Project Part 1

Christian Min, Mackenzie Fox, Shiloh Gelnett

2025-11-04

Introduction

The number of calories burned during a workout session is an important factor to measure to determine the intensity of a workout and to help achieve weight loss goals. There are various factors that can impact the number of calories burned in a workout session, such as the duration of the session, body composition, and genetics. In this research project we are looking to determine if different types of workouts impact the average number of calories burned in an exercise session. If we are able to determine if various workout types impact the average number of calories burned, this can assist in finding an ideal workout session/type based on the goal number of calories burned to help gym goers achieve their fitness goals.

Our research question is: Does workout type lead to differences in the average number of calories burned per session?

Data Summary

The data represents a sample of 20,000 people. We are unsure about how the data was selected and collected, along with any data modifications that occurred after collection as this data set comes from Kaggle.

Some potential issues with this data include that there may be variation in workout type and intensity between males and females. Males may prefer more intensive workout types which can increase the average number of calories burned or vice versa. This could lead us to believe that there are gender differences in calories burned, when in reality workout preferences differ by gender. Another issue is that the number of calories burned data is right skewed which would make the test unreliable. We would potentially need to log transform the data to make it more reliable.

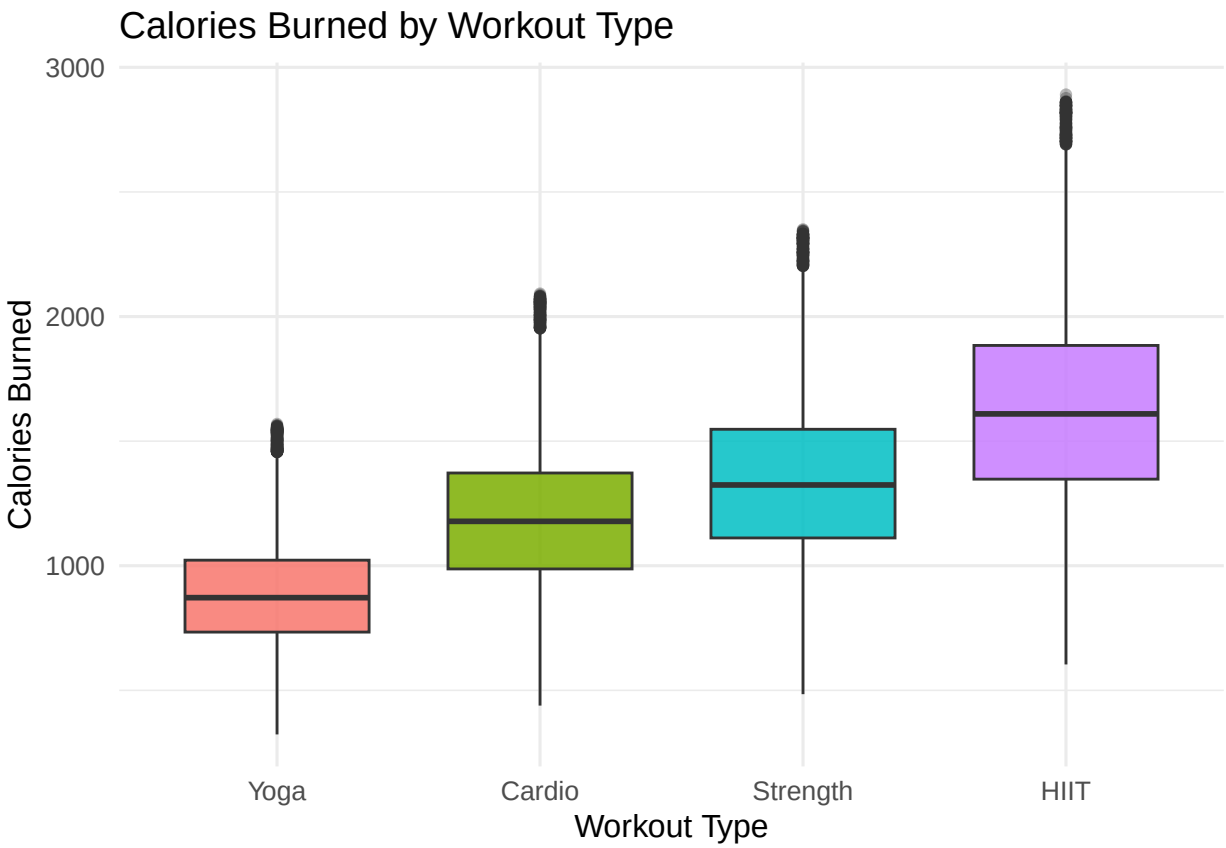
This data is appropriate to answer this research question because it contains a large sample

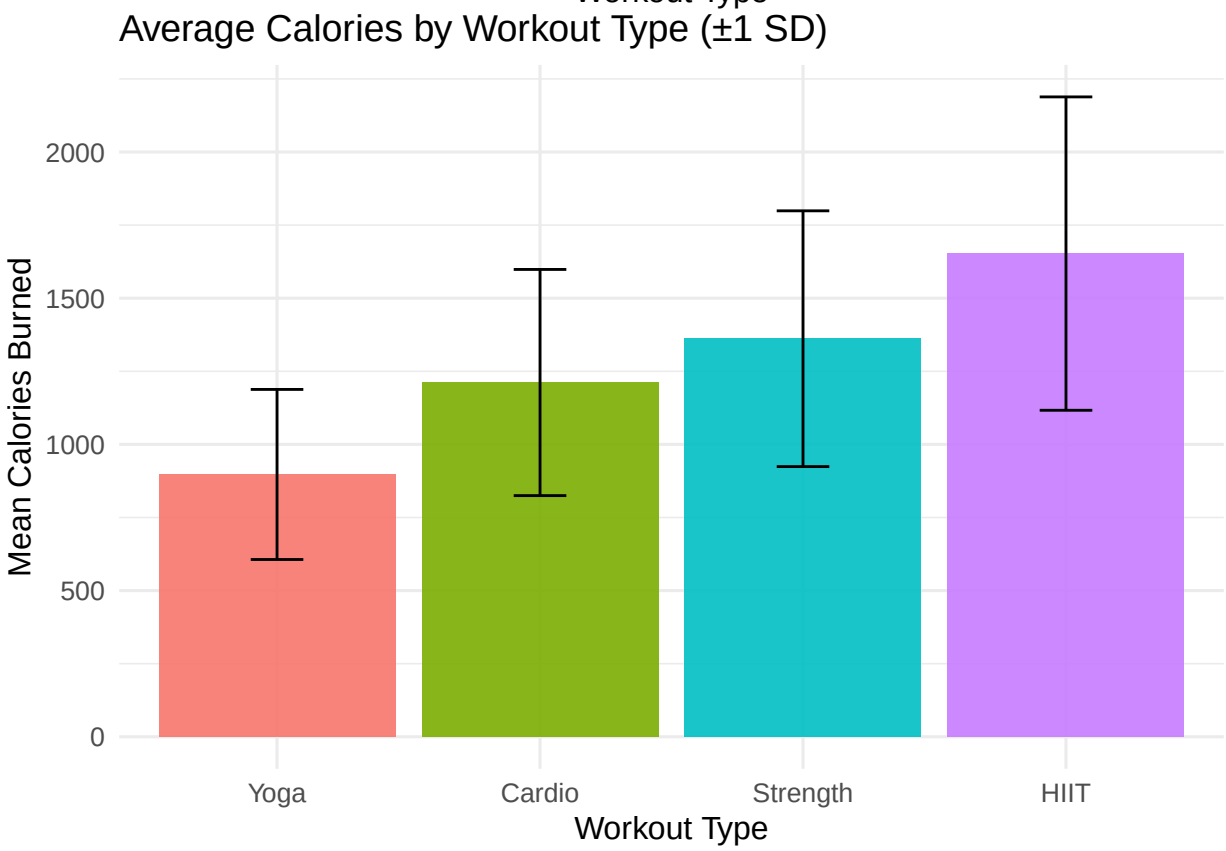
size of the number of calories burned per workout session and the data is split evenly between men and women. The data also contains a diverse number of different workouts completed across many different ages and levels of exercises which provides variability in the data. Lastly, we were given numerous other variables relating to attributes of the person and the workout that can help to control for confounding variables.

Exploratory Data Analysis

```
## # A tibble: 4 x 4
##   Workout_Type mean_cal sd_cal     n
##   <fct>           <dbl>  <dbl> <int>
## 1 Yoga             897.   291.  5032
## 2 Cardio          1212.   387.  4923
## 3 Strength        1361.   437.  5071
## 4 HIIT            1653.   536.  4974
```

```
## # A tibble: 4 x 3
##   Workout_Type median_cal iqr_cal
##   <fct>             <dbl>   <dbl>
## 1 Yoga               872.    289.
## 2 Cardio            1178.    385.
## 3 Strength          1324.    436.
## 4 HIIT              1609.    537.
```


Calories Burned by Workout Type
3000
2000
1000
Calories Burned
Yoga
Cardio
Strength
HIIT
Workout Type


Average Calories by Workout Type (±1 SD)
2000
1500
1000
500
0
Mean Calories Burned
Yoga
Cardio
Strength
HIIT
Workout Type

Conclusion

For our first numerical/graphical summary, we examined the mean, standard deviation, and the counts for each workout type . We can see that HIIT workouts have the highest calories burned (1652.5) as well as the highest standard deviation (535.9), while yoga has the lowest mean calories burned (897.1) and the lowest standard deviation (290.9). From this we conclude that on average you burn more calories from a HIIT workout, but because the SD is the highest, this means there is the highest variability for this workout also. This pattern implies that HIIT workout sessions are both the most calorie-intensive and the most variable. The large spread in HIIT could be because this type of workout depends the most on intensity, interval length, and rest duration, all of which differ significantly between different individuals and workout programs. For our next numerical/graphical summary, we see the same ranking in workout types versus calories burned: HIIT, strength, cardio, then yoga. However, in this summary we see that the median calories burned for all of the workout types are less than the mean. This means that for all workout types, there are upward outliers that are dragging the mean up. The highest difference between the mean and the median are HIIT workouts, with a spread of 43, most likely because of more intense or longer workout sessions.

Both the numerical and graphical summaries imply systematic differences in calorie burn across workout types, with greater variability linked to higher-intensity formats. In practical terms, these observations imply that choosing HIIT generally yields higher energy burn, but results depend greatly on how intensely the session is performed. On the other hand, yoga provides more predictable but lower calorie output.

Overall, our exploratory analysis supports the conclusion that workout type is a meaningful factor of calories burned, with intensity driving both the level and variability of energy expenditure.

References

https://www.kaggle.com/datasets/jockeroika/life-style-data?resource=download https://www.menshealth.com/fitness/a25424850/best-hiit-exercises-workout/ https://health.clevelandclinic.org/what-exercise-burns-the-most-calories https://www.careinsurance.com/blog/health-insurance-articles/benefits-and-ways-of-burning-calories-for-good-health https://github.com/cjjj2/STAT-3080-Project