APPLIED DATA SCIENCE 1 -STATISTICS AND TRENDS

GYM WORKOUT ANALAYSIS

Name: Daswadayalan Myladumparai Deenadayalan

Student Id: 23068427

GitHub repository: https://github.com/MDD25-Web-data-science/applied_data_science_1

Introduction:

In short span of years, The Data science has grown in various fields. It also now enhanced in work out and fitness. Using these datasets we can optimize the workout schedule. It also allows to track our workouts, performance and improve health outcomes. The report analyses the pattern of the workout data can make rightful decision in scheduling workout routines and overall fitness plans.

The aim of the report is to explain how data science can utilize the trend, minimised workout schedules and user experienced insights. Analysing the Gym Workout data, it gives clear decision on workout routine, Preferences and goals.

Data Collecting, Cleaning and Pre-Processing:

Collecting data from csv and data used in this analysis are workout type, session duration calories burned during work and genders involved in workout. Work out types are classified into **cardio**, **Hiit**, **strength**, **yoga** and time spent in workout is measured in hours and minute. Gender is classified into male and female. **Data cleaning** is to Remove empty values ,duplicated and formatting the column name in more readable way. **Categorize Data** is Categorizing the workout types, calories burned and duration to analyse the data in more efficient way.

Methods

Describing the data to get a Statistical analysis, which has minimum, maximum value, mean, quartile, Total count, median.

kurtosis can help to understand the variability in calories burned, session duration, BMI, Body Percentage, Heart Rate etc. **Kurtosis =0**, Where their most people have averagely data distributed workout. **Kurtosis > 0**, Where their people with extreme outliers E.g. Calories burned with intense workout with less session work time. **Kurtosis <0** Where their people heart rate is really close to mean E.g. Heart rate of workout people resting same or lesser than usual.

Skewness finds a statistical asymmetry of datasets around its mean.it is distributed evenly or in one direction. **Positive Skew** it shows the high calories burned due to intense workout in session than moderate people calories. **Negative Skew** it shows the Most of the induvial have less heart rate than Average heart rate and high heart rate people.

Correlating Data: Correlating all numerical values in the dataset using **Pearson** method to find whether each columns is falling positive, or negative. The Positive Correlating represents the

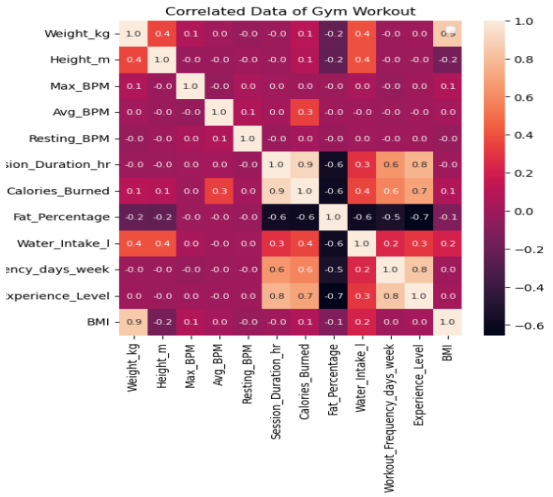

Fig 1.1 Correlated heat map

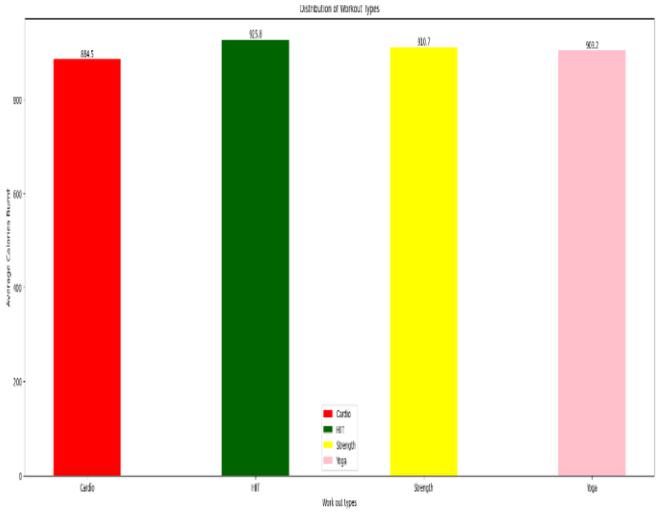

Fig 1.2 Average calories burnt and workout types

strong correlation between each column and Negative Correlating represents the Weaker Correlation between each column. **E.g. Fig 1.1 Correlated heat map**

Bar Graph is represented by average calories burnt by workout types, The Worked Out types were grouped into four types 'Yoga','Cardio','Hiit','Strength' and calories were burned into average and annotation of calories burnt also mentioned. Each workout type were represented with colours. **E.g. Fig 1.2 Average calories burnt and workout types**

Line Graph is represents Average Genders workout types and how much time spend to complete the workout. The genders workout session duration were averaged and framed into new data frame and then plotted a line graph for each genders x -axis represents workout types and y -axis represent the session in hr . **E.g. Fig.1.3 Workout duration and types**

Fig 1.3 Workout Duration and types

Conclusion

Therefore, the data approach to analysing gym workouts allows for optimized workout types, fitness plans, and improved health. By using data science, we can make informed decisions that enhance the workout types and efficiency of workout routines, ultimately leading to better fitness and health for individuals.