

Make a Plot

Kieran Healy

10-January-2020

Load Libraries

Tidy Data

```
# import Stata file by use `haven` package
findata <- read_dta(file = "Data/micro_world.dta")

class(findata)

## [1] "tbl_df"      "tbl"        "data.frame"

show(findata)

## # A tibble: 154,923 x 105
##   economy  economycode regionwb  pop_adult wpid_random  wgt  female  age
##   <chr>      <chr>      <chr>      <dbl>      <dbl> <dbl> <dbl+lbl> <dbl+lbl>
## 1 Afghanis~ AFG      South As~  19456060  139880545 1.94  1 [Male]  18 [18]
## 2 Afghanis~ AFG      South As~  19456060  111705622 0.333 2 [Femal~ 25 [25]
## 3 Afghanis~ AFG      South As~  19456060  128866217 1.34  2 [Femal~ 40 [40]
## 4 Afghanis~ AFG      South As~  19456060  138171768 0.276 2 [Femal~ 45 [45]
## 5 Afghanis~ AFG      South As~  19456060  125156190 0.629 1 [Male]  30 [30]
## 6 Afghanis~ AFG      South As~  19456060  209050144 0.773 1 [Male]  53 [53]
## 7 Afghanis~ AFG      South As~  19456060  167258084 0.482 1 [Male]  36 [36]
## 8 Afghanis~ AFG      South As~  19456060  147754514 0.565 1 [Male]  29 [29]
## 9 Afghanis~ AFG      South As~  19456060  171367518 1.51  1 [Male]  21 [21]
## 10 Afghanis~ AFG      South As~  19456060  145691120 0.676 1 [Male]  75 [75]
## # ... with 154,913 more rows, and 97 more variables: educ <dbl+lbl>,
## #   inc_q <dbl+lbl>, emp_in <dbl+lbl>, fin2 <dbl+lbl>, fin3 <dbl+lbl>,
## #   fin4 <dbl+lbl>, fin5 <dbl+lbl>, fin6 <dbl+lbl>, fin7 <dbl+lbl>,
## #   fin8 <dbl+lbl>, fin9 <dbl+lbl>, fin10 <dbl+lbl>, fin11a <dbl+lbl>,
## #   fin11b <dbl+lbl>, fin11c <dbl+lbl>, fin11d <dbl+lbl>, fin11e <dbl+lbl>,
## #   fin11f <dbl+lbl>, fin11g <dbl+lbl>, fin11h <dbl+lbl>, fin14a <dbl+lbl>,
## #   fin14b <dbl+lbl>, fin14c <dbl+lbl>, fin15 <dbl+lbl>, fin16 <dbl+lbl>,
## #   fin17a <dbl+lbl>, fin17b <dbl+lbl>, fin19 <dbl+lbl>, fin20 <dbl+lbl>,
## #   fin21 <dbl+lbl>, fin22a <dbl+lbl>, fin22b <dbl+lbl>, fin22c <dbl+lbl>,
## #   fin24 <dbl+lbl>, fin25 <dbl+lbl>, fin26 <dbl+lbl>, fin27a <dbl+lbl>,
## #   fin27b <dbl+lbl>, fin27c1 <dbl+lbl>, fin27c2 <dbl+lbl>, fin28 <dbl+lbl>,
## #   fin29a <dbl+lbl>, fin29b <dbl+lbl>, fin29c1 <dbl+lbl>, fin29c2 <dbl+lbl>,
## #   fin30 <dbl+lbl>, fin31a <dbl+lbl>, fin31b <dbl+lbl>, fin31c <dbl+lbl>,
## #   fin32 <dbl+lbl>, fin33 <dbl+lbl>, fin34a <dbl+lbl>, fin34b <dbl+lbl>,
## #   fin34c1 <dbl+lbl>, fin34c2 <dbl+lbl>, fin35 <dbl+lbl>, fin36 <dbl+lbl>,
## #   fin37 <dbl+lbl>, fin38 <dbl+lbl>, fin39a <dbl+lbl>, fin39b <dbl+lbl>,
## #   fin39c1 <dbl+lbl>, fin39c2 <dbl+lbl>, fin40 <dbl+lbl>, fin41 <dbl+lbl>,
## #   fin42 <dbl+lbl>, fin43a <dbl+lbl>, fin43b <dbl+lbl>, fin43c1 <dbl+lbl>,
```

```
## # fin43c2 <dbl+lbl>, fin44 <dbl+lbl>, fin45 <dbl+lbl>, fin46 <dbl+lbl>,
## # fin47a <dbl+lbl>, fin47b <dbl+lbl>, fin47c1 <dbl+lbl>, fin47c2 <dbl+lbl>,
## # fin47c3 <dbl+lbl>, fin47c4 <dbl+lbl>, fin47c5 <dbl+lbl>,
## # mobileowner <dbl+lbl>, fin48 <dbl+lbl>, account_fin <dbl+lbl>,
## # account_mob <dbl+lbl>, account <dbl+lbl>, saved <dbl+lbl>,
## # borrowed <dbl+lbl>, receive_wages <dbl+lbl>, receive_transfers <dbl>,
## # receive_pension <dbl>, receive_agriculture <dbl>, pay_utilities <dbl>,
## # remittances <dbl>, pay_onlne <dbl+lbl>, pay_onlne_mobintbuy <dbl+lbl>,
## # pay_cash <dbl+lbl>, pay_cash_mobintbuy <dbl+lbl>
```

```
get_label(x = findata)
```

```
## economy
## "Economy"
## economycode
## "Economy Code"
## regionwb
## "Region, WB"
## pop_adult
## "Population 15+, 2016, WDI"
## wpid_random
## "Gallup World Poll identifier"
## wgt
## "Weight"
## female
## "Respondent is female"
## age
## "Respondent age"
## educ
## "Respondent education level"
## inc_q
## "Within-economy household income quintile"
## emp_in
## "Respondent is in the workforce"
## fin2
## "Has a debit card"
## fin3
## "If has debit card: card in own name"
## fin4
## "If has debit card: used card in past 12 months"
## fin5
## "Used mobile phone or internet to access FI account"
## fin6
## "Used mobile phone or internet to check account balance"
## fin7
## "Has a credit card"
## fin8
## "If has credit card: used card in past 12 months"
## fin9
## "If has account: any deposit into account in past 12 months"
## fin10
## "If has account: any withdrawal from account in past 12 months"
## fin11a
## "If does not have account: b/c too far away"
## fin11b
```

```

##             "If does not have account: b/c too expensive"
##                                     fin11c
##             "If does not have account: b/c lack documentation"
##                                     fin11d
##             "If does not have account: b/c lack trust"
##                                     fin11e
##             "If does not have account: b/c religious reasons"
##                                     fin11f
##             "If does not have account: b/c lack of money"
##                                     fin11g
##             "If does not have account: b/c family member already has one"
##                                     fin11h
##             "If does not have account: b/c no need for financial services"
##                                     fin14a
##             "Made bill payments online using the Internet"
##                                     fin14b
##             "Bought something online using the Internet"
##                                     fin14c
##             "Paid online or in cash at delivery"
##                                     fin15
##             "Saved in past 12 months: for farm/business purposes"
##                                     fin16
##             "Saved in past 12 months: for old age"
##                                     fin17a
##             "Saved in past 12 months: using an account at a financial institution"
##                                     fin17b
##             "Saved in past 12 months: using an informal savings club"
##                                     fin19
##             "Has loan from a financial institution for home, apartment, or land"
##                                     fin20
##             "Borrowed in past 12 months: for medical purposes"
##                                     fin21
##             "Borrowed in past 12 months: for farm/business purposes"
##                                     fin22a
##             "Borrowed in past 12 months: from a financial institution"
##                                     fin22b
##             "Borrowed in past 12 months: from family or friends"
##                                     fin22c
##             "Borrowed in past 12 months: from an informal savings club"
##                                     fin24
##             "Possibility of coming up with emergency funds"
##                                     fin25
##             "Main source of emergency funds"
##                                     fin26
##             "Sent domestic remittances in past 12 months"
##                                     fin27a
##             "If sent domestic remittances: through a financial institution"
##                                     fin27b
##             "If sent domestic remittances: through a mobile phone"
##                                     fin27c1
##             "If sent domestic remittances: in cash"
##                                     fin27c2
##             "If sent domestic remittances: through an MTO"
##                                     fin28

```

```

##          "Received domestic remittances in past 12 months"
##                                     fin29a
## "If received domestic remittances: through a financial institution"
##                                     fin29b
##          "If received domestic remittances: through a mobile phone"
##                                     fin29c1
##          "If received domestic remittances: in cash"
##                                     fin29c2
##          "If received domestic remittances: through an MT0"
##                                     fin30
##          "Paid utility bills in past 12 months"
##                                     fin31a
##          "If paid utility bills: using an account"
##                                     fin31b
##          "If paid utility bills: through a mobile phone"
##                                     fin31c
##          "If paid utility bills: in cash"
##                                     fin32
##          "Received wage payments in past 12 months"
##                                     fin33
##          "If received wage payments: work in public sector"
##                                     fin34a
##          "If received wage payments: into an account"
##                                     fin34b
##          "If received wage payments: through a mobile phone"
##                                     fin34c1
##          "If received wage payments: in cash"
##                                     fin34c2
##          "If received wage payments: to a card"
##                                     fin35
##          "If received cashless wage payments: first account"
##                                     fin36
##          "If received cashless wage payments: opened to receive payments"
##                                     fin37
##          "Received government transfers in past 12 months"
##                                     fin38
##          "Received government transfers in past 12 months: pension"
##                                     fin39a
##          "If received government transfers: into an account"
##                                     fin39b
##          "If received government transfers: through a mobile phone"
##                                     fin39c1
##          "If received government transfers: in cash"
##                                     fin39c2
##          "If received government transfers: to a card"
##                                     fin40
##          "If received cashless government transfers: first account"
##                                     fin41
##          "If received cashless government transfers: opened to receive payments"
##                                     fin42
##          "Received agricultural payments in past 12 months"
##                                     fin43a
##          "If received agricultural payments: into an account"
##                                     fin43b

```

```

##          "If received agricultural payments: through a mobile phone"
##                                     fin43c1
##          "If received agricultural payments: in cash"
##                                     fin43c2
##          "If received agricultural payments: to a card"
##                                     fin44
## "If received agricultural payments: opened to receive payments"
##                                     fin45
## "If received agricultural payments: opened to receive payments"
##                                     fin46
##          "Received self-employment payments: in past 12 months"
##                                     fin47a
##          "If received self-employment payments: into an account"
##                                     fin47b
## "If received self-employment payments: through a mobile phone"
##                                     fin47c1
##          "If received self-employment payments: in cash"
##                                     fin47c2
##          "If received self-employment payments: to a card"
##                                     fin47c3
##          "Marginal: Used debit card in your own name"
##                                     fin47c4
## "Marginal: Used mobile phone or internet to access FI account"
##                                     fin47c5
## "Marginal: Used mobile phone or internet to check account balance"
##                                     mobileowner
##                                     "Owns a mobile phone"
##                                     fin48
##                                     "Has a national ID"
##                                     account_fin
##          "Has an account at a financial institution"
##                                     account_mob
##          "Has a mobile money account"
##                                     account
##          "Has an account"
##                                     saved
##          "Saved in the past year"
##                                     borrowed
##          "Borrowed in the past year"
##                                     receive_wages
##          "Payments: wage payments"
##                                     receive_transfers
##          "Payments: government transfers"
##                                     receive_pension
##          "Payments: government pension"
##                                     receive_agriculture
##          "Payments: selling agricultural goods"
##                                     pay_utilities
##          "Payments: utility bills"
##                                     remittances
##          "Domestic remittances"
##                                     pay_onlne
##          "Pay online [of adults]- F2F"
##                                     pay_onlne_mobintbuy

```

```
##                                "Pay online [of buyers]- F2F"
##                                pay_cash
##                                "Pay in cash when order is delivered [of adults]- F2F"
##                                pay_cash_mobintbuy
##                                "Pay in cash when order is delivered [of buyers]- F2F"
```

Mappings link data to things you see

```
gapminder
```

```
## # A tibble: 1,704 x 6
##   country    continent  year lifeExp      pop gdpPercap
##   <fct>      <fct>    <int>  <dbl>    <int>    <dbl>
## 1 Afghanistan Asia      1952   28.8  8425333    779.
## 2 Afghanistan Asia      1957   30.3  9240934    821.
## 3 Afghanistan Asia      1962   32.0 10267083    853.
## 4 Afghanistan Asia      1967   34.0 11537966    836.
## 5 Afghanistan Asia      1972   36.1 13079460    740.
## 6 Afghanistan Asia      1977   38.4 14880372    786.
## 7 Afghanistan Asia      1982   39.9 12881816    978.
## 8 Afghanistan Asia      1987   40.8 13867957    852.
## 9 Afghanistan Asia      1992   41.7 16317921    649.
## 10 Afghanistan Asia      1997   41.8 22227415    635.
## # ... with 1,694 more rows
```

```
p <- ggplot(data = gapminder)
```

```
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y = lifeExp))
```

```
p
```

```
p + geom_point()
```

Build your plots layer by layer

```
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y=lifeExp))
p + geom_smooth()
```

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

```
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y=lifeExp))
p + geom_point() + geom_smooth()
```

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

```
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y=lifeExp))
p + geom_point() + geom_smooth(method = "lm")
```

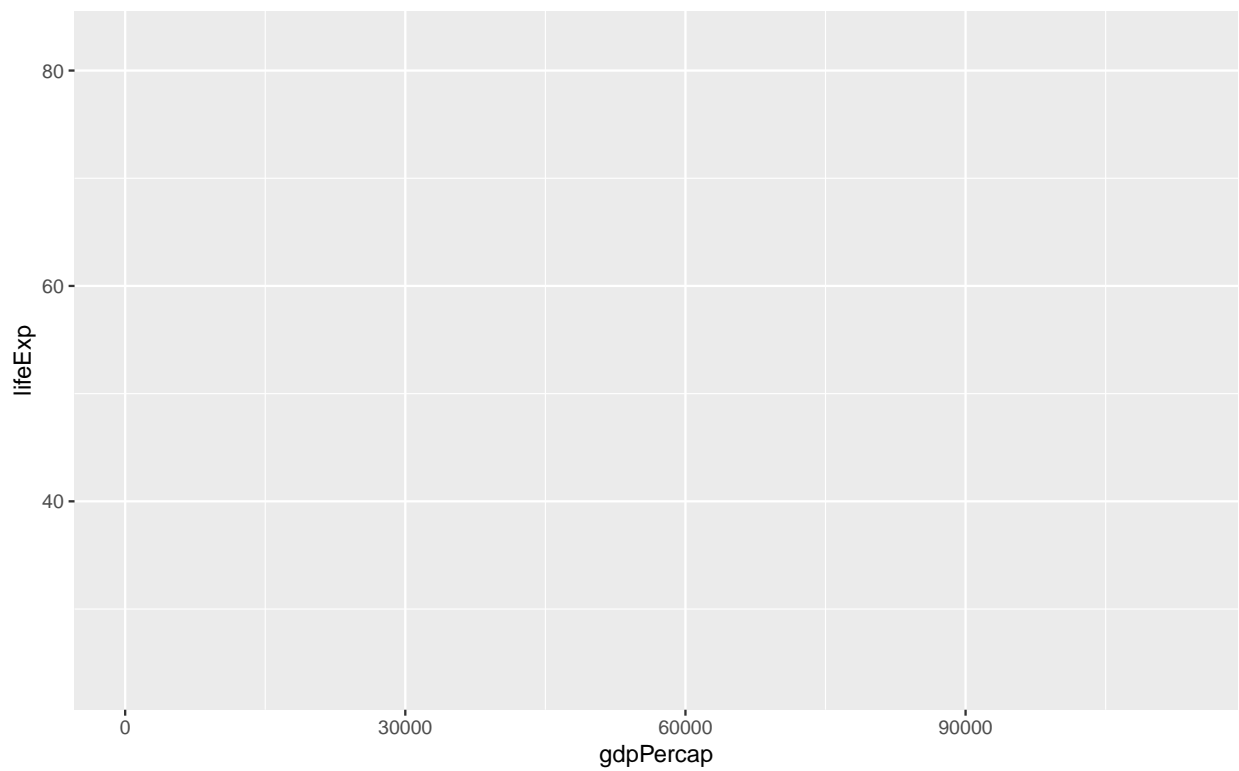


Figure 1: This empty plot has no geoms.

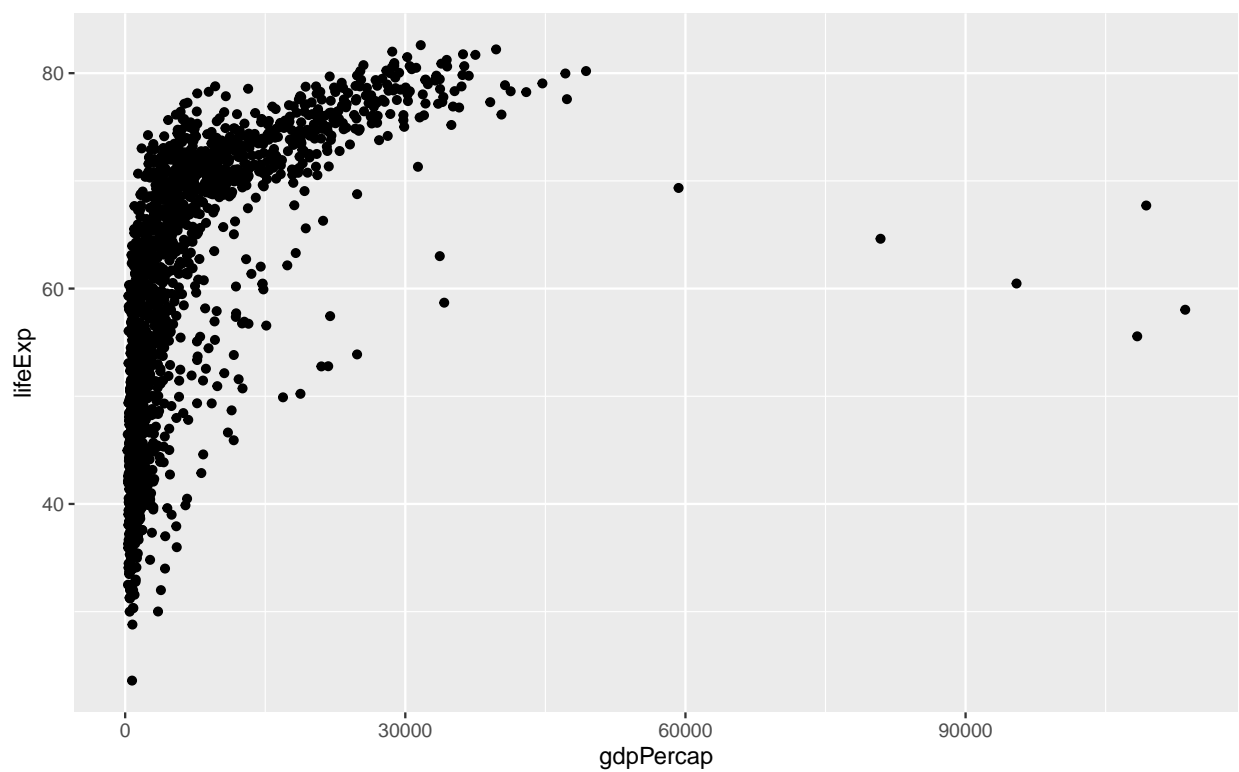


Figure 2: A scatterplot of Life Expectancy vs GDP

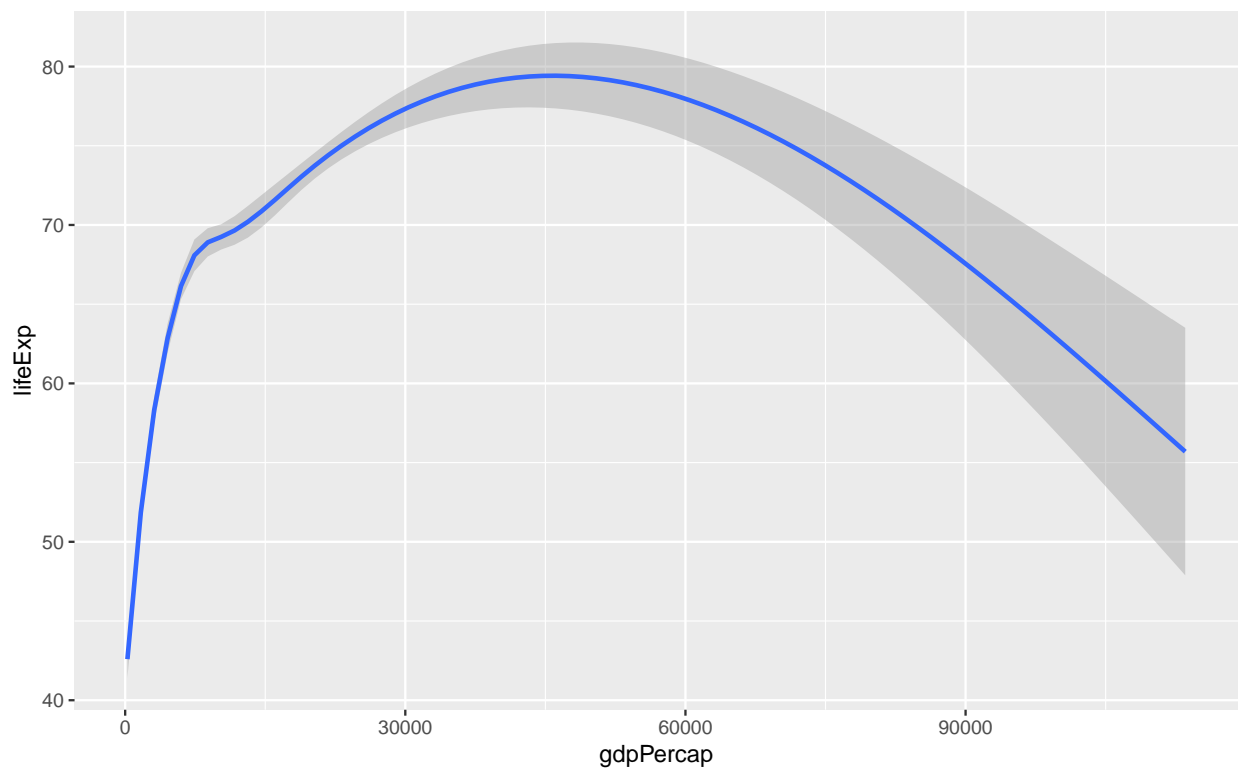


Figure 3: Life Expectancy vs GDP, using a smoother.

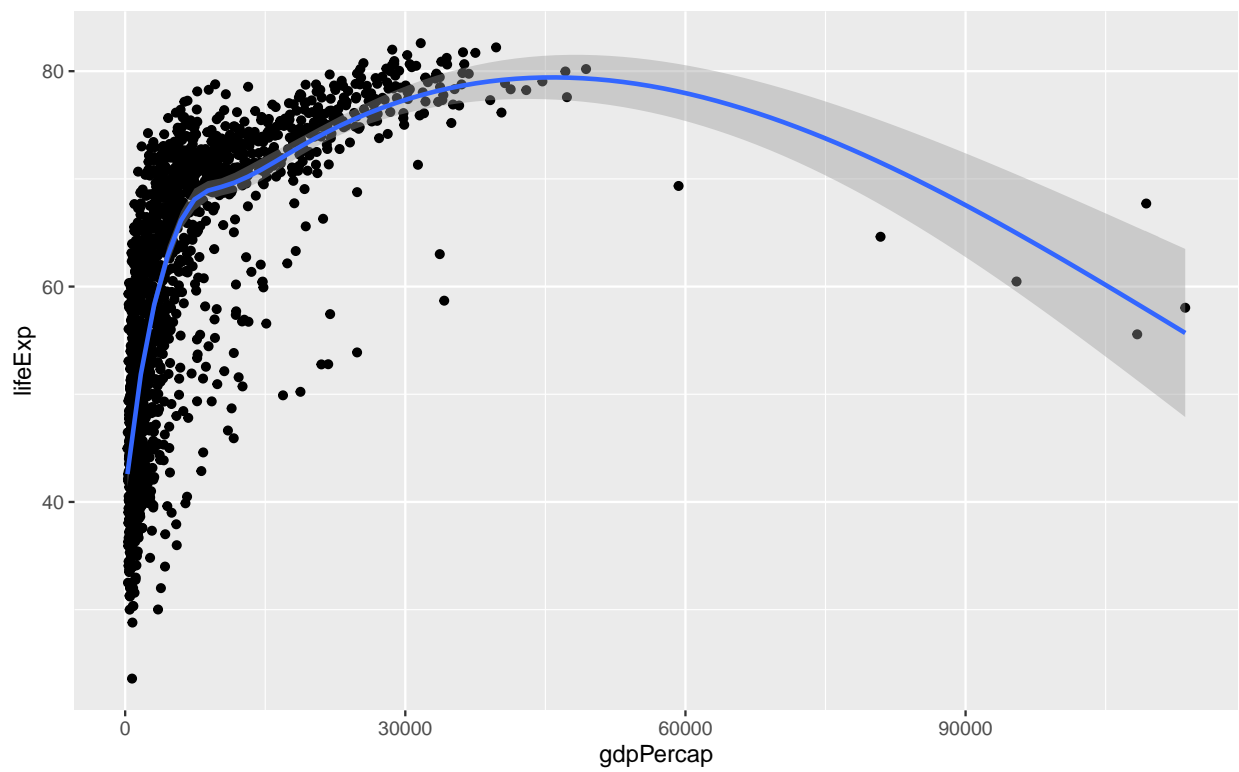


Figure 4: Life Expectancy vs GDP, showing both points and a GAM smoother.


```
## `geom_smooth()` using formula 'y ~ x'
```

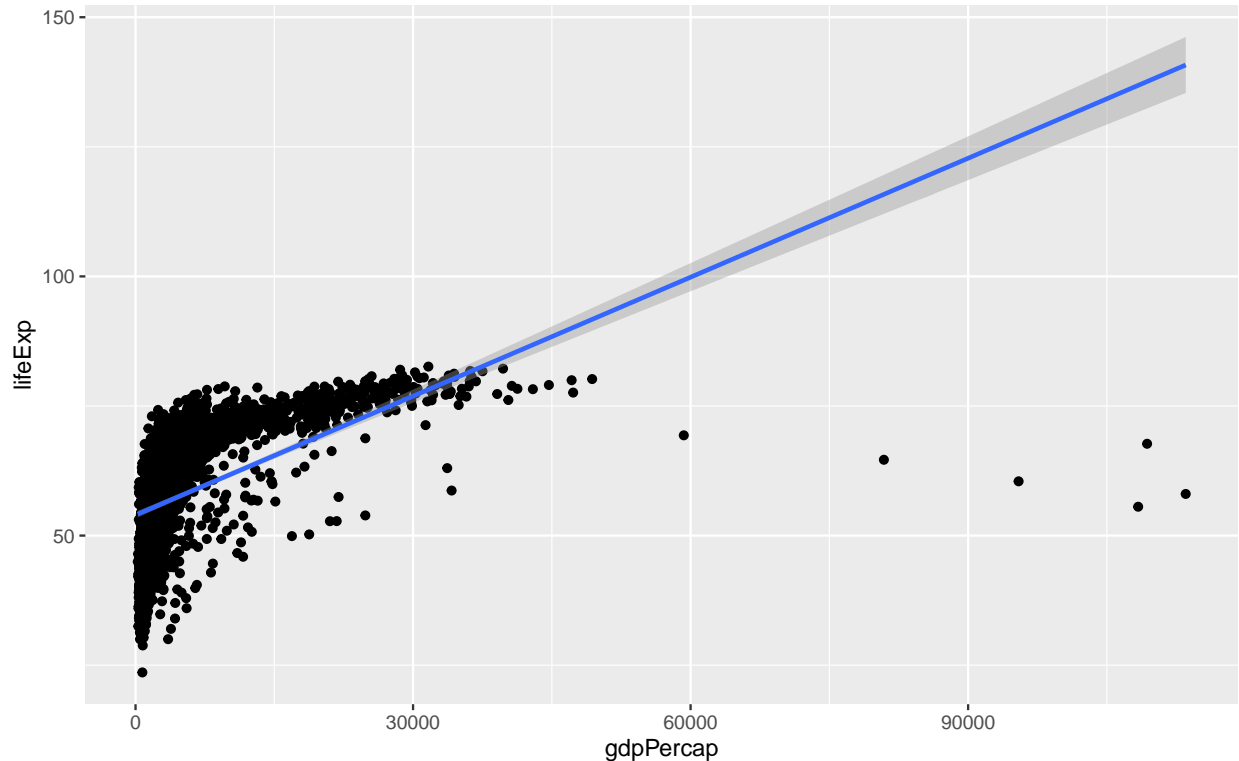


Figure 5: Life Expectancy vs GDP, points and an ill-advised linear fit.

```
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y=lifeExp))
p + geom_point() +
  geom_smooth(method = "gam") +
  scale_x_log10()
```

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

```
p <- ggplot(data = gapminder, mapping = aes(x = gdpPercap, y=lifeExp))
p + geom_point() +
  geom_smooth(method = "gam") +
  scale_x_log10(labels = scales::dollar)
```

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

Mapping aesthetics vs setting them

```
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y = lifeExp,
                          color = "purple"))
p + geom_point() +
  geom_smooth(method = "loess") +
  scale_x_log10()
```

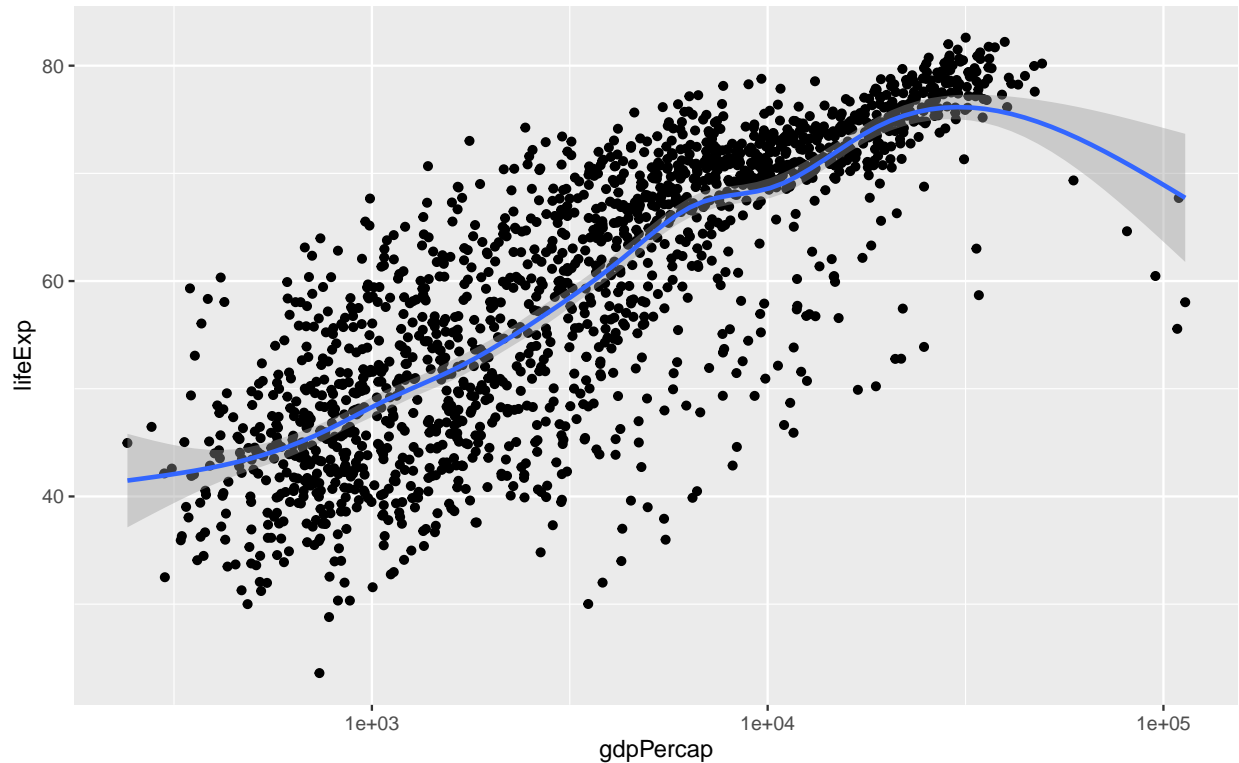


Figure 6: Life Expectancy vs GDP scatterplot, with a GAM smoother and a log scale on the x-axis.

```
## `geom_smooth()` using formula 'y ~ x'
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y = lifeExp))
p + geom_point(color = "purple") +
  geom_smooth(method = "loess") +
  scale_x_log10()

## `geom_smooth()` using formula 'y ~ x'
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y = lifeExp))
p + geom_point(alpha = 0.3) +
  geom_smooth(color = "orange", se = FALSE, size = 8, method = "lm") +
  scale_x_log10()

## `geom_smooth()` using formula 'y ~ x'
p <- ggplot(data = gapminder, mapping = aes(x = gdpPercap, y=lifeExp))
p + geom_point(alpha = 0.3) + geom_smooth(method = "gam") +
  scale_x_log10(labels = scales::dollar) +
  labs(x = "GDP Per Capita", y = "Life Expectancy in Years",
       title = "Economic Growth and Life Expectancy",
       subtitle = "Data points are country-years",
       caption = "Source: Gapminder.")
```

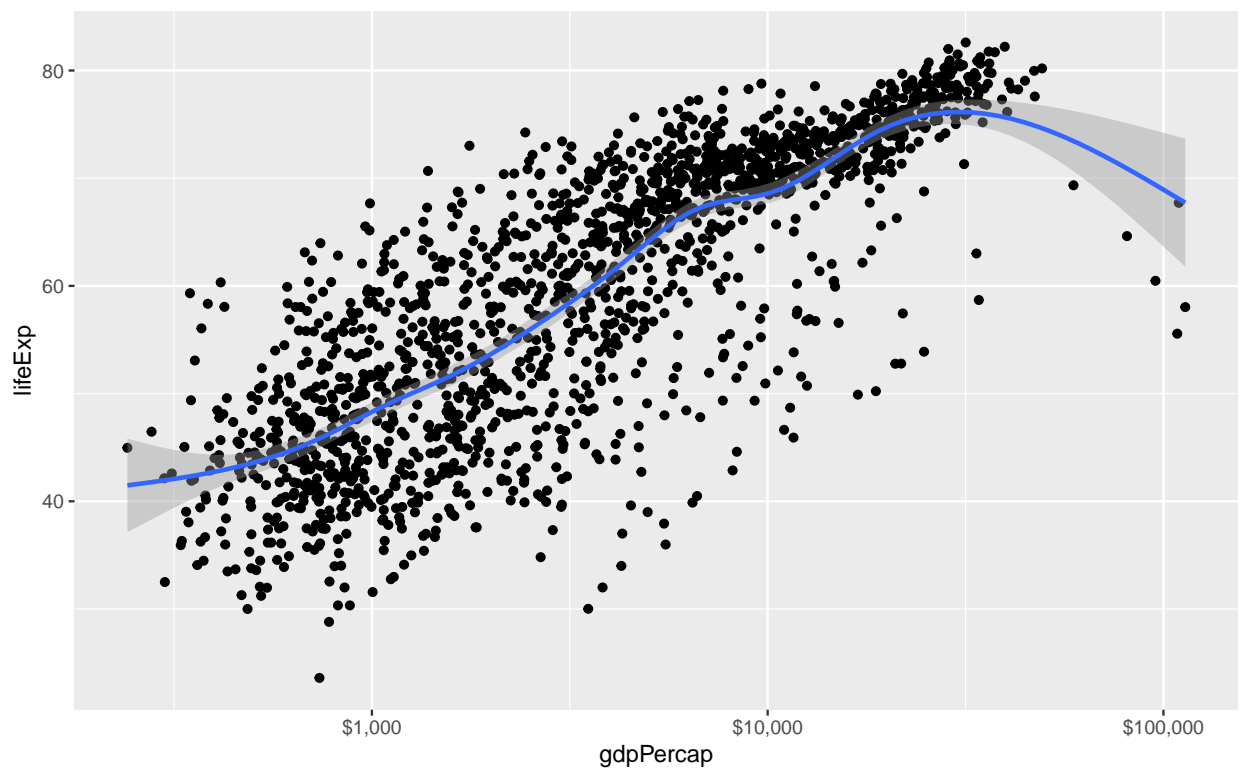


Figure 7: Life Expectancy vs GDP scatterplot, with a GAM smoother and a log scale on the x-axis, with better labels on the tick marks.

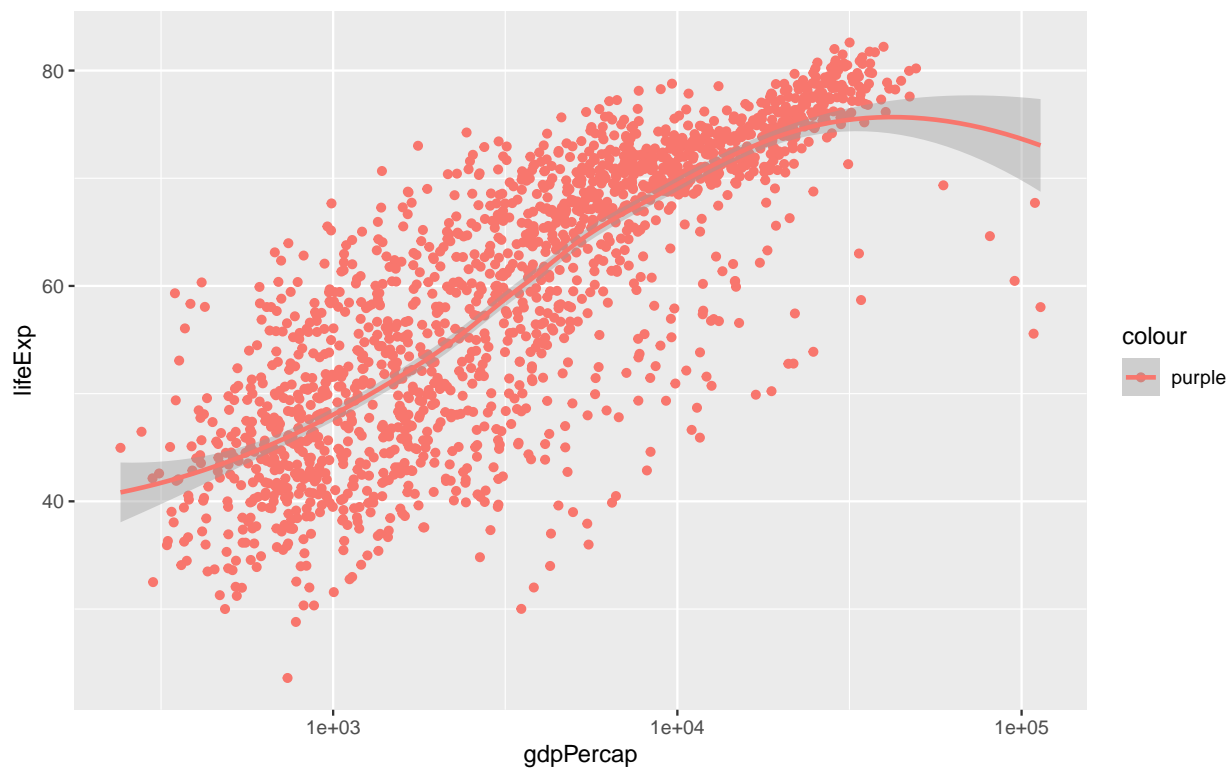


Figure 8: What has gone wrong here?

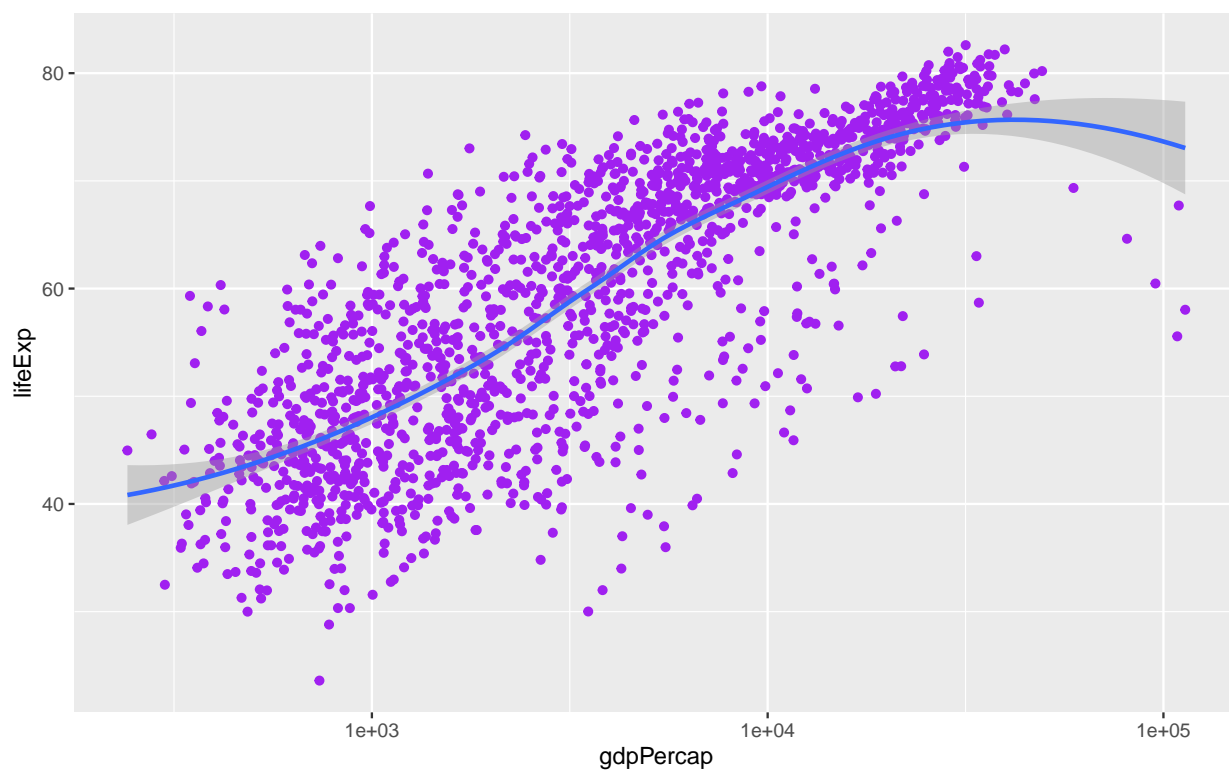


Figure 9: Setting the color attribute of the points directly.

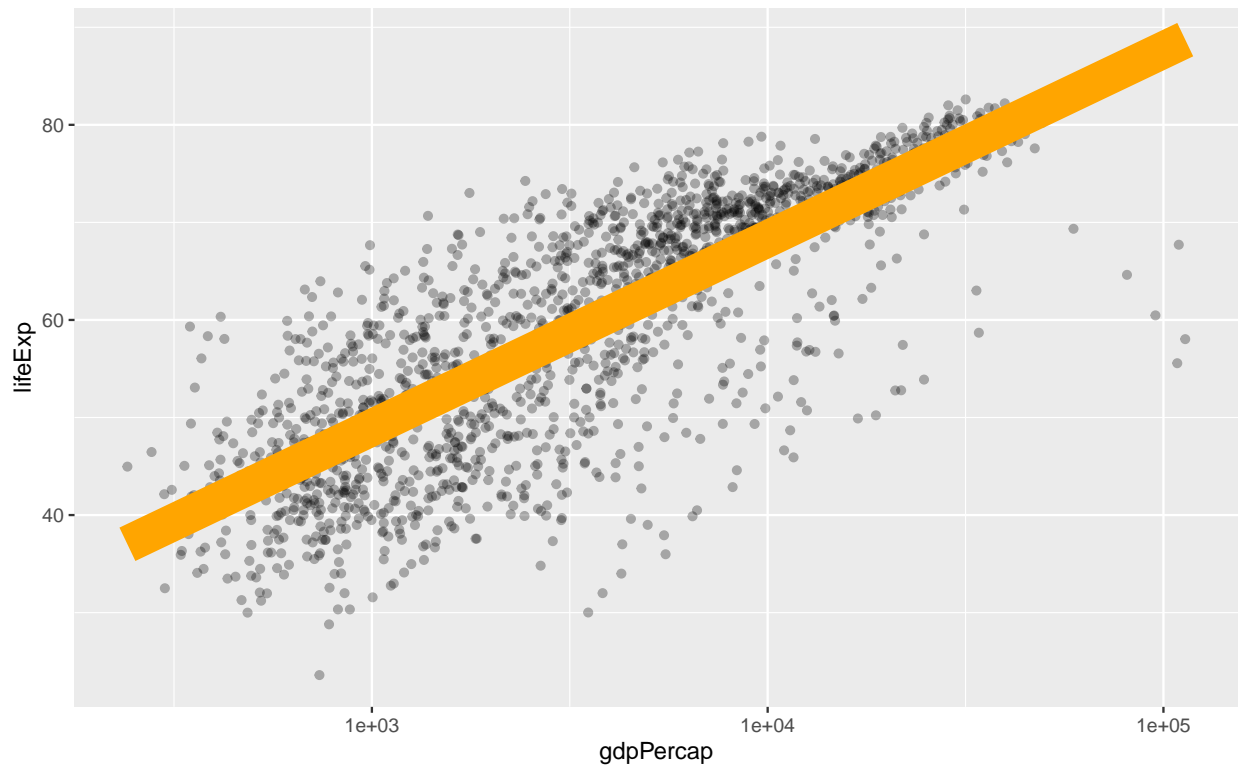


Figure 10: Setting some other arguments.

```
## `geom_smooth()` using formula 'y ~ s(x, bs = "cs")'
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPerCap,
                          y = lifeExp,
                          color = continent))
p + geom_point() +
  geom_smooth(method = "loess") +
  scale_x_log10()
```

```
## `geom_smooth()` using formula 'y ~ x'
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPerCap,
                          y = lifeExp,
                          color = continent,
                          fill = continent))
p + geom_point() +
  geom_smooth(method = "loess") +
  scale_x_log10()
```

```
## `geom_smooth()` using formula 'y ~ x'
```

Aesthetics can be mapped per geom

```
p <- ggplot(data = gapminder, mapping = aes(x = gdpPerCap, y = lifeExp))
p + geom_point(mapping = aes(color = continent)) +
```

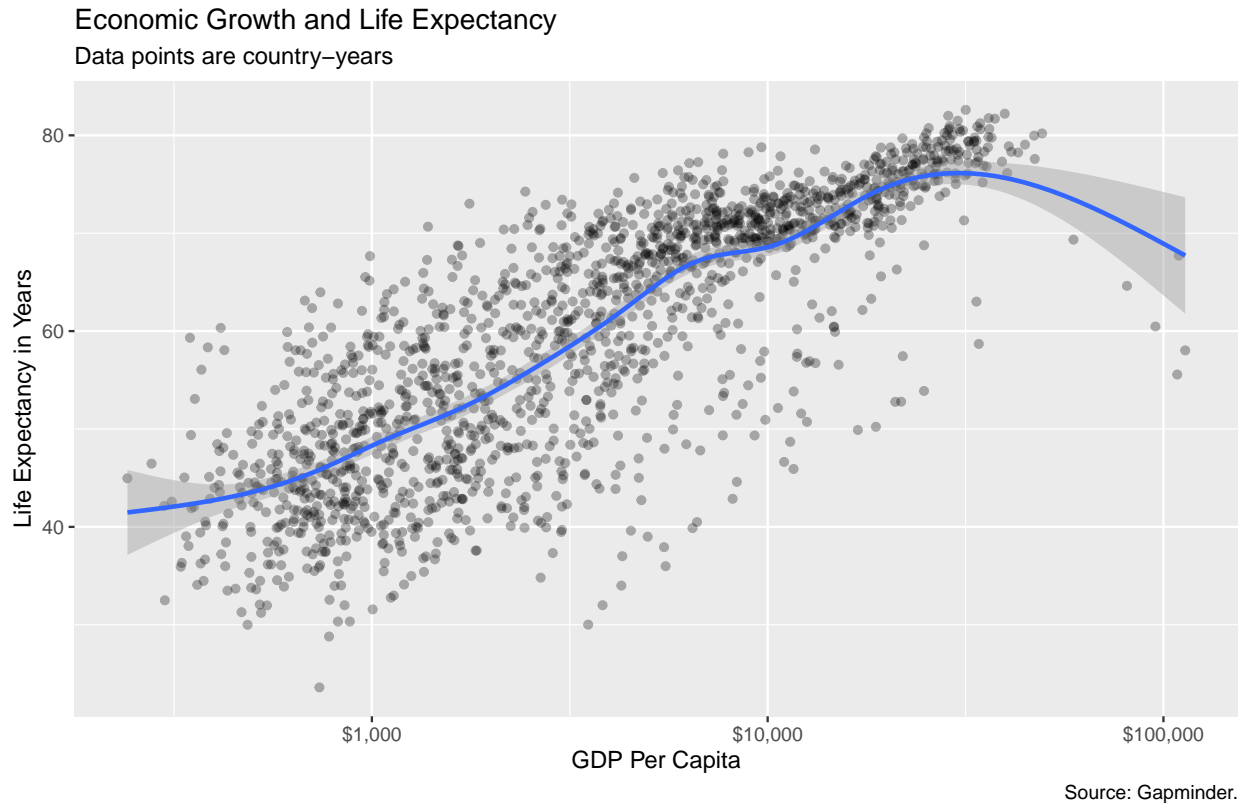


Figure 11: A more polished plot of Life Expectancy vs GDP.

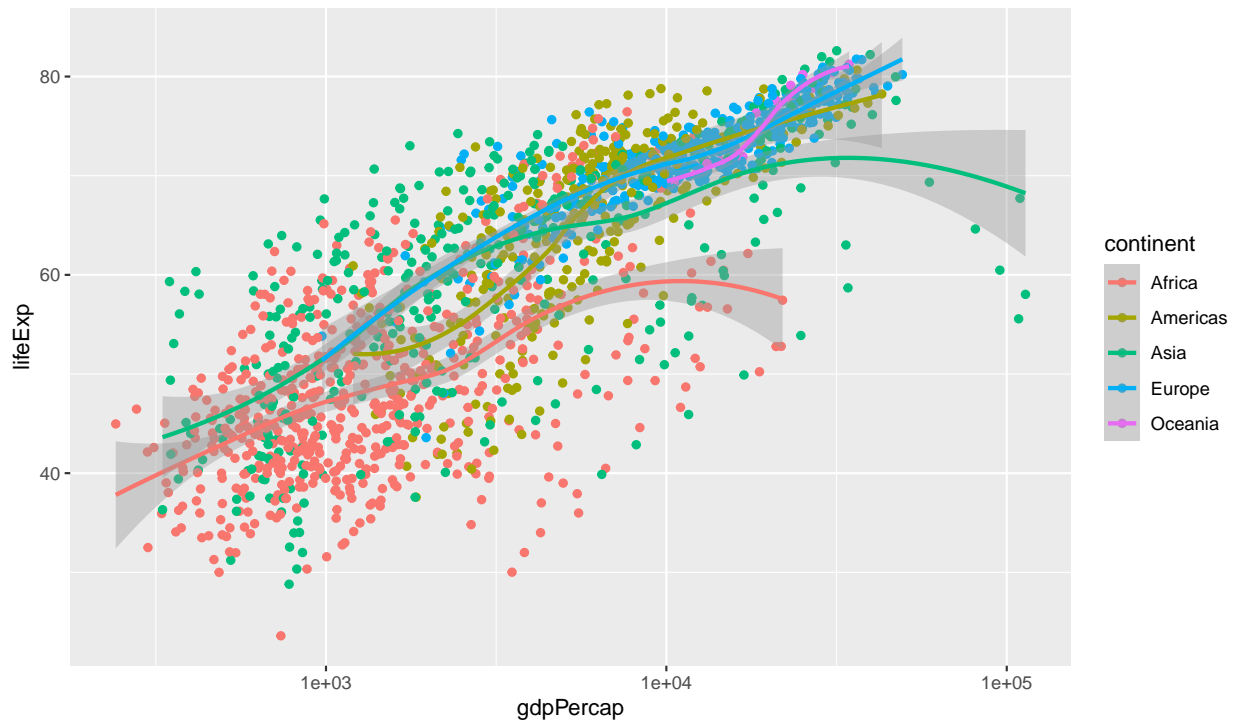


Figure 12: Mapping the continent variable to the color aesthetic.

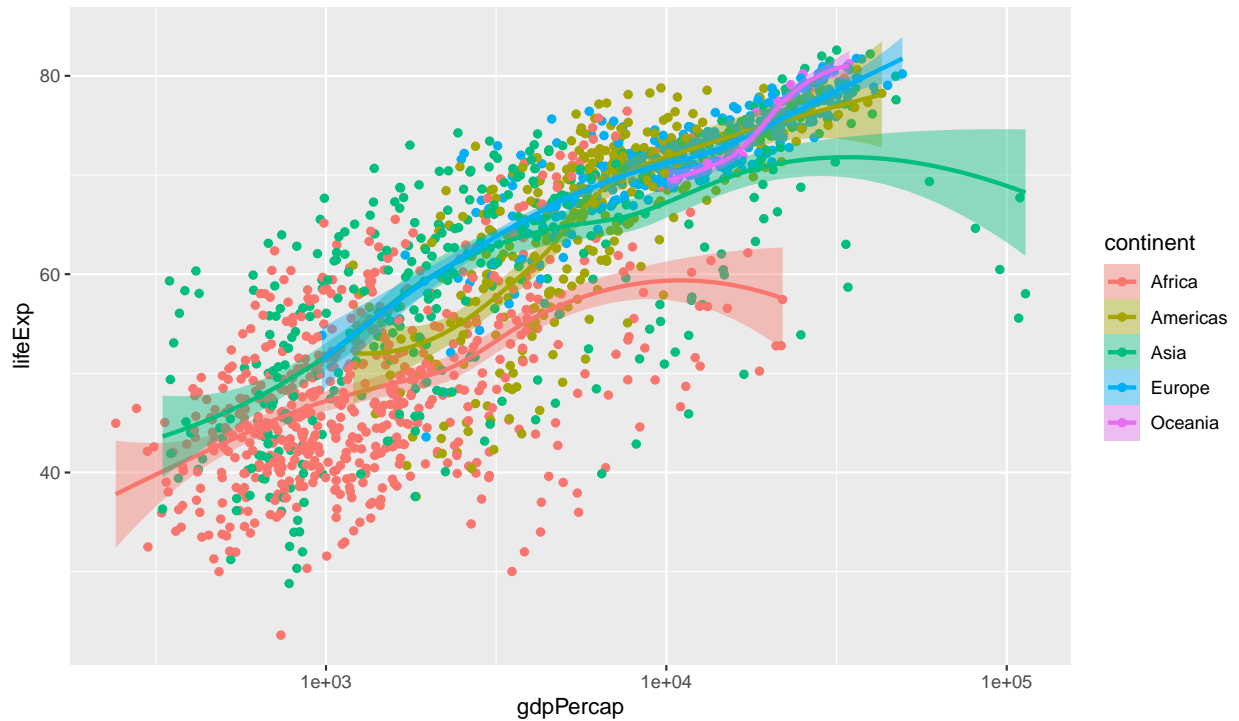


Figure 13: Mapping the continent variable to the color aesthetic, and correcting the error bars using the fill aesthetic.

```
geom_smooth(method = "loess") +
scale_x_log10()

## `geom_smooth()` using formula 'y ~ x'
p <- ggplot(data = gapminder,
            mapping = aes(x = gdpPercap,
                          y = lifeExp))
p + geom_point(mapping = aes(color = log(pop))) +
  scale_x_log10()
```

Save your work

```
knitr::opts_chunk$set(fig.width=8, fig.height=5)

ggsave(filename = "figures/my_figure.png")

here()

## [1] "C:/Users/Nith Kosal/Documents/Kosal Documents/Kosal Teaching/DataVisualization"
## "{r 03-make-a-plot-22}"
#p_out <- p + geom_point(mapping = aes(color = log(pop))) + # scale_x_log10()
#ggsave(here("figures", "lifexp_vs_gdp_gradient.pdf"), plot = p_out)
#ggsave(here("figures", "lifexp_vs_gdp_gradient.png"), plot = p_out) “
```

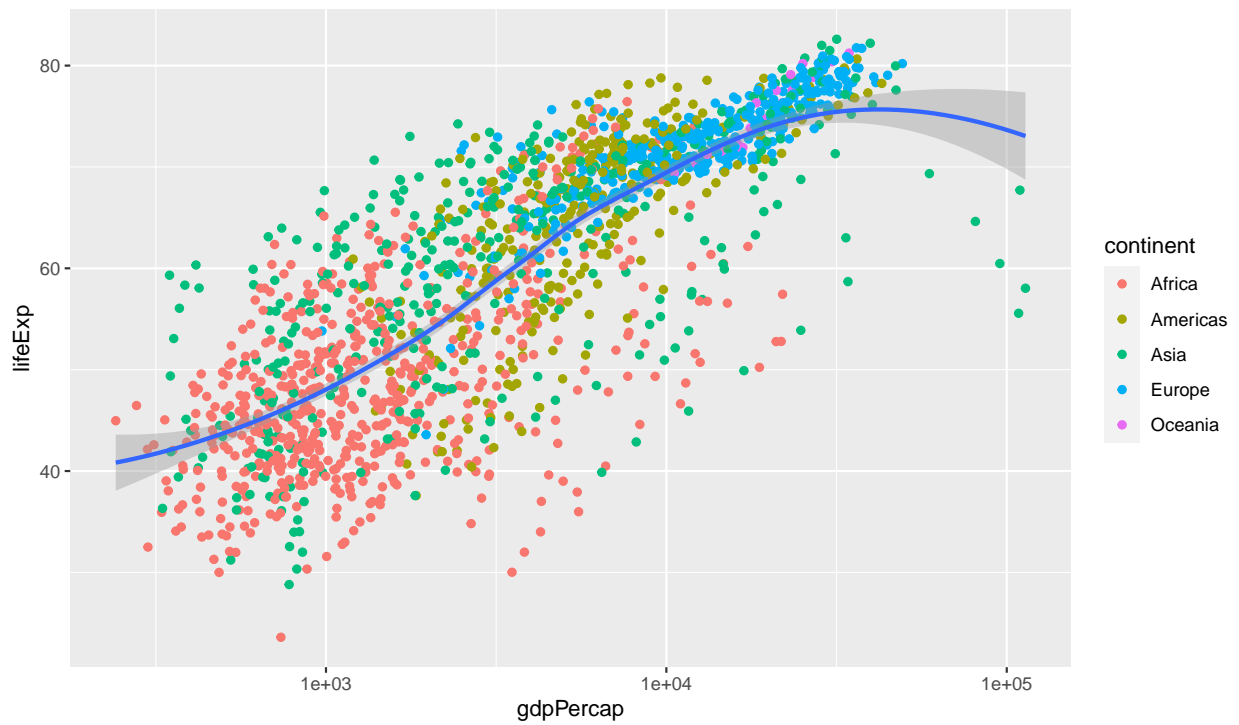


Figure 14: Mapping aesthetics on a per-geom basis. Here color is mapped to continent for the points but not the smoother.

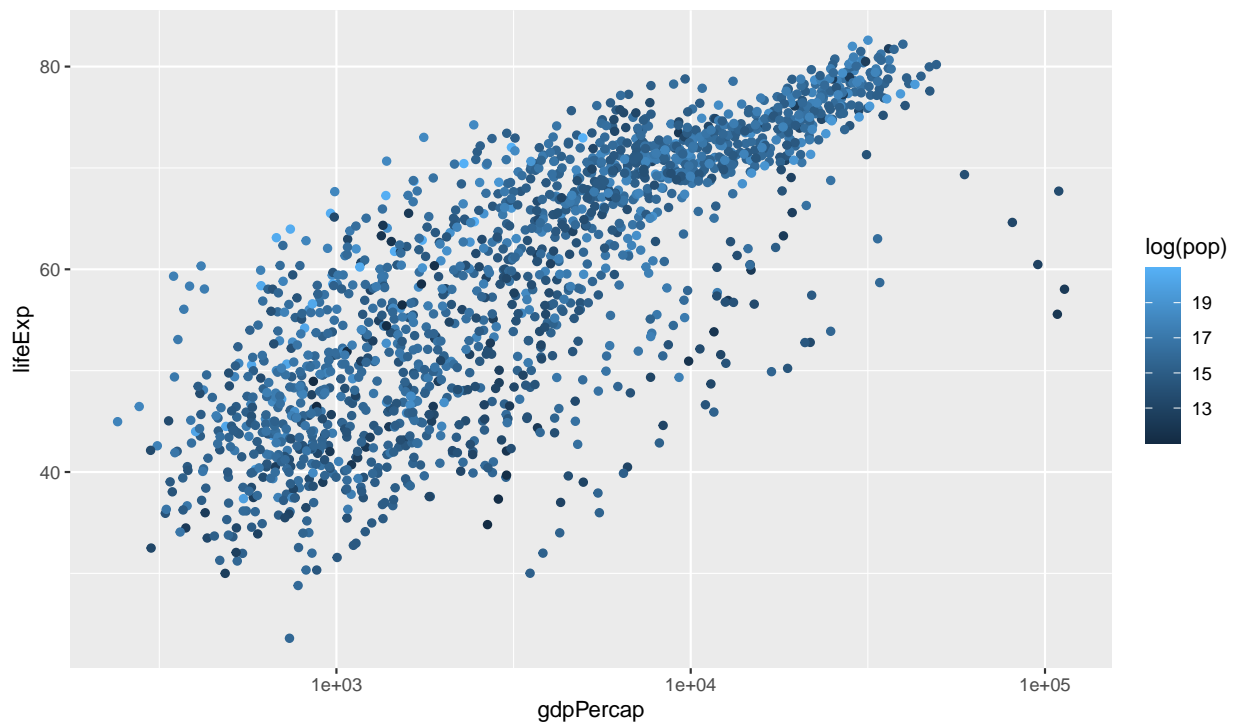


Figure 15: Mapping a continuous variable to color.