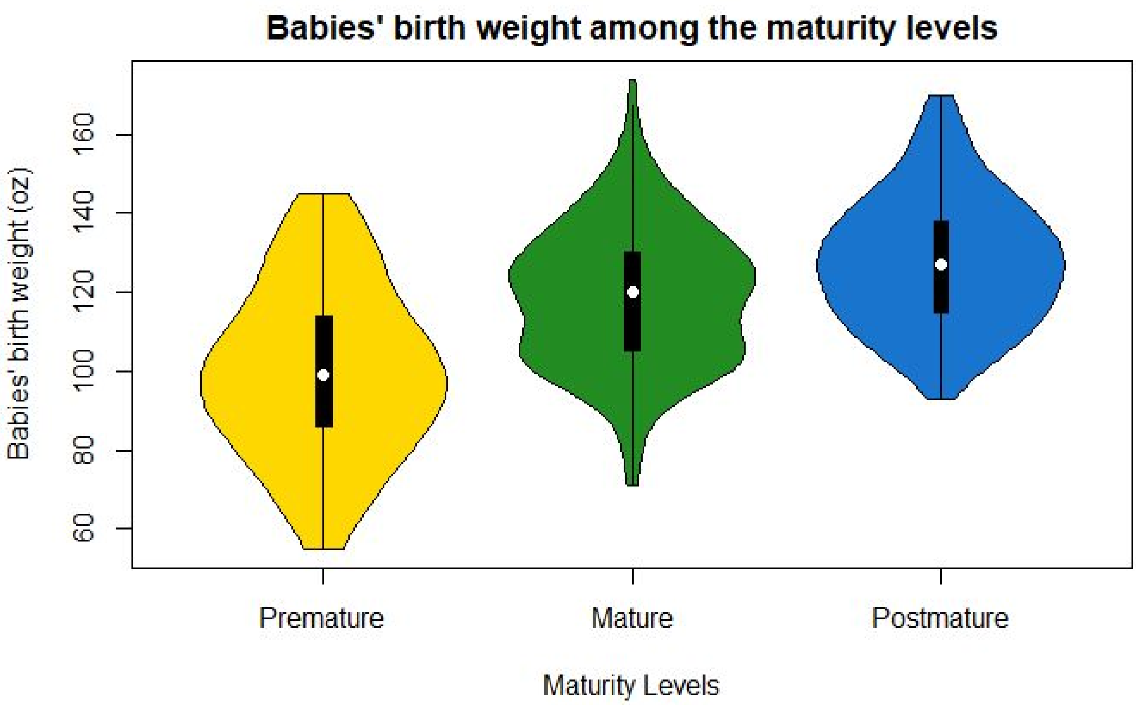
**Refer to Appendix A for test output.*

3 What factors affect a baby's birth weight?

3.1 What is the relationship between birth weight and maturity level?

From the 409 sampled babies:

- 97 were born prematurely
- 151 were born postmature
- 161 otherwise (mature)



In the graph above, the plots are around the same height for mature and postmature babies. However, the plot for premature babies is comparatively higher, suggesting a difference in birth weights between these groups and premature babies. This is evident when comparing the median (typical value) for each group. For premature babies, the median for birth weight approximates 100 oz. However, for mature and postmature babies, the median is around 120 oz and 125 oz, respectively.

Although a birth weight of 100 oz is still considered healthy, this is only considering the typical weight of a premature baby in the dataset. The median implies that 50% of these premature babies have a weight between 60 oz to 100 oz, which may put them at greater risk for health problems. Postmature babies are also born with this greater risk.

According to "Weight, fertility, and pregnancy" (2018), pre- and postmature babies can develop problems such as:

- Development problems
- Heart defects
- Obesity, type 2 diabetes, high cholesterol

Since the distribution for mature and postmature babies mainly range from 90 oz to 140 oz, many of these babies are considered to be within the healthy range for birth weight. However, it is important to also note that the distribution between mature and postmature babies are different. The plot for mature babies has a bimodal distribution suggesting that within this group, there may be many mature babies with a weight of either 105 oz or 125 oz. Since this is within the range of a healthy newborn's weight, it is not a cause of concern.

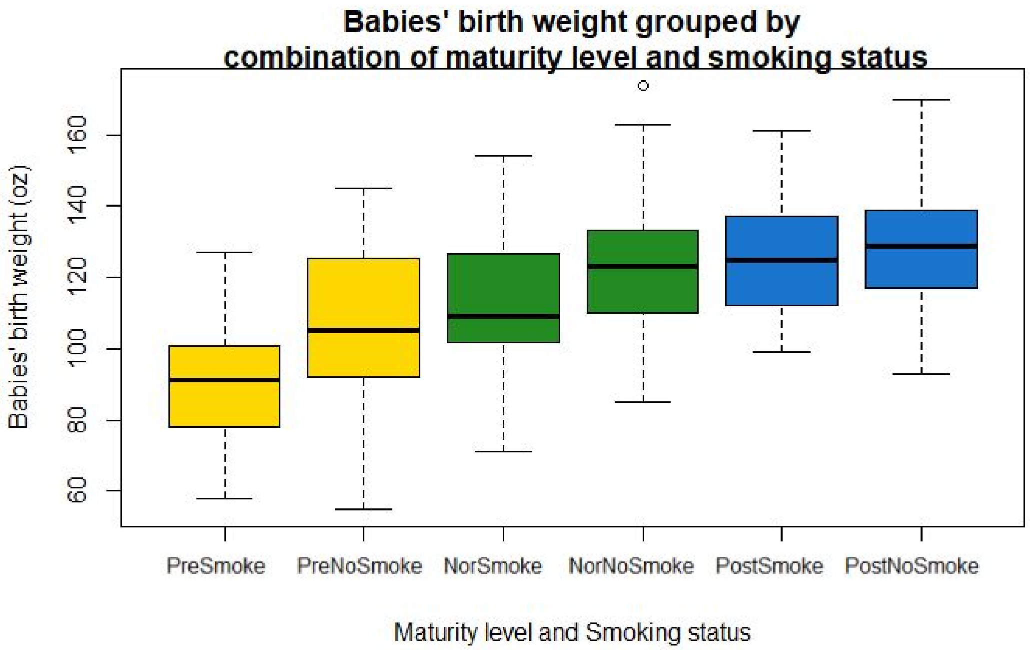
3.1.1 Technical analysis

Although the data contains different numbers of babies in the three maturity levels, Barlett's test shows no evidence of variance across groups being statistically significant ($p=0.096>0.05$). Thus, homogeneity of variances in different groups can be assumed, meaning each sample of babies can be considered equal.

Differences in mean birth weight among babies classified by gestational maturity can be checked by conducting a one-way analysis of variance (ANOVA) test. The data suggests that there is at least one category where mean birth weight is not equal to others ($p\text{-value}=2*10^{-16}<0.05$). The Bonferroni test is also used to further see which levels of maturity differ. Through pairwise comparisons, there is evidence that each pairing for maturity levels has a significant difference.

*Refer to Appendix B for test output.
*Refer to Appendix D for Barlett's test.

3.2 How do babies' birth weight compare when looking at maturity level and maternal smoking status?



In the graph above, birth weight is compared between the six categories of babies, grouped by the combination of their maturity level and maternal smoking status. The premature categories are indicated in yellow with the prefix "Pre-" in the label. Postmature categories are indicated in blue with the prefix "Post-". Otherwise, the category is shaded green and given the prefix "Nor-".

There are some notable differences. A key finding is that smoking in the pre-term resulted in the lowest average birth weight while no smoking in the post-term had the highest average birth weight. Generally, mothers who smoked (left box plot for each group) gave birth to babies with a lower birth weight than in the respective no smoke condition (e.g. PreSmoke and PreNoSmoke), suggesting a difference between the categories.

Knowing that maternal smoking plays a role in lowering their baby's birth weight, this information can be used to support other studies that have proposed that maternal smoking has stronger consequences on birth weight.

3.2.1 Technical analysis

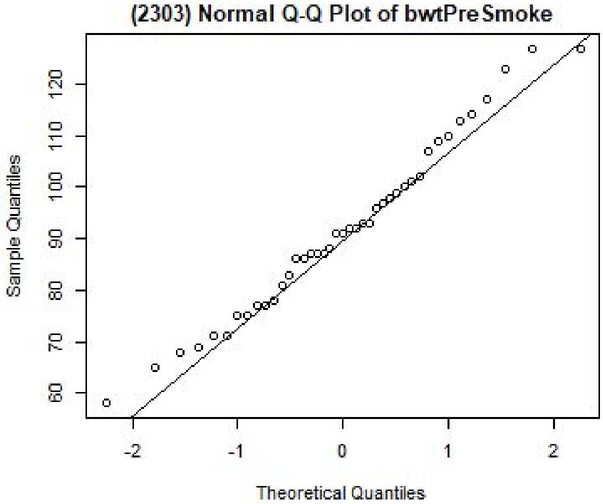
The one-way ANOVA conducted for the six categories of babies enforces previous findings that at least one average birth weight, classified by the combination of their maturity and maternal smoking status, is not equal to the others. Through pairwise comparisons, there is also evidence that the PostNoSmoke-NorNoSmoke and PreNoSmoke-NorSmoke pairings have similar average birth weights.

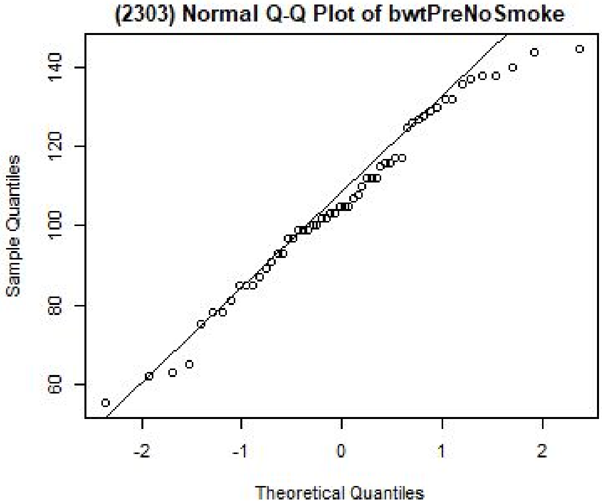
**Refer to Appendix C for test output.*

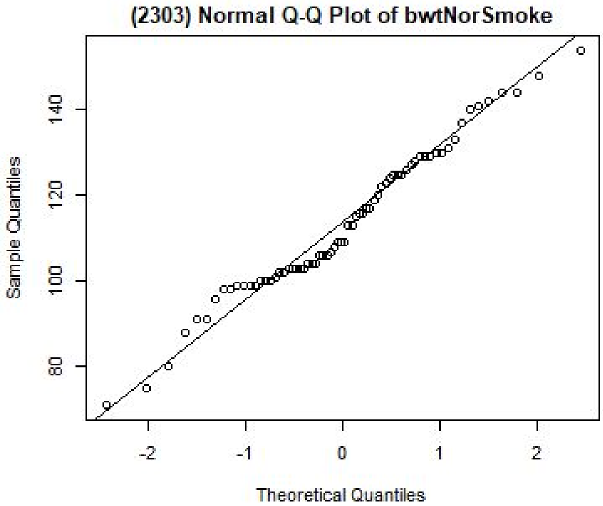
4. Checking validity of statistical tests

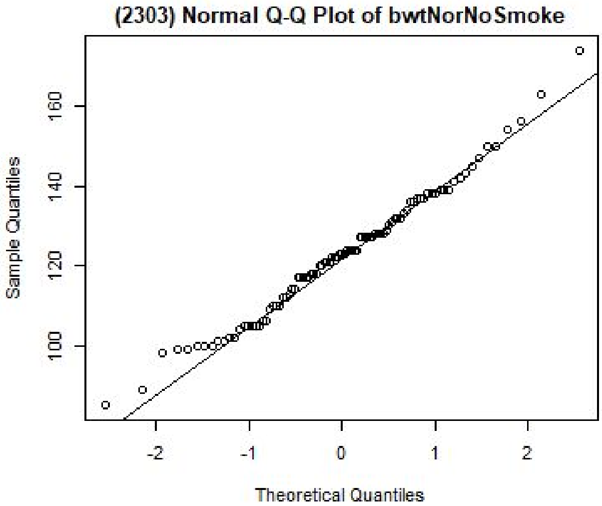
1. Since the data was based on an observational study and not an experiment, it can be assumed that there is no relationship within each group or between the groups themselves.

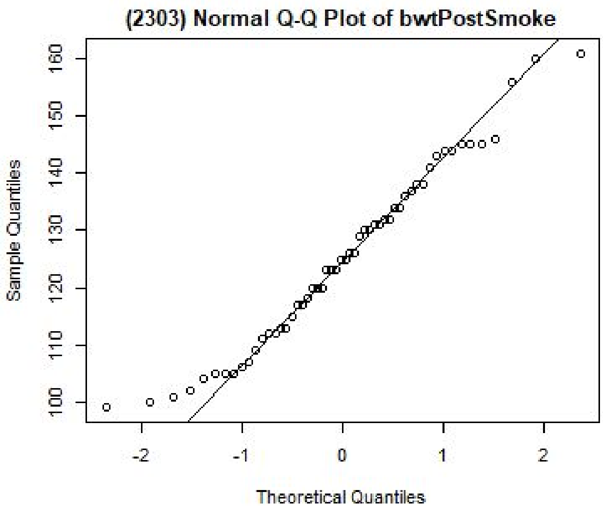
2. Normality for each population was also checked to ensure that results from tests used in this paper can be trusted.

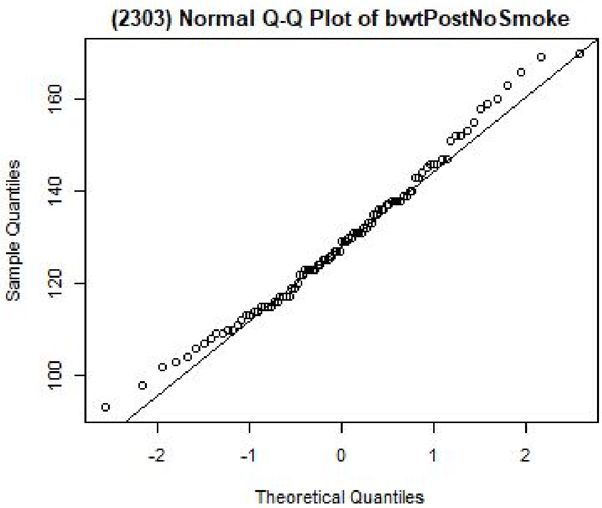












In the Q-Q plots, the points generally fall about a straight line. However, there are some slight deviations at the ends for the Q-Q plot of bwtPreNoSmoke, bwtNorNoSmoke, and bwtPostSmoke in particular. This implies the possibility of a deviation from normality.

However, most of the points are also on the line, which may be enough to not disqualify the data from being normal. The Shapiro-Wilks test supports this idea and suggests that there is no significant evidence to conclude departure from normality for any of the groups (p-value>0.05).

Shapiro-Wilks test

Data (Group)	**P-value**
PreSmoke	0.6843
PreNoSmoke	0.3812
NorSmoke	0.2334
NorNoSmoke	0.4507
PostSmoke	0.1956
PostNoSmoke	0.3702

Refer to Appendix D for test output.

5. Conclusion

Analysis of the baby weight data suggests the need for better prenatal education to provide mothers with info about the effects of smoking while pregnant. This is supported with evidence from Part 3, which shows that babies born to mothers who did not smoke consistently had healthier birth weights than their counterparts for each maturity level.

Improvement for tobacco cessation programs would also be helpful as it could provide better resources for mothers wanting to quit smoking. This could potentially lead to lower prevalence rates of maternal smoking, which in turn improves the health and physical development of future babies.

APPENDIX

All code and output: Here

Appendix A

```
        F test to compare two variances

data:  bwtSmoker and bwtNotSmoker
F = 1.1178, num df = 162, denom df = 245, p-value = 0.43
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.8469957 1.4873633
sample estimates:
ratio of variances
           1.11782


        Two Sample t-test

data:  bwtSmoker and bwtNotSmoker
t = -4.6937, df = 407, p-value = 3.672e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -13.748207  -5.631563
sample estimates:
mean of x mean of y
 111.8589  121.5488
```

Appendix B

```
[1] "1" "2" "3"
             Df Sum Sq Mean Sq F value Pr(>F)
maturity      2  46586   23293   71.28 <2e-16 ***
Residuals   406 132680     327
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

        Pairwise comparisons using t tests with pooled SD

data:  bwt and maturity

  2       1
1 1.4e-14 -
3 3.8e-05 < 2e-16

P value adjustment method: bonferroni
```

Appendix C

```
[1] "NorNoSmoke"  "NorSmoke"    "PostNoSmoke" "PostSmoke"   "PreNoSmoke"  "PreSmoke"
             Df Sum Sq Mean Sq F value Pr(>F)
MatSmoke      5  55448   11090   36.09 <2e-16 ***
Residuals   403 123818     307
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

        Pairwise comparisons using t tests with pooled SD

data:  bwt and MatSmoke

            PreSmoke NorNoSmoke NorSmoke PostNoSmoke PostSmoke
NorNoSmoke  < 2e-16  -          -        -           -
NorSmoke    1.7e-08  0.0114     -        -           -
PostNoSmoke < 2e-16  0.1625     2.4e-07  -           -
PostSmoke   < 2e-16  1.0000     0.0033   1.0000      -
PreNoSmoke  0.0015   3.2e-07    0.2824   2.4e-13     2.1e-07

P value adjustment method: bonferroni
```

Appendix D

```
        Shapiro-Wilk normality test

data:  bwtPreSmoke
W = 0.98023, p-value = 0.6843


        Shapiro-Wilk normality test

data:  bwtPreNoSmoke
W = 0.97762, p-value = 0.3812


        Shapiro-Wilk normality test

data:  bwtNorSmoke
W = 0.97675, p-value = 0.2334


        Shapiro-Wilk normality test

data:  bwtNorNoSmoke
W = 0.98639, p-value = 0.4507


        Shapiro-Wilk normality test

data:  bwtPostSmoke
W = 0.97014, p-value = 0.1956


        Shapiro-Wilk normality test

data:  bwtPostNoSmoke
W = 0.98557, p-value = 0.3702


        Bartlett test of homogeneity of variances

data:  bwt by MatSmoke
Bartlett's K-squared = 9.3393, df = 5, p-value = 0.09627
```

REFERENCES

Weight, fertility, and pregnancy. (2018, December 27). Retrieved from https://www.womenshealth.gov/healthy-weight/weight-fertility-and-pregnancy

Kataoka, M. C., Carvalheira, A. P. P., Ferrari, A. P., Malta, M. B., Maria Antonieta De Barros Leite Carvalhaes, & Cristina Maria Garcia De Lima Parada. (2018). Smoking during pregnancy and harm reduction in birth weight: a cross-sectional study. *BMC Pregnancy and Childbirth*, *18*(1). doi: 10.1186/s12884-018-1694-4