HW6Q1

Zachary Lazerick

11/4/2021

Load the Lawstat Library to Access the Brown-Forscythe-Levene Test

```
library(lawstat)
```

Read in the Data

```
Advertising = read.csv("Advertising.csv")
attach(Advertising)
```
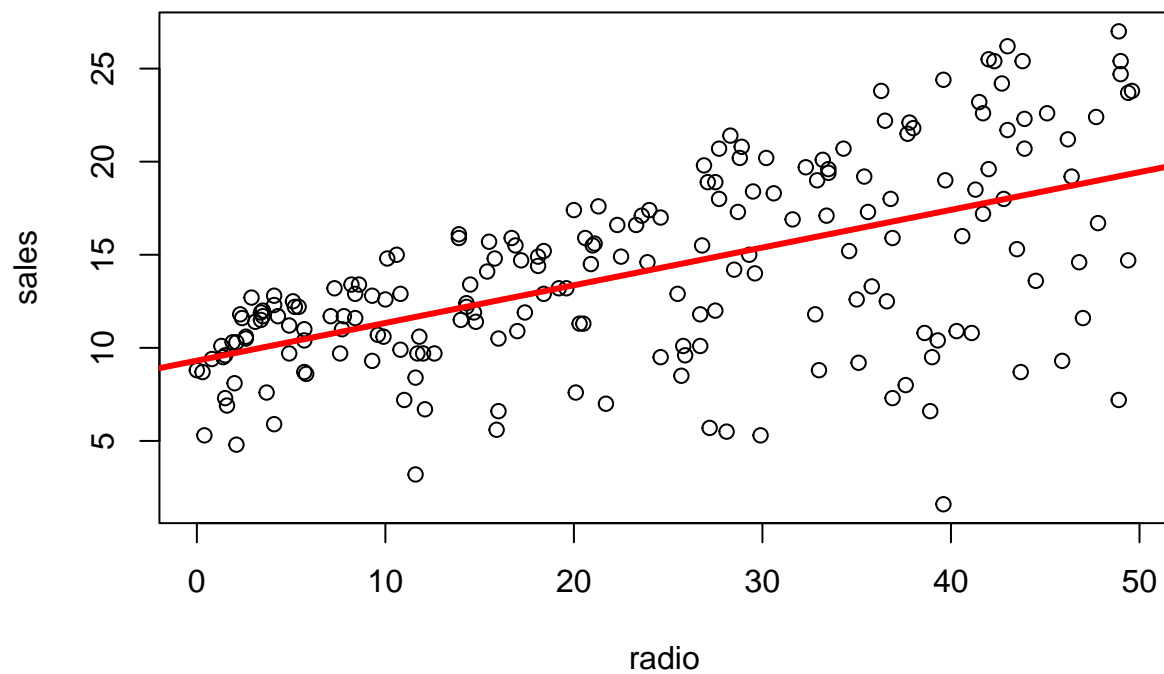
Question 1: Building an SLR Model Sales ~ Radio (Initial Results)

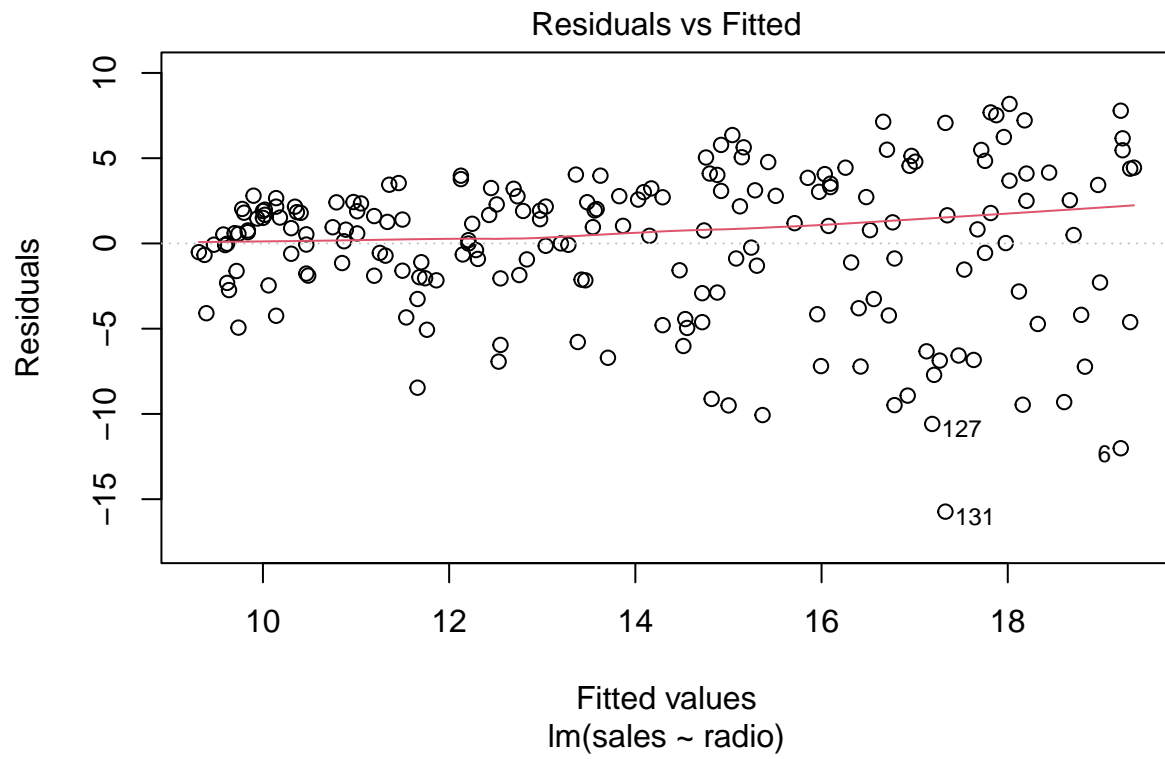Construct a Simple Linear Model with Predictor: Radio and Response: Sales

```
lm.fit.radio = lm(sales~radio, data = Advertising)
```
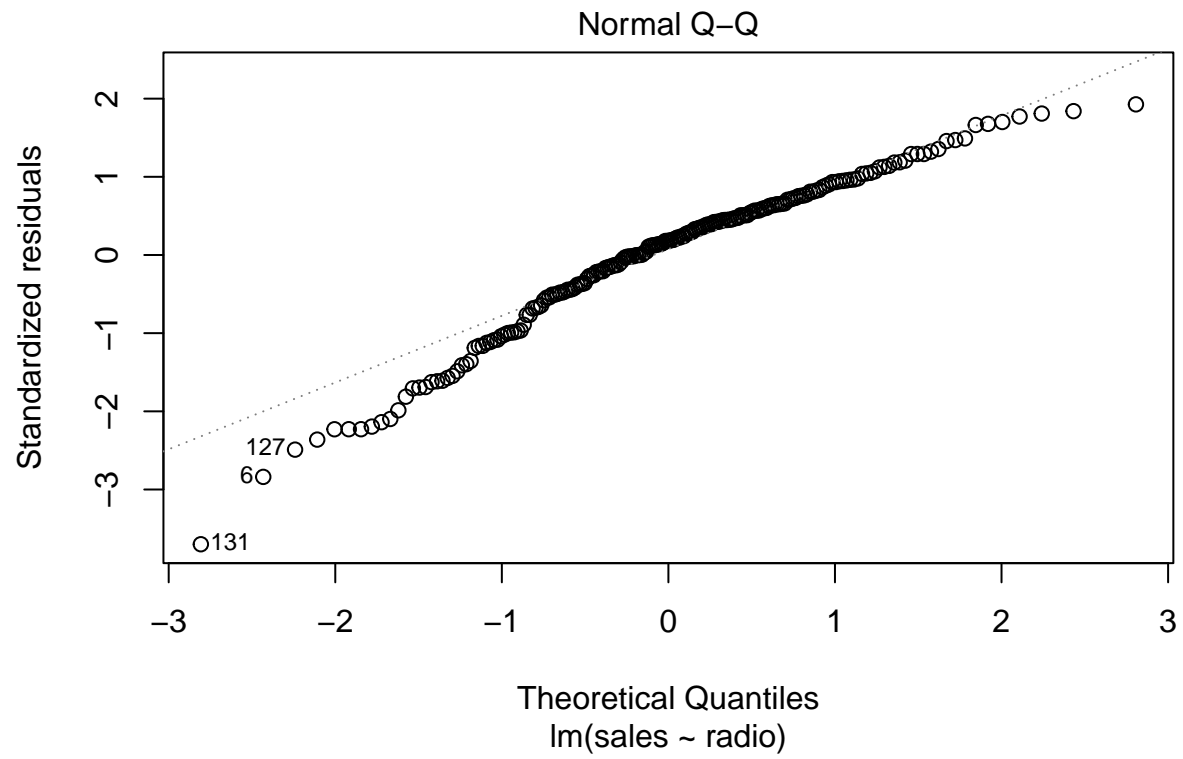
```
##
## Call:
## lm(formula = sales ~ radio, data = Advertising)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.7305  -2.1324   0.7707   2.7775   8.1810
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.31164    0.56290  16.542   <2e-16 ***
## radio        0.20250    0.02041   9.921   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.275 on 198 degrees of freedom
## Multiple R-squared:  0.332,  Adjusted R-squared:  0.3287
## F-statistic: 98.42 on 1 and 198 DF,  p-value: < 2.2e-16
```
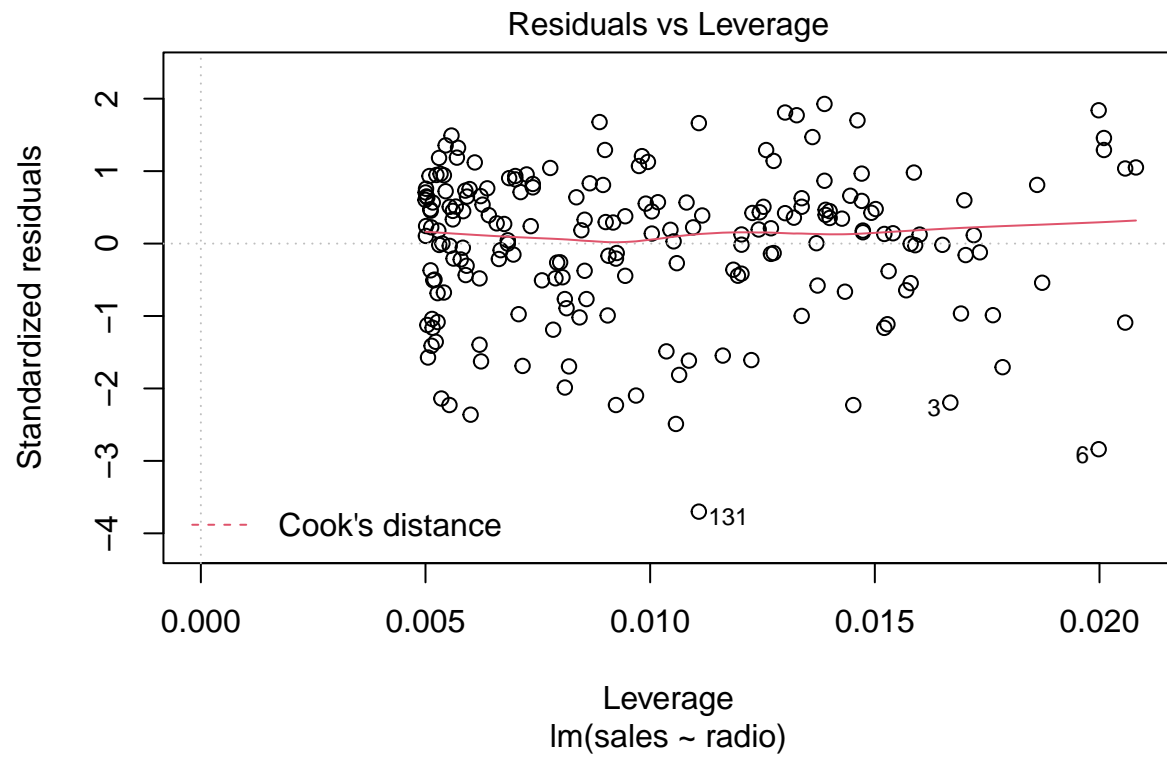
Construct a Scatter Plot with the Calculated Linear Model "Eye-Test" for Abnormalities (Non-linearity, etc.)

Construct Plots to Check Diagnostics

Residuals vs Fitted

Fitted values
lm(sales ~ radio)

Normal Q–Q

Standardized residuals

Theoretical Quantiles
lm(sales ~ radio)

Residuals vs Leverage

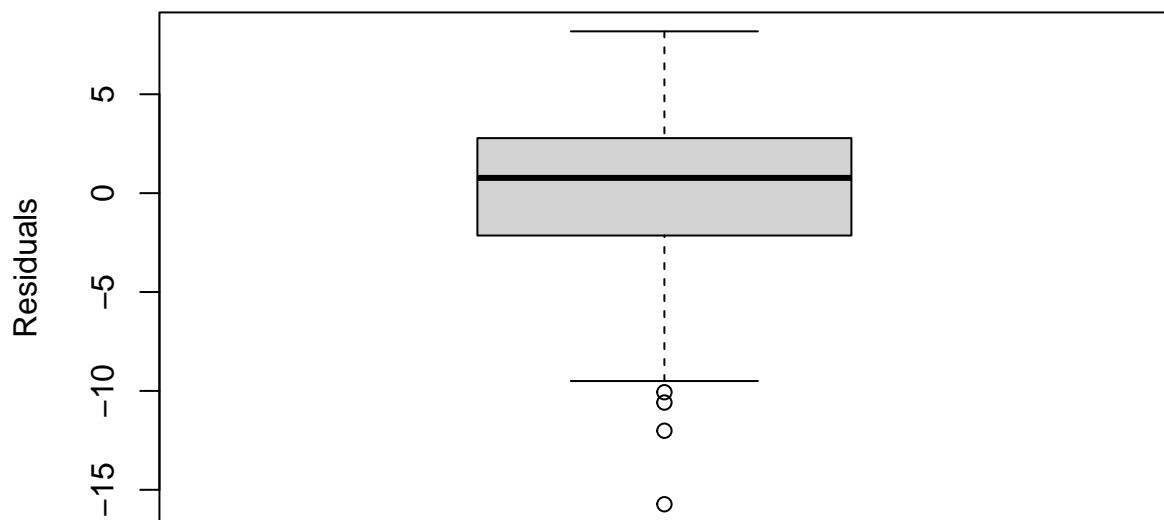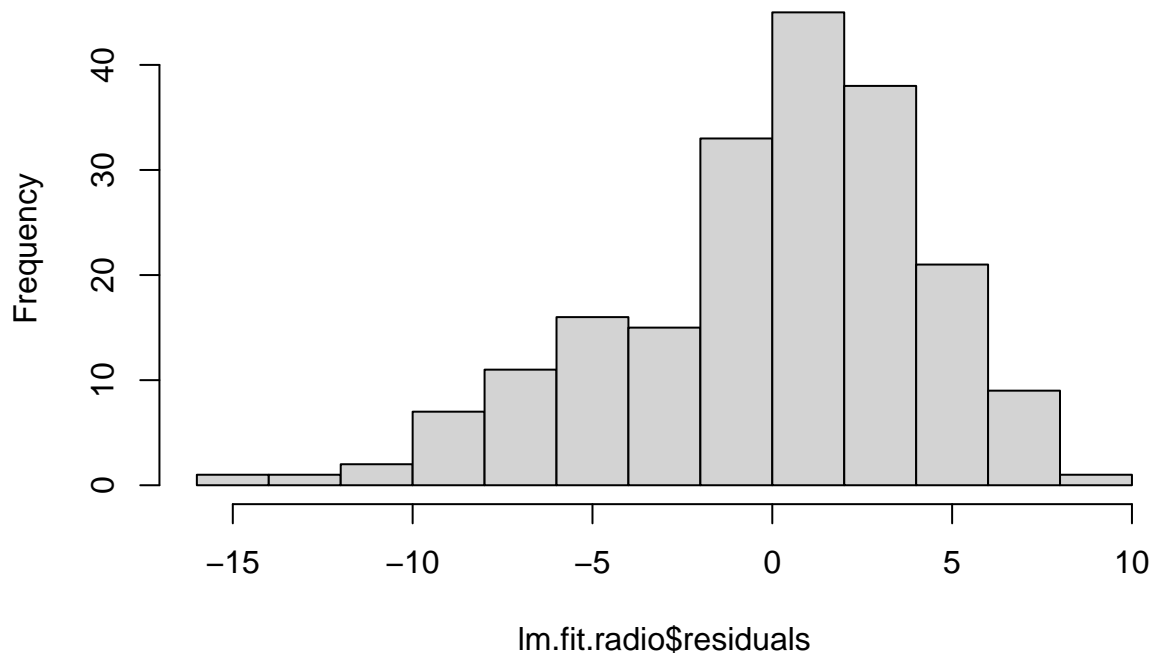lm(sales ~ radio)

Check Distribution of Residuals for Obvious Deviations

Histogram of lm.fit.radio$residuals

Run Shapiro-Wilk and Brown-Forscythe-Levene Tests

```
shapiro.test(lm.fit.radio$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  lm.fit.radio$residuals
## W = 0.96072, p-value = 2.367e-05
```

```
levene.test(lm.fit.radio$residuals, Advertising$Group, location = c("median"))
```

```
##
##  Modified robust Brown-Forsythe Levene-type test based on the absolute
##  deviations from the median
##
## data:  lm.fit.radio$residuals
## Test Statistic = 45.455, p-value = 1.671e-10
```

Conclusion From the Shapiro-Wilks Test, we conclude that the data is not from a normal distribution, and from the Levene Test, we conclude that the data is not homoscedastic. Therefore, we must transform the data in hopes of obtaining approximately normal and homoscedastic data to draw statistically significant conclusions from this data.

Data Transformation Test Results

Response ~ Predictor	Shapiro-Wilks P-Value	Levene P-Value
Sales ~ Radio	2.367e-05	1.671e-10
sqrt(Sales) ~ Radio	3.443e-09	4.465e-06
log10(Sales) ~ Radio	7.05e-14	.003309
(1/Sales) ~ Radio	2.2e-16	.3412
(1/Sales) ~ sqrt(Radio)	2.2e-16	.2836

Conclusion Upon transforming our Response Variable Sales to (1/Sales), we are able to obtain homoscedastic data. However, we are unable to transform the corresponding Predictor Radio, to obtain approximately normal data in conjunction with (1/Sales) to draw statistically significant data because one of the levels of Radio is 0. Therefore, transformations like (log10Radio) and (1/Radio) yield undefined values, and thus a simple linear regression model, with Shapiro-Wilks and Levene test results, cannot be obtained from these transformations.