

Week 1 seminar

Anastasia Chanbour

01/10/2021

Aim: Visualising the gapminder data.

Task 1: Gapminder analysis

1.1: Loading the required packages

```
library("gapminder")
library("tidyverse")

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr   0.3.4
## v tibble  3.1.5      v dplyr  1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.0.2      v forcats 0.5.1

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

1.2

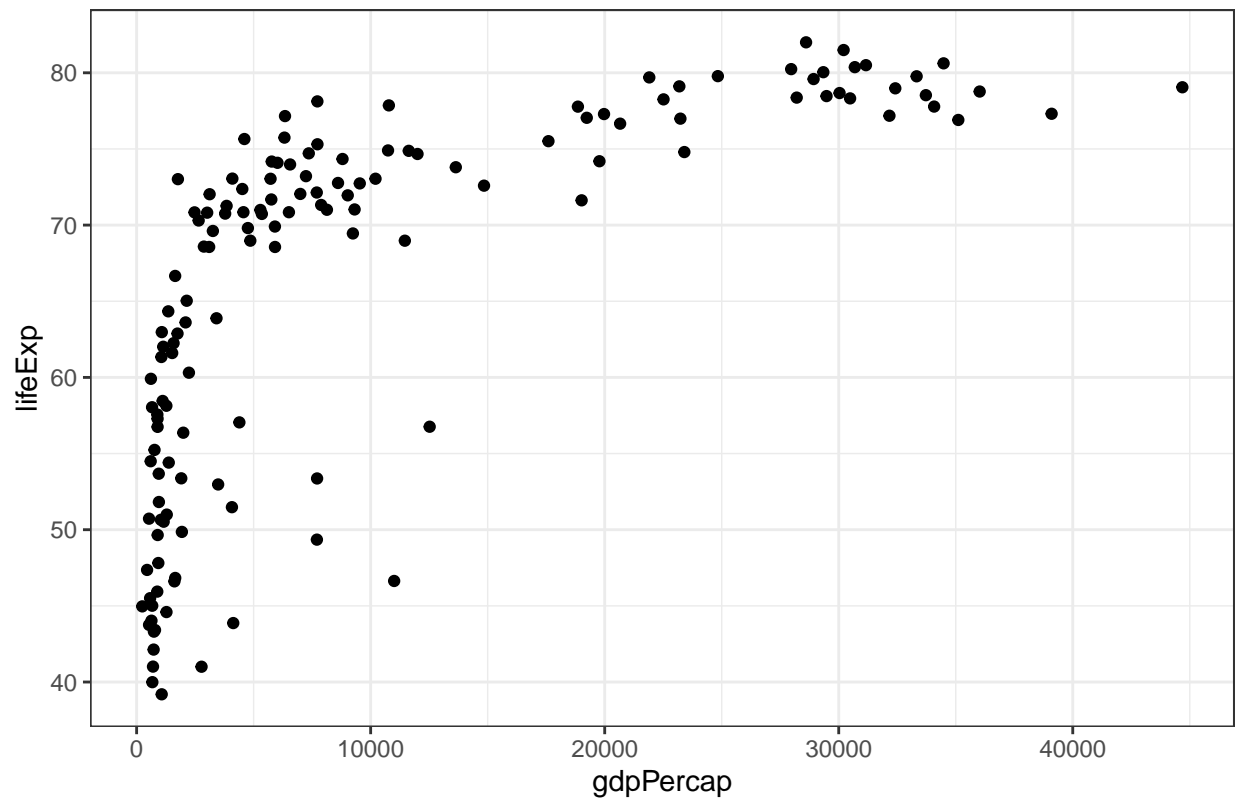
Generating a scatterplot with x=gdpPercap, and y=lifeExp

- we call the data object from 2002 gm2002
- we call the scatterplot gm_scatterplot
- we display the plot

```
gm2002 <- gapminder %>% filter(year == 2002)
gm_scatterplot <- ggplot(data=gm2002, aes(x=gdpPercap, y=lifeExp)) +
  geom_point()+
  labs(title="Scatter plot")+
  theme_bw()

gm_scatterplot
```

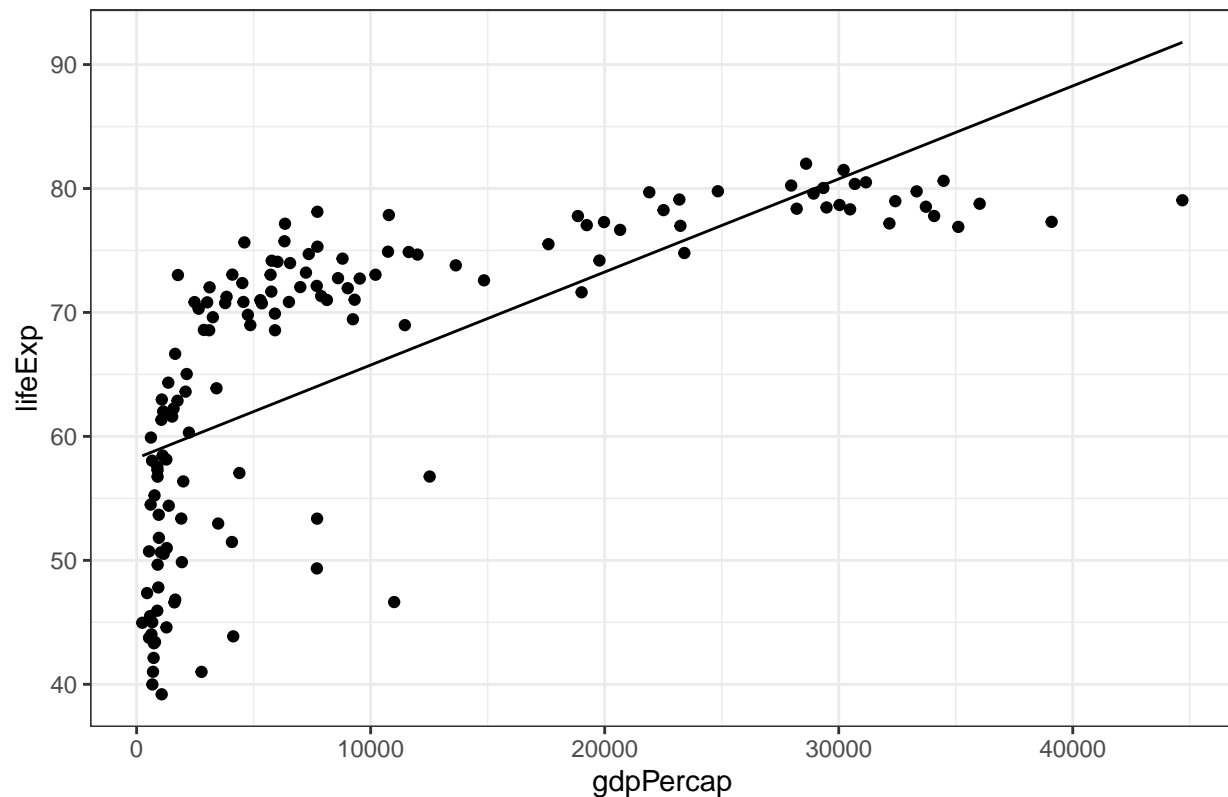
Scatter plot



1.3: Adding lines to the scatterplot to show the linear model and loess model

```
model_lm <- lm(lifeExp ~ gdpPerCap, data = gm2002)
predictions_lm <- broom::augment(model_lm)
gm_scatterplot + geom_line(data = predictions_lm, aes(y=.fitted))
```

Scatter plot



1.3.2 loess model to predict lifeExp

```
model_loess <- loess(lifeExp ~ gdpPercap,
  data = gm2002,
  span = 0.1)
```

```
predictions_loess <- broom::augment(model_loess)
summary(predictions_loess)
```

	lifeExp	gdpPercap	.fitted	.resid
## Min.	:39.19	Min. : 241.2	Min. :44.15	Min. : -22.32928
## 1st Qu.	:55.52	1st Qu.: 1409.6	1st Qu.:56.26	1st Qu.: -2.17004
## Median	:70.83	Median : 5319.8	Median :67.55	Median : 0.49155
## Mean	:65.69	Mean : 9917.9	Mean :65.63	Mean : 0.06093
## 3rd Qu.	:75.46	3rd Qu.:13359.5	3rd Qu.:74.03	3rd Qu.: 3.89777
## Max.	:82.00	Max. :44684.0	Max. :79.88	Max. : 13.04562

Now all three models together

```
gm_scatterplot + geom_line(data = predictions_lm,
  aes(y=.fitted), color = "blue", linetype = "dashed") +
  geom_line(data = predictions_loess, aes(y=.fitted), color = "green")
```

Scatter plot

