

Pokemon Analysis

Connor Bryson, Nathen Byford, Miguel Iglesias

10/12/2021

Introduction

Questions and Findings

Speed

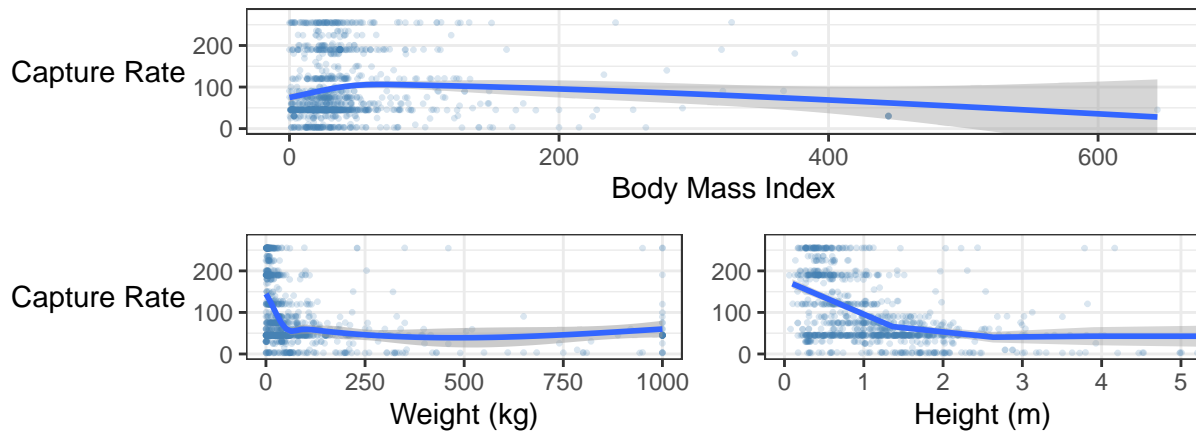
We want to begin with looking into what factors effect the speed of the Pokemon.

Capture Rate

We also wanted to look into how capture rate is effected by variables such as height, weight, gender, and legendary.

```
#Create a plot that shows the relationship between capture_rate and BMI  
#Limit to not include a lot of empty space  
BMI <- ggplot(pokemon_new, aes(BMI, capture_rate)) +  
  geom_jitter(alpha = 0.2, color = "Steel Blue", size = .5) +  
  geom_smooth() +  
  coord_cartesian(ylim = c(0,275)) +  
  labs(x = "Body Mass Index", y = "Capture Rate")  
  
#Create a plot that shows the relationship between capture_rate and Height  
#Limit included to allow for best view of relationship  
Height <- ggplot(pokemon_new, aes(height/10, capture_rate)) +  
  geom_jitter(alpha = 0.2, color = "Steel Blue", size = .5) +  
  geom_smooth() +  
  coord_cartesian(xlim = c(0,5), ylim = c(0,275)) +  
  labs(x = "Height (m)", y = NULL)  
  
#Create a plot that shows the relationship between capture_rate and weight  
#Limit included to allow for best view of relationship  
Weight <- ggplot(pokemon_new, aes(weight/10, capture_rate)) +  
  geom_jitter(alpha = 0.2, color = "Steel blue", size = .5) +  
  geom_smooth(alpha = 0.5) +  
  coord_cartesian(ylim = c(0,275)) +  
  labs(x = "Weight (kg)", y = "Capture Rate")  
  
BMI / (Weight | Height)
```

```
## 'geom_smooth()' using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## 'geom_smooth()' using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## 'geom_smooth()' using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



```
p1 <- ggplot(pokemon_new) +
  geom_boxplot(aes(reorder(shape, capture_rate, FUN = median), capture_rate)) +
  labs(x = "Shape", y = "Capture Rate")

p2 <- ggplot(pokemon_new, aes(mythical, capture_rate)) +
  geom_boxplot() +
  labs(x = "Mythical", y = "Capture Rate")

p3 <- ggplot(pokemon_new, aes(legendary, capture_rate)) +
  geom_boxplot() +
  labs(x = "Legendary", y = NULL)

p4 <- ggplot(pokemon_new, aes(genderless, capture_rate)) +
  geom_boxplot() +
  labs(x = "Genderless", y = NULL)

p5 <- ggplot(pokemon_new, aes(baby_pokemon, capture_rate)) +
  geom_boxplot() +
  labs(x = "Baby Pokemon", y = NULL)

p6 <- ggplot(pokemon_new, aes(as.factor(female_rate*100), capture_rate)) +
  geom_boxplot() +
  labs(x = "Percent Female", y = "Capture Rate")

(p2 | p3 | p4 | p5) / p6
```



Happiness

Lastly we wanted to see what factors effected the happiness of the Pokemon. This variable was just something we thought would be fun to look into. Our main questions were what other variables effect the Pokemon's happiness. We began with looking at if the Pokemon is a legendary and if it's a mythical Pokemon.

We were also curious if the female rate effected happiness of the Pokemon. Converting the happiness variable to a likert scale we can easily see the amount of data points in each bin of female rate in the following plot. As seen in the plots before there is a smaller amount of Pokemon with a high female rate. Notice how it appears that depending of the female rate there are different levels of happiness that are observed. What stood out to us was that Pokemon with a female rate of 0.25 or 0.875 are only observed with neutral happiness and there are no very unhappy Pokemon with a rate of 1. There is no apparent correlation between happiness and female rate.

```
pokemon_new |> ggplot(aes(as.factor(female_rate), happiness)) +
  geom_bin2d(alpha = .9) +
  scale_fill_viridis_c() +
  labs(title = "Happiness of Pokemon by Female Rate", x = "Female Rate", y = "Happiness") +
  theme(panel.grid.major.x = element_blank())
```

