R: ggplot2 Exercises

Ai Yukino

Import R packages

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.7      v dplyr   1.0.9
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Import data
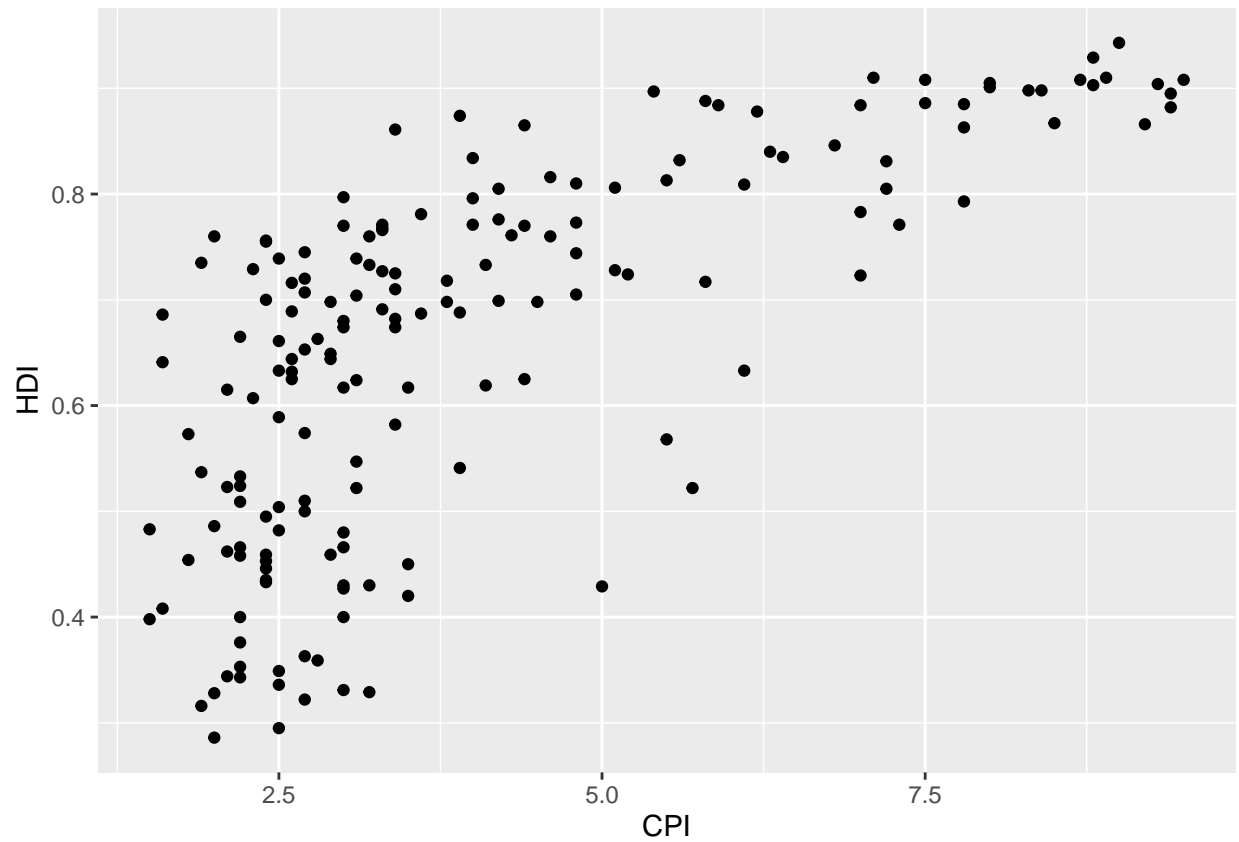
```
tb <- read_csv("data/EconomistData.csv")
```
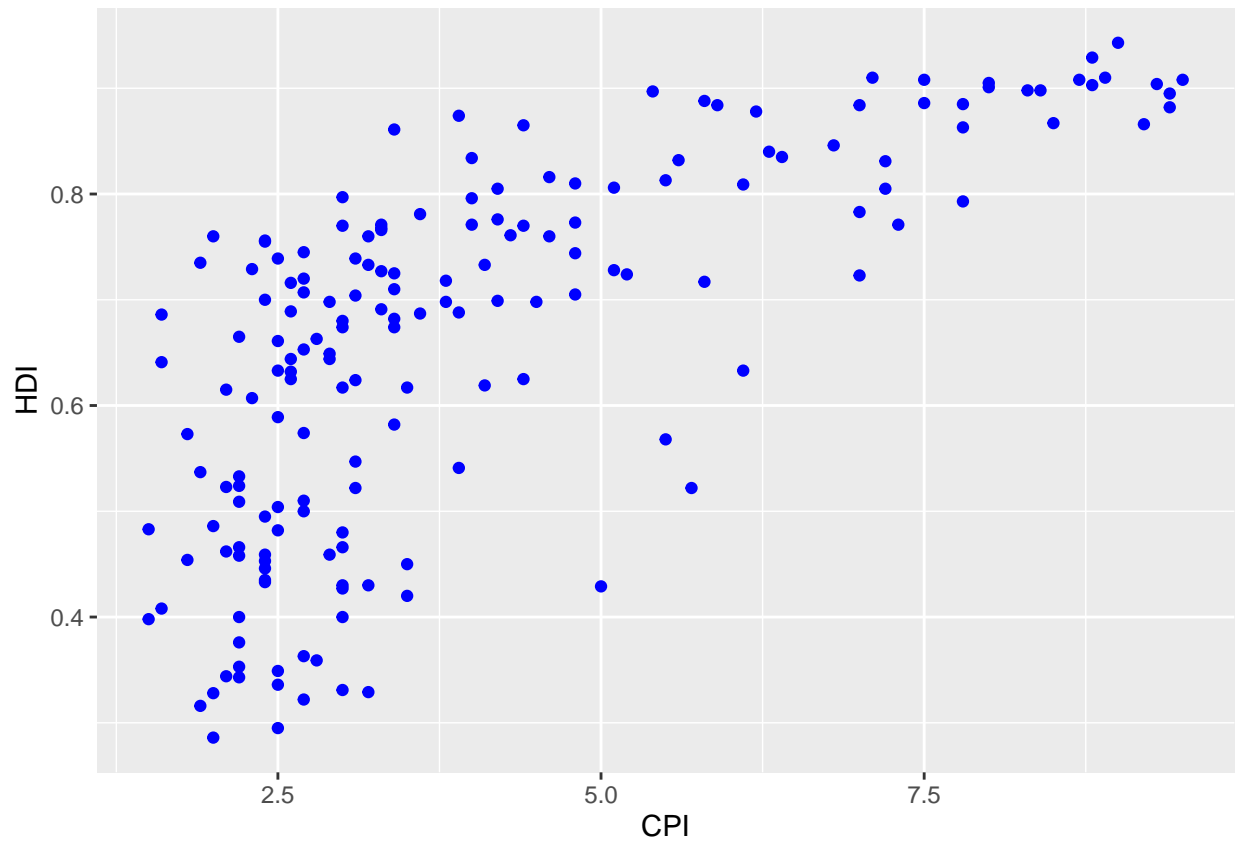
Exercise I

(1) Create a scatter plot with CPI on the x axis and HDI on the y axis

```
ggplot(data = tb, aes(x = CPI, y = HDI)) +
  geom_point()
```

(2) Color the points blue

```
ggplot(data = tb, aes(x = CPI, y = HDI)) +
  geom_point(color="blue")
```
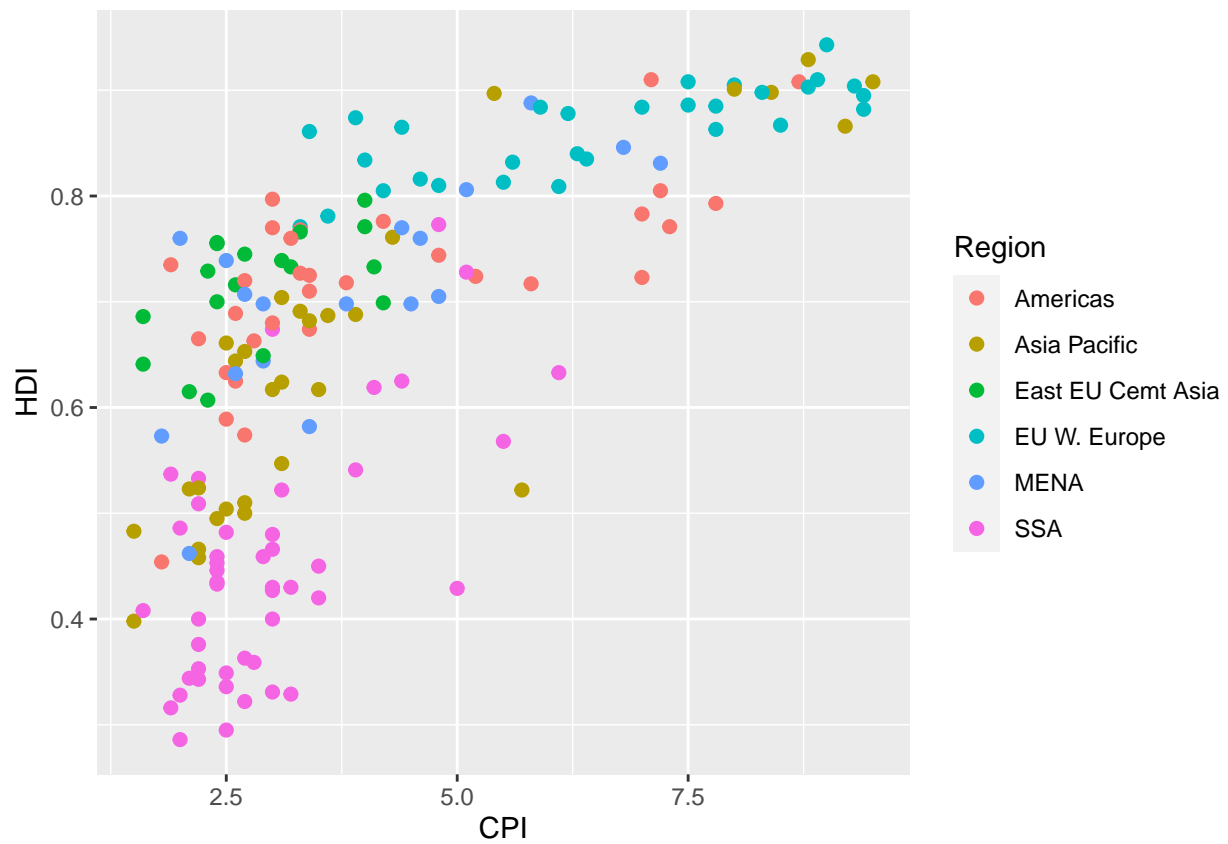
(3) Map the color of the the points to Region.

```
ggplot(data = tb, aes(x = CPI, y = HDI, color = Region)) +
  geom_point()
```
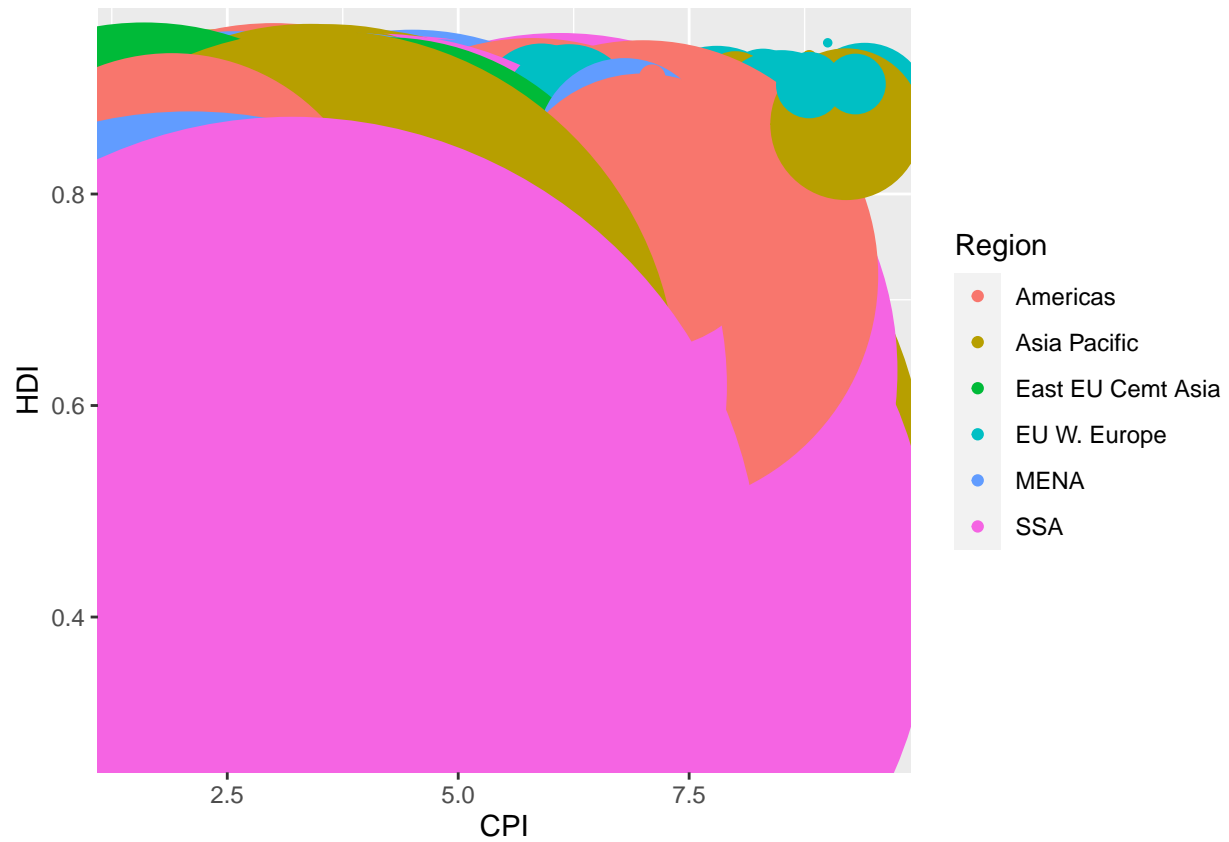
(4) Make the points bigger by setting size to 2

```
ggplot(data = tb, aes(x = CPI, y = HDI, color = Region)) +
  geom_point(size = 2)
```
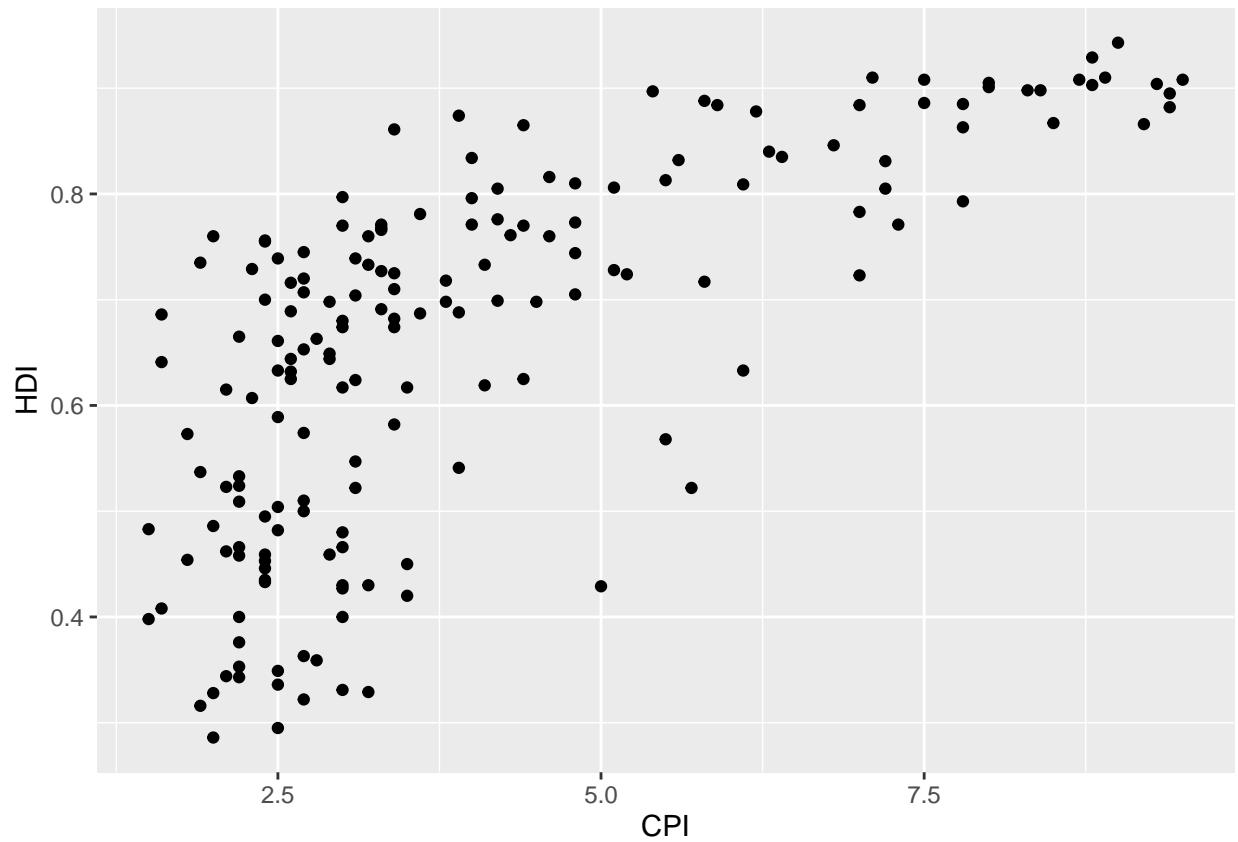
(5) Map the size of the points to HDI.Rank

```
ggplot(data = tb, aes(x = CPI, y = HDI, color = Region)) +
  geom_point(size = tb$HDI.Rank)
```

Exercise II

(1) Re-create a scatter plot with CPI on the x axis and HDI on the y axis (as you did in the previous exercise).
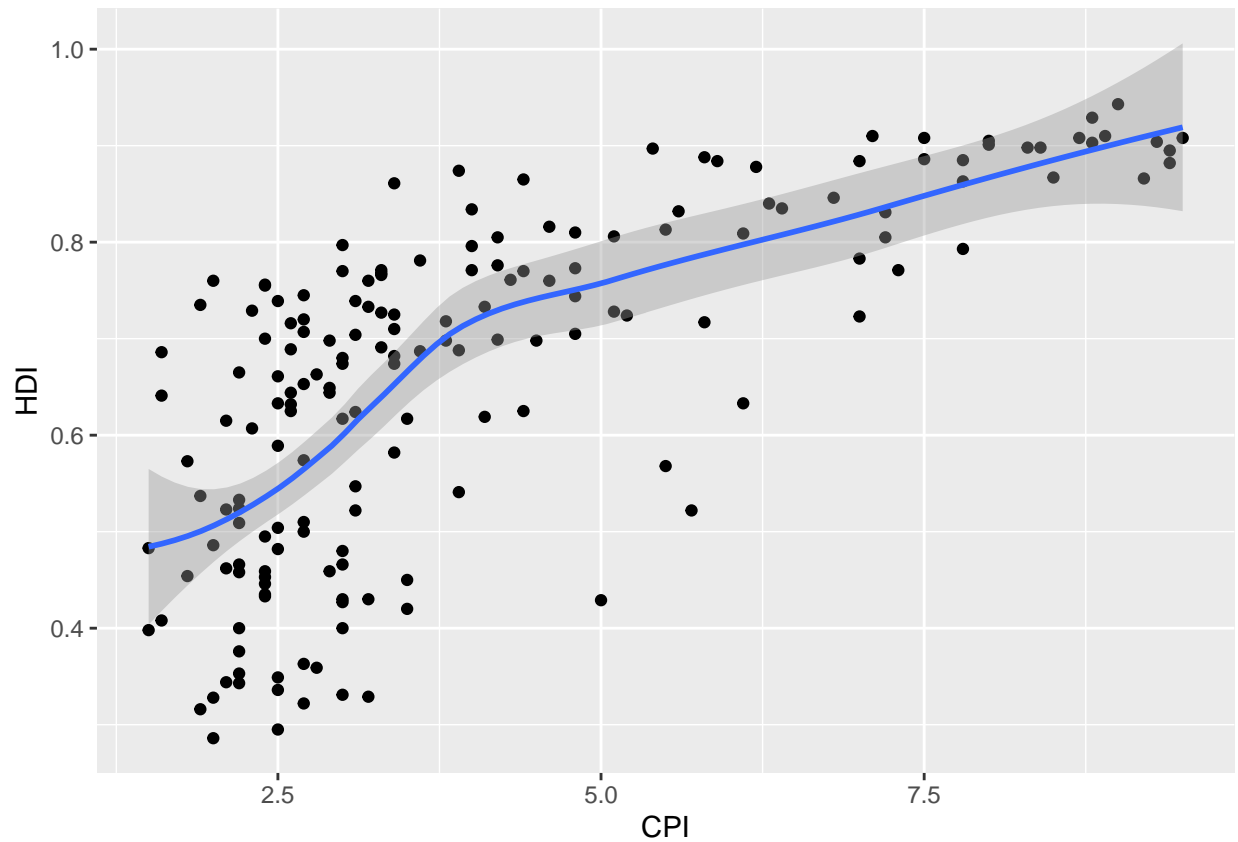
```
ggplot(data = tb, aes(x = CPI, y = HDI)) +
  geom_point()
```

(2) Overlay a smoothing line on top of the scatter plot using `geom_smooth`.

```
ggplot(data = tb, aes(x = CPI, y = HDI)) +
  geom_point() +
  geom_smooth()
```
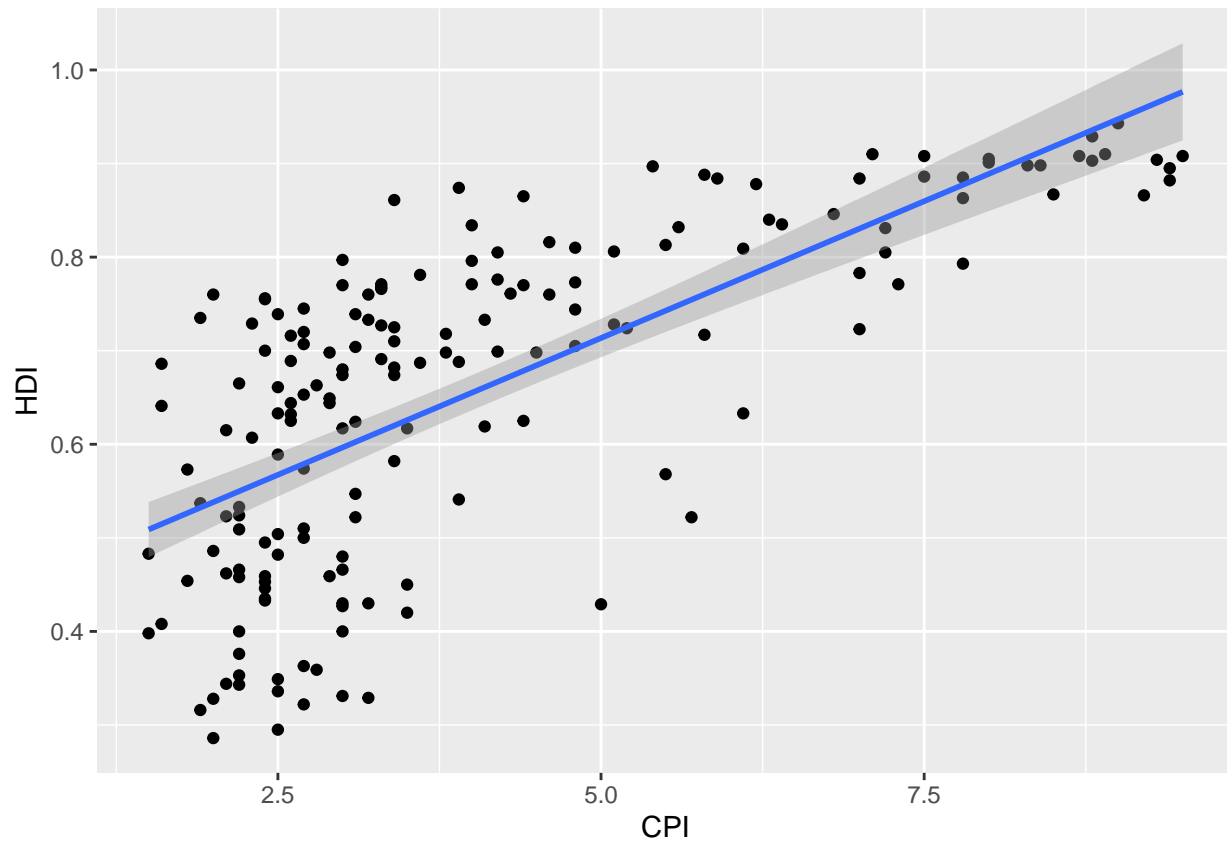
`geom_smooth()` using method = 'loess' and formula 'y ~ x'

(3) Overlay a smoothing line on top of the scatter plot using `geom_smooth`,

but use a linear model for the predictions. Hint: see `?stat_smooth`.

```
ggplot(data = tb, aes(x = CPI, y = HDI)) +
  geom_point() +
  geom_smooth(method = lm)
```
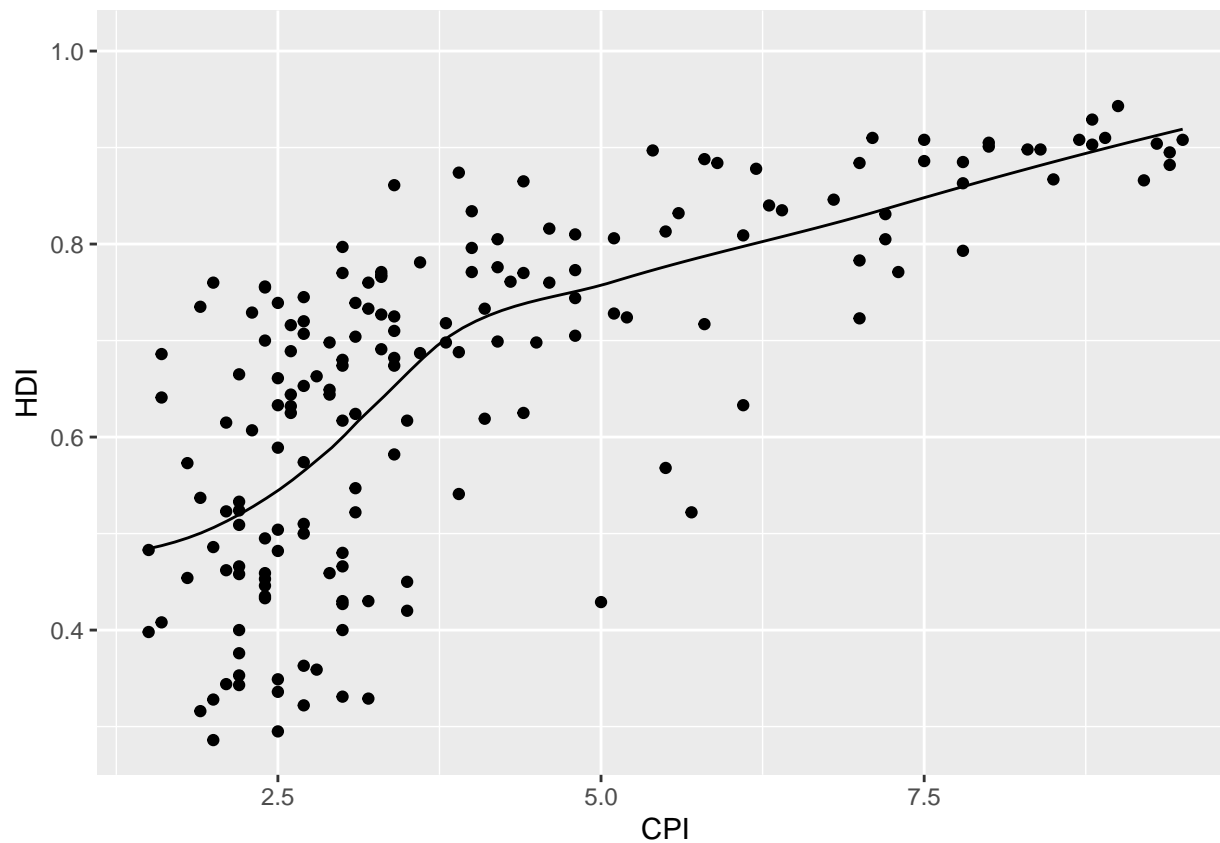
`geom_smooth()` using formula 'y ~ x'

(4) Overlay a smoothing line on top of the scatter plot using geom_line. Hint: change the statistical transformation.

```
ggplot(data = tb, aes(x = CPI, y = HDI)) +
  geom_point() +
  geom_line(stat = "smooth")
```
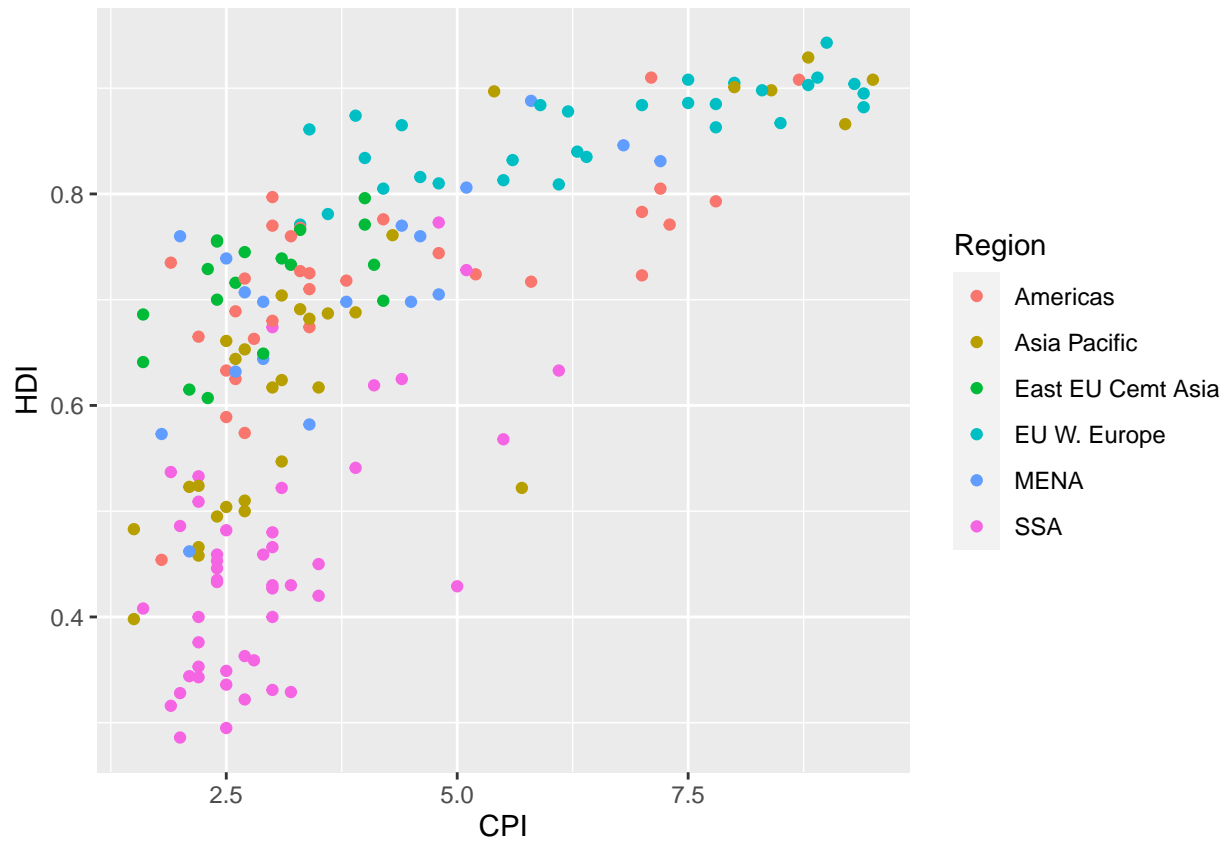
`geom_smooth()` using method = 'loess' and formula 'y ~ x'

Exercise III

(1) Create a scatter plot with CPI on the x axis and HDI on the y axis. Color the points to indicate region.

```
ggplot(data = tb, aes(x = CPI, y = HDI, color = Region)) +
  geom_point()
```

(2) Modify the x, y, and color scales so that they have more easily-understood names (e.g., spell out "Human development Index" instead of "HDI").

```
ggplot(data = tb, aes(x = CPI, y = HDI, color = Region)) +
  geom_point() +
  ylab("Human development index") +
  xlab("Consumer price index")
```
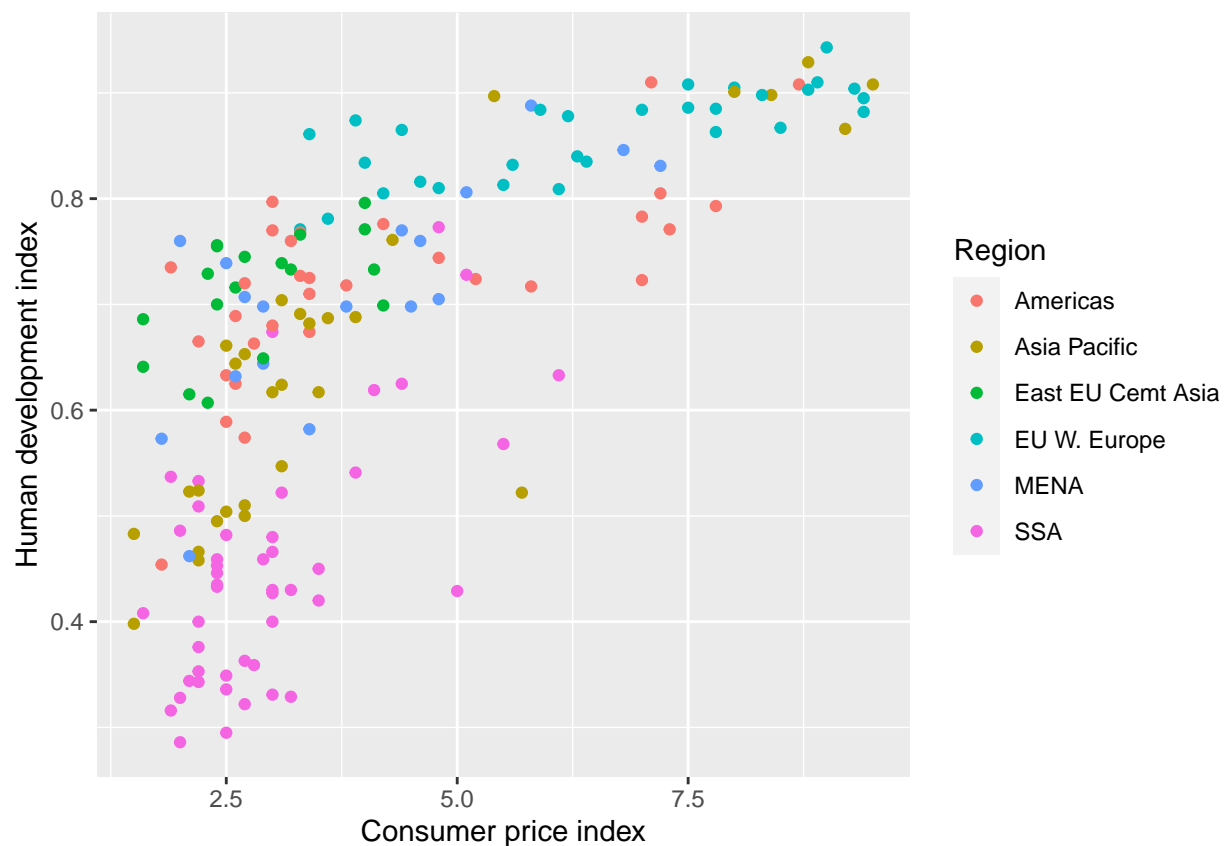
(3) Modify the color scale to use specific values of your choosing. Hint: see `?scale_color_manual`.

```
ggplot(data = tb, aes(x = CPI, y = HDI, color = Region)) +
  geom_point() +
  scale_color_manual(name = "Region",
                     values = c("Americas" = "#db006e",
                                "Asia Pacific" = "#9a3fa1",
                                "East EU Cemt Asia" = "#3a3da8",
                                "EU W. Europe" = "#e836a1",
                                "MENA" = "#ffef07",
                                "SSA" = "#0496df")) +
  ylab("Human development index") +
  xlab("Consumer price index")
```
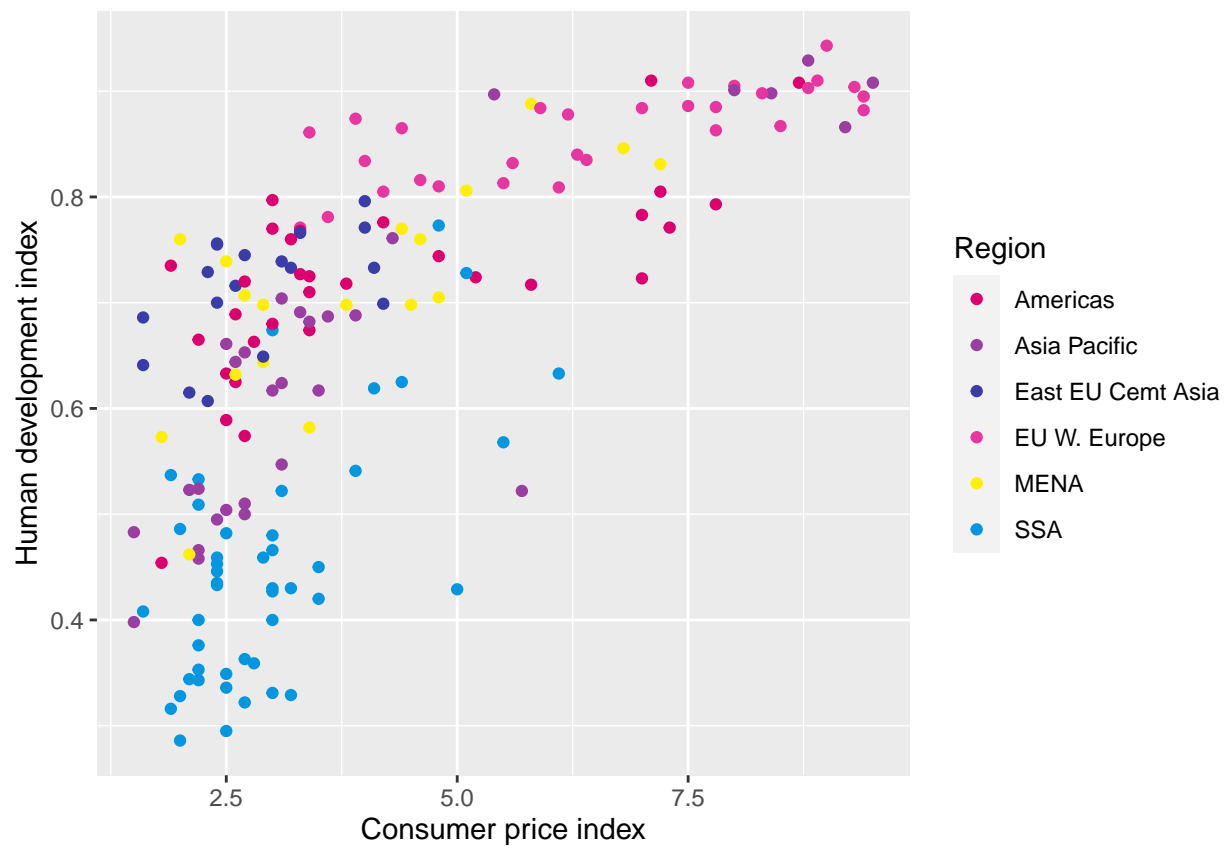
References

- Bi color-scheme
- Pan color scheme