

Opgaver lektion 3

Simon

19/11/2022

Inden opgaven læs følgende pakker!

```
library(tidyverse)
library(ggplot2)
```

Opgave 1:

Brug de følgende steps til at udregne BMI for Starwars karaktere. baseret på deres hår farve:

Step 1

Vælg de 4 variable vi ønsker at bruge (Vi beholder “name”)

Fejrn “NA” værdier med funktionen “drop_na()”

```
starwars %>%
  select(name, hair_color, mass, height)%>%
  drop_na()
```

```
## # A tibble: 54 x 4
##   name          hair_color    mass height
##   <chr>         <chr>      <dbl> <int>
## 1 Luke Skywalker blond          77    172
## 2 Darth Vader   none         136    202
## 3 Leia Organa   brown         49    150
## 4 Owen Lars     brown, grey   120    178
## 5 Beru Whitesun lars brown         75    165
## 6 Biggs Darklighter black         84    183
## 7 Obi-Wan Kenobi auburn, white 77    182
## 8 Anakin Skywalker blond         84    188
## 9 Chewbacca     brown        112    228
## 10 Han Solo      brown         80    180
## # ... with 44 more rows
```

Step 2

Brug mutate til at lave en ny kolonne der viser BMI:

Formlen for BMI er følgende: $BMI = \frac{mass(kg)}{height(M)^2}$

Da vi skal have height i meter skal vi også bruge mutate til at ændre denne kolonne:

Husk vi stadig skal bruge ovenstående steps

```
starwars %>%
  select(name, hair_color, mass, height)%>%
  drop_na() %>%
  mutate(height= height/100)
```

```
## # A tibble: 54 x 4
##   name          hair_color    mass height
##   <chr>         <chr>      <dbl> <dbl>
## 1 Luke Skywalker blond        77  1.72
## 2 Darth Vader   none       136  2.02
## 3 Leia Organa   brown        49  1.5
## 4 Owen Lars     brown, grey  120  1.78
## 5 Beru Whitesun lars brown        75  1.65
## 6 Biggs Darklighter black        84  1.83
## 7 Obi-Wan Kenobi auburn, white 77  1.82
## 8 Anakin Skywalker blond        84  1.88
## 9 Chewbacca     brown       112  2.28
## 10 Han Solo     brown        80  1.8
## # ... with 44 more rows
```

Vi tilføjer nu BMI:

```
starwars %>%
  select(name, hair_color, mass, height)%>%
  drop_na() %>%
  mutate(height= height/100) %>%
  mutate(BMI= mass/(height^2))
```

```
## # A tibble: 54 x 5
##   name          hair_color    mass height  BMI
##   <chr>         <chr>      <dbl> <dbl> <dbl>
## 1 Luke Skywalker blond        77  1.72 26.0
## 2 Darth Vader   none       136  2.02 33.3
## 3 Leia Organa   brown        49  1.5  21.8
## 4 Owen Lars     brown, grey  120  1.78 37.9
## 5 Beru Whitesun lars brown        75  1.65 27.5
## 6 Biggs Darklighter black        84  1.83 25.1
## 7 Obi-Wan Kenobi auburn, white 77  1.82 23.2
## 8 Anakin Skywalker blond        84  1.88 23.8
## 9 Chewbacca     brown       112  2.28 21.5
## 10 Han Solo     brown        80  1.8  24.7
## # ... with 44 more rows
```

Hurtig Øvelse, Hvem har den højeste og mindste BMI? Brug Arrange() funktionen

Step 3

Vi kan nu udregne gennemsnit af BMI for hver hår farve:

Vi bruger en ny funktion kaldet “group_by()” Den opdeller kategoriske variable op i hver deres kategori og fungerer derfor godt sammen med summarise funktionen:

```
starwars %>%
  select(name, hair_color, mass, height)%>%
  drop_na() %>%
  mutate(height= height/100) %>%
  mutate(BMI= mass/(height^2)) %>%
  group_by(hair_color) %>%
  summarise(mean(BMI))
```

```
## # A tibble: 9 x 2
##   hair_color    'mean(BMI)'
##   <chr>         <dbl>
## 1 auburn, white    23.2
## 2 black           22.8
## 3 blond           24.9
## 4 blonde          19.5
## 5 brown           24.5
## 6 brown, grey     37.9
## 7 grey            26.0
## 8 none            24.4
## 9 white           27.1
```

Så det ser ud til Starwars karaktere med brunt/gråt hår har den højeste BMI!