Maternal Influences on Infant Birth Weight

EEMB 146 Final Project

Sharon Nguyen

6/8/2022

Abstract

Several studies have reported that mother's a part of racial groups that experience racism and their smoking habits can cause low birth weight in their babies. The birth weight data set is composed of two categorical and two quantitative variables all randomly sampled. To discover the significant difference among groups the one-way ANOVA test is executed on the response variable birth weight and the categorical variable, race. The one-way ANOVA is repeated on the categorical variable, smoke. The linear regression model is used to predict whether log pre-pregnancy weight of mothers has an affect on their baby's birth weight. The result is race and smoking habits do play a significant role in predicting birth weight where as log mother's pre-pregnancy weight does not predict birth weight. The results show that the largest group differences among race was between white and black. The largest difference in maternal smoking habits are mother's who smoke now and ones who have never smoked

Introduction

In the quest to find out the affects of high and low birth weights, data was taken from the Child Health and Development Studies in 1961 and 1962 to address the question of whether there is a link between mother's prepregnancy weight, mother's race, and maternal smoking with birth weight of their babies. The Gestation dataset can be found in the R package: mosaicData. Low birth weights (< 5 lbs. 8 oz.) can result in infants with breathing problems, bleeding in the brain, infections, etc (March of Dimes). High birth weights (> 9 lbs. 15 oz.), on the other hand, can have problems at birth like longer delivery times, injury to the baby such as broken collar bones or damaged nerves in the arms, and more (Large for Gestational Age). With the `birthwt` dataset, we will discover maternal influences on infant birth weight and find which independent variables play a significant role in predicting birth weight. I hypothesize that a mother's pre-pregnancy weight, race, and smoking habits will have a significant effect on their infant's birth weight.

Exploratory Data Analysis

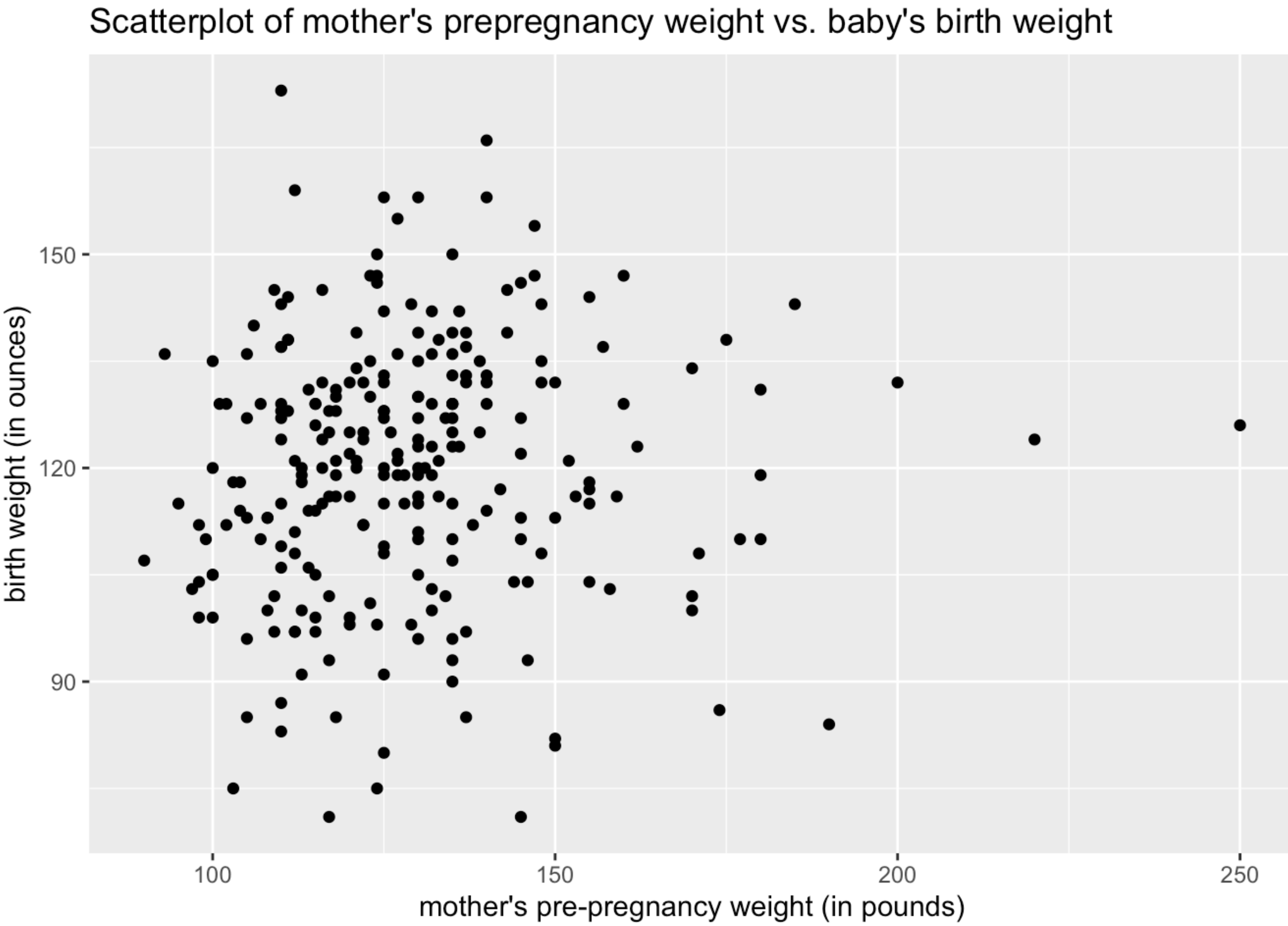

Figure 1. Scatterplot of mother’s prepregnancy weight (in pounds) vs. baby’s birth weight (in ounces). It seems to be a large cluster of data in the middle left of the plot and possibly some outliers for mother’s prepregnancy weight. Overall, there does not seem to be a relationship between the variables.

The Central Limit Theorem states that data with a sufficiently large random sample can be considered approximately normal, and this sample contains 250 observations.

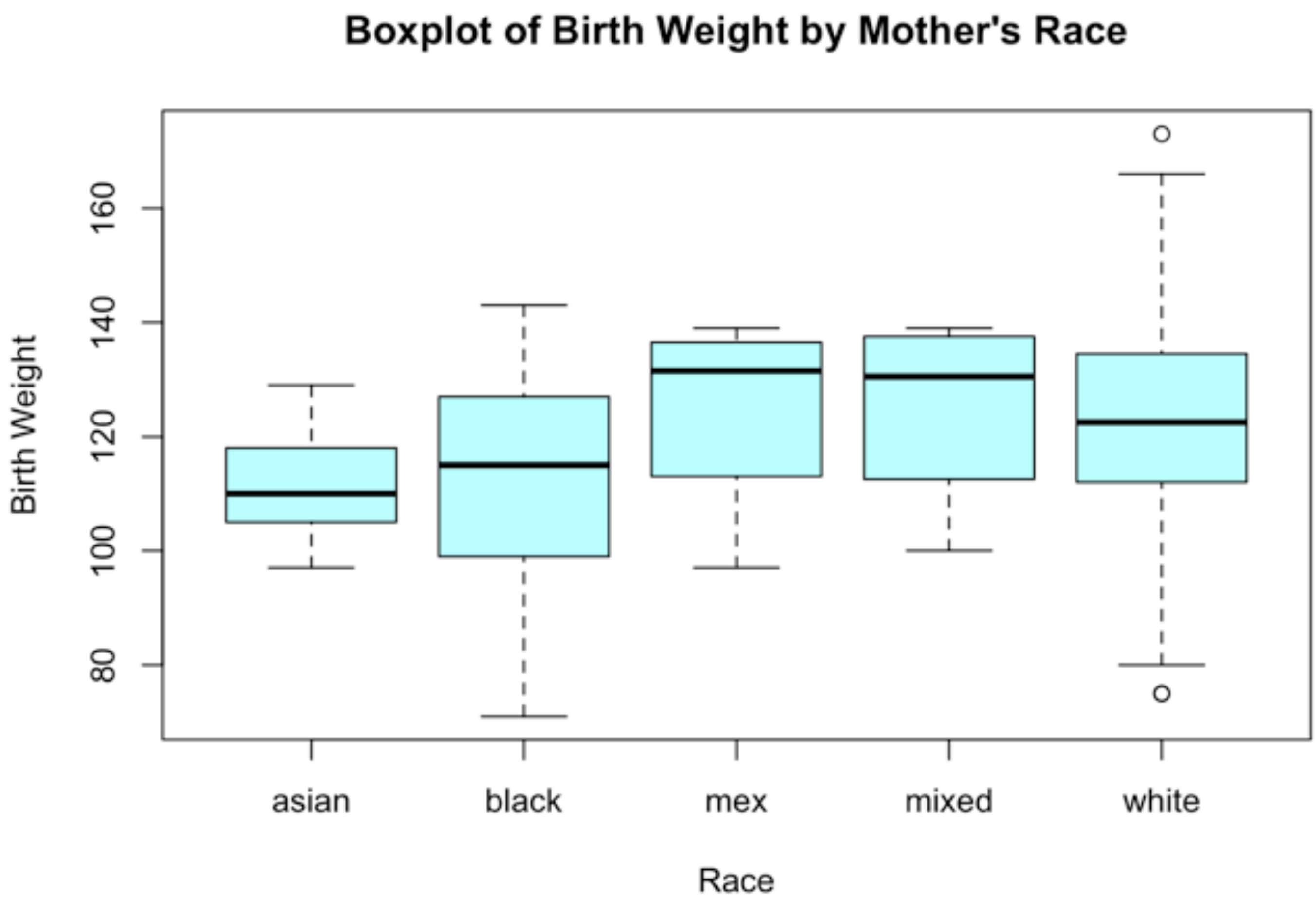

Figure 2. Boxplot of birth weights by race. There does not seem to be significant differences between birth weight and race, but there are outliers for the white group. The range for the white group is large and the second largest is black, in comparison, to the other races.

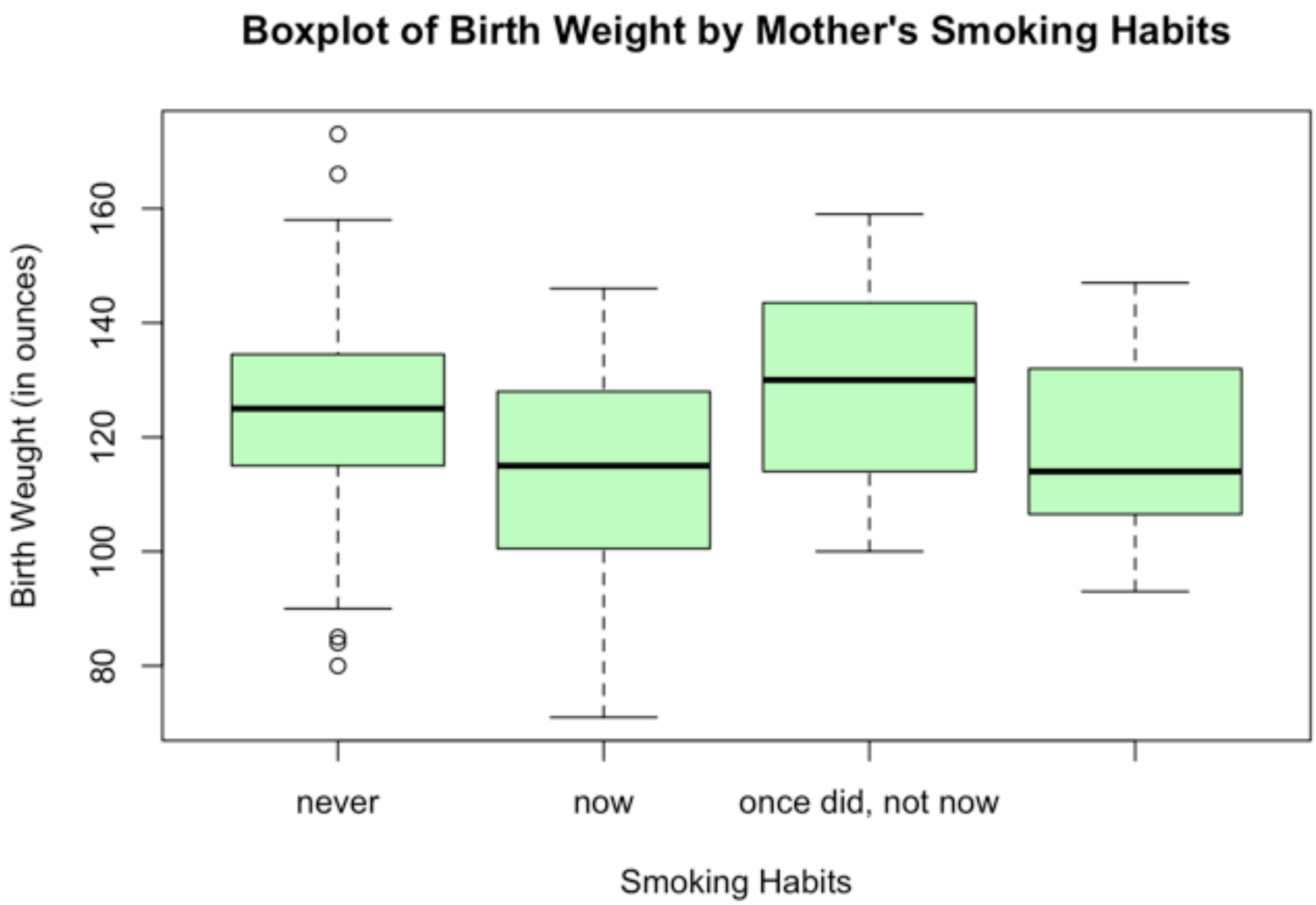

Figure 3. Boxplot of Birth Weight by Mother's Smoking Habits. There does not seem to be significant differences between birth weight and smoking habits during pregnancy. The never smoked group has some outliers. The range of the never group and now group are also larger than the other groups.

Figure 4. Histogram and QQ-plot of Birth Weight are both normally distributed.

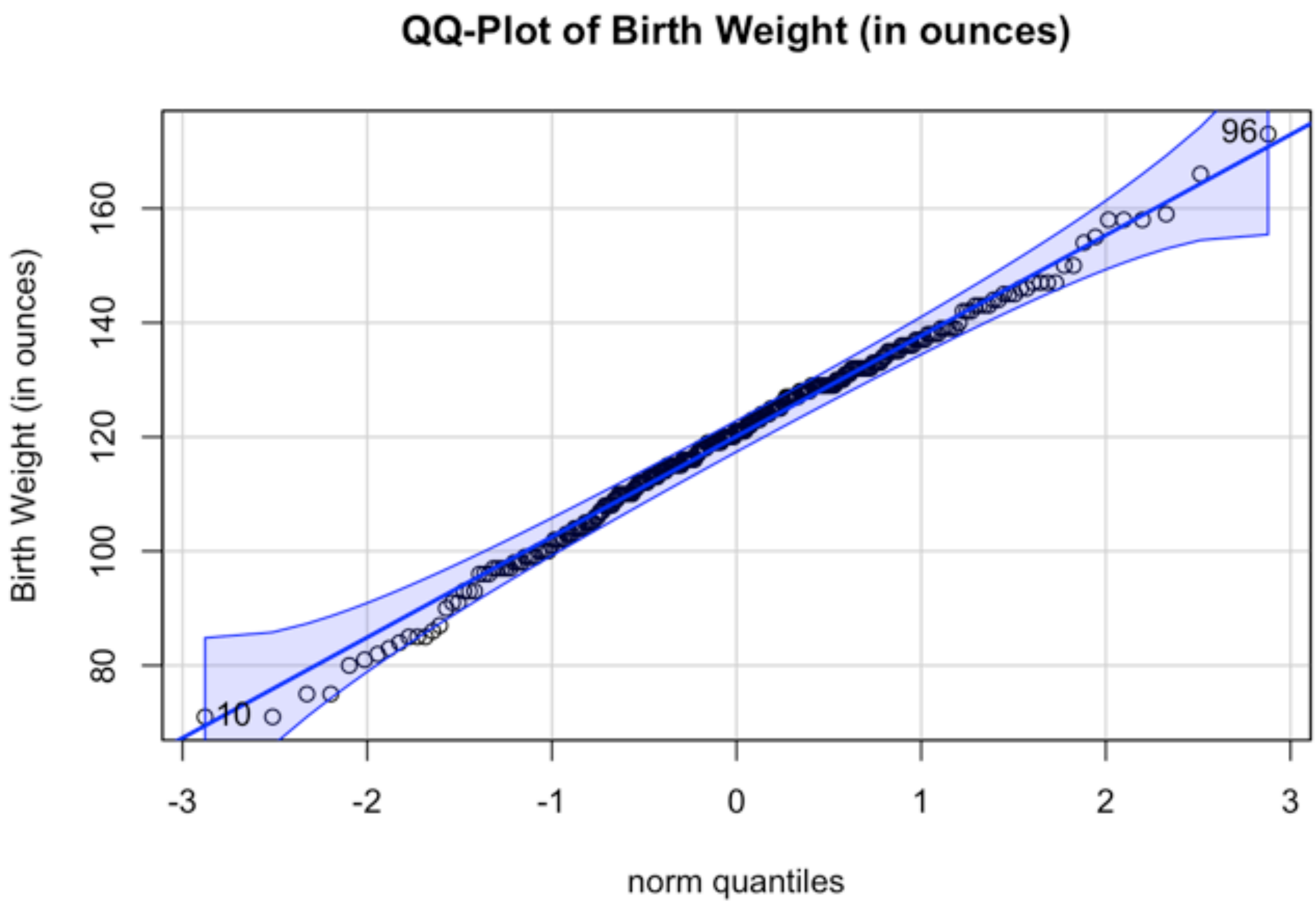

Figure 4. Histogram and QQ-plot of Birth Weight are both normally distributed.

```
## [1] 96 10
```

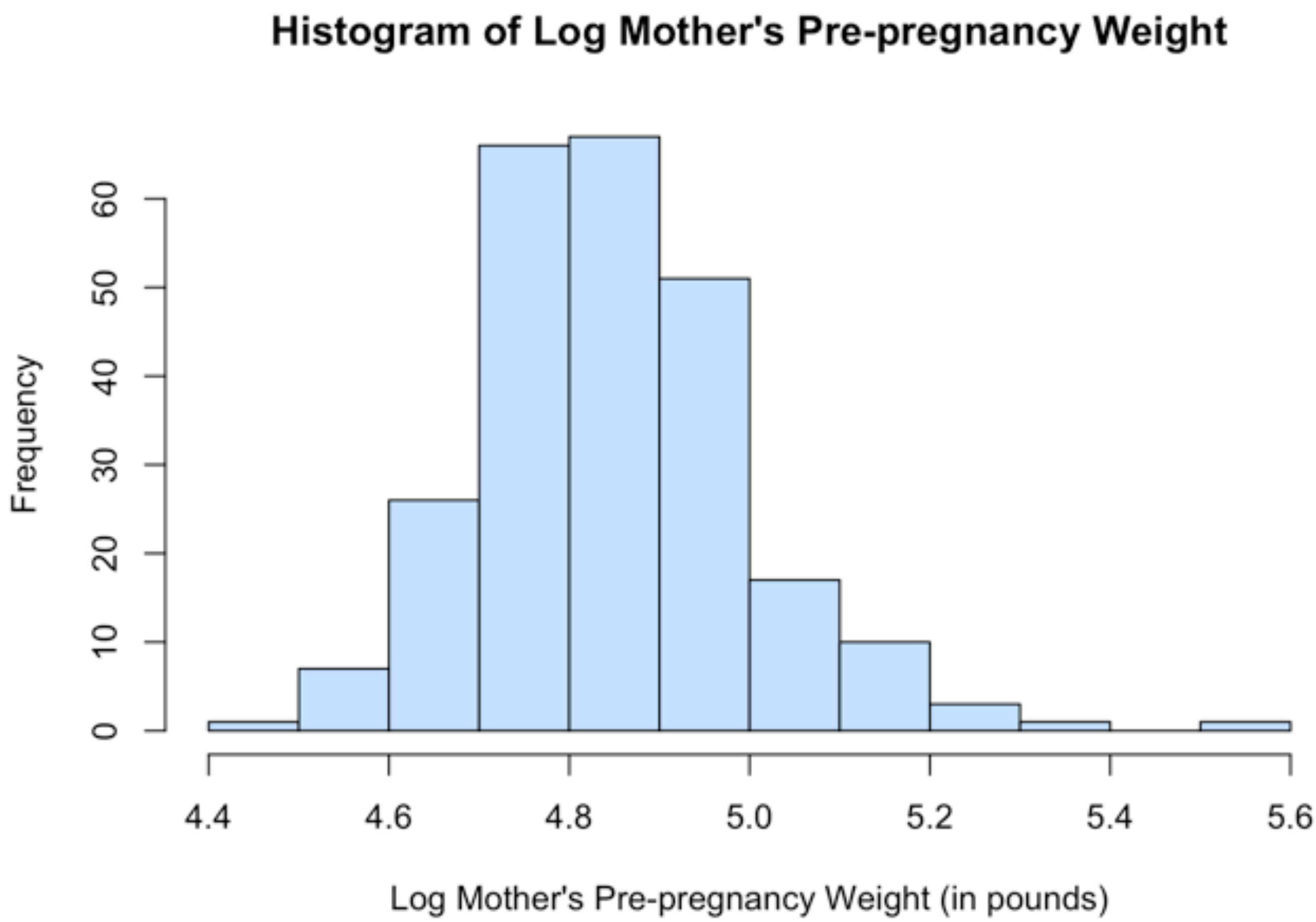

Figure 5. Histogram of Log Mother's Pre-pregnancy Weight was used to explore the distribution of the mother's pre-pregnancy weight (in pounds). It showed that the data is right skewed and the log of the data makes it more normal.

The histogram and QQ-plot both show that our data is normally distributed. In the QQ-plot most of the data follows the linear line and falls within the 95% confidence interval. The data also meets the criteria for Central Limit Theorem which means our data can be considered approximately normal. According to the Shapiro-Wilk normality test, we Fail to Reject H_0 that the data is normally distributed because our p-value is greater than 0.05. In other words, the data is normally distributed according to Shapiro-Wilk normality test. Our W is W = 0.9937.

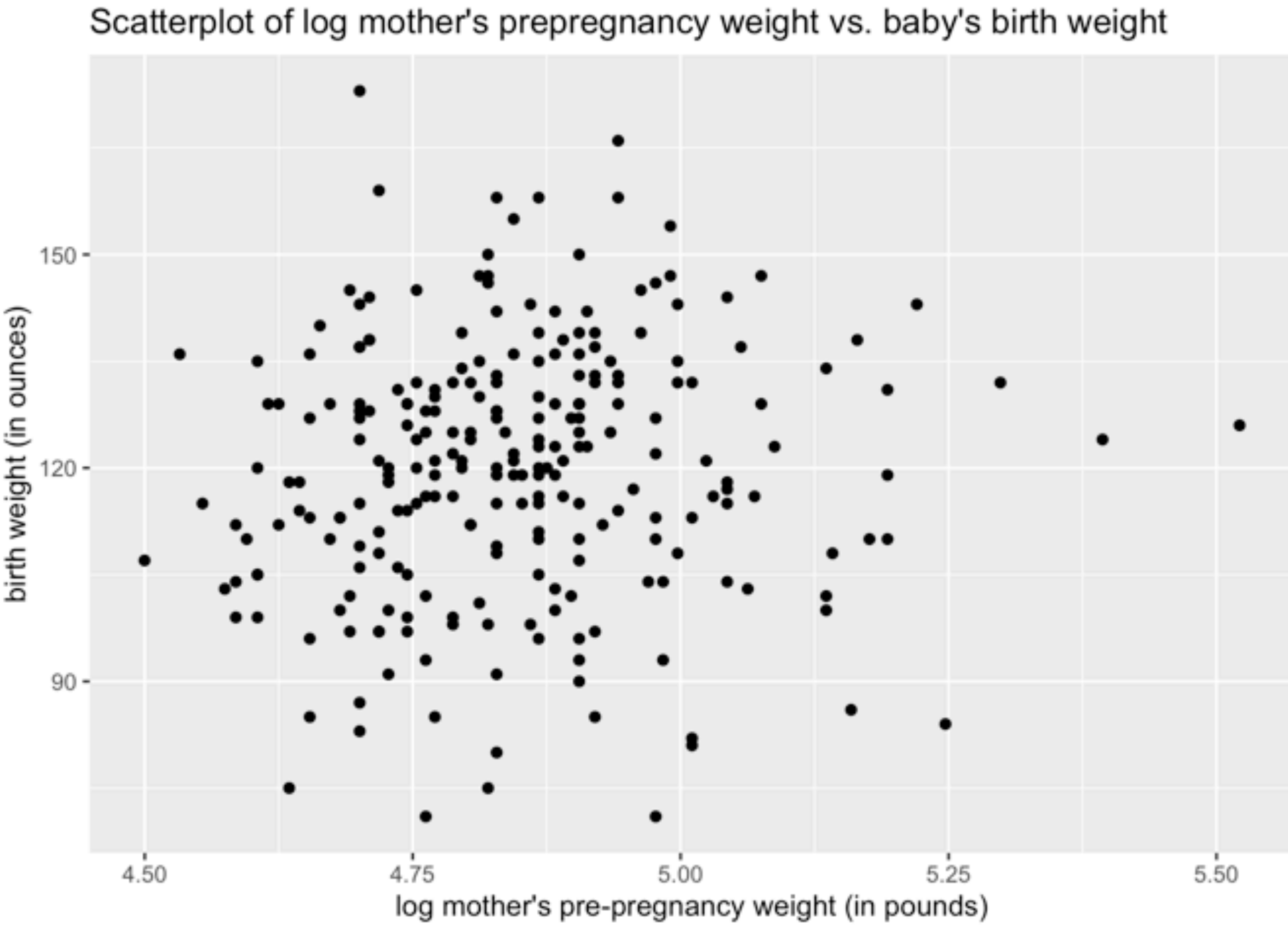

Figure 6. Scatterplot of log mother's prepregnancy weight (in pounds) vs. baby's birth weight (in ounces). It seems to be a large cluster of data in the middle left of the plot and possibly an outlier for mother's prepregnancy weight. Overall, there does not seem to be a relationship between the variables.

Statistical Methods

ANOVA

An ANOVA test is a statistical test that assesses whether there are any statistical differences between the means of three or more independent groups. It shows how a quantitative independent variable changes by categorical variables. In a one-way ANOVA, we look at one treatment with multiple levels (i.e., Birth Weight with different types of races and Birth Weight with different types of maternal smoking habits). The ANOVA finds if birth weight changes according to race and if it changes according to maternal smoking habits. Birth weight is the dependent variable. Race and smoke are the independent categorical variables. Race has 5 levels: "asian", "black", "mex", "mixed", or "white". Smoke has 4 levels: never, smokes now, until current pregnancy, once did and not now.

ANOVA for Birth Weight and Race

$$H_0 : \mu_{asian} = \mu_{black} = \mu_{mex} = \mu_{mixed} = \mu_{white} \text{ (all means are the same)}$$
$$H_A : \mu_i \neq \mu_j \text{ (at least one mean is different)}$$

where i = “asian”, “black”, “mex”, “mixed”, or “white”

where j = “asian”, “black”, “mex”, “mixed”, or “white”

The null hypothesis states that there is no difference in the mean birth weight between the five races.

The alternative hypothesis states that there is at least one mean different in the mean birth weight between the five races.

ANOVA Assumptions

The three ANOVA assumptions are random sample, normally distributed residuals, and equality of variance. Homogeneity of variance can be checked by looking at the residuals. Residuals are the difference between experimental and theoretical data represented by the red trend line in the Residuals vs. Fitted plot.

The random sampling assumption is met because all birth weight observations were randomly sampled from the Gestation data set and is a large sample size.

Based on the Residuals vs. Fitted plot, homogeneity of variance may be met because at the beginning of the plot the residual spread is small and at the end of the trend line the residual spread is only slightly wider. The trend line shows no distinct pattern.

Based on the QQ-plot, the data seems to be normal where most of the points fall within the 95% confidence interval. There are a few outliers in the data.

Based on Shapiro-Wilk’s test, the p-value 0.4476 is greater than 0.05 so we Fail to Reject H_0. The data does come from a normal distribution. Our W is W = 0.99417.

Levene’s Test was also used to find homogeneity of variance and our p-value 0.3454 is greater than 0.05 so we fail to Reject H_0. The variances are equal. Our F is 1.1247.

ANOVA for Birth Weight and Mother Smoking Habits

$$H_0 : \mu_{\text{never}} = \mu_{\text{smokes now}} = \mu_{\text{until current pregnancy}} = \mu_{\text{once did \& not now}} \text{ (all means are the same)}$$
$$H_A : \mu_i \neq \mu_j \text{ (at least one mean is different)}$$

where i = never, smokes now, until current pregnancy, once did & not now

where j = never, smokes now, until current pregnancy, once did & not now

The null hypothesis states that there is no difference in the mean birth weight between the 4 types of maternal smoking habits.

The alternative hypothesis states that there is at least one mean different in the mean birth weight between the 4 types of maternal smoking habits.

ANOVA Assumptions

The three ANOVA assumptions are random sample, normally distributed residuals, and equality of variance. Homogeneity of variance can be checked by looking at the residuals. Residuals are the difference between experimental and theoretical data represented by the red trend line in the Residuals vs. Fitted plot.

The random sampling assumption is met because all birth weight observations were randomly sampled from the Gestation data set and is a large sample size.

Based on the Residuals vs. Fitted plot, homogeneity of variance may be met because at the beginning of the plot the residual spread is small and at the end of the trend line the residual spread is only slightly wider. The trend line shows no distinct pattern.

Based on the QQ-plot, the data seems to be normal where most of the points fall within the 95% confidence interval. There are a few outliers in the data.

Based on Shapiro-Wilk's test, the p-value 0.3511 is greater than 0.05 so we Fail to Reject H_0. The data does come from a normal distribution. Our W is W = 0.9935.

Levene's Test was also used to find homogeneity of variance and our p-value 0.2917 is greater than 0.05 so we fail to Reject H_0. The variances are equal. Our F is 1.2514.

Multivariate Linear Model

Linear regression models the relationship between a scalar response variable and one or more explanatory variables. Creating a linear model, finds whether birth weight can be predicted by a combined mother's pre-pregnancy weight, race, and smoke. We already know from ANOVA test that race and smoke habits play a role in predicting birth weights.

Linear Regression Assumptions

According to Pearson's product-moment correlation, the correlation between birth weight (in ounces) and log mother's pre-pregnancy weight (in pounds) is 0.08 which is a high positive correlation.

$$H_0 : \beta_1 = 0$$
$$H_A : \beta_1 \neq 0$$

H_0: There is no statistically significant relationship between log mother's pre-pregnancy weight (in pounds) and their baby's birth weight (in ounces). The slope of the linear regression line is 0.

H_A : There is statistically significant relationship between log mother's pre-pregnancy weight (in pounds) and their baby's birth weight (in ounces). The slope of the linear regression line is not 0.

The linear regression assumptions include: random sample, Y is normally distributed with equal variance for all X, and residuals are normally distributed plus equal variance.

The random sampling assumption is met because all birth weight observations were randomly sampled from the Gestation data set and is a large sample size.

In Residuals vs. Fitted plot, we see that the red line is almost perfectly horizontal and that there is no distinct pattern in the residuals about the 0 line. Therefore, we can say our equal variance assumption is met.

In the Normal Q-Q plot shows that our residuals do not deviate much from the straight line aside from a few outliers. We may be able to conclude that our assumption of normality is met.

Based on Shapiro-Wilk's test, the p-value 0.4476 is greater than 0.05 so we Fail to Reject H_0. The data does come from a normal distribution. Our W is W = 0.99417.

Results

ANOVA

ANOVA for Birth Weight and Race

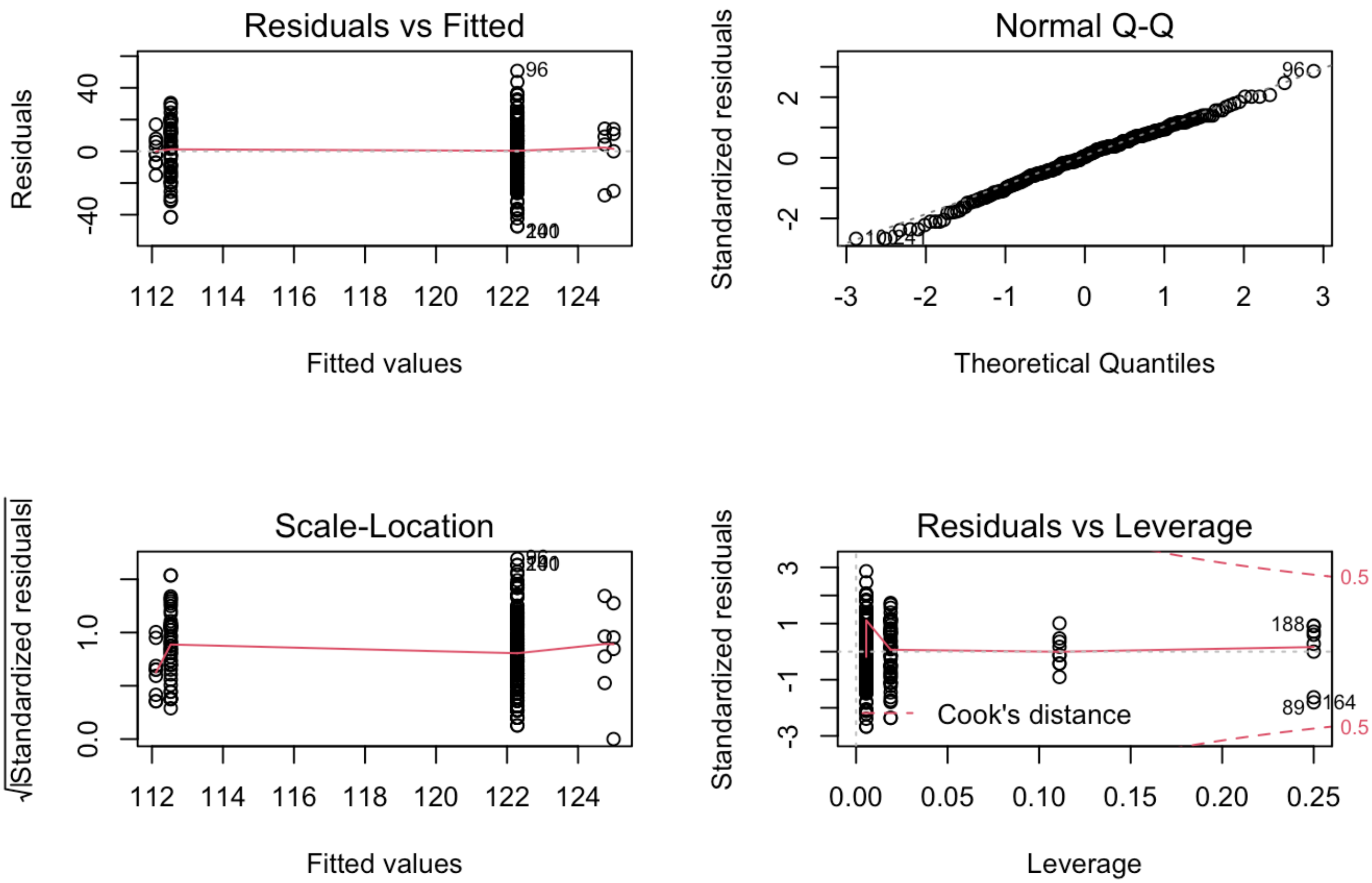

Figure 7. Diagnostic Plots for Birth Weight and Race. The residuals vs. fitted plot exhibits slight heteroskedasicity due to the slight cone shape of the distribution of data points. The dataset is large and meets Central Limit Theorem criteria so the data can assume normality. The trend line also shows no distinct pattern. The normal Q-Q plot shows normality where majprity of the data points fall within the 95% confidence interval.

```
##         Df   Sum Sq  Mean Sq  F value Pr(>F)
## 1 race   4 4651.152 1162.788 3.685239  0.006  **
```

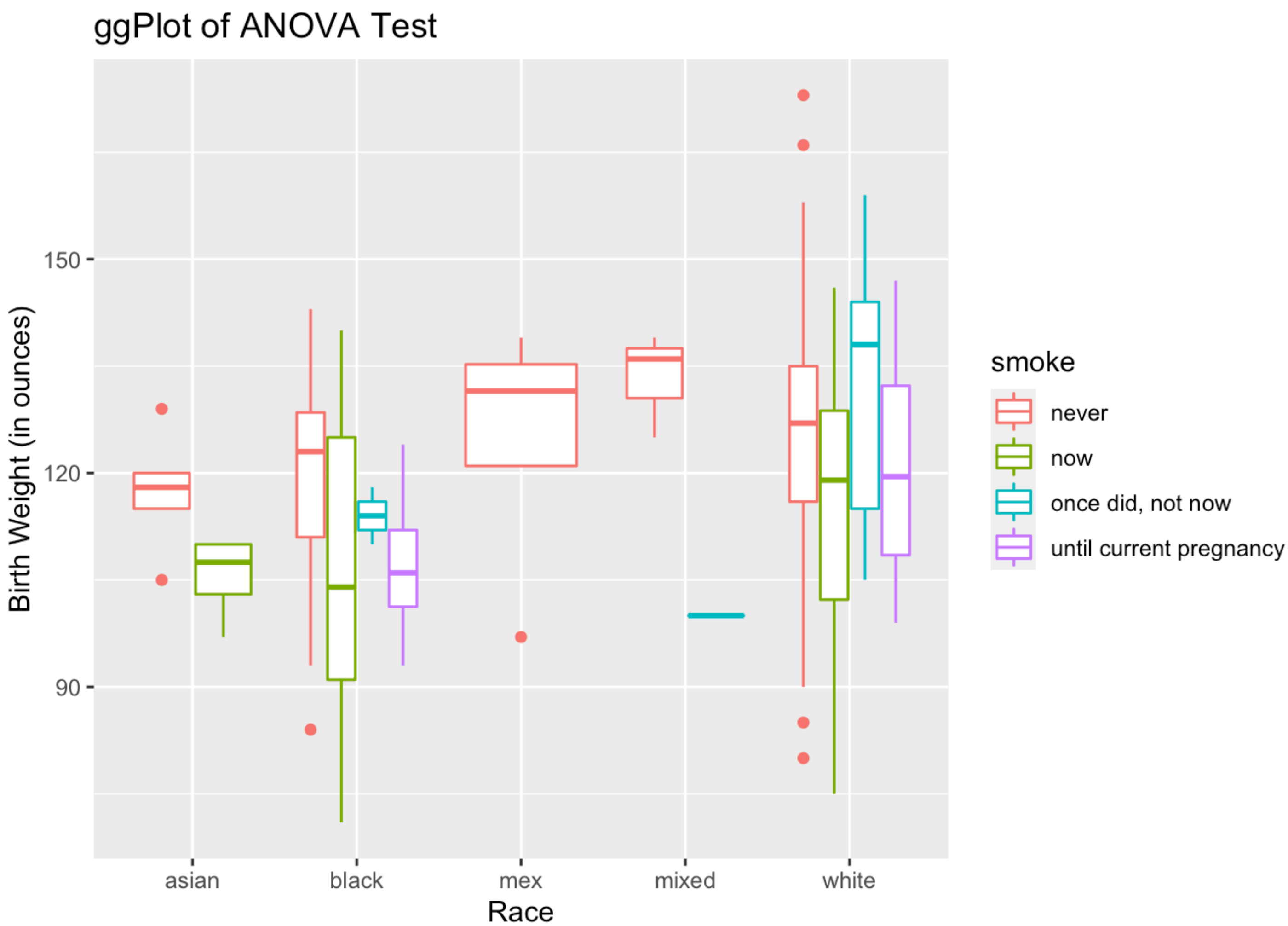

Figure 8. ANOVA and Boxplot of birth weight by race and colored by smoking habits. Some racial groups seem to be less likely to have certain groups than others. The birth weight range also varies significantly for white and black more than the other races. The ANOVA shows small p-value meaning race is a significant predictor of birth weight.

The p-value 0.006196 is less than 0.05 so we Reject H_0 and conclude that at least one mean birth weight between the five races is different from the other races. Our F is 3.685239.

The race group of the mother significantly affects the mean birth weight.

Post-hoc Test

We know that race makes a significant difference on birth weights. According to the Tukey multiple comparisons of means (alpha = 0.05), the race groups that are the most significantly different from each other are the race white and the race black with p-value 0.005.

ANOVA for Birth Weight and Mother Smoking Habits

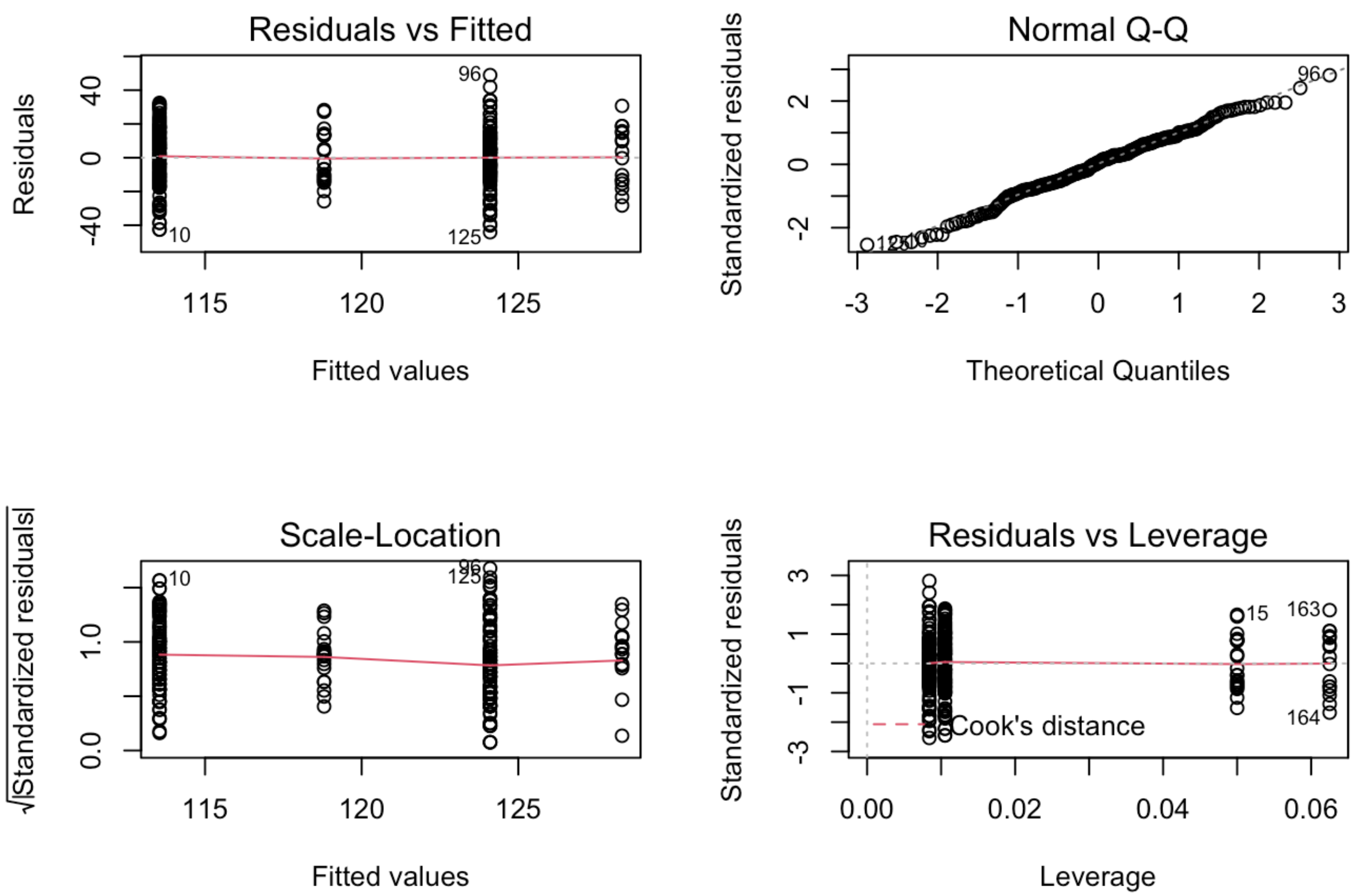

Figure 9. Diagnostic Plots for Birth Weight and Smoke. The residuals vs. fitted plot exhibits homogenity of variance. The trend line also shows no distinct pattern. The normal Q-Q plot shows normality where majprity of the data points fall within the 95% confidence interval.

Based on the residual plot, homogeneity of variance may be met because the residual spread is about the same more or less. Based on the QQ-plot, the data seems to be normal where most of the points fall within the 95% confidence interval. There are a few outliers in the data. Based on Shapiro-Wilk's test, the p-value 0.3511 is greater than 0.05 so we Fail to Reject H_0. Our W is W = 0.9935. The data does come from a normal distribution. Based on the Levene's Test, our p-value 0.2917 is greater than 0.05 so we Fail to Reject H_0. The variances are equal. Our F is 1.2514.

```
##          Df   Sum Sq  Mean Sq  F value Pr(>F)
## 1 smoke  3 7090.012 2363.337 7.765728 <0.001 ***
```

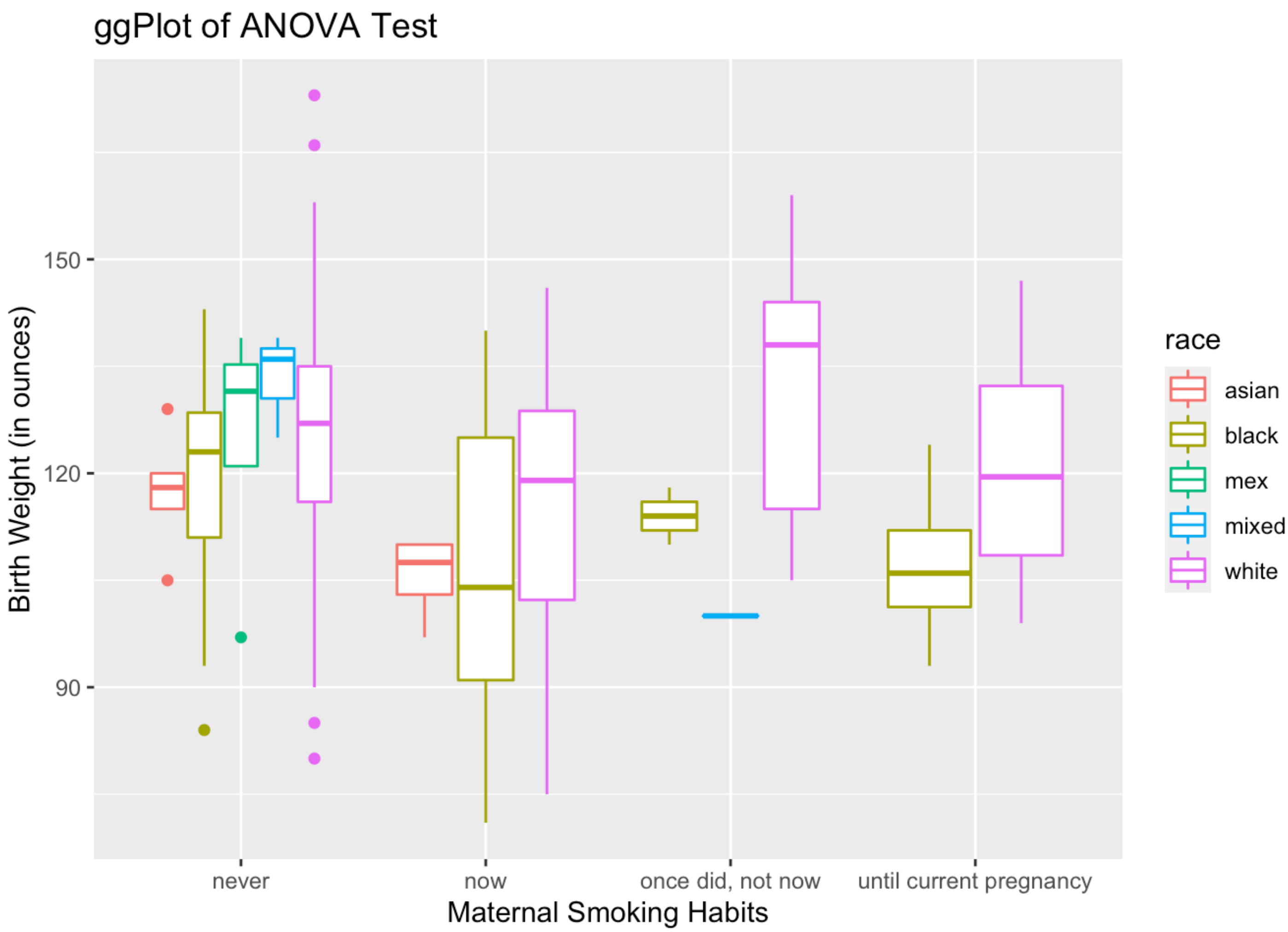

Figure 10. ANOVA and Boxplot of birth weight by smoke and colored by race. Some racial groups seem to be less likely to have certain groups than others. The birth weight range also varies significantly for smokers now and non-smokers more than the other groups. The ANOVA shows small p-value meaning maternal smoke habits are a significant predictor of birth weight.

The p-value 5.639e-05 is less than 0.05 so we Reject H_0 and conclude that at least one mean birth weight between the 4 types of maternal smoking habits is different from the others. The F value for test is 7.765728.

The maternal smoking habits significantly affect the mean birth weight.

Post-hoc Test

We know that maternal smoking habits make a significant difference on birth weights. According to the Tukey multiple comparisons of means (alpha = 0.05), the maternal smoking habits that are the most significantly different from each other are mother's who smoke now and ones who have never smoked with p-value 0.00. The next most significantly different are mother's who once did, not now smoke compared to those who now smoke with p-value 0.10.

Multivariate Linear Model

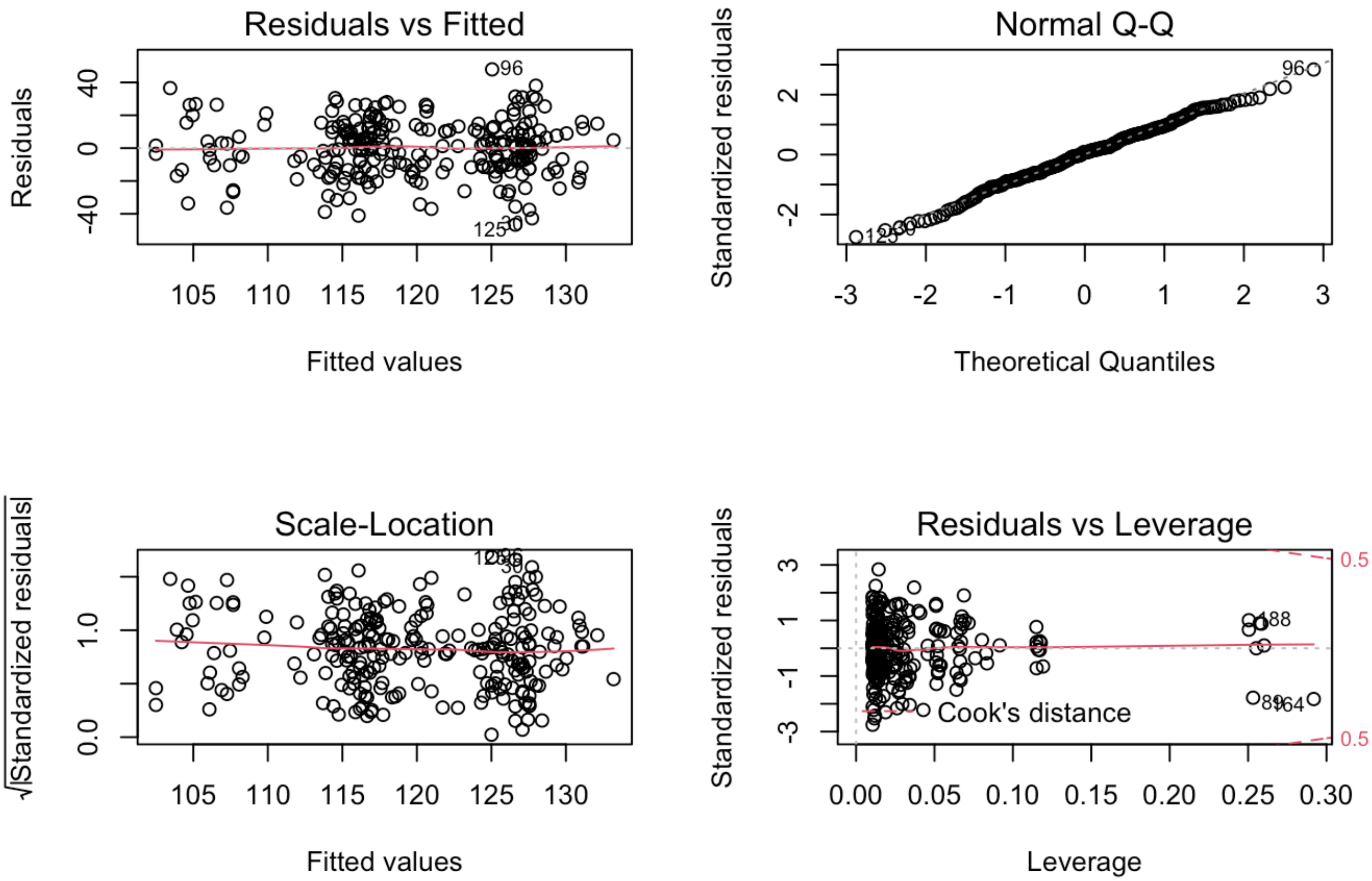

Figure 11. Diagnostic Plots for linear model birth weight with explanatory variables race and smoke. The residuals vs. fitted plot exhibits homogenity of variance. The trend line also shows no distinct pattern. The normal Q-Q plot shows normality where majprity of the data points fall within the 95% confidence interval.

```
## `geom_smooth()` using formula 'y ~ x'
```

Figure 12. Scatterplot shows not much correlation between Birth Weight (in ounces) and log maternal pre-pregnancy weight (in pounds)

The p-value 0.1062 is greater than alpha = 0.05. We fail to reject our H_0 and say that log mother's pre-pregnancy weight (in pounds) is not a significant predictor of birth weight. (i.e. the slope is 0). More specifically, as we increase the log mother's pre-pregnancy weight (in pounds), we do not see a significant increase in birth weight. When doing this test we hold race and maternal smoking habits all the same after discovering from the ANOVA tests that they are significant predictors of birth weight. The t statistic is 1.714.

Discussion

The results of the ANOVA shows that different racial groups and maternal smoking habits play a significant role in predicting birth weights. This supports the information that pregnant women who smoke are more likely to have babies with low birth weight (March of Dimes). On top of that, it is said that being a member of a group that experiences the effects of racism can also lead to having babies with low birth weights (March of Dimes). The linear regression model test proves that log mother's pre-pregnancy weight does is not a significant predictor in predicting birth weights.

There are limitations to the linear regression model such that adding too many variables could over fit the sample data. I would have liked to compare more variables considering we know that race and smoking habits are significant predictors of birth weight. After finding out which races and which smoking habits were significant factors in maternal influence on their baby's birth weight, I would have liked to adjust the data and make analyses on these specific factors. I would also like to find other factors like why there is such a big gap between the white group and black group when it comes to birth weight. Also, it would be interesting to discover the risk factors tied to smoking while pregnant. While pre-pregnancy weight does not play a significant role in predicting birth weights, I would have liked to consider whether mother's present weight at childbirth is a significant predictor or possibly the weight gained during pregnancy (March of Dimes & Large for Gestational Age).

As a result, infant birth weights can be predicted by their mother's racial group and smoking habits.

References

B. Hofner (2021). papeR: A Toolbox for Writing Pretty Papers and Reports, R package version 1.0-5,

https://CRAN.R-project.org/package=papeR. (https://CRAN.R-project.org/package=papeR.)

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Hadley Wickham, Jim Hester and Jennifer Bryan (2022). readr: Read Rectangular Text Data. R package version 2.1.2. https://CRAN.R-project.org/package=readr (https://CRAN.R-project.org/package=readr)

Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2022). dplyr: A Grammar of Data Manipulation. R package version 1.0.8. https://CRAN.R-project.org/package=dplyr (https://CRAN.R-project.org/package=dplyr)

John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/ (https://socialsciences.mcmaster.ca/jfox/Books/Companion/)

Large for Gestational Age - Health Encyclopedia - University of Rochester Medical Center. (n.d.). University of Rochester Medical Center. Retrieved June 9, 2022, from https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=90&ContentID=P02383 (https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=90&ContentID=P02383)

March of Dimes. (n.d.). Low Birthweight. Retrieved June 9, 2022, from https://www.marchofdimes.org/complications/low-birthweight.aspx (https://www.marchofdimes.org/complications/low-birthweight.aspx)

Randall Pruim, Daniel Kaplan and Nicholas Horton (2021). mosaicData: Project MOSAIC Data Sets. R package version 0.20.2. https://CRAN.R-project.org/package=mosaicData (https://CRAN.R-project.org/package=mosaicData)

Revelle, W. (2022) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych (https://CRAN.R-project.org/package=psych) Version = 2.2.5.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686 (https://doi.org/10.21105/joss.01686)

Appendix

```
knitr::opts_chunk$set(echo = FALSE)

#important packages!
library(readr)
library(ggplot2)
library(car)
library(tidyverse)
library(mosaicData)
library(psych)
library(dplyr)
library(papeR)
# load in your data here
data(Gestation)

# dimensions (1236 rows, 23 columns)
dim(Gestation)

# structure of data
str(Gestation)

gest <- Gestation %>%
  dplyr::select(wt, race, wt.1, smoke) %>% # subset columns
  mutate(across(c(race, smoke), factor)) %>% # mutate variabnles into a factor
  drop_na() # drops NA values
# sample Gestation data
set.seed(7)
birthwt <- sample_n(gest, 250)

# row 250, columns 4
dim(birthwt)
str(birthwt)
summary(birthwt)

# scatterplot
ggplot(birthwt, aes(wt.1, wt)) +
  geom_point() +
  ggtitle("Scatterplot of mother's prepregnancy weight vs. baby's birth weight") +
  xlab("mother's pre-pregnancy weight (in pounds)") +
  ylab("birth weight (in ounces)")
# Boxplot
boxplot(wt~race, data=birthwt,
        main = "Boxplot of Birth Weight by Mother's Race",
        xlab = "Race",
        ylab = "Birth Weight",
        col = "paleturquoise1")
# Boxplot
boxplot(wt~smoke, data=birthwt,
        main = "Boxplot of Birth Weight by Mother's Smoking Habits",
        xlab = "Smoking Habits",
        ylab = "Birth Weught (in ounces)",
```

```
        col = "darkseagreen1")
#check normality of your y variable here
hist(birthwt$wt,
     main = "Histogram of Birth Weight",
     xlab = "Birth Weight (in ounces)",
     col = "thistle1")
qqPlot(birthwt$wt,
       main="QQ-Plot of Birth Weight (in ounces)",
       ylab = "Birth Weight (in ounces)")
hist(log(birthwt$wt.1),
     main = "Histogram of Log Mother's Pre-pregnancy Weight",
     xlab = "Log Mother's Pre-pregnancy Weight (in pounds)",
     col = "slategray1")
# adding log mother's weight to data
birthwt$log_wt.1 <- log(birthwt$wt.1)
birthwt <- birthwt[-3]
#shapiro.test
shapiro.test(birthwt$wt)
# scatterplot
ggplot(birthwt, aes(log_wt.1, wt)) +
  geom_point() +
  ggtitle("Scatterplot of log mother's prepregnancy weight vs. baby's birth weight")
        +
  xlab("log mother's pre-pregnancy weight (in pounds)") +
  ylab("birth weight (in ounces)")
fit <- lm(wt~race, data=birthwt)

par(mfrow=c(1,1)) # resets output view
res <- fit$residuals

# TEST NORMALITY
qqPlot(res)
shapiro.test(res)

# Levene's Test
leveneTest(wt~race, data=birthwt)

summary(fit)
fit <- lm(wt~smoke, data=birthwt)

par(mfrow=c(1,1)) # resets output view
res <- fit$residuals

# TEST NORMALITY
qqPlot(res)
shapiro.test(res)

# Levene's Test
leveneTest(wt~smoke, data=birthwt)

summary(fit)
```

```
fit <- lm(wt~race, data=birthwt)

cor.test(birthwt$wt, birthwt$log_wt.1, method="pearson", alternative="two.sided")

shapiro.test(fit$residuals) #yep its normo
fit <- lm(wt~race, data=birthwt)

par(mfrow=c(2,2))

# Diagnostic Plots- residual plot and qqplot
plot(fit)
# One-Way ANOVA
prettify(anova(fit))

#ggPlot The Data
ggplot(birthwt, aes(x=race, y = wt, color=smoke))+
geom_boxplot() +
labs(title="ggPlot of ANOVA Test", x="Race", y="Birth Weight (in ounces)")
#step 2: run the anova

raceaov <- aov(wt ~ race, data=birthwt)
summary(raceaov)

#step 3: post-hoc analyses

TukeyHSD(raceaov)
fit <- lm(wt~smoke, data=birthwt)

par(mfrow=c(2,2))

# Diagnostic Plots- residual plot and qqplot
plot(fit)
# One-Way ANOVA
prettify(anova(fit))

#ggPlot The Data
ggplot(birthwt, aes(x=smoke, y=wt, color=race))+
geom_boxplot() +
labs(title="ggPlot of ANOVA Test", x="Maternal Smoking Habits", y = "Birth Weight (in
        ounces)")
#step 2: run the anova

smokeaov <- aov(wt ~ smoke, data=birthwt)
summary(smokeaov)

#step 3: post-hoc analyses

TukeyHSD(smokeaov)
wt.lm <- lm(wt~log_wt.1+race+smoke, data=birthwt)

par(mfrow=c(2,2))
```

```
plot(wt.lm)
# ggPlot
ggplot(birthwt, aes(x=log_wt.1, y=wt)) +
  geom_point() +
  geom_smooth(method="lm") +
labs(title="ggPlot of Linear Regression", x="log maternal pre-pregnancy weight (in po
        unds)", y ="Birth Weight (in ounces)")
summary(wt.lm)
# Appendix: All code for this report
#Cite Packages
citation('readr')
citation('ggplot2')
citation('car')
citation('tidyverse')
citation('mosaicData')
citation('psych')
citation('dplyr')
citation("papeR")
```

```
##
## To cite package 'readr' in publications use:
##
##   Hadley Wickham, Jim Hester and Jennifer Bryan (2022). readr: Read
##   Rectangular Text Data. R package version 2.1.2.
##   https://CRAN.R-project.org/package=readr
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {readr: Read Rectangular Text Data},
##     author = {Hadley Wickham and Jim Hester and Jennifer Bryan},
##     year = {2022},
##     note = {R package version 2.1.2},
##     url = {https://CRAN.R-project.org/package=readr},
##   }
```

```
##
## To cite ggplot2 in publications, please use:
##
##   H. Wickham. ggplot2: Elegant Graphics for Data Analysis.
##   Springer-Verlag New York, 2016.
##
## A BibTeX entry for LaTeX users is
##
##   @Book{,
##     author = {Hadley Wickham},
##     title = {ggplot2: Elegant Graphics for Data Analysis},
##     publisher = {Springer-Verlag New York},
##     year = {2016},
##     isbn = {978-3-319-24277-4},
##     url = {https://ggplot2.tidyverse.org},
##   }
```

```
##
## To cite the car package in publications use:
##
##   John Fox and Sanford Weisberg (2019). An {R} Companion to Applied
##   Regression, Third Edition. Thousand Oaks CA: Sage. URL:
##   https://socialsciences.mcmaster.ca/jfox/Books/Companion/
##
## A BibTeX entry for LaTeX users is
##
##   @Book{,
##     title = {An {R} Companion to Applied Regression},
##     edition = {Third},
##     author = {John Fox and Sanford Weisberg},
##     year = {2019},
##     publisher = {Sage},
##     address = {Thousand Oaks {CA}},
##     url = {https://socialsciences.mcmaster.ca/jfox/Books/Companion/},
##   }
```

```
##
##   Wickham et al., (2019). Welcome to the tidyverse. Journal of Open
##   Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
##
## A BibTeX entry for LaTeX users is
##
##   @Article{,
##     title = {Welcome to the {tidyverse}},
##     author = {Hadley Wickham and Mara Averick and Jennifer Bryan and Winston Chang
and Lucy D'Agostino McGowan and Romain François and Garrett Grolemund and Alex Hayes
and Lionel Henry and Jim Hester and Max Kuhn and Thomas Lin Pedersen and Evan Miller
and Stephan Milton Bache and Kirill Müller and Jeroen Ooms and David Robinson and Dan
a Paige Seidel and Vitalie Spinu and Kohske Takahashi and Davis Vaughan and Claus Wil
ke and Kara Woo and Hiroaki Yutani},
##     year = {2019},
##     journal = {Journal of Open Source Software},
##     volume = {4},
##     number = {43},
##     pages = {1686},
##     doi = {10.21105/joss.01686},
##   }
```

```
##
## To cite package 'mosaicData' in publications use:
##
##   Randall Pruim, Daniel Kaplan and Nicholas Horton (2021). mosaicData:
##   Project MOSAIC Data Sets. R package version 0.20.2.
##   https://CRAN.R-project.org/package=mosaicData
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {mosaicData: Project MOSAIC Data Sets},
##     author = {Randall Pruim and Daniel Kaplan and Nicholas Horton},
##     year = {2021},
##     note = {R package version 0.20.2},
##     url = {https://CRAN.R-project.org/package=mosaicData},
##   }
##
## ATTENTION: This citation information has been auto-generated from the
## package DESCRIPTION file and may need manual editing, see
## 'help("citation")'.
```

```
##
## To cite the psych package in publications use:
##
##   Revelle, W. (2022) psych: Procedures for Personality and
##   Psychological Research, Northwestern University, Evanston, Illinois,
##   USA, https://CRAN.R-project.org/package=psych Version = 2.2.5.
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {psych: Procedures for Psychological, Psychometric, and Personality Re
search},
##     author = {William Revelle},
##     organization = { Northwestern University},
##     address = { Evanston, Illinois},
##     year = {2022},
##     note = {R package version 2.2.5},
##     url = {https://CRAN.R-project.org/package=psych},
##   }
```

```
##
## To cite package 'dplyr' in publications use:
##
##   Hadley Wickham, Romain François, Lionel Henry and Kirill Müller
##   (2022). dplyr: A Grammar of Data Manipulation. R package version
##   1.0.8. https://CRAN.R-project.org/package=dplyr
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {dplyr: A Grammar of Data Manipulation},
##     author = {Hadley Wickham and Romain François and Lionel Henry and Kirill Mülle
r},
##     year = {2022},
##     note = {R package version 1.0.8},
##     url = {https://CRAN.R-project.org/package=dplyr},
##   }
```

```
##
## To cite package 'papeR' in publications use:
##
##   B. Hofner (2021). papeR: A Toolbox for Writing Pretty Papers and
##   Reports, R package version 1.0-5,
##   https://CRAN.R-project.org/package=papeR.
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {{papeR}: A Toolbox for Writing Pretty Papers and Reports},
##     author = {Benjamin Hofner},
##     year = {2021},
##     note = {{R} package version 1.0-5},
##     url = {https://CRAN.R-project.org/package=papeR},
##   }
```