# Gym Data

By Benjamin Taraba, Maureen Carpenter, and Sammy Motlak

# Questions

How does session duration affect calories burned?

How does BMI/Weight affect calories burned?

What affects someone's BMI?

# How does session duration affect calories burned?

Regardless of experience, type of workout, frequency of workout, height, weight, and gender.

# Proving a Relationship with Chi-Squared

Breaking down data:

- Session Duration: Short or Long
- Calories Burned: Low (303–800), Medium (801–1300), or High (1301–1783)

# Terms

- Short sessions — sessions shorter than 1 hour
- Long sessions — sessions longer than 1 hour

# Proving a Relationship with Chi-Squared

Then create a table:

| | Low | Medium | High |
|---|---|---|---|
| Long | 159 | 543 | 85 |
| Short | 185 | 1 | 0 |

# Proving a Relationship with Chi-Squared

Using a Pearson’s Chi-square test, P value is found to be less than 2.2e-16.

Reject the Ho (null hypothesis)!!

# A Hypothesis Test

Null Hypothesis ($H_0$): The mean calories burned for short sessions is equal to the mean calories burned for long sessions.

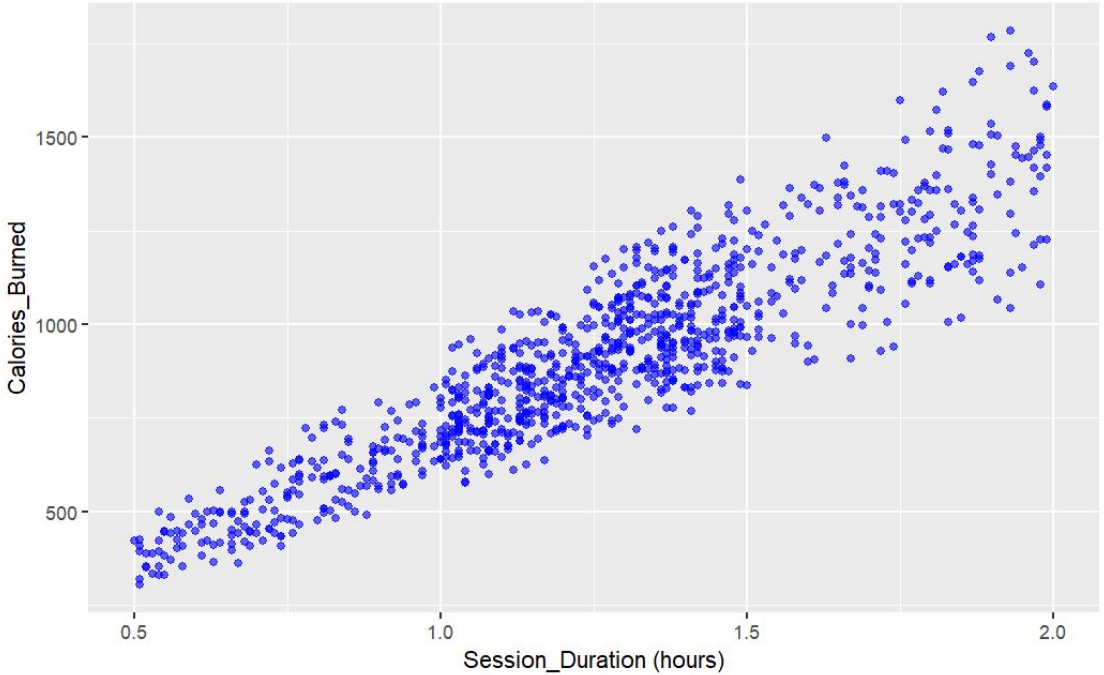
Alternative Hypothesis ($H_1$): The mean calories burned for short sessions is different from that for long sessions.

# Results

Using an Alpha of 0.05, we find the P-value returned from a Welch Two Sample t-test.

P-value < 2.2e-16. Both confidence intervals are negative, meaning the mean calories burned from long sessions is larger than the mean for smaller sessions

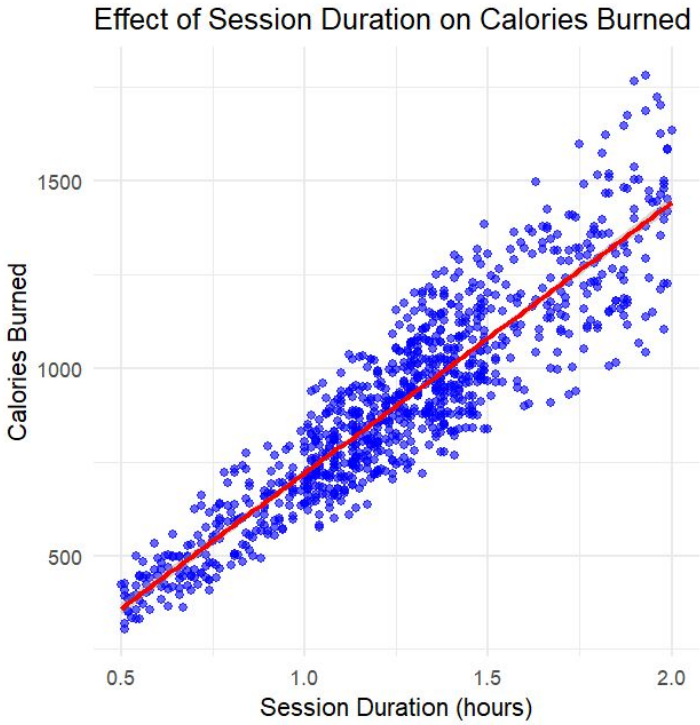
# Graph of Calories Burned and Session Duration



# Linear Regression Model

When creating a linear regression model on the effect session duration has on calories burned, we find a p-value less than 2e-16 and the R squared of the model is 0.8245.

For every additional hour in the Session Duration, Calories Burned increases by 721.79.

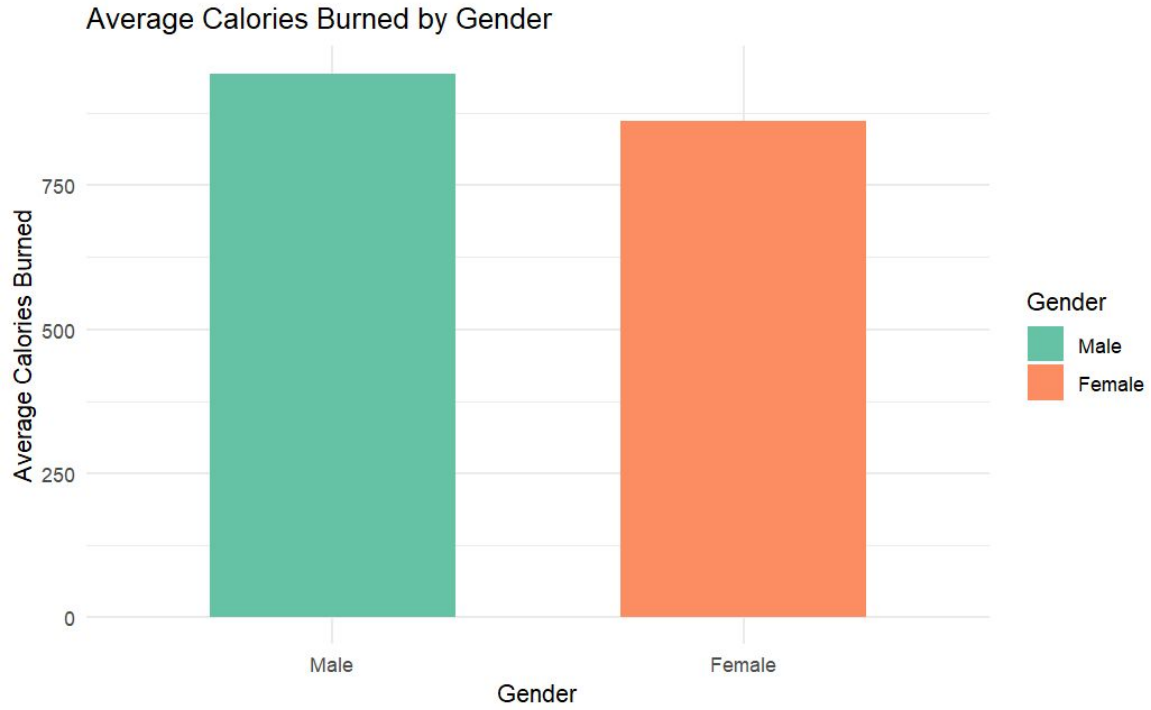


# Multiple Regression Model — Adding Gender

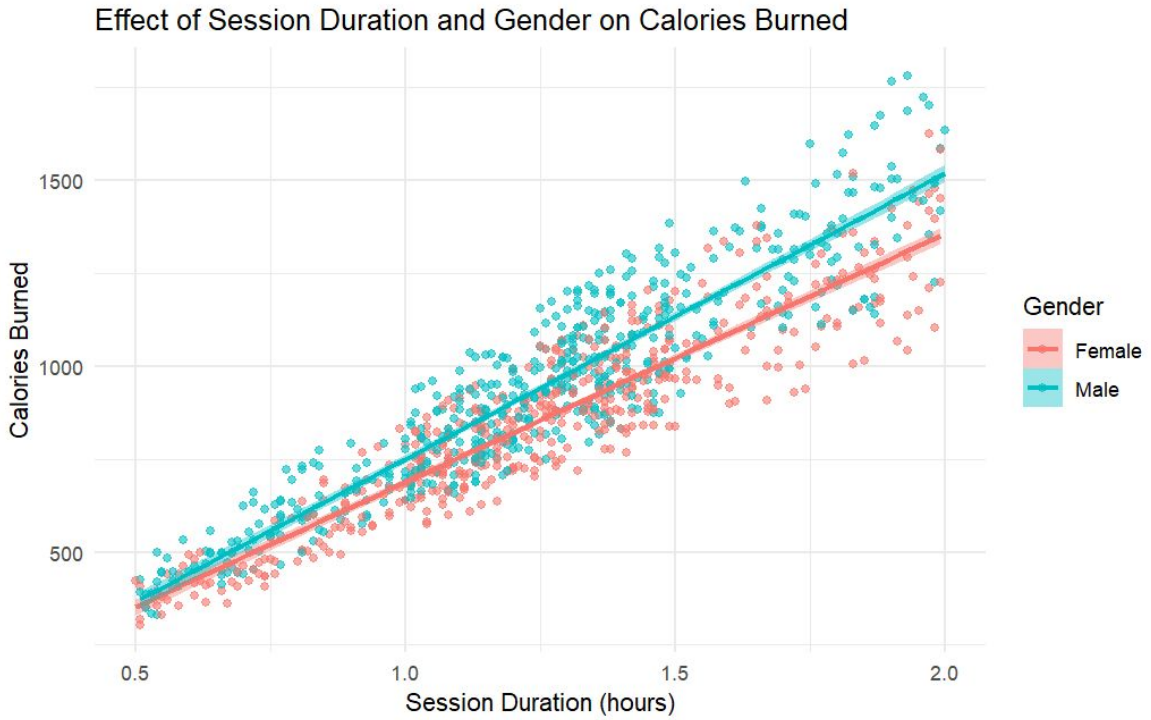
When creating a multiple regression model on the effect session duration and gender have on calories burned, we find the p-value for both gender and session duration is less than 2e-16 and the R squared of the model is 0.8506.

For every additional hour in the Session Duration, Calories Burned increases by 723.36, and if the gender is male the calories burned go up by an additional 88.26.

# Average Calories Burned by Gender



# Multiple Regression Model — Adding Gender



# Multiple Regression Model — All Variables Included

Includes: Age, Gender, Weight (kg), Height (m), Max_BPM, Avg_BPM, Resting_BPM, Session_Duration (hours), Calories_Burned, Workout_Type, Fat_Percentage, Water_Intake (liters), Workout_Frequency (days/week), Experience_Level, and BMI.

Not all of these ended up being significant.

## Multiple Regression Model — Significant Variables

Using an alpha of 0.05, only age, gender, weight, height, Average BPM, Resting BPM, Session Duration, and BMI were significant.

Rerunning the model with only significant variables gives us…

# Multiple Regression Model — Significant Variables

If age goes up by 1 year, Calories burned goes down by 3.421.

If the gender is male, Calories burned goes up by 82.81.

If weight goes up by 1 kg, Calories burned goes down by 1.145.

If height goes up by 1 meter, Calories burned goes up by 116.1. (useless)

If Avg BPM goes up by 1, Calories burned goes up by 6.238.

If Resting BPM goes up by 1, Calories burned goes down by 0.3894.

If Session Duration goes up by 1 hour, Calories burned goes up by 714.

If BMI goes up by 1, Calories burned goes up by 3.811.

# How does session duration affect calories burned?

Ignoring all other factors:

- Proven a relationship between session duration and calories burned.
- Created a linear model for session duration and its effect on calories burned and when session duration increases by 1 hour, calories burned increases by 721.79.

Including Gender:

- Males in the sample burn more calories.

Other factors that also affect calories burned:

- age, weight, height, avg BPM, resting BPM, and BMI.

# How does BMI/Weight affect calories burned?

# BMI and Calories Burned with Single Regression

- P-value is at 0.06
  - indicates BMI is not a good predictor of Calories Burned
- Low R-squared value at 0.003
  - a low variation in Calories Burned is being explained.

# Relationship between BMI and Calories Burned with Chi-Squared

Breaking down data:

- Calories Burned: Low (303–800), Medium/High (801–1783)

BMI:

- Underweight<18.5
- Normal<24.9
- Overweight<29.9
- Obese

# Table of BMI Vs. Calories Burned


Then create a table:

| | Underweight | Normal | Overweight | Obese |
|---|---|---|---|---|
| Low | 75 | 117 | 84 | 67 |
| Med/High | 93 | 249 | 158 | 130 |

# Proving a Relationship with Chi-Squared

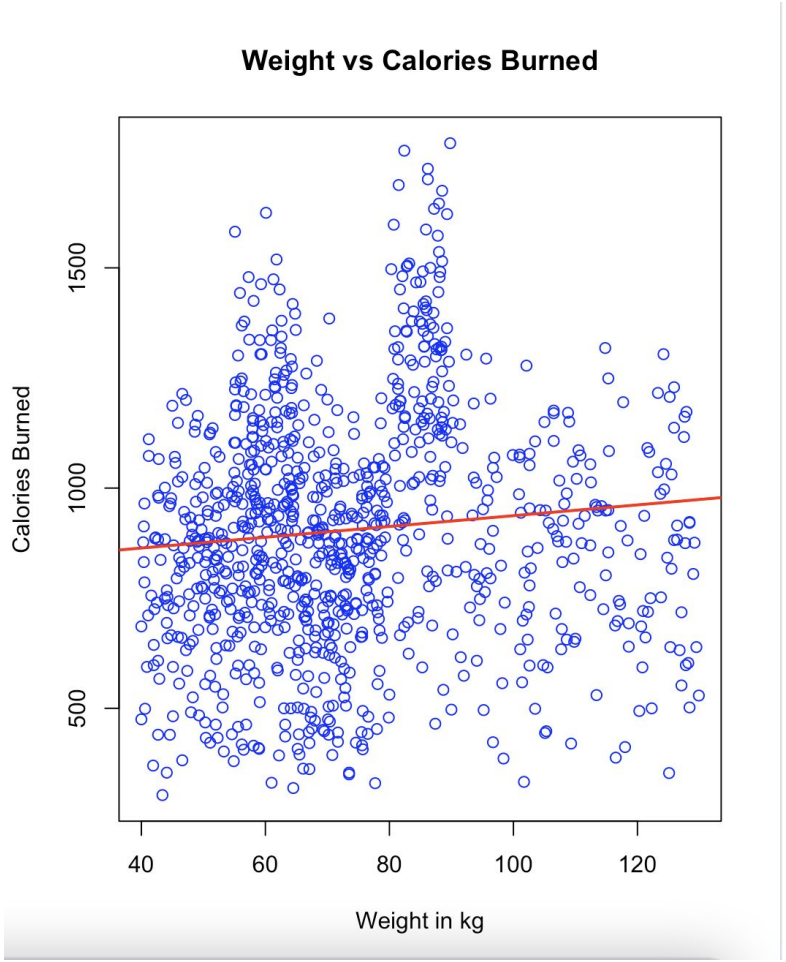
The p-value is at 0.039, which is less than the threshold of 0.05. Therefore, there is a relationship between calories burned and BMI, when the categories are broken down.

# Single Regression

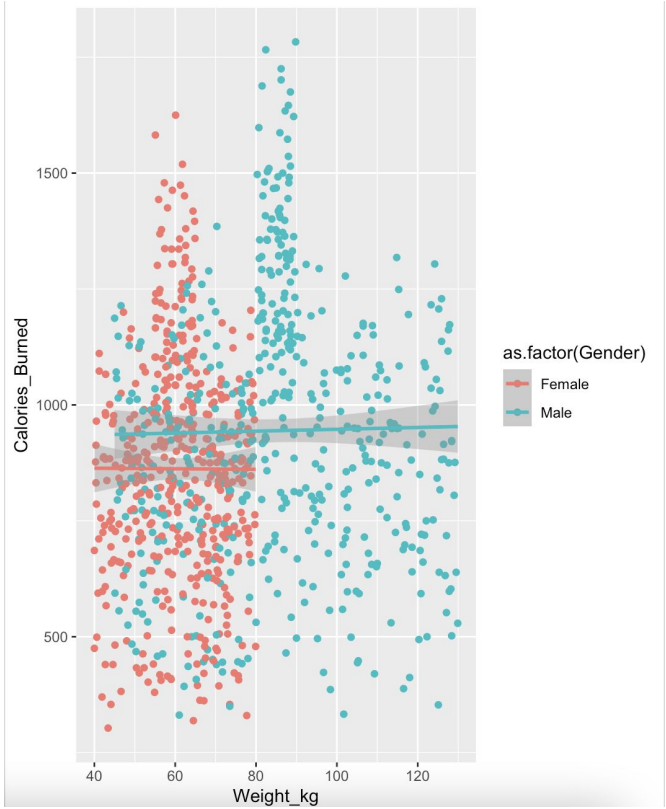
## Weight vs. Calories Burned

**Finding:** There is a relationship between weight and calories burned, since the p value is low 0.003.

**Interpretation:** As weight increases by 1kg, calories burned increases by 1.23. The heavier someone weighs, the more calories they burn.



# Does the effect of Weight on Calories Burned depend on Gender?




No, the p value for the interaction term is large, at 0.7. Therefore, the effect of weight on calories burned does not depend on gender.

# What affects someone's BMI?

What explains someone's BMI? Does our workout data help show a connection between working out and BMI?

# Hypothesis Test Male

Ho= The average male BMI in America is 29.1

Ha=  male BMI is not equal to 29.1

# Hypothesis Test Male

P-value = 1.356e-10

With a 95% CI from 26.22 to 27.55

Mean of 26.89

You would Reject the Ho

```
data:  MBMI
t = -6.5561, df = 510, p-value = 1.356e-10
alternative hypothesis: true mean is not equal to 29.1
95 percent confidence interval:
 26.22380 27.55013
sample estimates:
mean of x
 26.88697
```

# Hypothesis Test Female

Ho= The average female BMI in America is 29.8

Ha=  female BMI is not equal to 29.8

# Hypothesis Test Female

p-value < 2.2e-16

With a 95% CI from 22.32 to 23.14

Mean of 22.727

Reject the Ho

```
data:  FBMI
t = -33.966, df = 461, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 29.8
95 percent confidence interval:
 22.31867 23.13700
sample estimates:
mean of x
 22.72784
```

# Multi Regression

- Calories Burned
- Water Intake
- Session Duration
- Max BPM
- Avg BPM
- Age
- Resting BPM

```
Residuals:
     Min       1Q   Median       3Q      Max
-15.2199  -4.6420  -0.1858   3.7995  23.7601

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      34.018183   5.635488   6.036 2.24e-09 ***
CaloriesBurned    0.022096   0.004055   5.450 6.41e-08 ***
WaterIntake       1.410398   0.418598   3.369 0.000783 ***
SessionDuration -16.673835   2.860664  -5.829 7.61e-09 ***
MaxBPM            0.036329   0.017833   2.037 0.041904 *
AvgBPM           -0.125440   0.029266  -4.286 2.00e-05 ***
RestingBPM       -0.044957   0.028117  -1.599 0.110163
Age               0.062795   0.021929   2.864 0.004280 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.394 on 965 degrees of freedom
Multiple R-squared:  0.0852,    Adjusted R-squared:  0.07856
F-statistic: 12.84 on 7 and 965 DF,  p-value: 7.608e-16
```

# Multi Regression

- All variables besides Resting BPM are significant at alpha .05
- R^2 is 0.07856 -> .78.56% of BMI is explained by the variables here
- When Session Duration(hr) goes up by 1, BMI goes down by 16.67

```
Coefficients:
                    Estimate S
(Intercept)       34.018183
CaloriesBurned     0.022096
WaterIntake        1.410398
SessionDuration  -16.673835
MaxBPM             0.036329
AvgBPM            -0.125440
RestingBPM        -0.044957
Age                0.062795
```

# What did we find?