How General Health Index (GHI) performs better to predict health levels*

A comparitive analysis of GHI and BMI against body measurements

Aamishi Avarsekar

November 27, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

BMI is a common indicator for general health but it has its limitations. It does not take into consideration the actual body composition of an individual and other important factors such as sex and age. Through this paper, I would like to propose a new way categorising people based on other body measurements such as the waist-to-hip ratio and wrist and ankle circumferences measurements as they are more indicative of a person's visceral fat, this proving a more accurate representation of general health. I would like to propose a newer formula for General Health Index (GHI) through the BodyM dataset. I plan on using the Isaac Kuzmar dataset that has accurate measurements of body fat to create a Confusion Matrix to see how my model compares in providing more accurate body fat estimations through easlily measurable data such as waist and hip circumference.

2 Data

2.1 Overview

2.2 BMI Classification:

*Code and data are available at: https://github.com/aamishi/ImprovedGeneralHealthIndex/

```r
bmi_canada_gov <- data.frame(
  Category = c("Underweight", "Normal weight", "Overweight", "Obese - Class I", "Obese - Clas
  `BMI Range` = c("Below 18.5", "18.5-24.9", "25.0-29.9", "30.0-34.9", "35.0-39.9", "40.0 and
)

# Create a table using knitr::kable
knitr::kable(
  bmi_canada_gov,
  col.names = c("Category", "BMI Range"),
  caption = "BMI Categories as per the Government of Canada Guidelines"
)
```

Table 1: BMI Categories as per the Government of Canada Guidelines

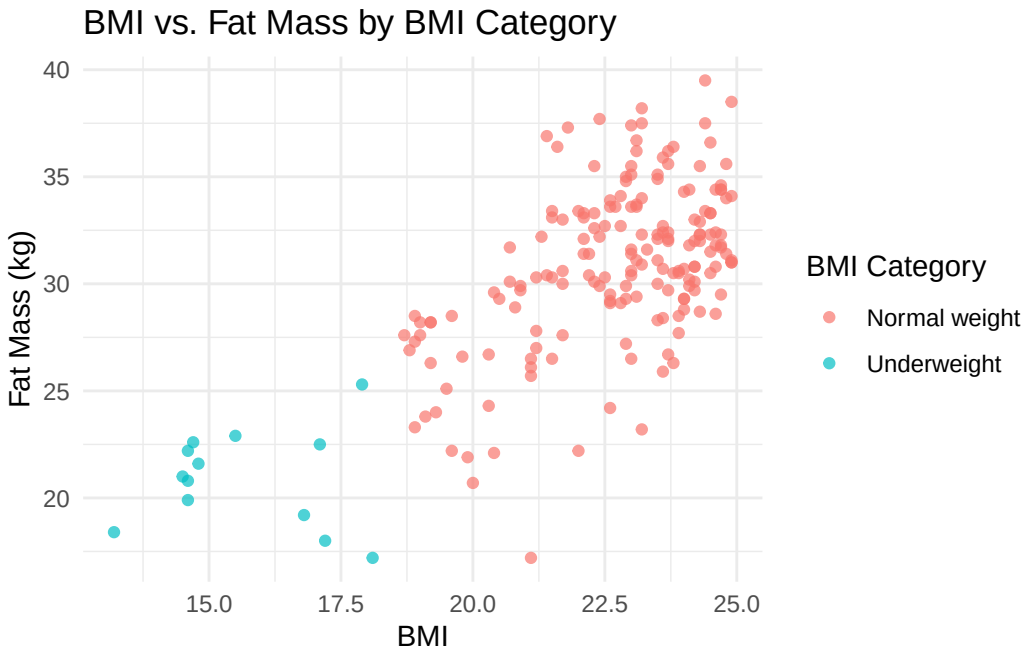
Category	BMI Range
Underweight	Below 18.5
Normal weight	18.5–24.9
Overweight	25.0–29.9
Obese - Class I	30.0–34.9
Obese - Class II	35.0–39.9
Obese - Class III	40.0 and above

2.2.1 The difference between an underweight and normal person:

Let us take a closer look at the BMI levels below 25 for underweight and normal weight adults.

```r
body_mass_data_raw <- read_parquet("../data/raw_data/body_mass_data.parquet")
body_mass_clean <- body_mass_data_raw %>%
  clean_names() %>%
  select(gender_m1f2, age, height_cm, w_kg, bmi, fat_mass_kg, fm, ffm_kg, bone_mass_kg, muscl
  mutate(
    bmi_category = case_when(
      bmi < 18.5 ~ "Underweight",
      bmi >= 18.5 & bmi <= 24.9 ~ "Normal weight",
      bmi >= 25.0 & bmi <= 29.9 ~ "Overweight",
      bmi >= 30.0 & bmi <= 34.9 ~ "Obese - Class I",
      bmi >= 35.0 & bmi <= 39.9 ~ "Obese - Class II",
      bmi >= 40.0 ~ "Obese - Class III"
    ))
```

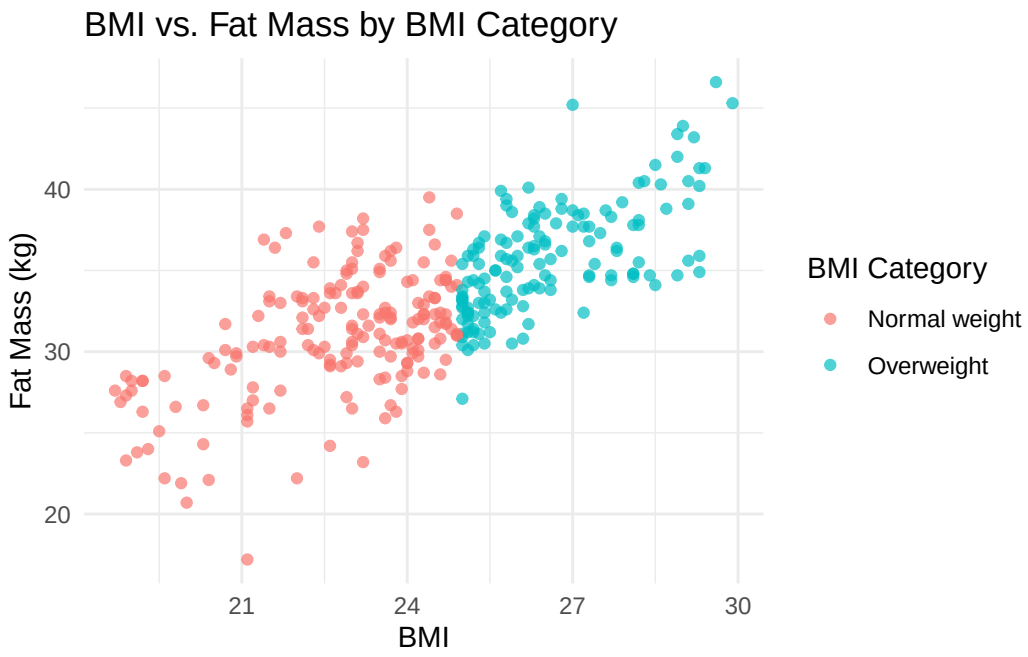
```r
# Data Section: Part 1: What is the difference between an underweight / athletic person / nor
bmi_under_25 <- body_mass_clean %>%
  filter(bmi < 25)

ggplot(bmi_under_25, aes(x = bmi, y = fm)) +
  geom_point(aes(color = bmi_category), alpha = 0.7) +
  labs(
    title = "BMI vs. Fat Mass by BMI Category",
    x = "BMI",
    y = "Fat Mass (kg)",
    color = "BMI Category"
  ) +
  theme_minimal()
```



2.2.2 The difference between a normal and overweight person:

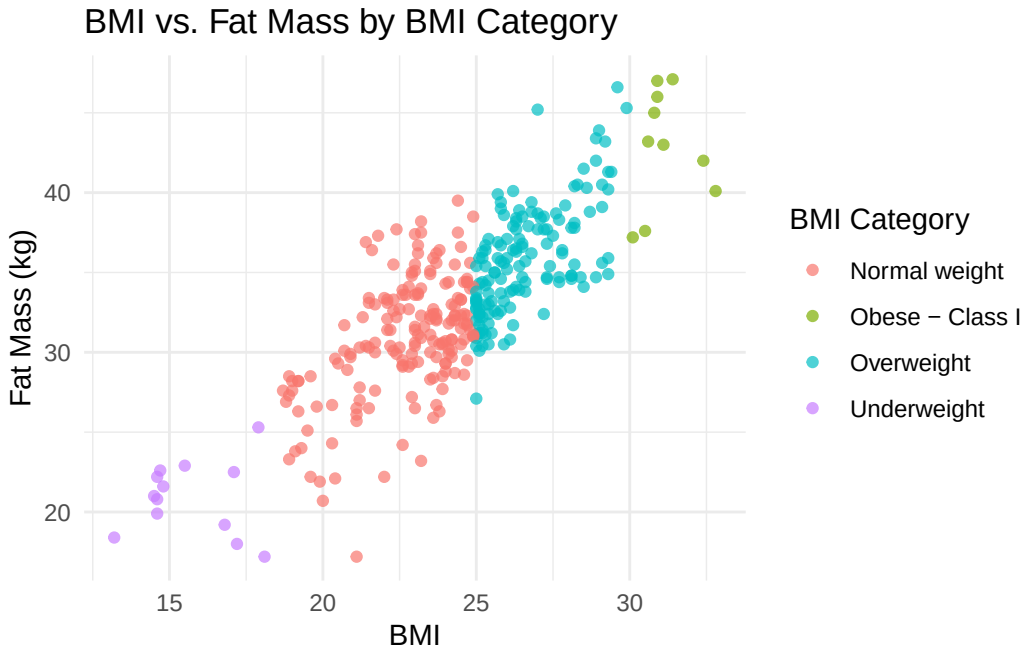
```r
bmi_normal_over <- body_mass_clean %>%
  filter(18.5 < bmi & bmi < 30)

# this is what i chose
ggplot(bmi_normal_over, aes(x = bmi, y = fm)) +
```

```
  geom_point(aes(color = bmi_category), alpha = 0.7) +
  labs(
    title = "BMI vs. Fat Mass by BMI Category",
    x = "BMI",
    y = "Fat Mass (kg)",
    color = "BMI Category"
  ) +
  theme_minimal()
```



2.2.3 All categories

```
# this is what i chose
ggplot(body_mass_clean, aes(x = bmi, y = fm)) +
  geom_point(aes(color = bmi_category), alpha = 0.7) +
  labs(
    title = "BMI vs. Fat Mass by BMI Category",
    x = "BMI",
    y = "Fat Mass (kg)",
    color = "BMI Category"
  ) +
  theme_minimal()
```



Overview text

2.3 Measurement

Some paragraphs about how we go from a phenomena in the world to an entry in the dataset.

Talk more about it.

Talk way more about it.

2.4 Predictor variables

Add graphs, tables and text.

Use sub-sub-headings for each outcome variable and feel free to combine a few into one if they go together naturally.

3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

Table 2: Explanatory models of flight time based on wing width and wing length

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

4 Results

Our results are summarized in Table 2.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

we compare the posterior with the prior. This shows...

References

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.