adventr3

Nival Kolambage

5/27/2020

Loading the needed libraries

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.0     v purrr   0.3.4
## v tibble  3.0.1     v dplyr   0.8.5
## v tidyr   1.0.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Loading data file

```
ha_tib <- readr::read_csv("../data/ais_ch03_ha.csv")
```

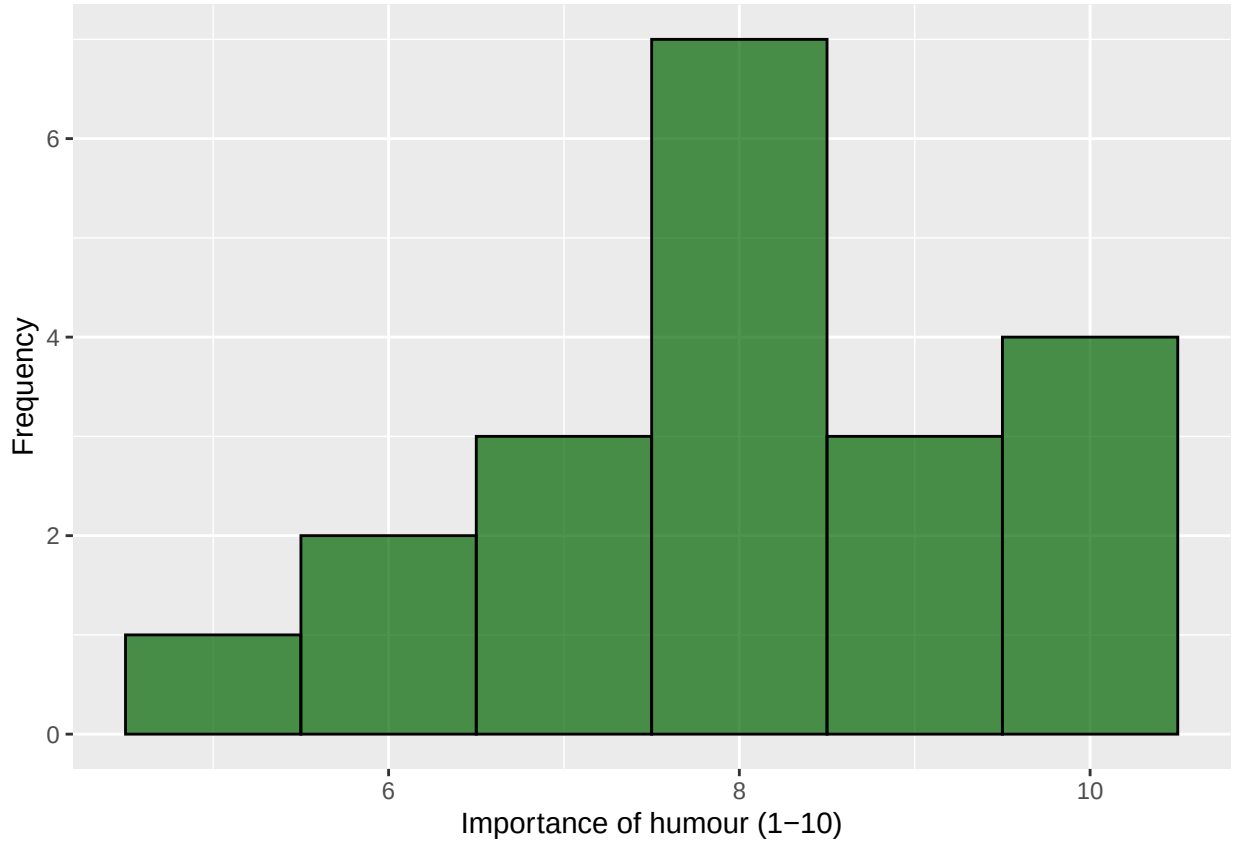
```
## Parsed with column specification:
## cols(
##   ID = col_double(),
##   ageYears = col_double(),
##   Sex = col_character(),
##   HighSalary = col_double(),
##   FinishedEducation = col_double(),
##   Kind = col_double(),
##   Humour = col_double(),
##   Ambitious = col_double(),
##   WantsChildren = col_double(),
##   Romantic = col_double(),
##   AttractiveAppearance = col_double(),
##   Creativity = col_double(),
##   Honest = col_double()
## )
```

```
ha_tib
```

```
## # A tibble: 20 x 13
##       ID ageYears Sex   HighSalary FinishedEducati~  Kind Humour Ambitious
##    <dbl>    <dbl> <chr>      <dbl>             <dbl> <dbl>  <dbl>     <dbl>
##  1 90308       16 Fema~          4                 9    10      8         7
##  2 90220       15 Fema~          5                 8    10     10         7
##  3 10723       16 Fema~          9                 5     8      8         8
##  4 40502       16 Fema~          4                 4     9      8         5
##  5 60622       17 Fema~          5                 7    10     10         8
##  6 10625       16 Fema~          3                 3    10      8         7
##  7 40716       18 Fema~         10                10    10     10        10
##  8 50433       15 Fema~         10                 7     9      9         6
##  9 90505       15 Fema~          6                 6     8      7         7
## 10 80416       15 Fema~          1                 4     7      6         7
## 11 70218       14 Fema~          3                 4     8      6         4
## 12 50421       15 Fema~          8                 8     9      8         6
## 13 20607       14 Fema~          8                 9     9      7         5
## 14 30208       15 Fema~          2                 1    10      8         9
## 15 10811       16 Fema~          6                 7     9      7         7
## 16 90420       15 Fema~          4                 3     9      8         6
## 17 60111       16 Fema~          8                 7     9      5         7
## 18 30618       16 Fema~          4                 6    10      9         8
## 19 80117       15 Fema~          7                10    10     10         5
## 20 10505       16 Fema~          7                 9    10      9         6
## # ... with 5 more variables: WantsChildren <dbl>, Romantic <dbl>,
## #   AttractiveAppearance <dbl>, Creativity <dbl>, Honest <dbl>
```

Plotting 'humour' variable in a histogram :

```
humour_hist <- ggplot(ha_tib, aes(Humour))
humour_hist +
  geom_histogram(binwidth = 1, color = 'black', fill = 'dark green', alpha = 0.7) +
  labs(y = "Frequency", x = "Importance of humour (1-10)")
```

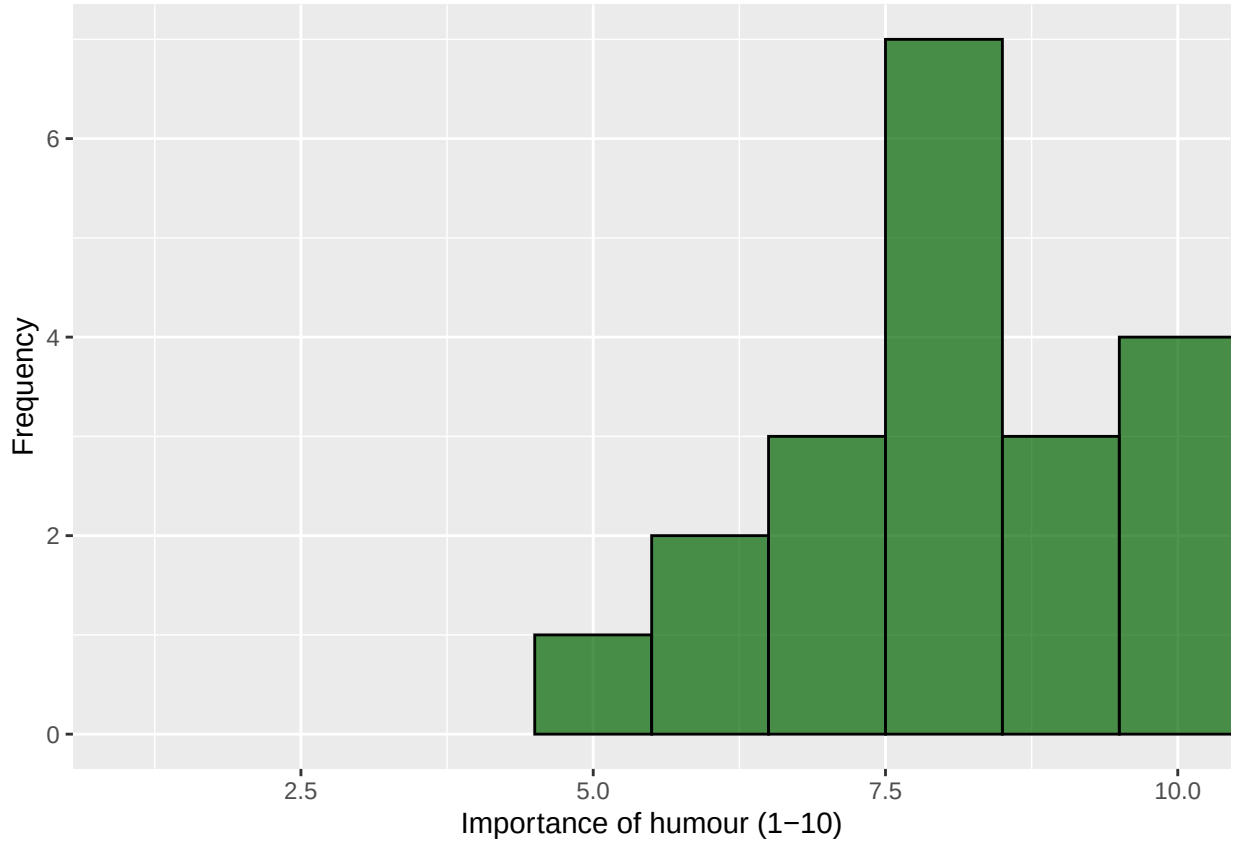


Customizing x scale : At the moment the x-axis is scaled from 5 to 10. Let's show the full range of the scale. To do this we need to set the limits of the x-axis using the coord_cartesian() function:

coord_cartesian(xlim = c(begin, end), ylim = c(begin, end))

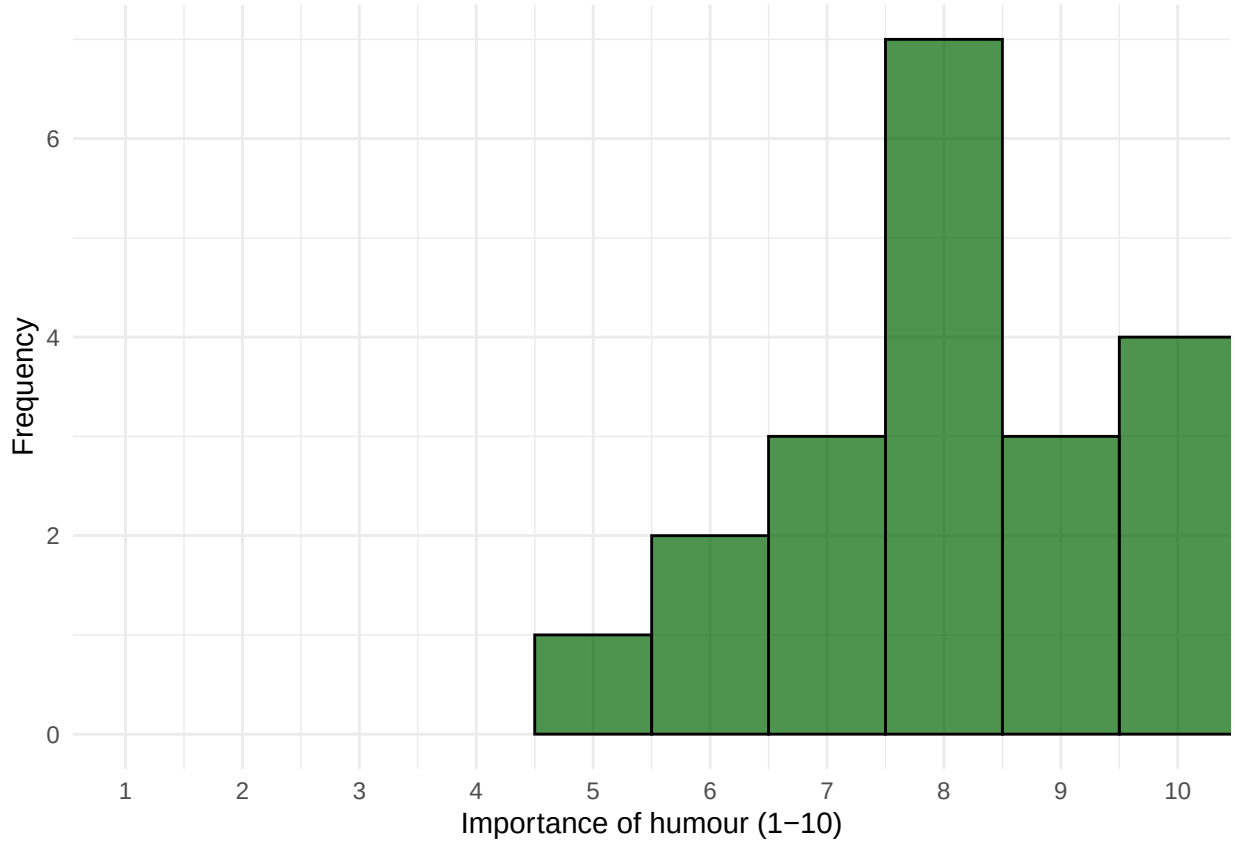
You set the limits of the x-axis using xlim and the limits of the y-axis with ylim. After each you specify numbers representing the start and end values for the axis. You need to collect these values into a single object by enclosing them in c().

```
humour_hist <- ggplot(ha_tib, aes(Humour))
humour_hist +
  geom_histogram(binwidth = 1, color = 'black', fill = 'dark green', alpha = 0.7) +
  labs(y = "Frequency", x = "Importance of humour (1-10)")+
  coord_cartesian(xlim = c(1, 10))
```



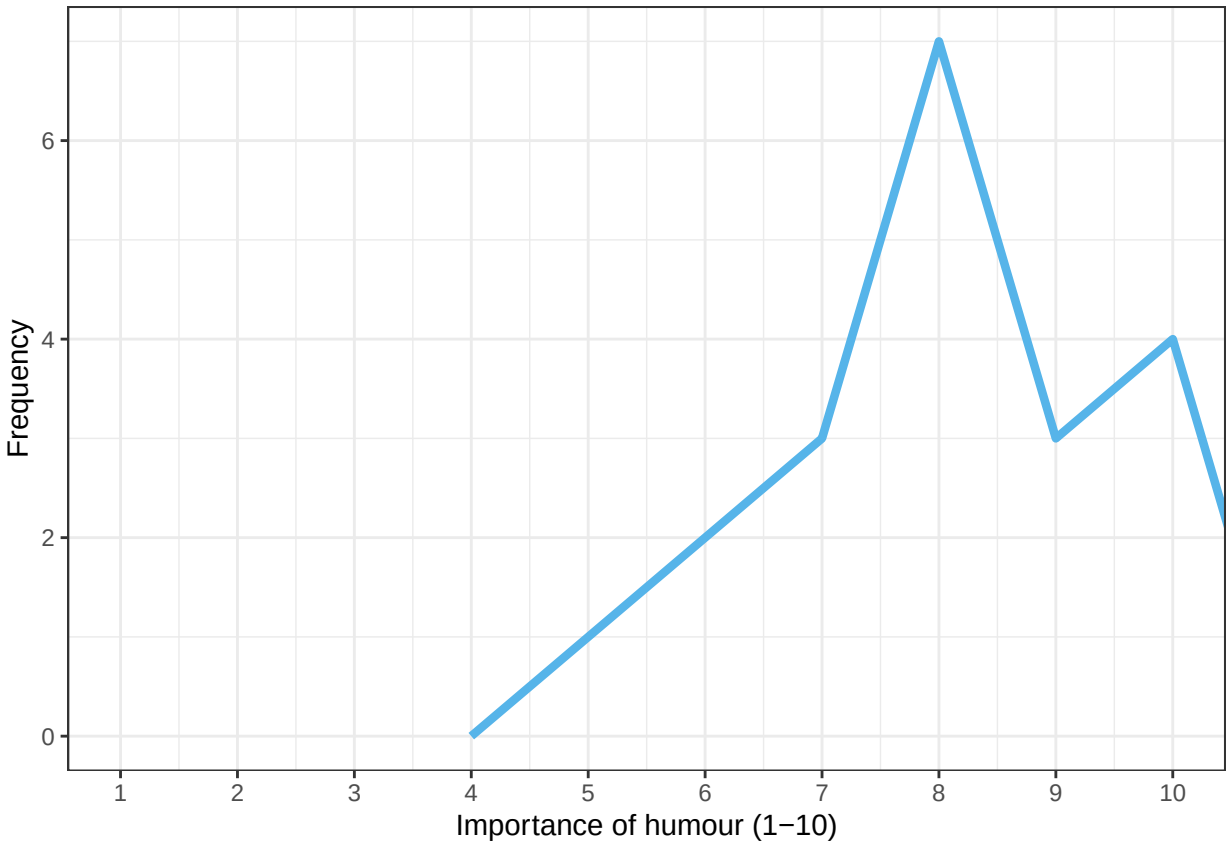
Overriding the default breaks and specifying the breaks for x-axis. Last line applies a theme to change the look. Other themes include theme_classic, theme_bw, theme_dark.

```
humour_hist <- ggplot(ha_tib, aes(Humour))
humour_hist +
  geom_histogram(binwidth = 1, color = 'black', fill = 'dark green', alpha = 0.7) +
  labs(y = "Frequency", x = "Importance of humour (1-10)")+
  coord_cartesian(xlim = c(1, 10)) +
  scale_x_continuous(breaks = 1:10) +
  theme_minimal()
```



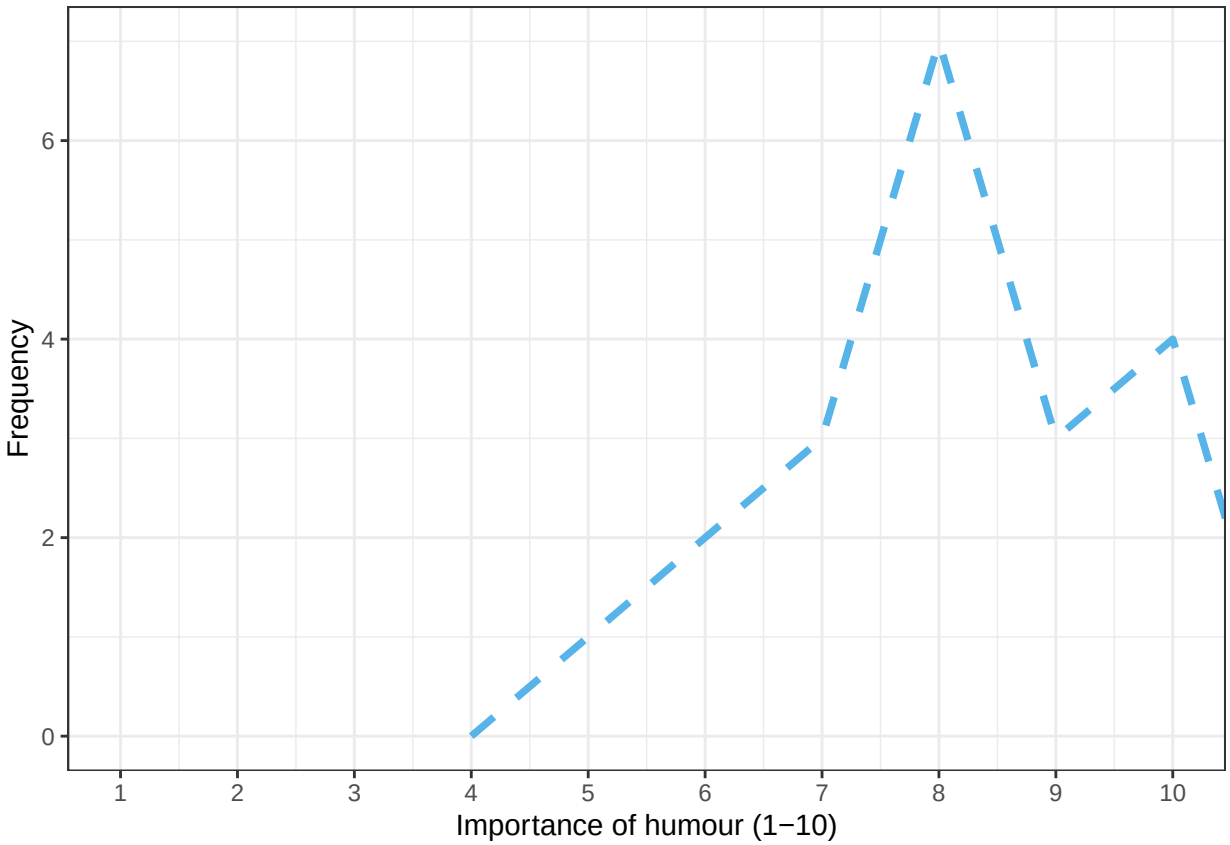
Drawing the same above graph as a frequency polygon. You can change the line size by adding "size" arg to geom_freqpoly():

```
humour_poly <- ggplot2::ggplot(ha_tib, aes(Humour))
humour_poly +
  geom_freqpoly(binwidth = 1, colour = "#56B4E9",size = 1.5) +
  labs(y = "Frequency", x = "Importance of humour (1-10)") +
  coord_cartesian(xlim = c(1, 10)) +
  scale_x_continuous(breaks = 1:10) +
  theme_bw()
```



We can change the style of the line by including the linetype = option in the geom_freqpoly() function. Line types can either be defined using numbers (0 = blank, 1 = solid (default), 2 = dashed, 3 = dotted, 4 = dotdash, 5 = longdash, 6 = twodash) or as text ("blank", "solid", "dashed", "dotted", "dotdash", "longdash", or "twodash").

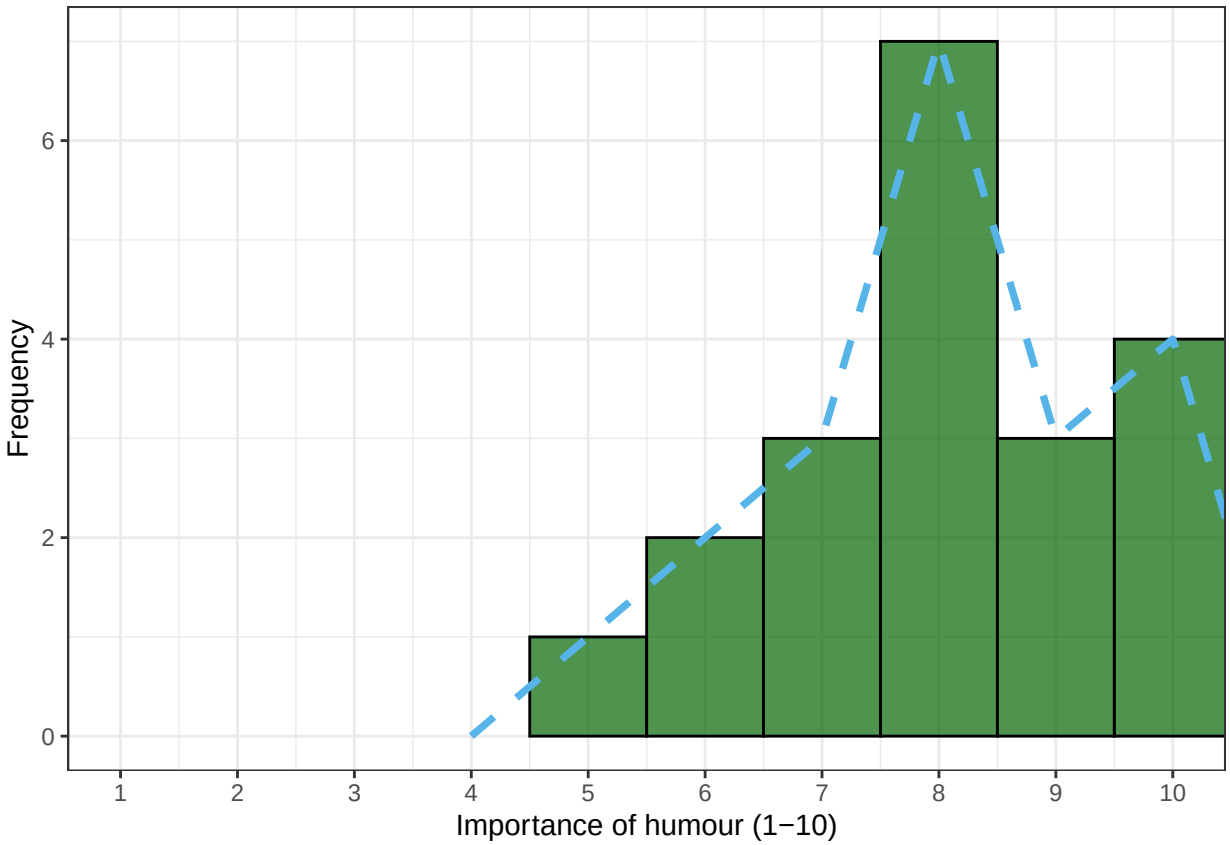
```
humour_poly <- ggplot2::ggplot(ha_tib, aes(Humour))
humour_poly +
  geom_freqpoly(binwidth = 1, colour = "#56B4E9",size = 1.3, linetype = 'dashed') +
  labs(y = "Frequency", x = "Importance of humour (1-10)") +
  coord_cartesian(xlim = c(1, 10)) +
  scale_x_continuous(breaks = 1:10) +
  theme_bw()
```



You can draw more tha one plots in one diagram, but R will process the plots in the order they are executed:

```
humour_poly <- ggplot2::ggplot(ha_tib, aes(Humour))

humour_poly +
  geom_histogram(binwidth = 1, color = 'black', fill = 'dark green', alpha = 0.7) +
  geom_freqpoly(binwidth = 1, colour = "#56B4E9",size = 1.3, linetype = 'dashed') +
  labs(y = "Frequency", x = "Importance of humour (1-10)") +
  coord_cartesian(xlim = c(1, 10)) +
  scale_x_continuous(breaks = 1:10) +
  theme_bw()
```


Frequency
0
2
4
6
1
2
3
4
5
6
7
8
9
10
Importance of humour (1−10)