
UTM

UNIVERSITI TEKNOLOGI MALAYSIA

**SCHOOL OF COMPUTING, FACULTY OF ENGINEERING**

**UNIVERSITY TECHNOLOGY MALAYSIA**

# SECI2143-09 PROBABILITY AND STATISTICAL DATA ANALYSIS

**SEMESTER 2 2020/2021**

# PROJECT 2: CAR WEIGHT AND ACCELERATION ANALYSIS REPORT



| | |
|---|---|
| **GROUP MEMBERS** | 1. NUR AISYAH BT AZMI (A20EC0111)<br>2. AIDATUL KHAIRINA BT AZMI (A20EC0007)<br>3. HASIN ARAF (A20EC4024) |
| **SECTION** | **09** |
| **LECTURER** | **DR. AZURAH ABU SAMAH** |

# 1.0 INTRODUCTION

Understanding vehicle acceleration performance is important to transportation analysts and researchers for several reasons. First, acceleration can be traded off against other vehicle attributes, including fuel economy. All else being equal, larger improvements in acceleration performance over time mean smaller improvements in fuel economy, leading to higher energy consumption. Second, the acceleration performance of a vehicle can affect its utility to consumers, influencing purchase decisions. From the data chosen, we expected to see the acceleration capabilities of vehicles to influence how aggressively they end up being driven and lastly affecting in-use fuel consumption.

# 2.0 DATASET

We conducted data analysis using a dataset obtained from the internet (Project datasets (telecom-paristech.fr)). 70 out of 408 data are chosen randomly from the available dataset. Variables used in this project test are Car Weight and Car Acceleration. These two variables are chosen as we want to know if one of the variable values actually can affect the other variable value. At first, we assumed that there is a negative relationship between Car Weight and Car Acceleration. We also wanted to see whether in average, does a car weigh more than 3000 pounds and to know whether the dataset chosen has a strong relationship between two variables chosen.

# 3.0 DATA ANALYSIS

## 3.1 HYPOTHESIS TESTING OF 1 SAMPLE TEST

One sample test is used to support the claim "mean of car weight is more than 3000 pounds".

The null hypothesis, $H_0$ and alternative hypothesis, $H_1$ are denoted as:

$$H_0: \mu = 3000 \text{ pounds}$$

$$H_1: \mu > 3000 \text{ pounds}$$

The significance level, $a = 0.05$ is used to test the claim.

```
> t.test(data$`CAR WEIGHT`, mu=3000, alternative = "greater")

        One Sample t-test

data:  data$`CAR WEIGHT`
t = 0.72507, df = 69, p-value =
0.2354
alternative hypothesis: true mean is greater than 3000
95 percent confidence interval:
 2896.492      Inf
sample estimates:
mean of x
 3079.657
```

*Figure 1: One sample t-test result using R*

Degree of freedom, df = 69

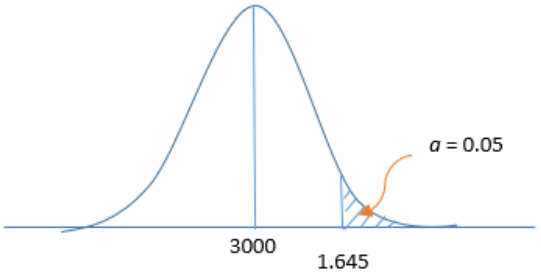
Critical value, $Z_{0.05}$ = 1.645

P-value = 0.2354

Variance, σ = 89.3912

Mean, $\bar{x}$ = 3079.657

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = 7.456$$




Since $Z > Z_{0.05}$, we reject the null hypothesis, H0 at $a$ = 0.05. We have sufficient evidence that the mean of car weight is greater than 3000 pounds.

### 3.2 CORRELATION TEST

In this test, a random 70 sample dataset from a total of 408 are selected. The measurement level of this dataset is ratio, hence we used Pearson's Product-Moment Correlation Coefficient to conduct the test. The purpose of this test is to find the relationship between Car Weight (in pounds) and Car Acceleration.

The null hypothesis, $H_0$ and alternative hypothesis, $H_1$ are denoted as:

$H_0$: ρ = 0 (no linear correlation)

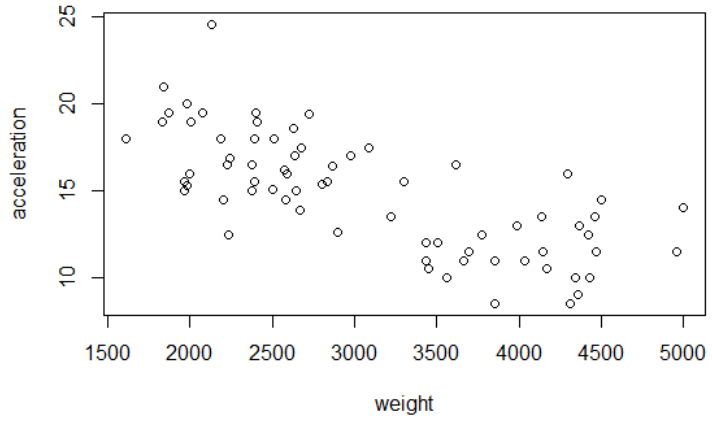
$H_1$: ρ ≠ 0 (linear correlation exists)

The significance level, $a$ = 0.05 is used to test the claim.

```
        Pearson's product-moment correlation

data:  weight and acceleration
t = -8.684, df = 68, p-value = 1.258e-12
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.8203430 -0.5908494
sample estimates:
       cor
-0.7251481
```

*Figure 2: Pearson's Product Moment Correlation Coefficient test result using R*

In Figure 2, $t$ = -8.684 is the value for the test statistic while the p-value obtained is 1.258e-12. Since p-value is less than significance level, $a$ = 0.05, $H_0$ is rejected. There is sufficient evidence of a linear relationship between the Car Weight and Car Acceleration at the 95% level of significance.




*Graph 1: Scatter plot graph for this dataset*

Graph 1 illustrates the relationship between Car Weight and Car Acceleration with a negative linear relationship. From the test conducted in Figure 2, the correlation coefficient value, *r* calculated is -0.7251481. Both scatter plot and *r* value indicate that there is a moderate negative linear relationship between Car Weight and Car Acceleration.

### 3.3 REGRESSION TEST

In this test, a random 70 sample dataset from a total of 408 are selected. The chosen independent variable, *x* is Car Weight while dependent variable, *y* is Car Acceleration.

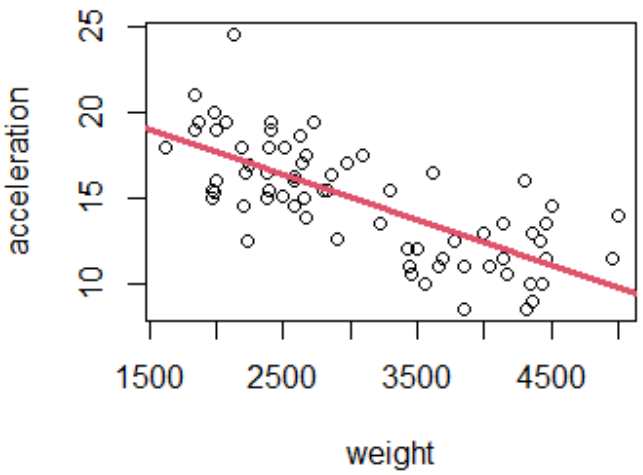
The null hypothesis, $H_0$ and alternative hypothesis, $H_1$ with $a = 0.05$, are denoted as:

$$H_0: \boldsymbol{\beta} = 0 \text{ (no relationship)}$$

$$H_1: \boldsymbol{\beta} \neq 0 \text{ (relationship exists)}$$

```
Call:
lm(formula = car$acceleration ~ car$weight)

Residuals:
    Min      1Q  Median      3Q     Max
-4.5791 -1.7174 -0.1657  1.5735  7.2457

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 22.9903554  0.9787189  23.490  < 2e-16 ***
car$weight  -0.0026461  0.0003047  -8.684 1.26e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.326 on 68 degrees of freedom
Multiple R-squared:  0.5258,    Adjusted R-squared:  0.5189
F-statistic: 75.41 on 1 and 68 DF,  p-value: 1.258e-12

> #y-intercept and slope coefficient
> mod$coef
(Intercept)  car$weight
22.99035541 -0.00264605
```

*Figure 3: Linear Regression model test result using R*




*Graph 2: Scatter plot graph for regression*

From the test, $\beta_0$ obtained is 22.99035541 while $\beta_1$ is -0.00264605. The graph illustrates a negative linear relationship with $y$ = 22.99035541 - 0.00264605$x$. The p-value with significance level $a$ = 0.05 is 1.256e-12. Since p-value calculated is less than $a$, null hypothesis is rejected. There is enough evidence of a relationship between the Car Weight and Car Acceleration with 0.05 as confidence interval.

### 3.4 CHI-SQUARE TEST OF INDEPENDENCE

This analysis is used to know whether there is a relationship between variable Car Weight and Car Acceleration.

Test Hypothesis:

$H_0$ = Car Weight and Car Acceleration are independent

$H_1$ = Car Weight and Car Acceleration has a relationship

```
> library("MASS")
> car.data <- data.frame(data$`CAR WEIGHT`,data$`CAR ACCELERATION`)
> print(car.data)
> print(chisq.test(car.data))

        Pearson's Chi-squared test

data:  car.data
X-squared = 261.5, df = 69, p-value < 2.2e-16

> alpha <- 0.05
> cv.alpha <- qchisq(alpha,df=69,lower.tail = FALSE)
> cv.alpha
[1] 89.39121
```

*Figure 4: Chi-Square Test of Independence*

Significance level, $a = 0.05$

Test statistic, $x^2 = 261.5$

Critical value, $x^2_{69,0.05} = 89.39121$

Since the test statistic value is greater than critical value, then we reject the null hypothesis at $a = 0.05$. There is significant evidence that car weight and car acceleration have a relationship.

## 4.0 CONCLUSION

We have successfully managed to complete this project about the relationship between car weight and its acceleration. We have learned to choose the dataset as we need to consider some factors such as source and complete dataset. After that, we make sure that the dataset is complete for every sample. During the analysis process, we mostly used R language to get accurate results and make our work easier. As we never learned how to use the R language for statistics, this project has given us new experience and opportunity especially by learning the R language,

From the several tests conducted, there is a relationship between car weight and the acceleration. We have conducted tests on the car weight and acceleration. By doing correlation and regression tests, we found that there is a relationship between car weight and acceleration. The negative relationship indicates that car acceleration is dependent on car weight where the car acceleration can decrease as the car weight increases. However, we discovered that a car's power also can influence how fast a car can move. In this case, the car will move faster with higher power. This explains the reason why some dataset readings have a huge difference with other dataset. At the end of this project, we think that we gained many good experiences and added new skills that we can use in future works.

## 5.0 APPENDIX

1. Link for Project Video Presentation (Youtube): https://youtu.be/0Pn09EUTx_Y
2. Some of the sample of dataset:
   (source:https://perso.telecom-paristech.fr/eagan/class/igr204/datasets)

| Weight | Acceleration |
|---|---|
| 3504 | 12 |
| 3693 | 11.5 |
| 3436 | 11 |
| 3433 | 12 |

| | |
|---|---|
| 3449 | 10.5 |
| 4341 | 10 |
| 4354 | 9 |
| 4312 | 8.5 |
| 4425 | 10 |
| 3850 | 8.5 |