mtCars: Vehicle Peformance

Alberta “Albi” Kovatcheva

Background

- The information in the ‘mtcars’ data set was extracted from the 1974 Motor Trend US magazine and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973-74 models).

Scenario

- A researcher is using the mtcars data frame to examine the effect that engine horsepower (hp) and vehicle weight (wt), measured in thousands of pounds) have on the time necessary to travel one quarter mile from a standing start (qsec).

Requirements

- For the results of Pearson' product-moment correlation to be valid, all variables involved in this analysis must be normally distributed. The researcher verified that they are:
  - Engine horsepower, vehicle weight, and quarter mile time.

Task

1. Create a scatter plot with a trend line where the horizontal axis is engine horsepower, and the vertical axis is quarter mile time. What is the relationship between time and engine horsepower: positively correlated, negatively correlated, or uncorrelated?

2. Compute the linear regression for time and engine horsepower. What is the equation of the line? What is the R-squared value? Is this what you would expect?

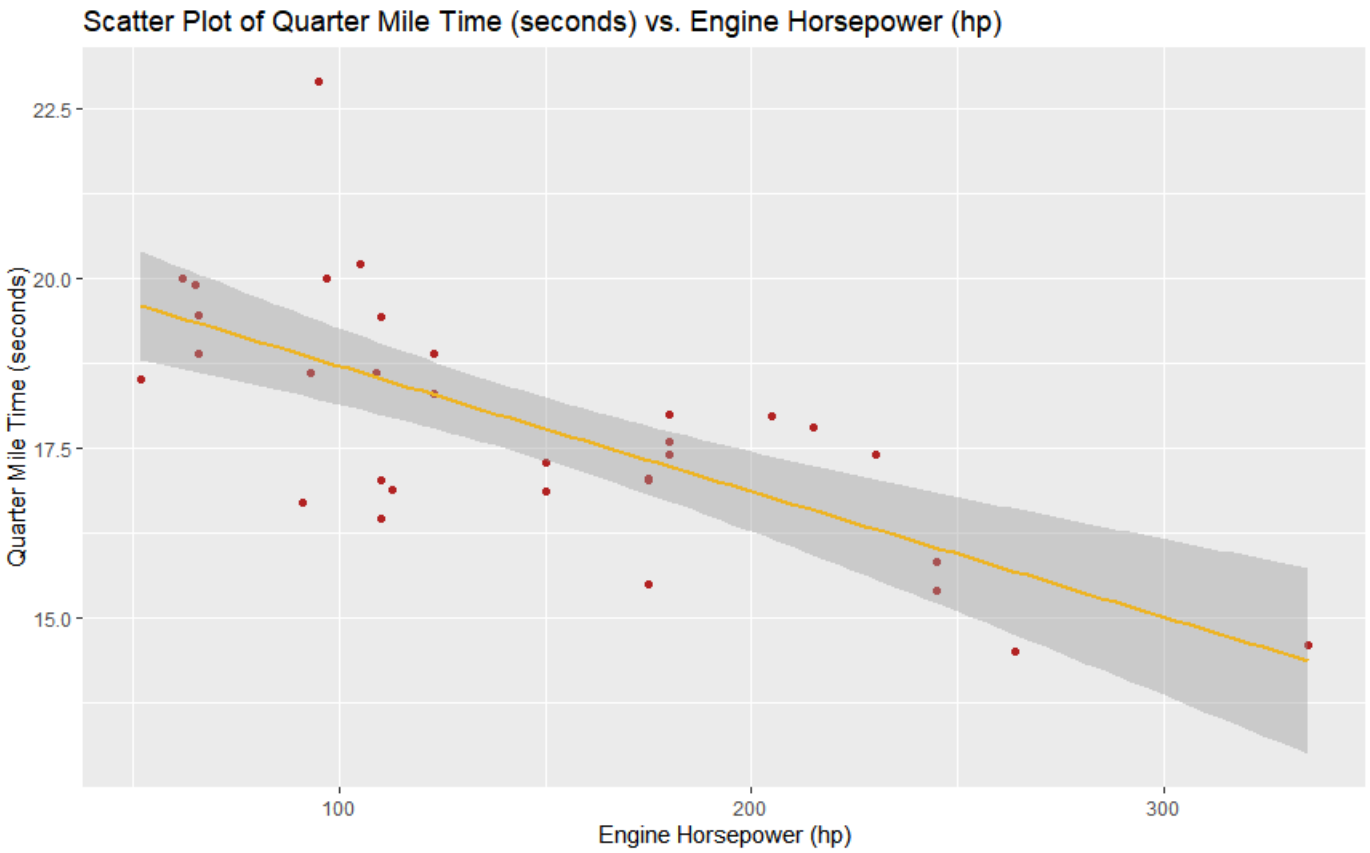
3. Create a scatter plot with a trend line where the horizontal axis is vehicle weight, and the vertical axis is quarter mile time. What is this relationship: positively correlated, negatively correlated, or uncorrelated?

4. Compute the linear regression for these two variables. What is the equation of the line? What is the R-squared value? Is this what you would expect?

5. Show your results and the code you used to generate the results. Please include your interpretation of the data included and answer all the questions posed above.

Part 1: Quarter Mile Time vs. Engine Performance

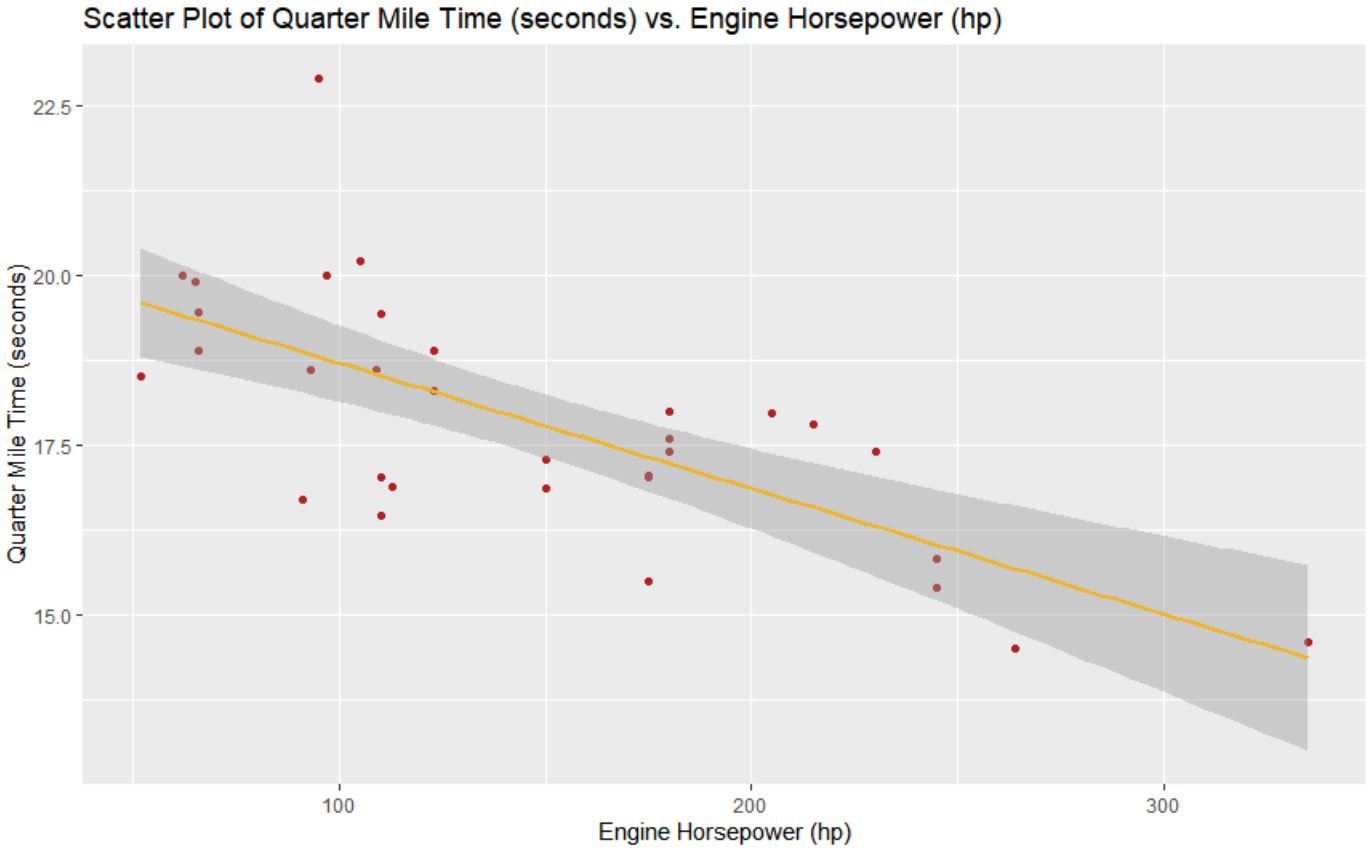
Scatter Plot of Quarter Mile Time vs. Engine Performance



Code to create this scatter plot:

```
ggplot(mtcars, aes(x = hp, y = qsec))
+
  geom_point(color = "firebrick") +
  ggtitle("Scatter Plot of Quarter Mile
Time (seconds) vs. Engine
Horsepower (hp)") +
  xlab("Engine Horsepower (hp)") +
  ylab("Quarter Mile Time
(seconds)") +
  geom_smooth(method = lm, se =
TRUE, color = "goldenrod2")
```

Analysis

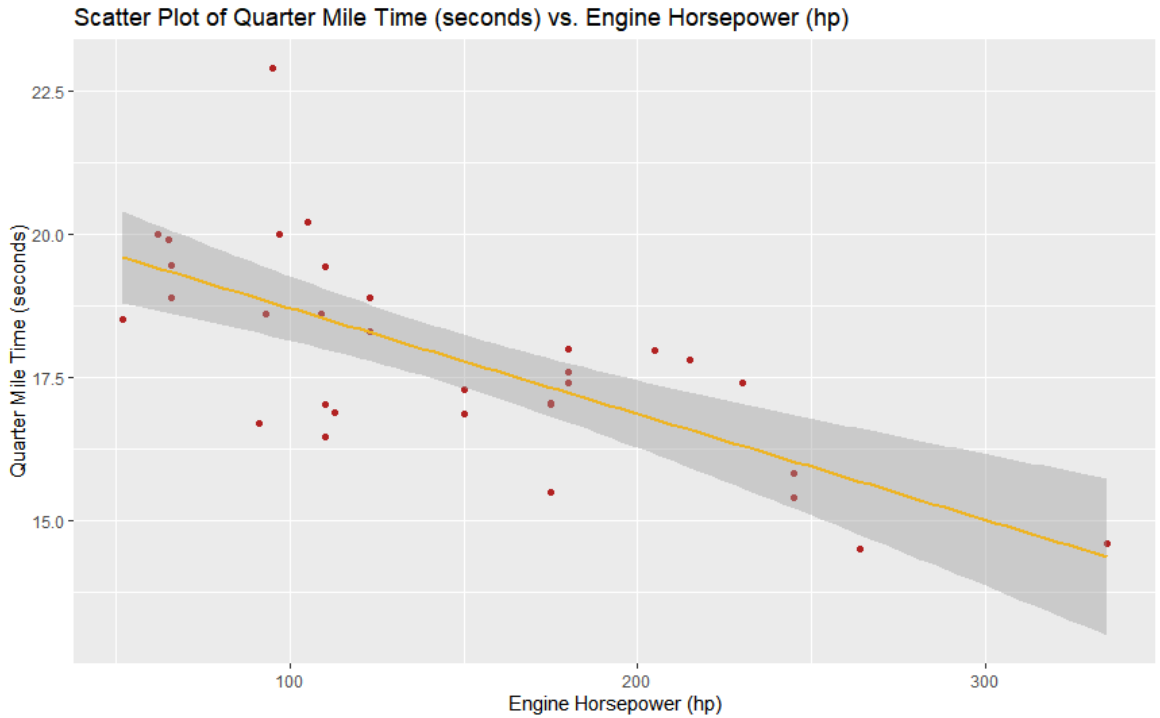


- There seems to be a negative correlation between quarter mile time and engine horsepower. As horsepower increases, quarter mile time decreases.

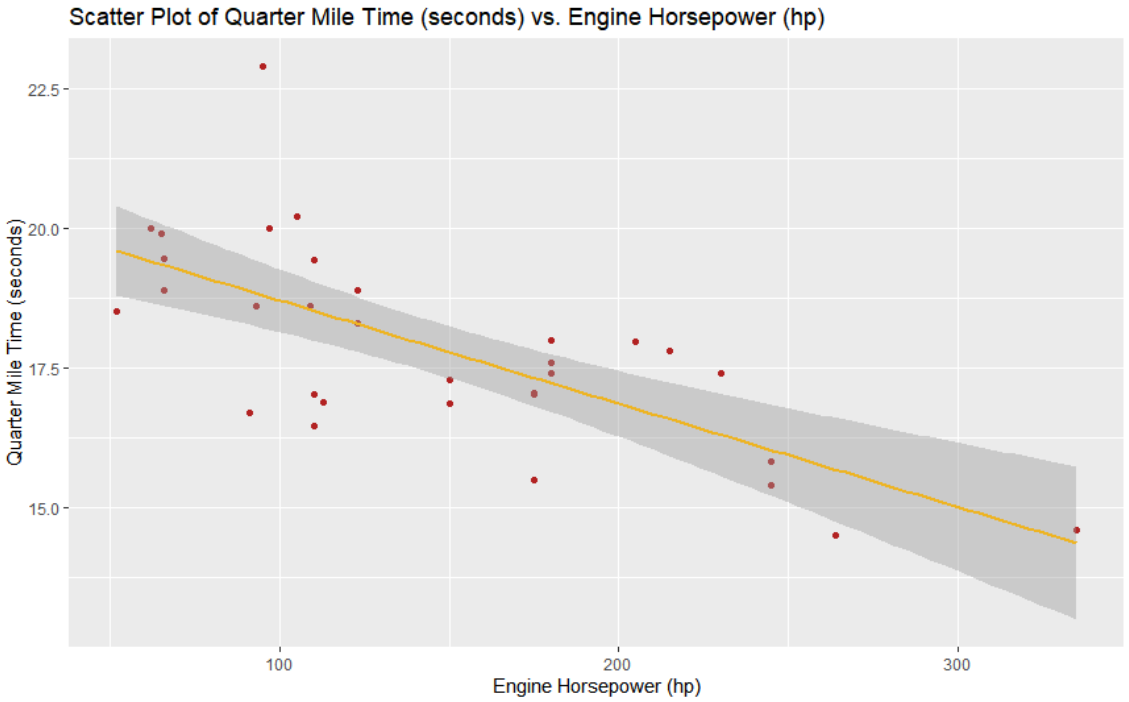
Correlation of Engine Performance with Quarter Mile Time

Code to compute Correlation:

```
cor.test(mtcars$hp, mtcars$qsec, method = "pearson", use = "complete.obs")
```



		Meaning
p-value	5.766e-06	Statistically Significant
R (Correlation Coefficient)	-0.7082234	Strong Negative Correlation

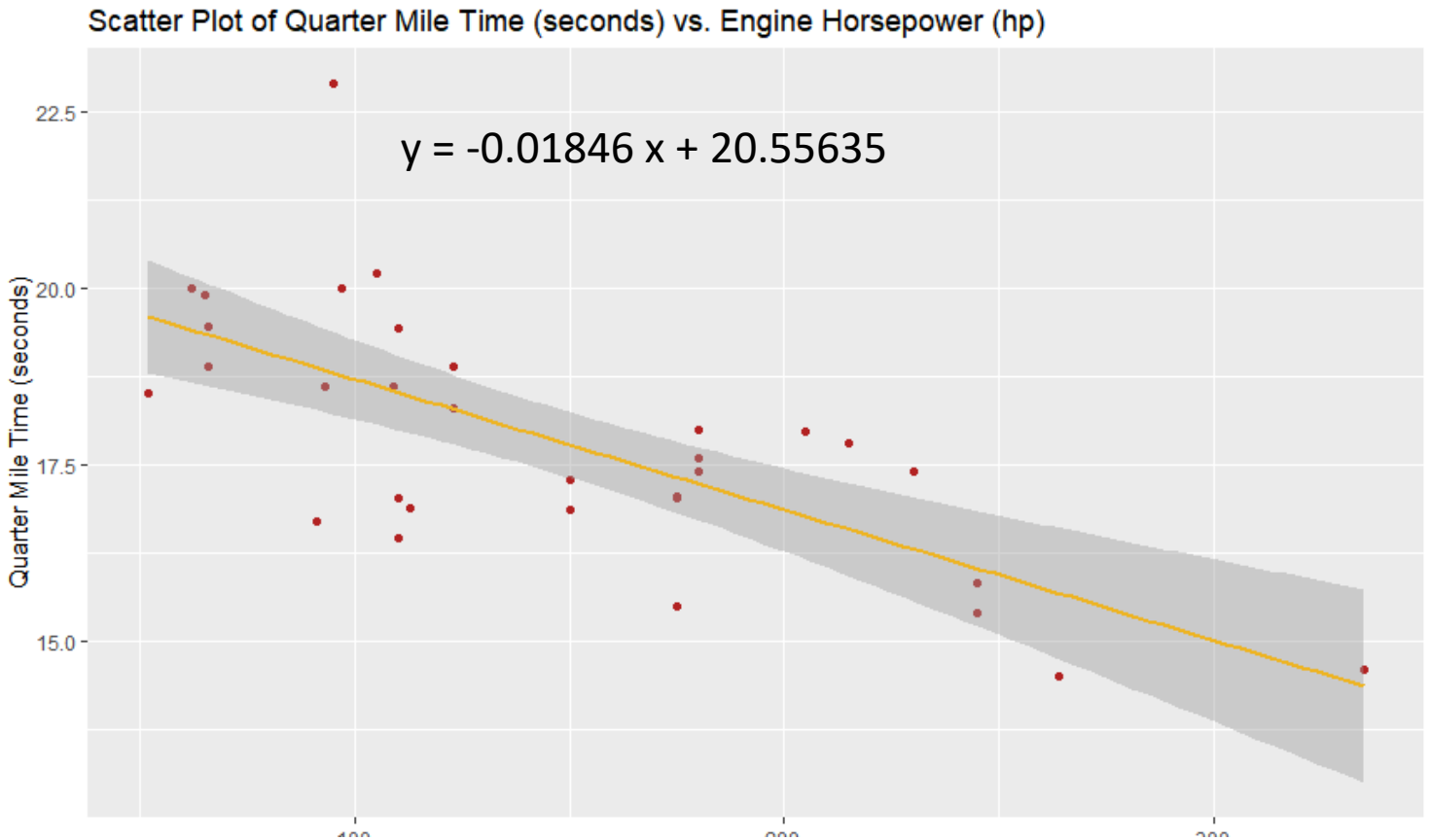


		Meaning
p-value	5.766e-06 < 0.5	Statistically Significant
R (Correlation Coefficient)	-0.7082234	Strong Negative Correlation

Analysis

- Because the absolute value of the correlation coefficient, r, exceeds 0.7, the negative correlation between quarter mile time and engine horsepower is strong.
- Because $p < 0.5$, the correlation is statistically significant.

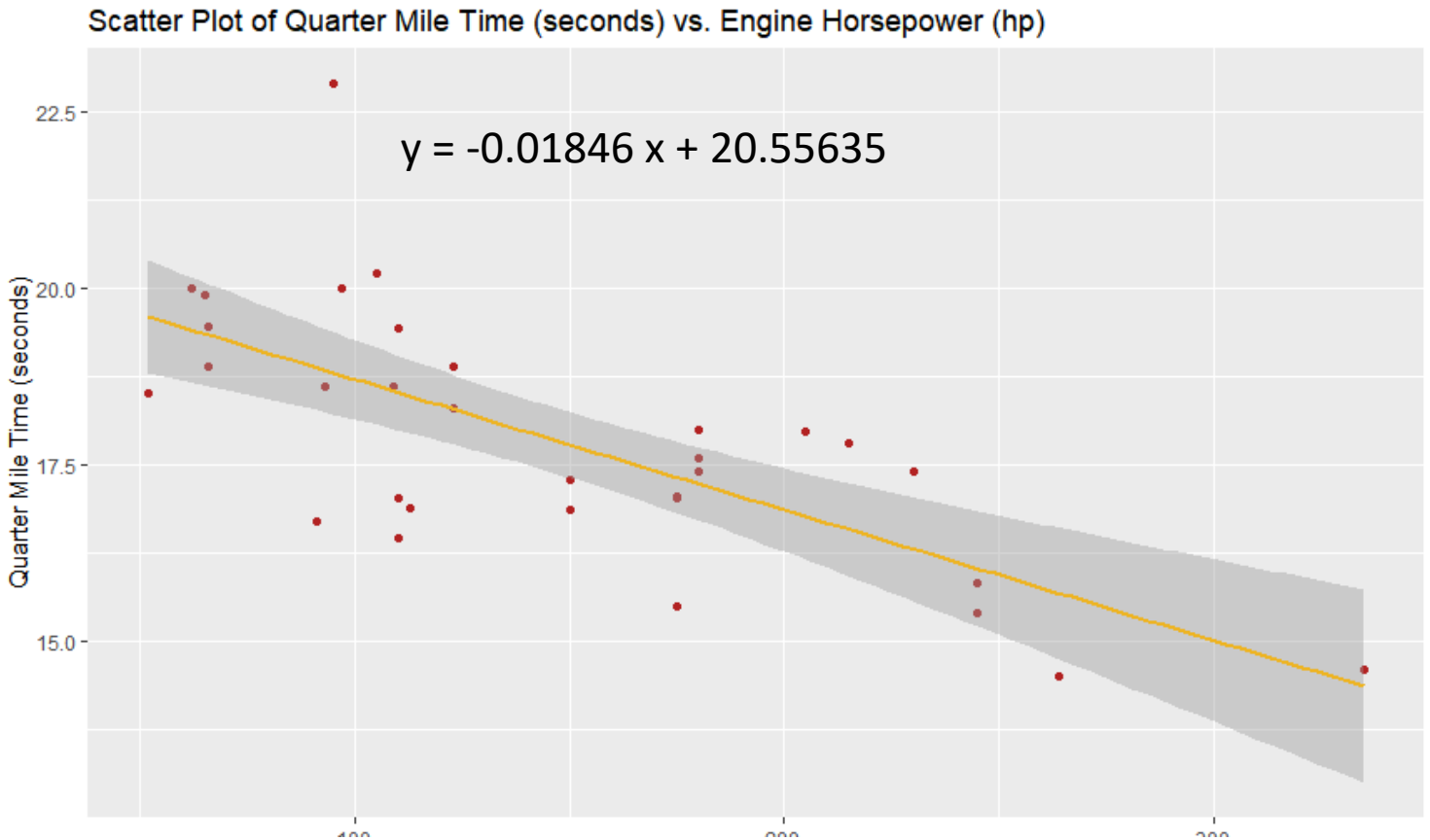
Linear Regression of Quarter Mile Time vs. Engine Performance



Code to compute this linear regression:

```
linreg1 <- lm(qsec ~ hp, mtcars)
print(linreg1)
```

Analysis



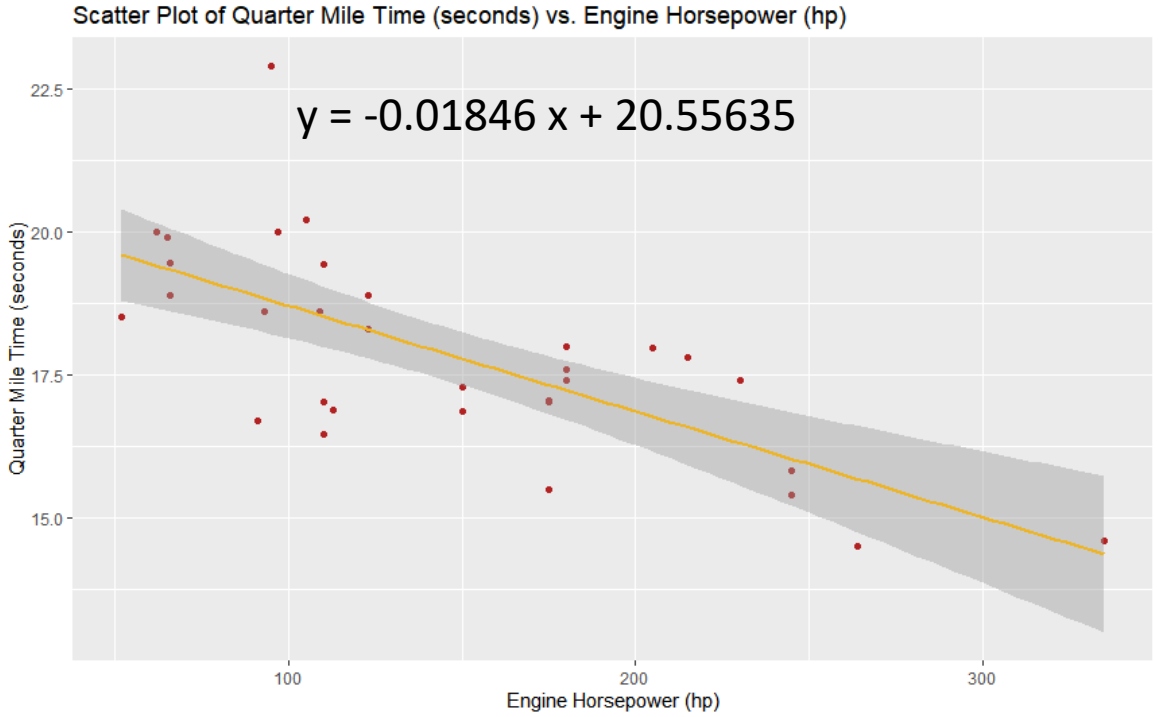
- From the mtcars data set, the greatest possible quarter mile time is approximately 20.6 seconds at about 50 hp.
- For every 100 hp, the quarter mile time decreases by approximately 1.8 seconds.

Summary of Linear Regression of Quarter Mile Time vs. Engine Performance

Code to summarize linear regression:

summary(linreg1)

```
> summary(linreg1)

Call:
lm(formula = qsec ~ hp, data = mtcars)

Residuals:
    Min      1Q  Median      3Q     Max
-2.1766 -0.6975  0.0348  0.6520  4.0972

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 20.556354   0.542424  37.897  < 2e-16 ***
hp          -0.018458   0.003359  -5.495 5.77e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.282 on 30 degrees of freedom
Multiple R-squared:  0.5016,    Adjusted R-squared:  0.485
F-statistic: 30.19 on 1 and 30 DF,  p-value: 5.766e-06
```



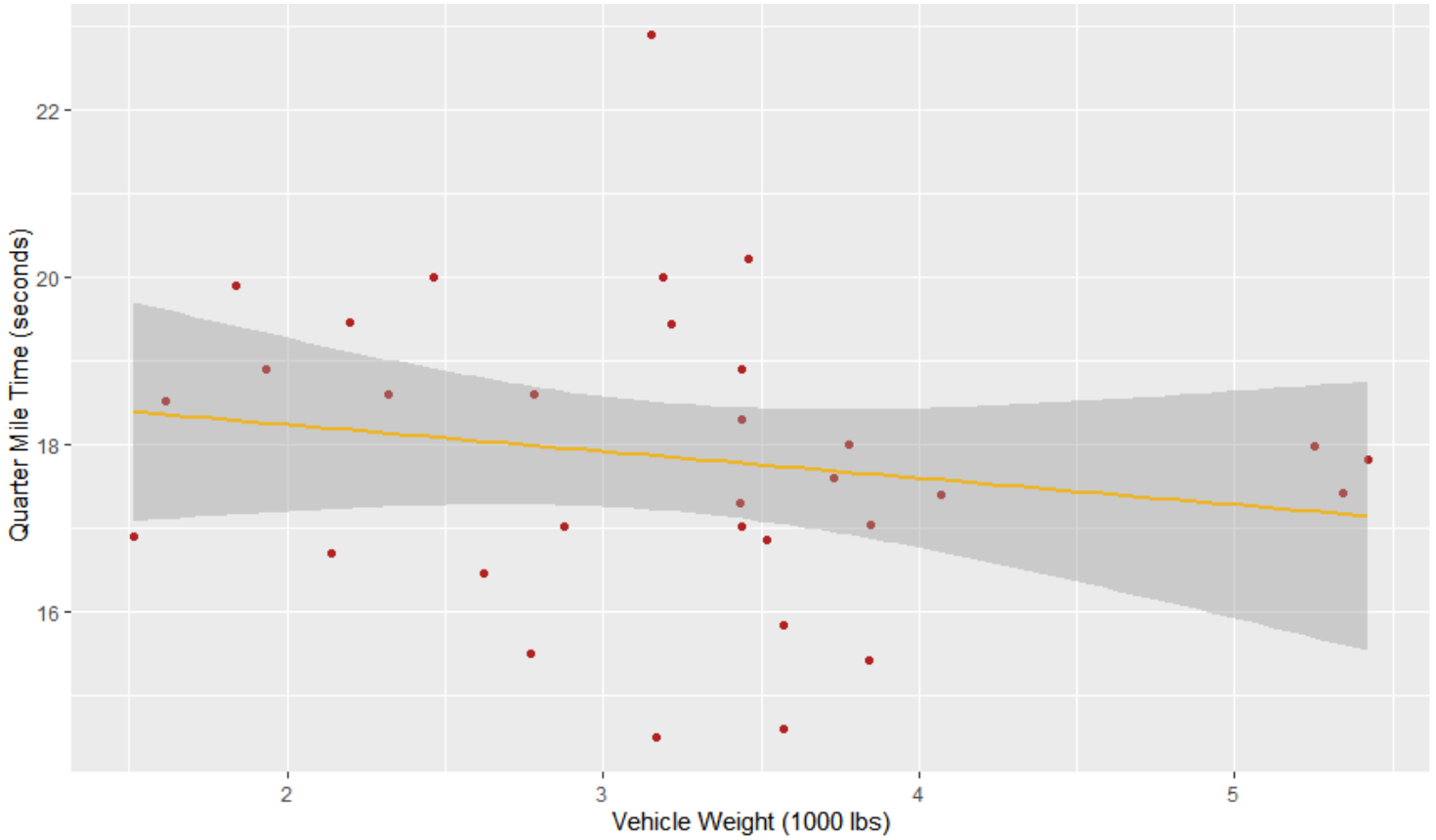
Analysis

```
> summary(linreg1)

Call:
lm(formula = qsec ~ hp, data = mtcars)

Residuals:
    Min      1Q  Median      3Q     Max
-2.1766 -0.6975  0.0348  0.6520  4.0972

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 20.556354   0.542424  37.897  < 2e-16 ***
hp          -0.018458   0.003359  -5.495 5.77e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’

Residual standard error: 1.282 on 30 degrees of freedom
Multiple R-squared:  0.5016,    Adjusted R-squared:  0.485
F-statistic: 30.19 on 1 and 30 DF,  p-value: 5.766e-06
```

- Pr(>|t|) is the p-value of the t-test conducted on horsepower as an individual predictor of quarter mile time.
- Pr(>|t|) = 5.77e-06 < 0.5. Because the results of the t-test are statistically significant, horsepower is a significant predictor of quarter mile time.
- The Adjusted R-squared value is 0.485, which means that horsepower 48.5% of the factors which affect quarter mile time.
- This outcome meets the researcher’s expectation and shows that horsepower is significantly related to quarter mile distance.

Part 2: Quarter Mile Time vs. Vehicle Weight

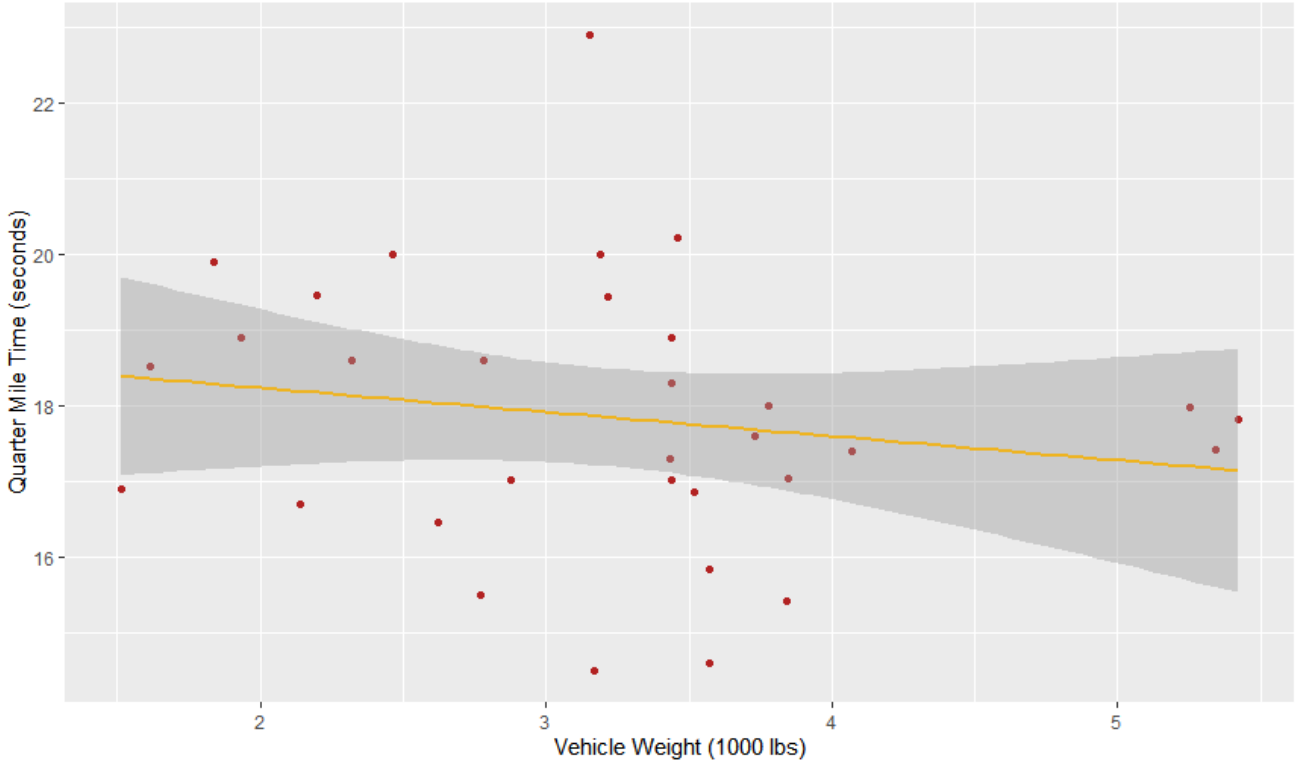
Scatter Plot of Quarter Mile Time vs. Vehicle Weight



Code to create this scatter plot:

```
ggplot(mtcars, aes(x = wt, y = qsec)) +
  geom_point(color = "firebrick") +
  ggtitle("Scatter Plot of Quarter Mile Time (seconds) vs. Vehicle Weight (1000 lbs)") +
  xlab("Vehicle Weight (1000 lbs)") +
  ylab("Quarter Mile Time (seconds)") +
  geom_smooth(method = lm, se = TRUE, color = "goldenrod2")
```

Analysis

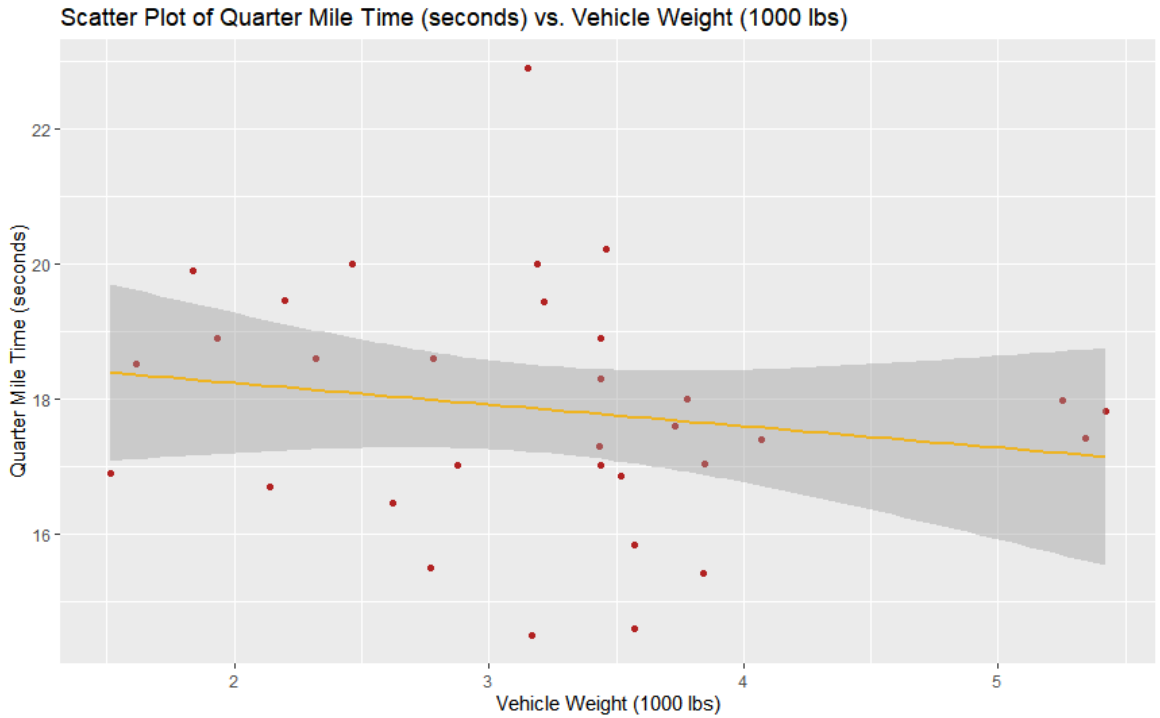


- There seems to be a negative correlation between quarter mile time and vehicle weight. As vehicle weight increases, quarter mile time decreases.

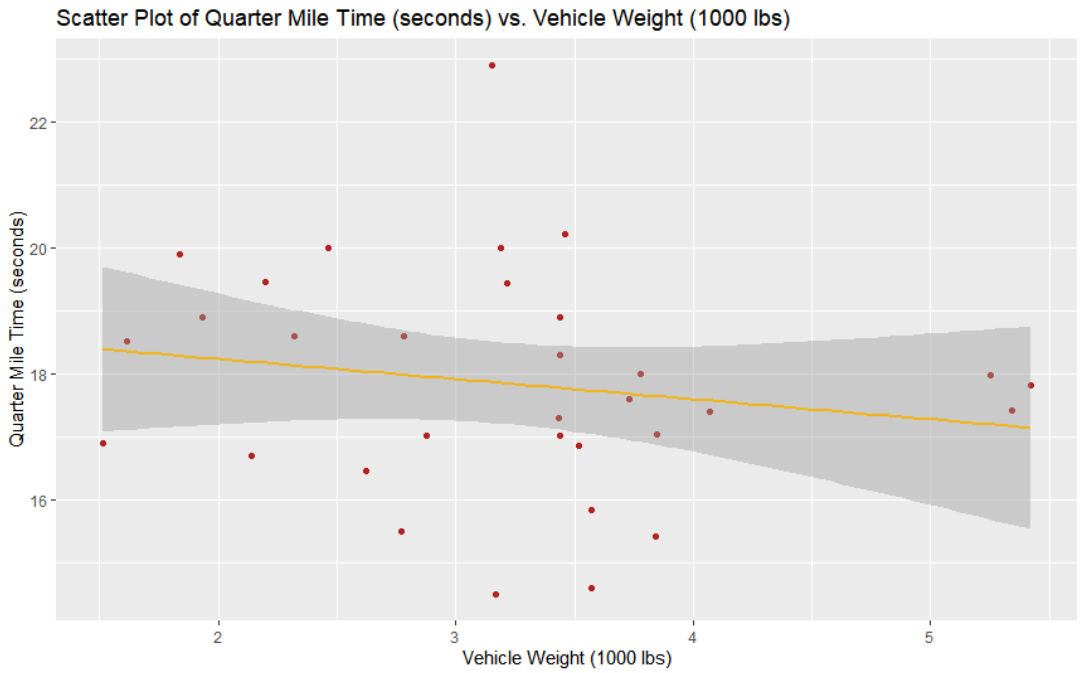
Correlation of Engine Performance with Quarter Mile Time

Code to compute correlation:

```
cor.test(mtcars$wt, mtcars$qsec, method = "pearson", use = "complete.obs")
```



		Meaning
p-value	0.3389	NOT Statistically Significant
R (Correlation Coefficient)	-0.1747159	Weak Negative Correlation

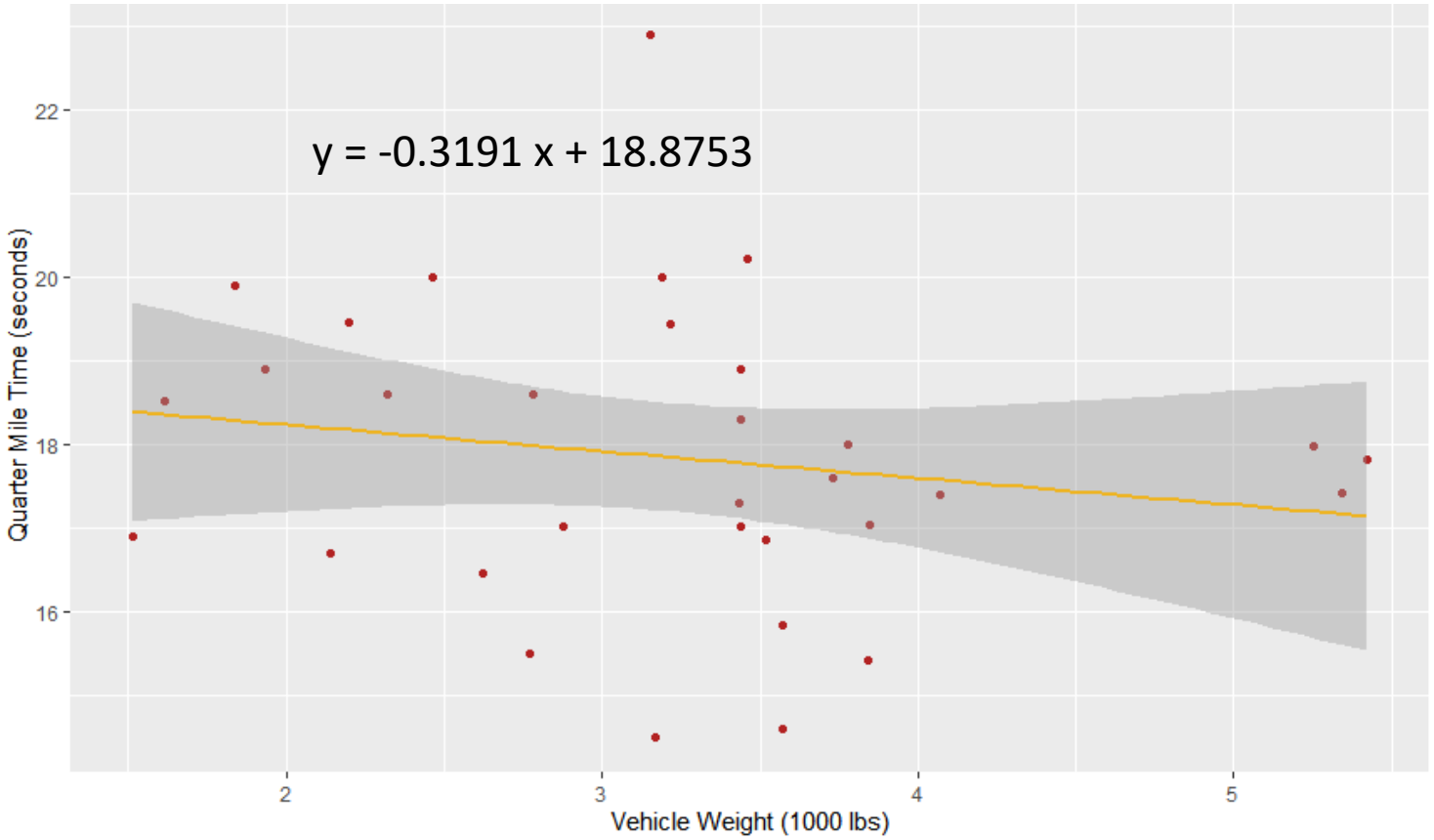


		Meaning
p-value	0.3389	NOT Statistically Significant
R (Correlation Coefficient)	-0.1747159	Weak Negative Correlation

Analysis

- Because the absolute value of the correlation coefficient, r, is less than 0.29, the negative correlation between quarter mile time and vehicle weight is weak.
- Because p = 0.3389 > 0.5, the correlation is NOT statistically significant.

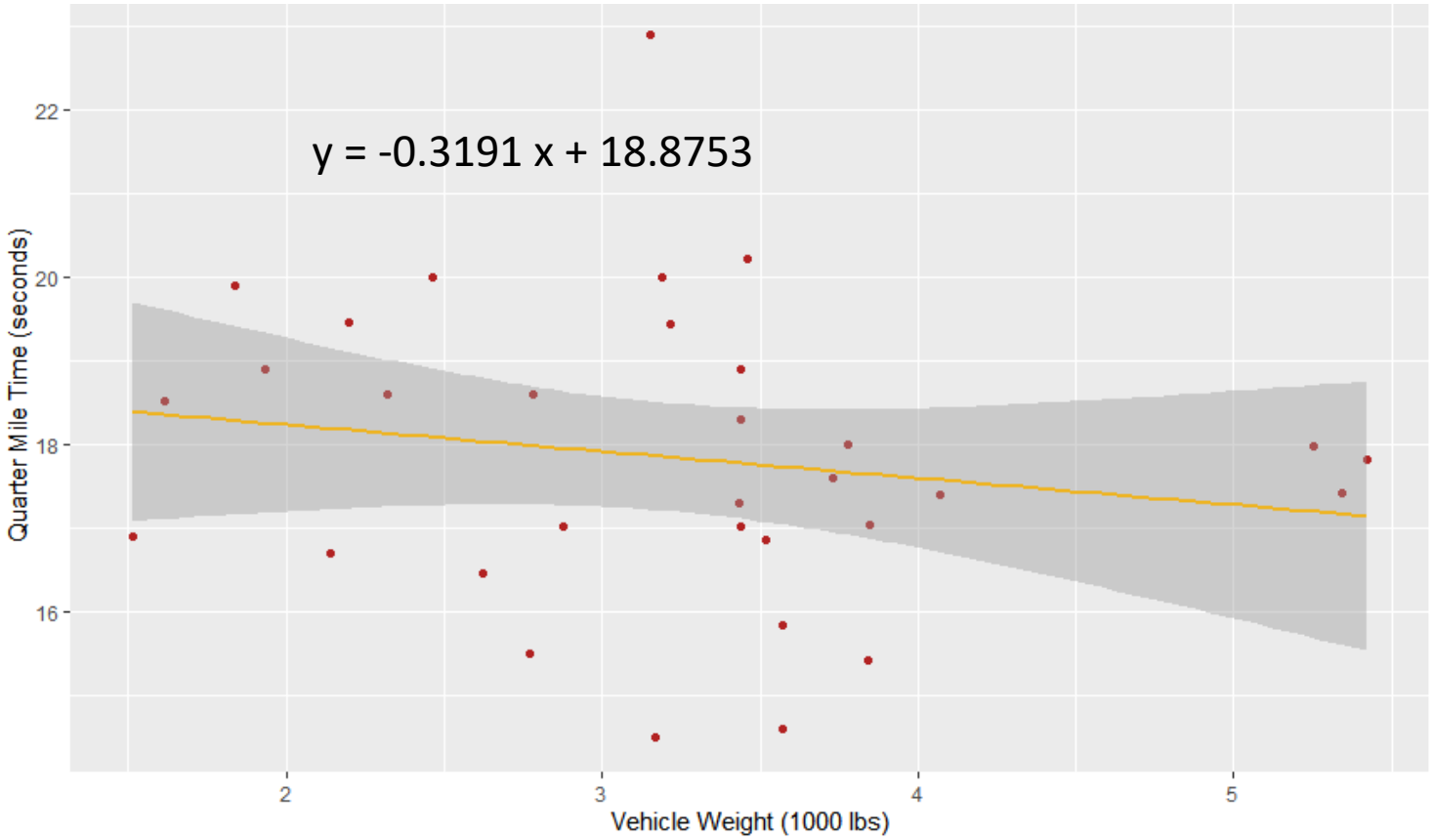
Linear Regression of Quarter Mile Time vs. Vehicle Weight



Code to compute this linear regression:

```
linreg2 <- lm(qsec ~ wt, mtcars)
print(linreg2)
```

Analysis

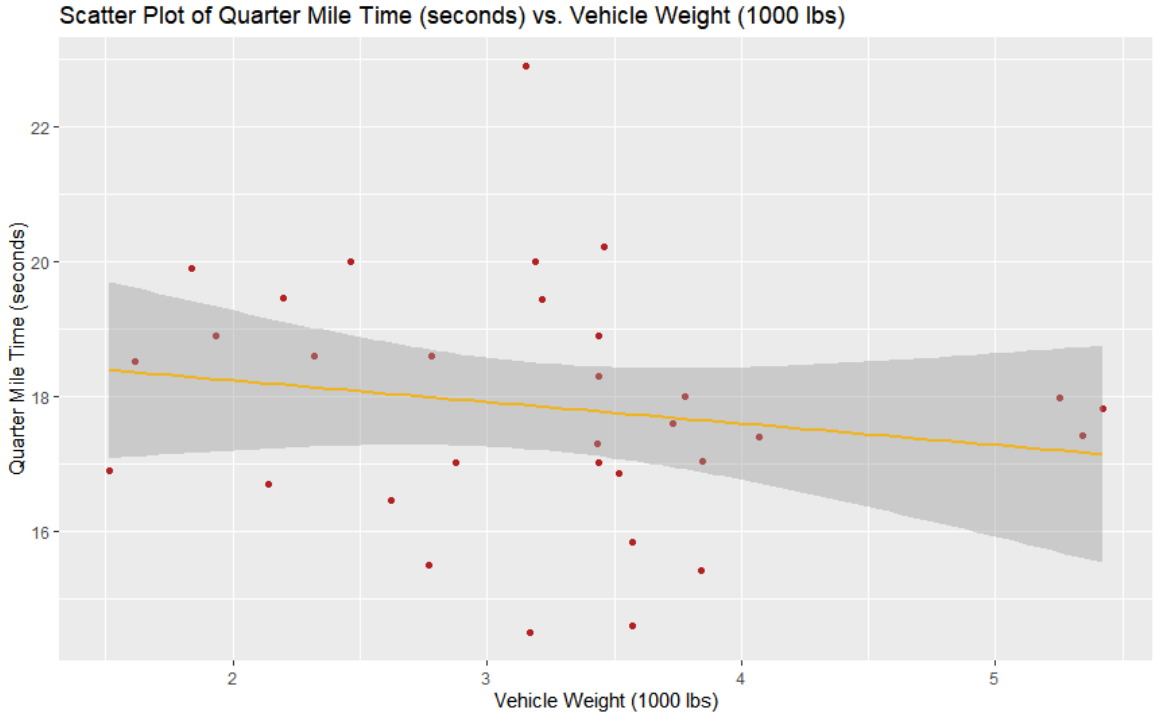


- From the mtcars data set, the greatest possible quarter mile time is approximately 19 seconds at about 1000 lbs.
- For every 1000 lbs in vehicle weight, the quarter mile time decreases by approximately 0.32 seconds.

Summary of Linear Regression of Quarter Mile Time vs. Vehicle Weight

- Code to summarize linear regression:
- summary(linreg2)

```
> summary(linreg2)

Call:
lm(formula = qsec ~ wt, data = mtcars)

Residuals:
    Min      1Q  Median      3Q     Max
-3.3638 -1.0766  0.2051  0.8655  5.0298

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  18.8753     1.1025  17.120   <2e-16 ***
wt           -0.3191     0.3283  -0.972    0.339
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.789 on 30 degrees of freedom
Multiple R-squared:  0.03053,   Adjusted R-squared:  -0.00179
F-statistic: 0.9446 on 1 and 30 DF,  p-value: 0.3389
```



Analysis

```
> summary(linreg2)

Call:
lm(formula = qsec ~ wt, data = mtcars)

Residuals:
    Min      1Q  Median      3Q     Max
-3.3638 -1.0766  0.2051  0.8655  5.0298

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  18.8753     1.1025  17.120   <2e-16 ***
wt           -0.3191     0.3283  -0.972    0.339
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.789 on 30 degrees of freedom
Multiple R-squared:  0.03053,   Adjusted R-squared:  -0.00179
F-statistic: 0.9446 on 1 and 30 DF,  p-value: 0.3389
```

- Pr(>|t|) is the p-value of the t-test conducted on horsepower as an individual predictor of quarter mile time.
- Pr(>|t|) = 0.339 > 0.5. Because the results of the t-test are NOT statistically significant, vehicle weight is NOT a significant predictor of quarter mile time.
- The Adjusted R-squared value is -0.00179, which means that vehicle weight does not adequately explain which quarter mile time.
- This result meet’s the researcher’s expectation of a weak relationship between quarter mile distance and vehicle weight.

Conclusion

- There is a strong, statistically significant, negative correlation between quarter mile time and engine horsepower: A 100 hp increase in horsepower would decrease quarter mile time by approximately 1.8 seconds. Horsepower is a significant predictor of quarter mile time.
- In contrast, there is a weak negative correlation between quarter mile time and vehicle weight. However, this correlation is NOT statistically significant. Unlike horsepower, vehicle weight does not adequately explain quarter mile time.
- These results meet the researcher's expectations and show that cars with more powerful engines can travel at higher speeds and therefor show a lower quarter mile time. The results also show that vehicle weight does not have a strong relationship with quarter mile time.