

# Class05 Lab Report

Thisha Thiagarajan, PID: A15474979

10/12/2021

```
# Class 05 Data Visualization
#comments for packages that were installed
  #install.packages("ggplot2")
  #install.packages("dplyr")
  #install.packages("gifski")
  #install.packages("patchwork")
  #install.packages("ggridge")
  #install.packages("gapminder") to access ALL the data
  #used url for genes and gapminder

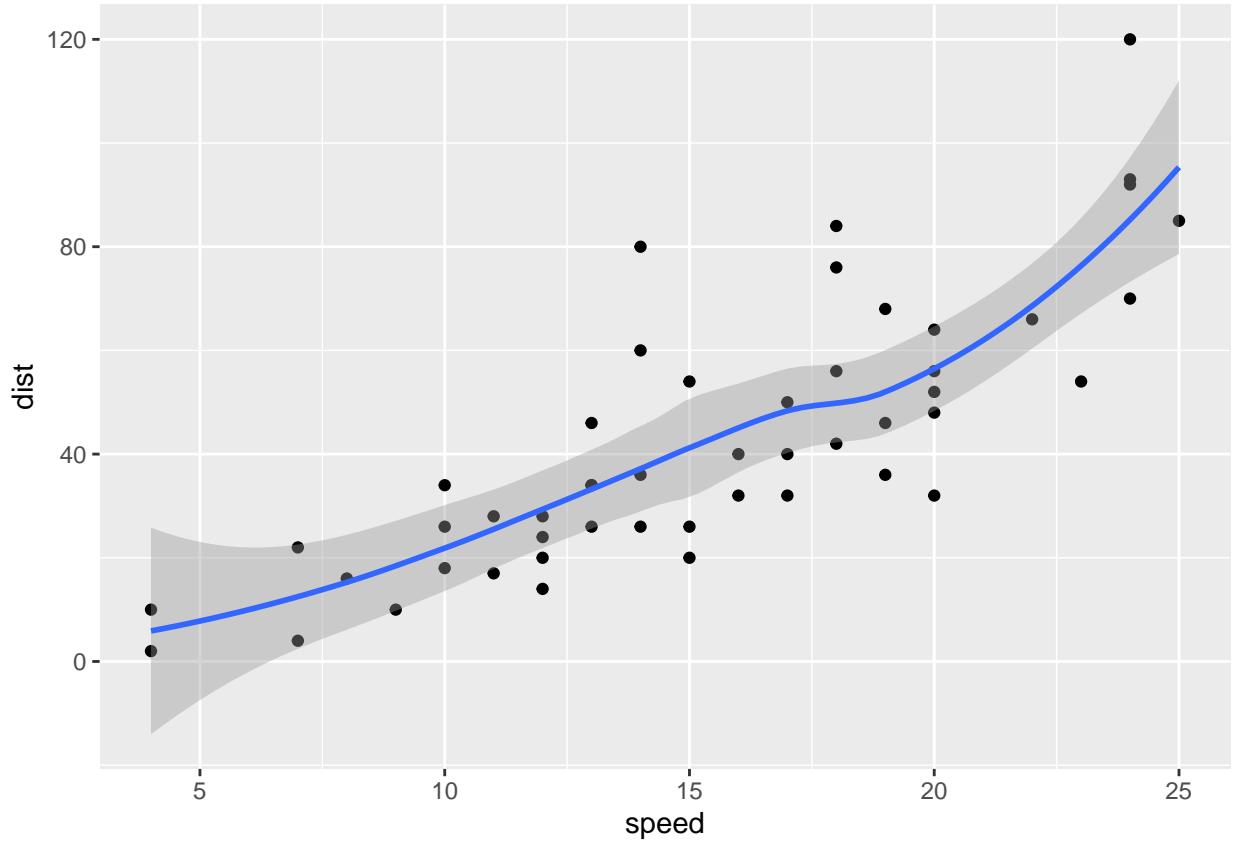
#scatterplot

# loading ggplot (install is done one time in console)
library(ggplot2)

#just sets up the plot, haven't added data + aes + geoms yet (plot = blank)
# ggplot(cars)

#data + aes + geom added, we can see scatterplot
ggplot(data = cars)+ 
  aes(x = speed, y = dist) +
  geom_point()+
  geom_smooth()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



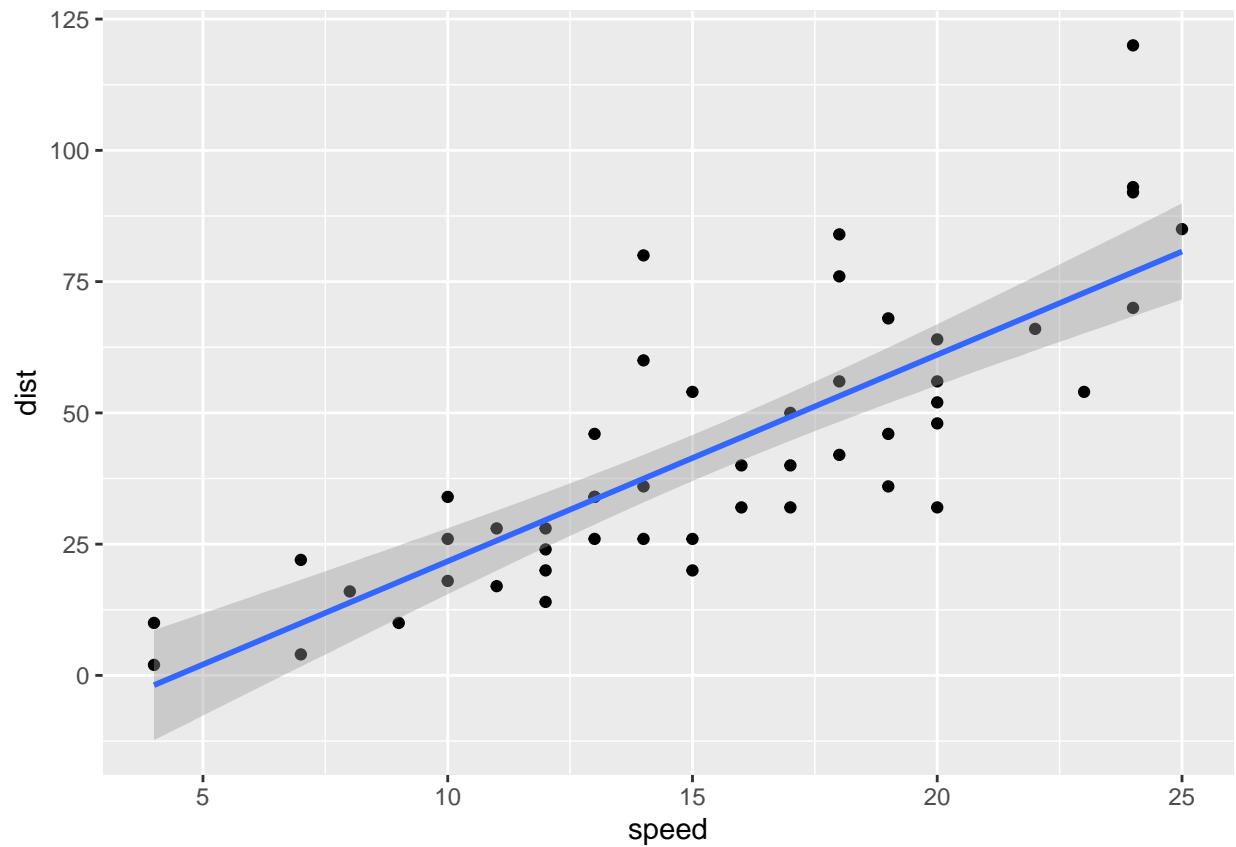
```
#we can stack multiple geom layers, geom_smooth used for trendlines
```

```
#change to linear model, add labels
```

```
carsplot<- ggplot(data = cars)+  
  aes(x = speed, y = dist)+  
  geom_point()+  
  geom_smooth(method = "lm")
```

```
carsplot
```

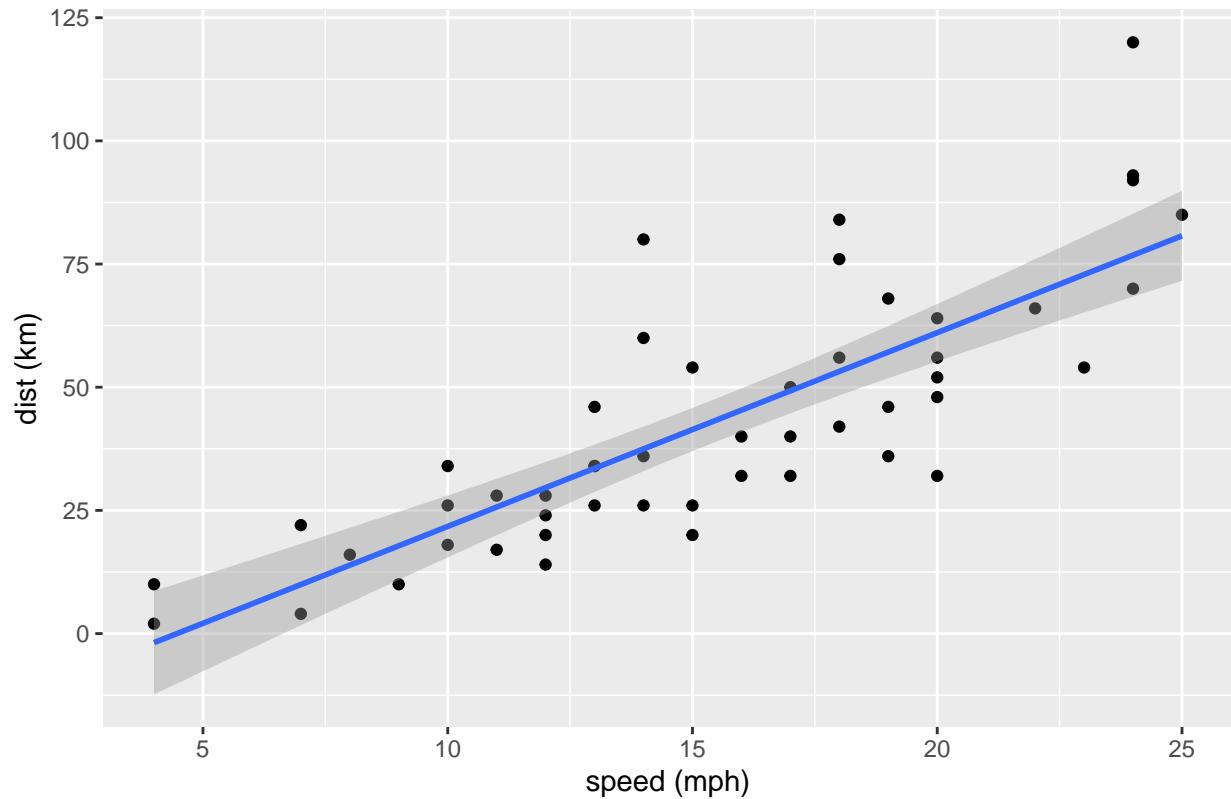
```
## `geom_smooth()` using formula 'y ~ x'
```



```
carsplot+ labs(title="My nice plot", x="speed (mph)", y="dist (km)")
```

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

My nice plot



```
url <- "https://bioboot.github.io/bimm143_S20/class-material/up_down_expression.txt"
genes <- read.delim(url)
head(genes)
```

```
##           Gene Condition1 Condition2      State
## 1       A4GNT -3.6808610 -3.4401355 unchanging
## 2        AAAS  4.5479580  4.3864126 unchanging
## 3       AASDH  3.7190695  3.4787276 unchanging
## 4        AATF  5.0784720  5.0151916 unchanging
## 5        AATK  0.4711421  0.5598642 unchanging
## 6 AB015752.4 -3.6808610 -3.5921390 unchanging
```

```
#numgenes = 5196
nrow(genes)
```

```
## [1] 5196
```

```
#numcolumns = 4
ncol(genes)
```

```
## [1] 4
```

```
#names of columns = "Gene" "Condition1" "Condition2" "State"
colnames(genes)
```

```
## [1] "Gene"      "Condition1" "Condition2" "State"
```

```
#how many genes are upregulated = 127
table(genes[, "State"])
```

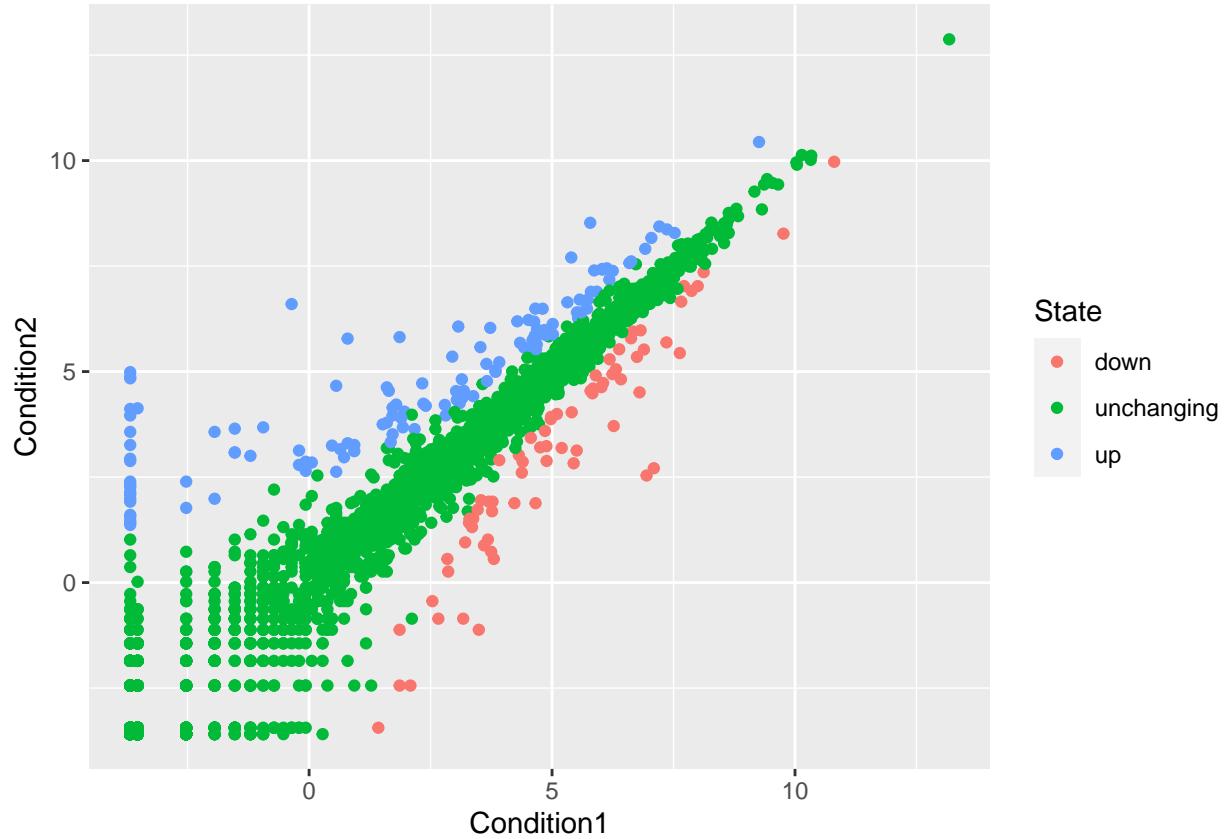
```
##
##      down unchanged      up
##      72        4997     127
```

```
#percentage of upregulated genes = 2.44%
percent <- table(genes[, "State"]) / nrow(genes) *100
round(percent, 2)
```

```
##
##      down unchanged      up
##      1.39        96.17    2.44
```

```
#scatterplot for this new data set
genesplot <- ggplot(genes) +
  aes(x=Condition1, y=Condition2, col= State) +
  geom_point()
```

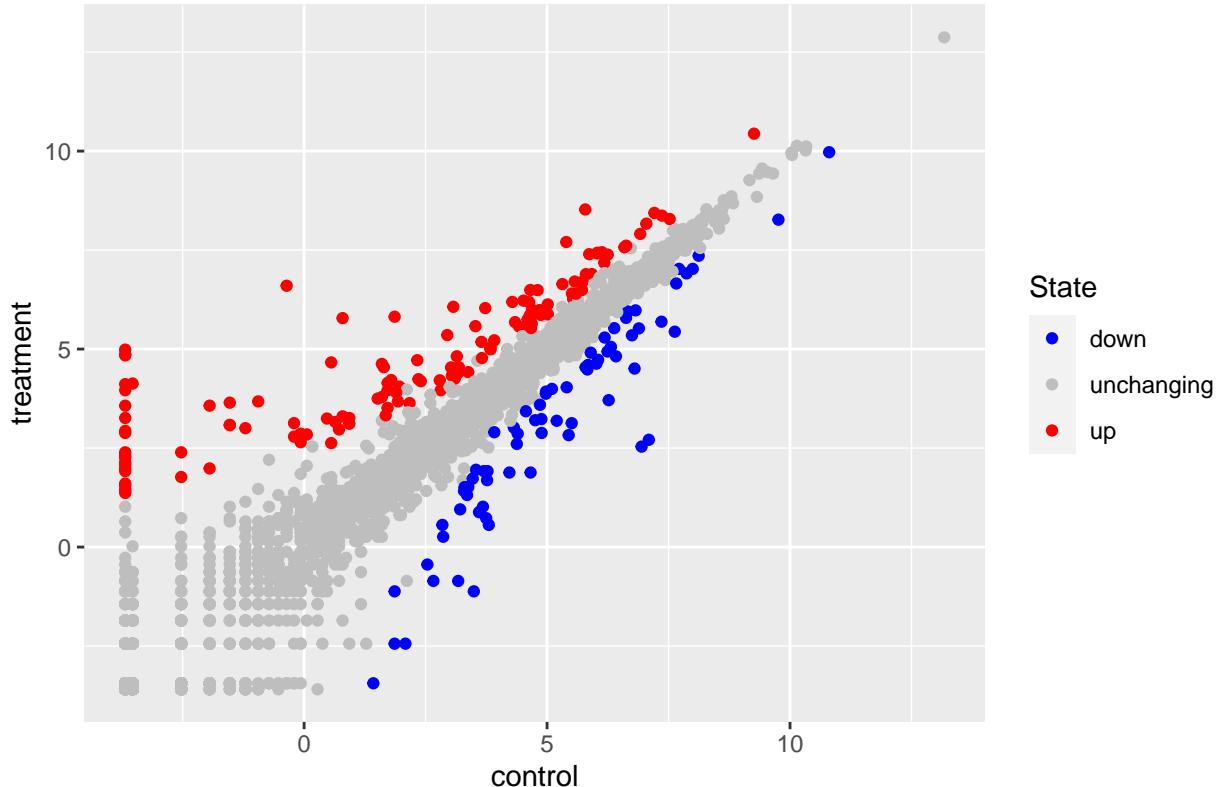
```
genesplot
```



```
#change color scale and add labels

genesplot +
  scale_colour_manual( values=c("blue","gray","red") ) +
  labs(title= "gene expression", x= "control", y="treatment")
```

## gene expression



```
#new dataset = gapminder
url <- "https://raw.githubusercontent.com/jennybc/gapminder/master/inst/extdata/gapminder.tsv"
gapminder <- read.delim(url)

#will learn abt dplyr next class i believe
library(dplyr)

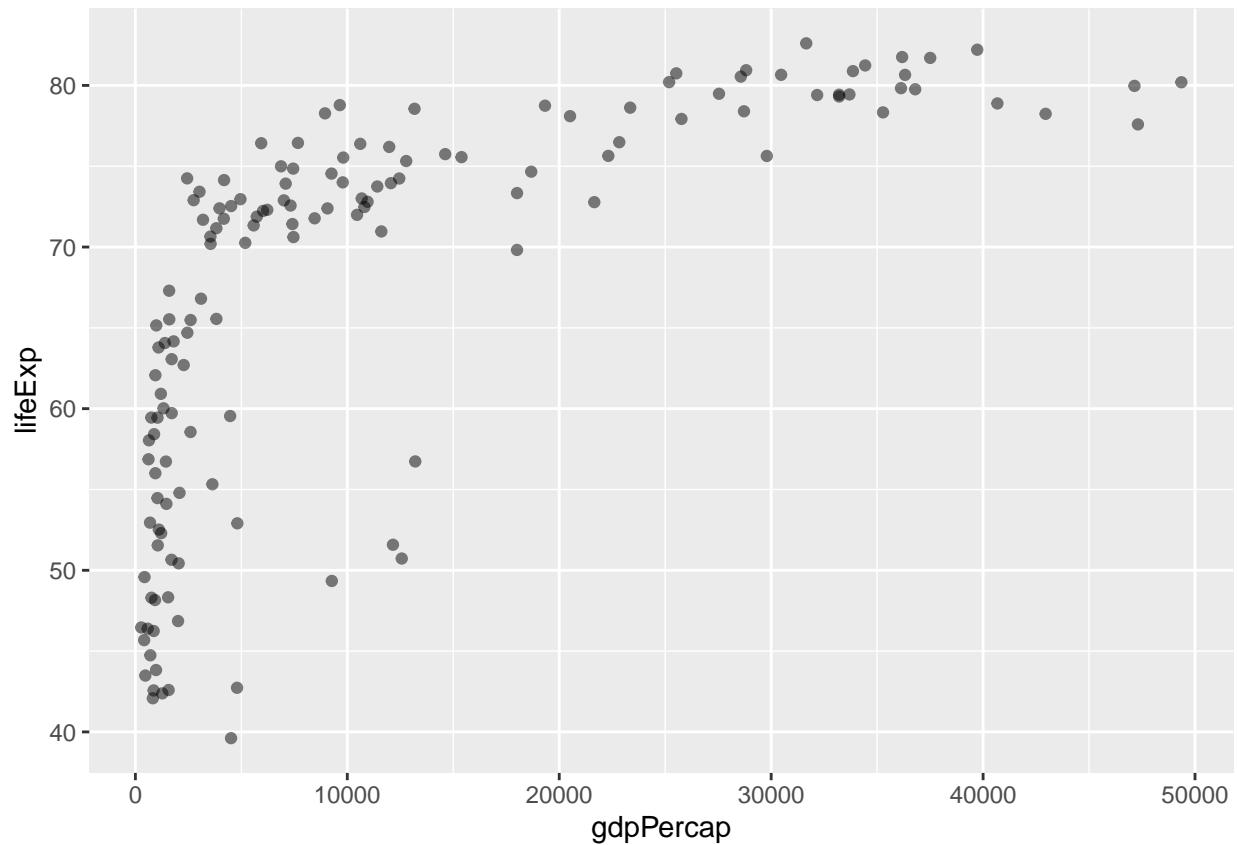
## 
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
## 
##     filter, lag

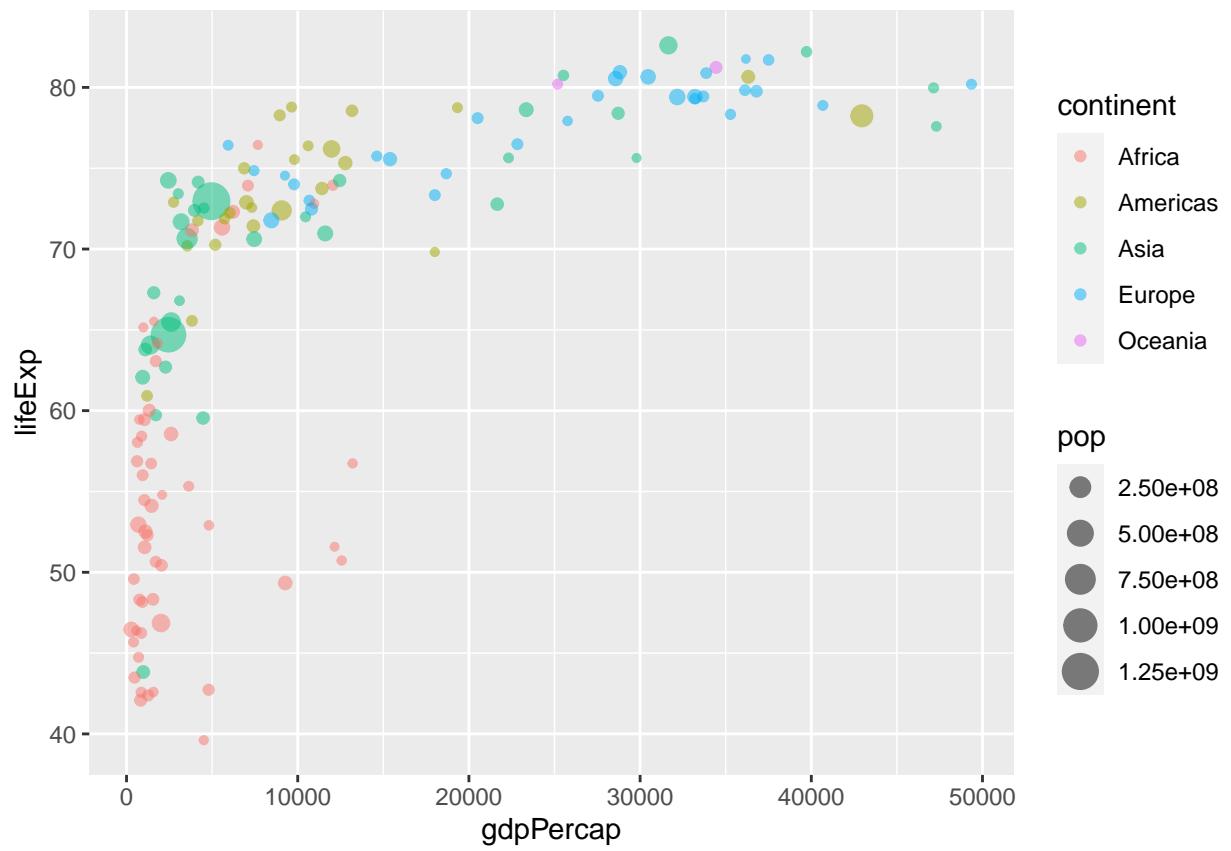
## The following objects are masked from 'package:base':
## 
##     intersect, setdiff, setequal, union

gapminder_2007 <- gapminder %>% filter(year==2007)

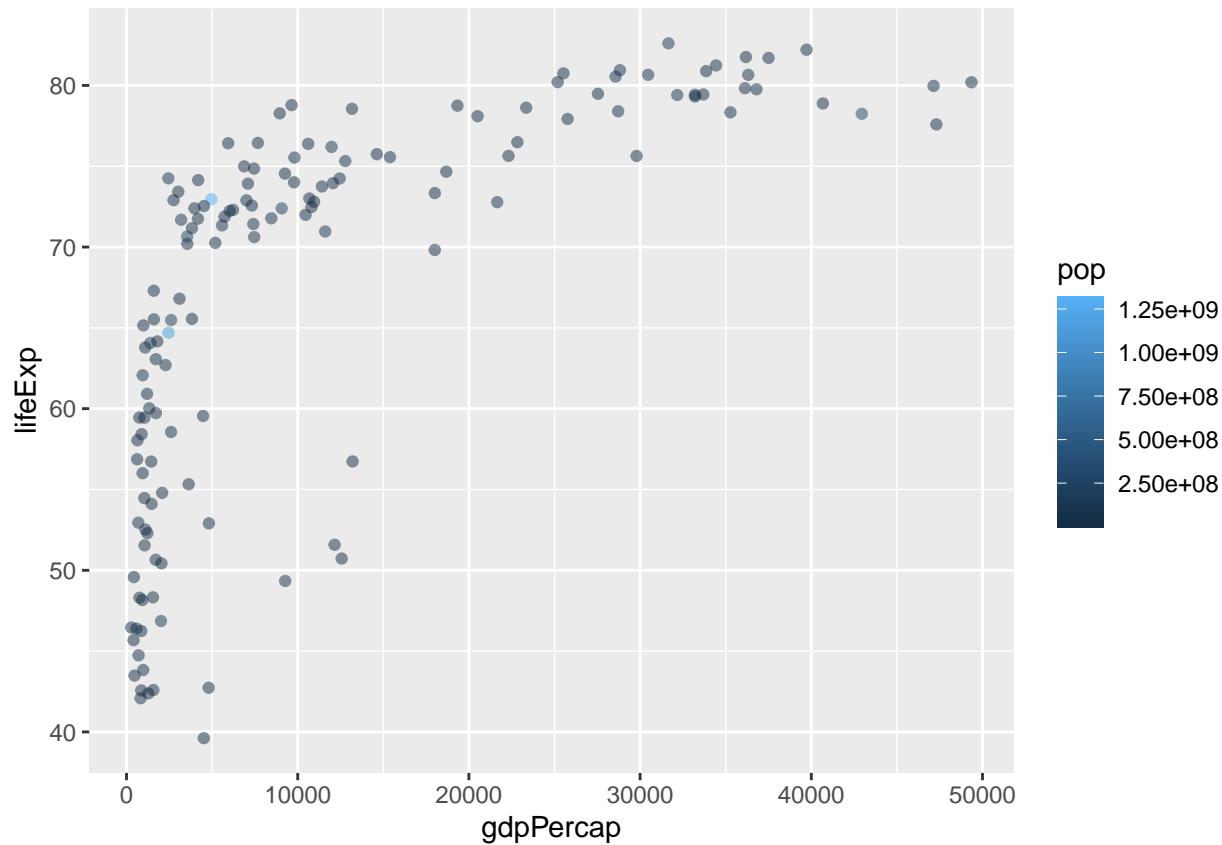
# create a scatterplot (alpha controls transparency)
ggplot(gapminder_2007) +
  aes(x=gdpPercap, y=lifeExp) +
  geom_point(alpha = 0.5)
```



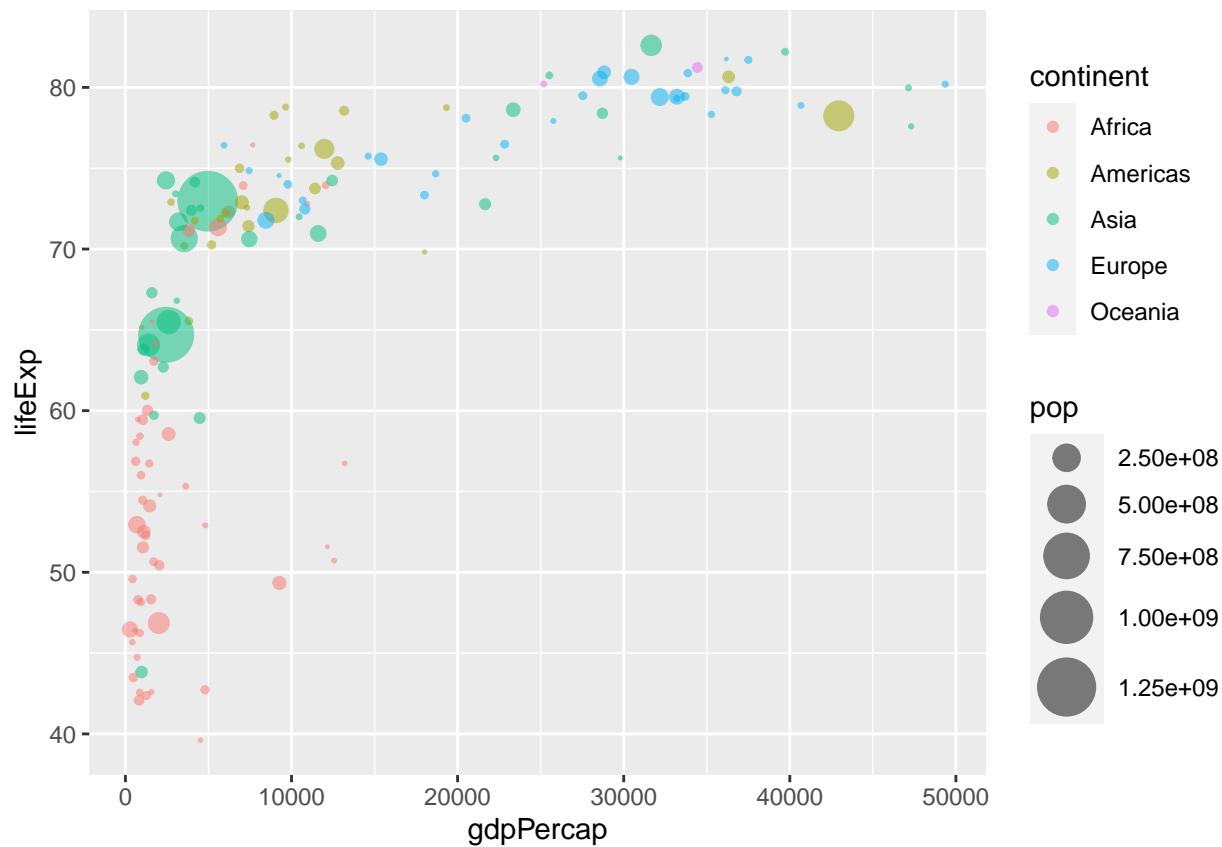
```
#add more aes, in this color is categorical value  
ggplot(gapminder_2007) +  
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +  
  geom_point(alpha = 0.5)
```



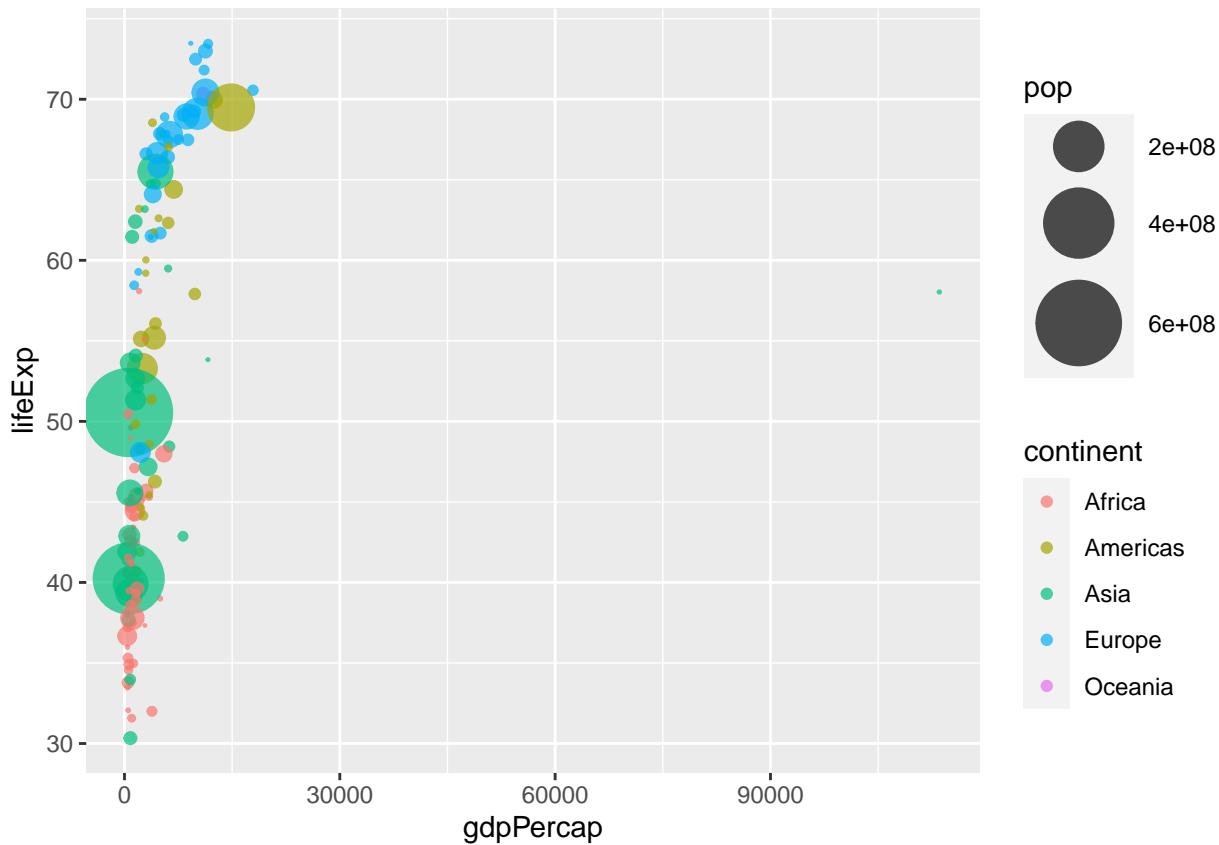
```
#try a new aes, making color a numerical value (continuous)
ggplot(gapminder_2007) +
  aes(x=gdpPercap, y=lifeExp, color=pop) +
  geom_point(alpha = 0.5)
```



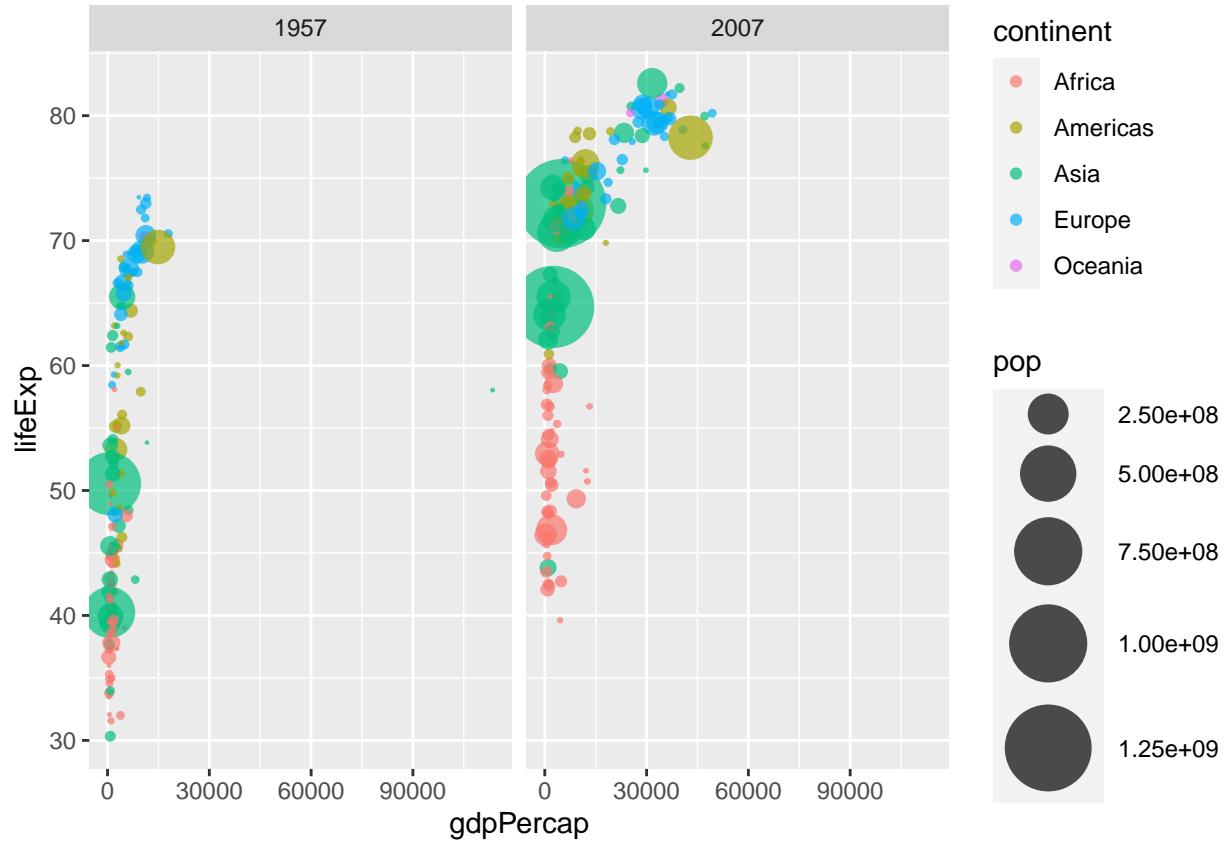
```
#make sizes proportional to pop, size=pop does not make them proportional, can't
#compare two points accurately
ggplot(gapminder_2007) +
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point(alpha = 0.5) +
  scale_size_area(max_size = 10)
```



```
#now do this for a diff year 1957
gapminder_1957 <- gapminder %>% filter(year==1957)
ggplot(gapminder_1957) +
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point(alpha = 0.7) +
  scale_size_area(max_size = 15)
```



```
#now show both 2007 and 1957 side by side
gapminder_both <- gapminder %>% filter(year==1957 | year==2007)
ggplot(gapminder_both) +
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point(alpha = 0.7) +
  scale_size_area(max_size = 15) +
  facet_wrap(~year)
```



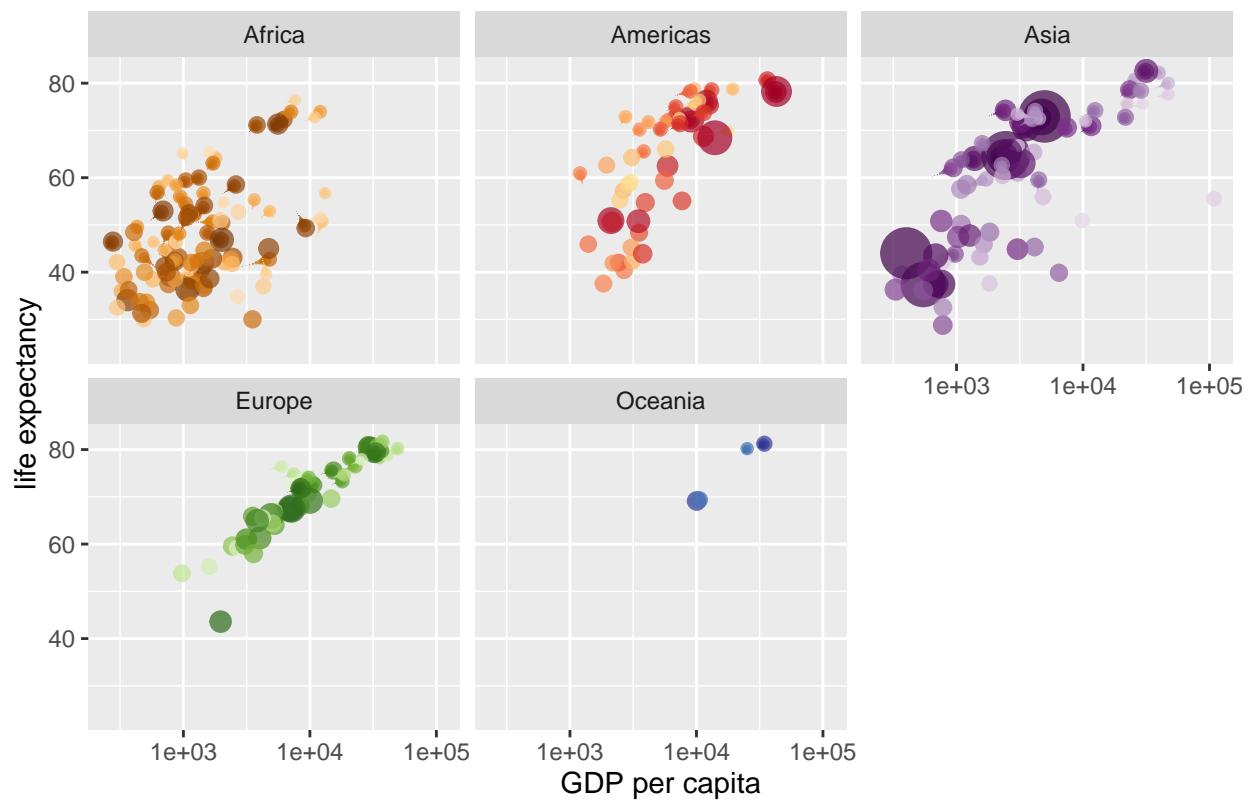
```
#plot animation
library(gganimate)
library(gapminder)

## 
## Attaching package: 'gapminder'

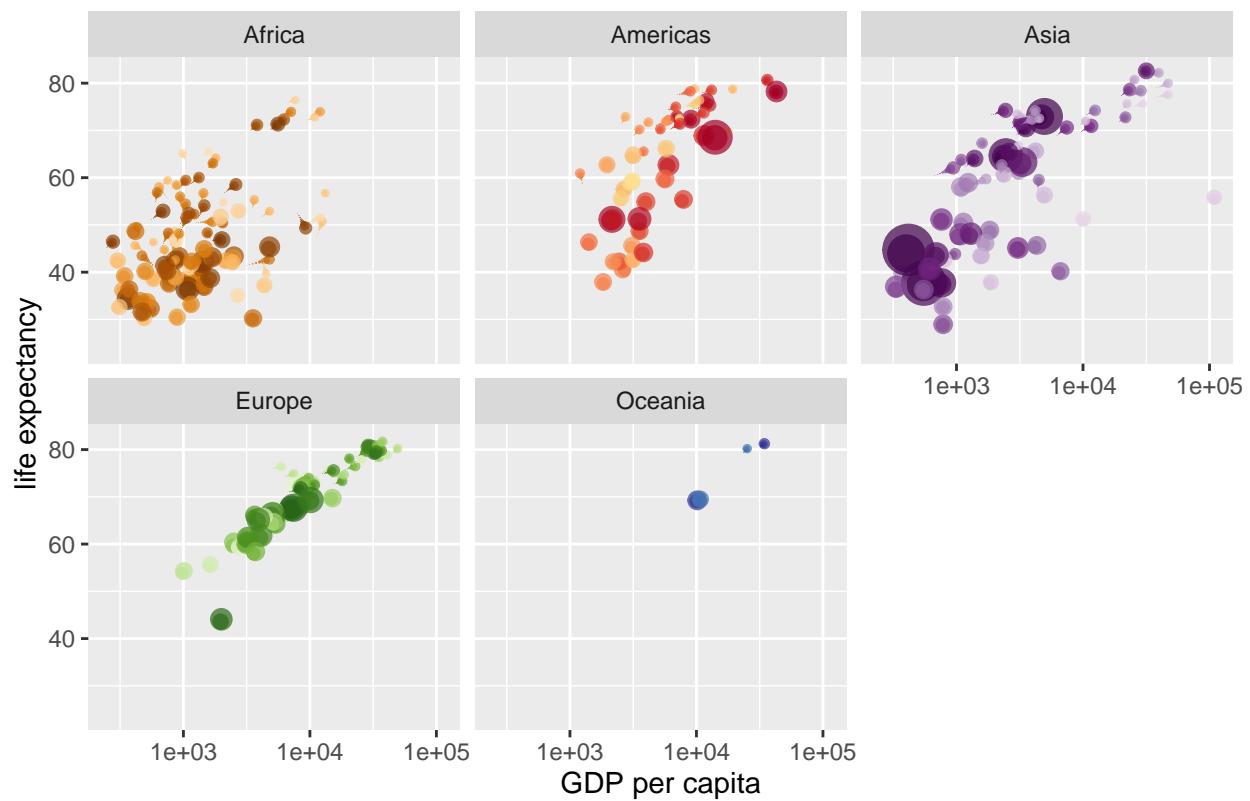
## The following object is masked _by_ '.GlobalEnv':
## 
##     gapminder

ggplot(gapminder, aes(gdpPercap, lifeExp, size = pop, colour = country)) +
  geom_point(alpha = 0.7, show.legend = FALSE) +
  scale_colour_manual(values = country_colors) +
  scale_size(range = c(2, 12)) +
  scale_x_log10() +
  # facet wrap separates into diff graphs by continent
  facet_wrap(~continent) +
  # new layers from gganimate
  labs(title = 'Year: {frame_time}', x = 'GDP per capita', y = 'life expectancy') +
  transition_time(year) +
  shadow_wake(wake_length = 0.1, alpha = FALSE)
```

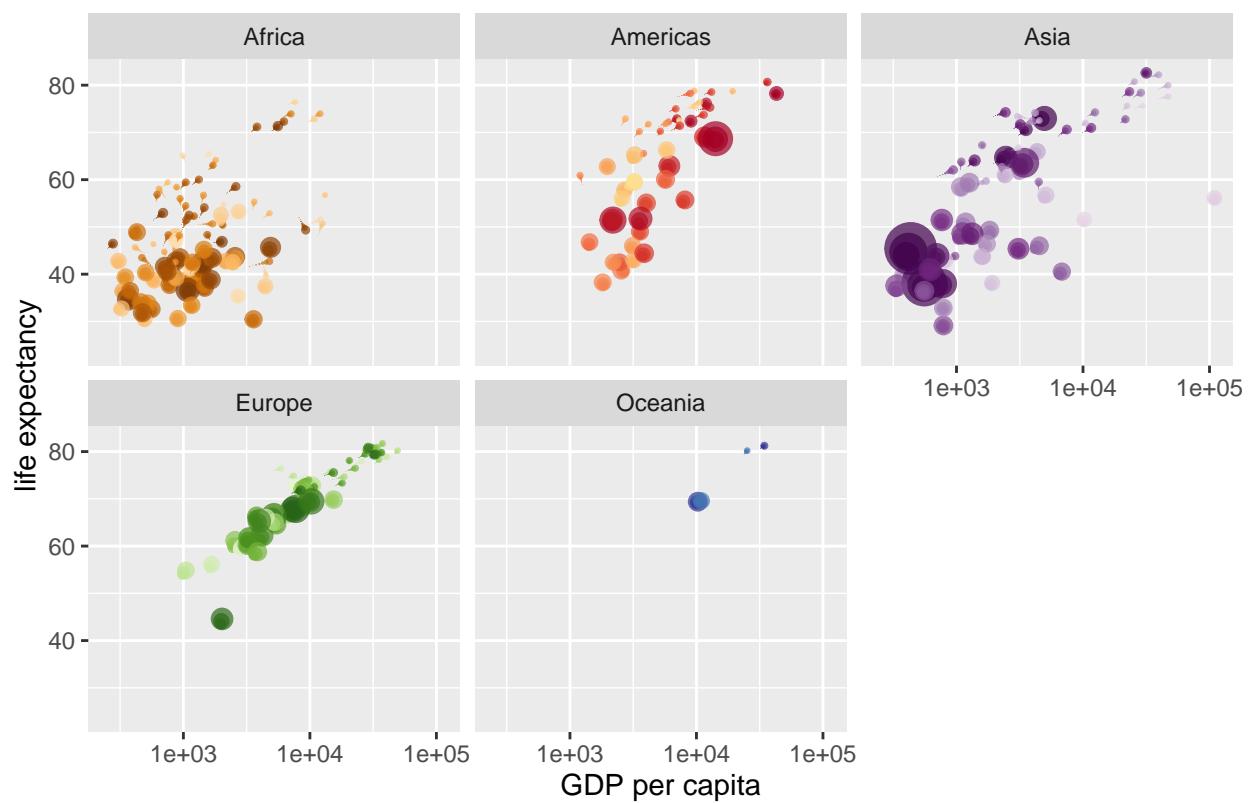
Year: 1952



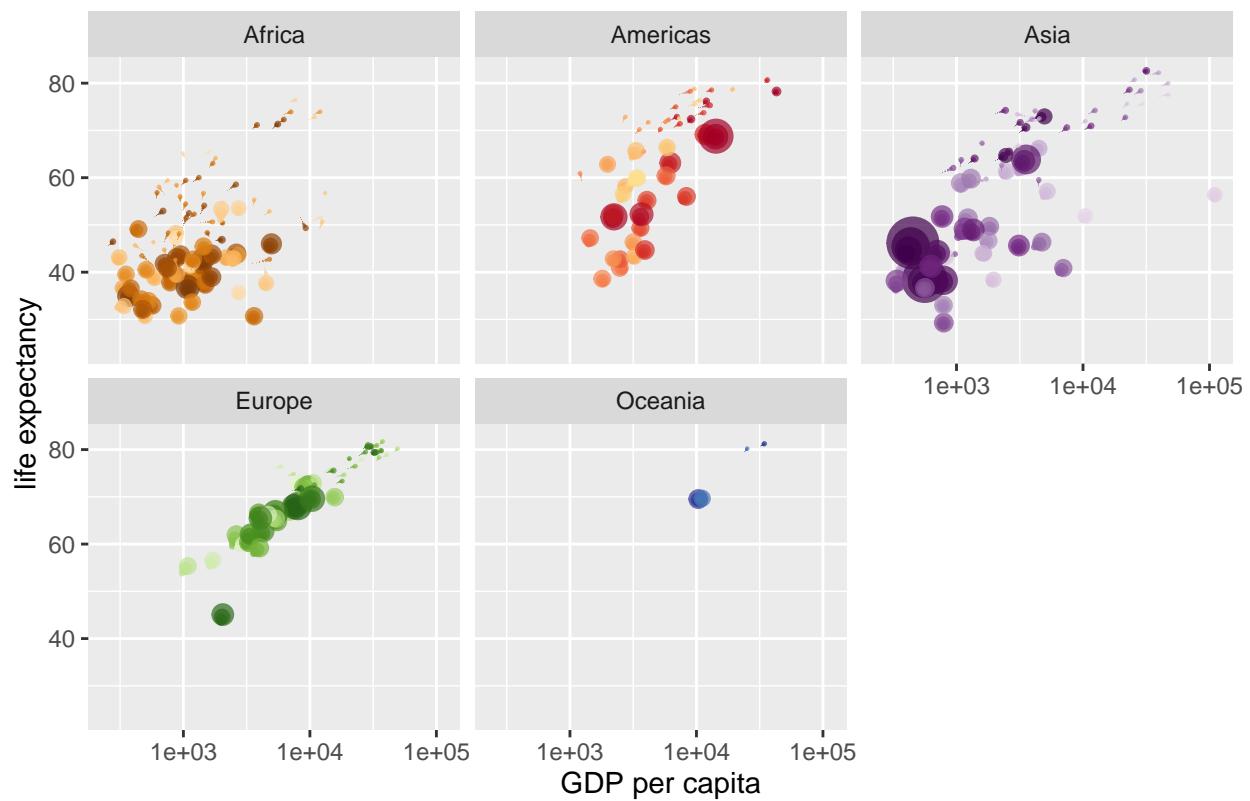
Year: 1953



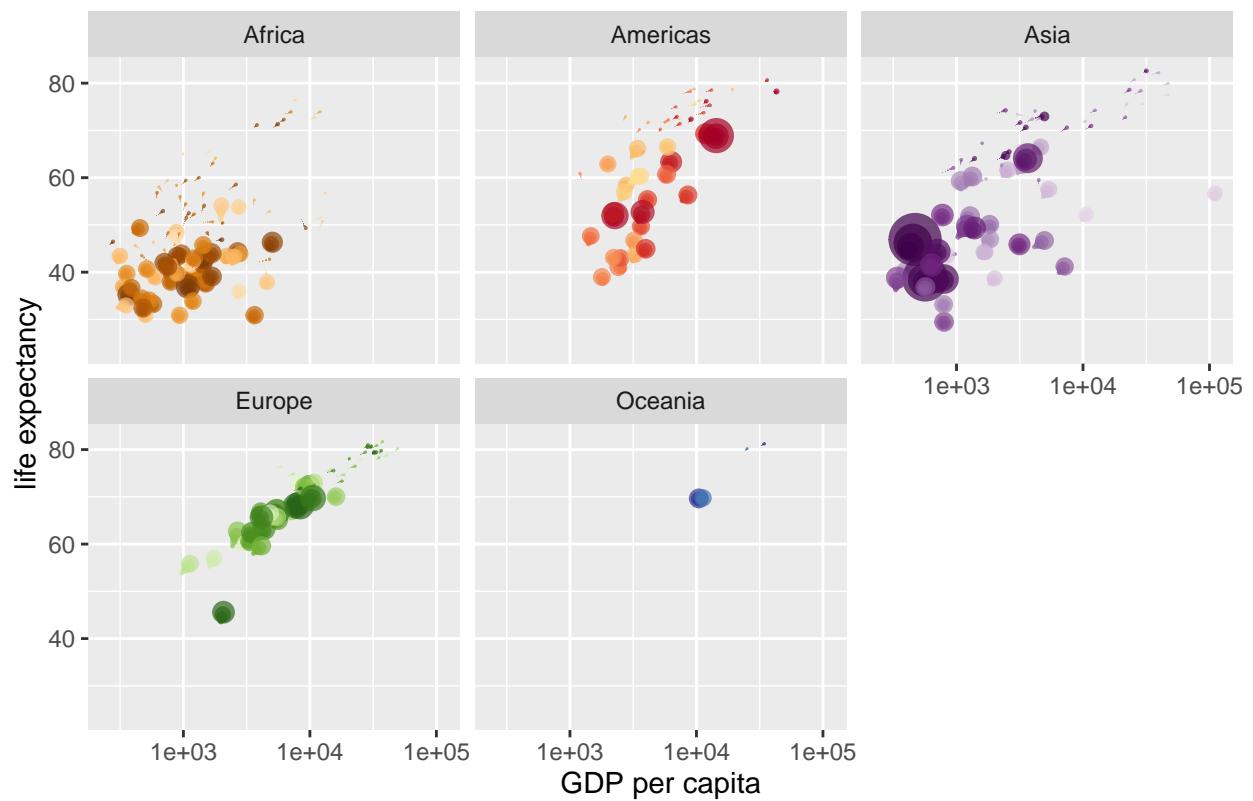
Year: 1953



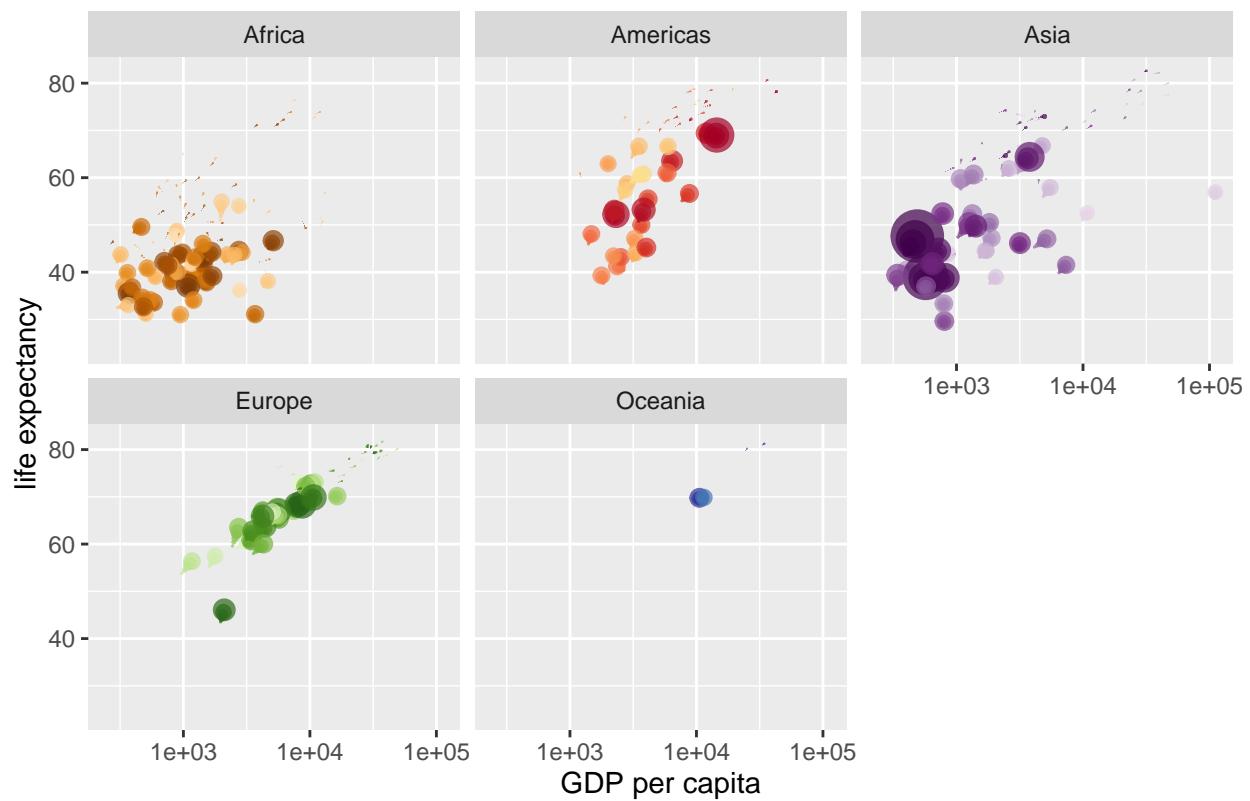
Year: 1954



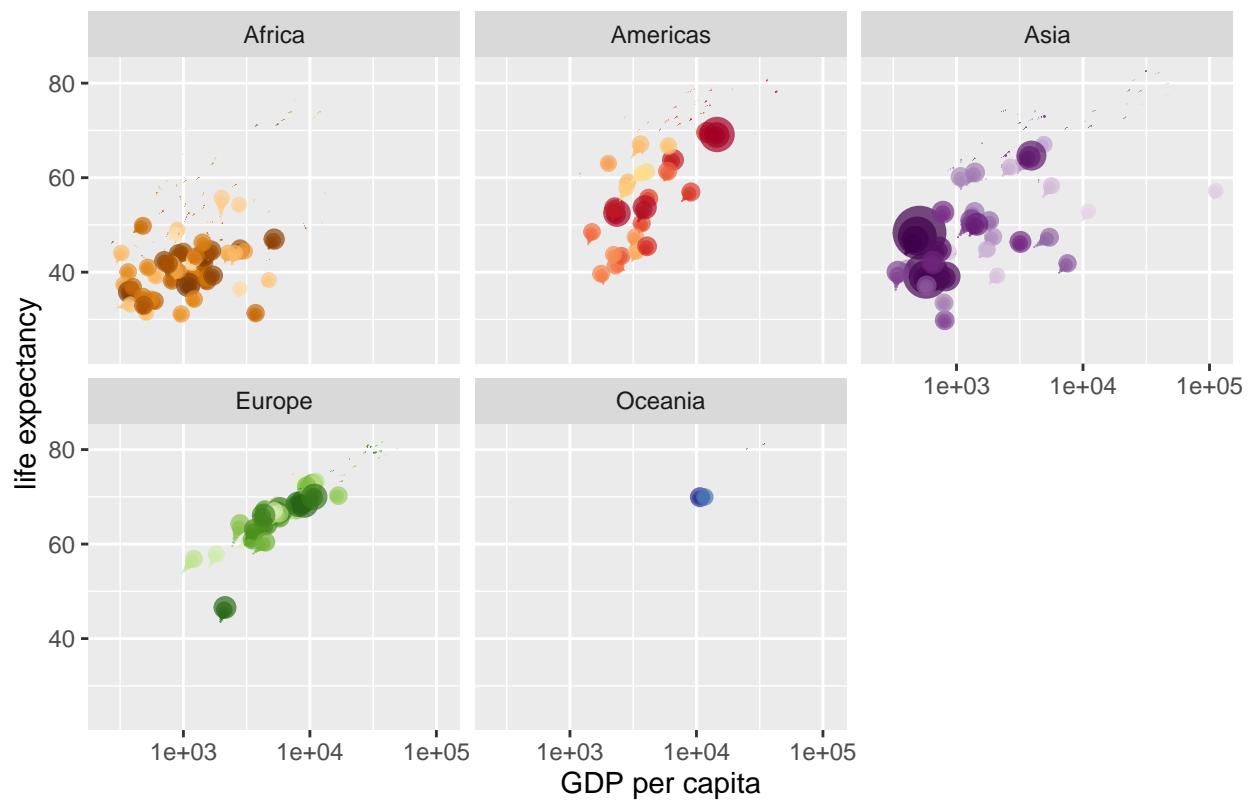
Year: 1954



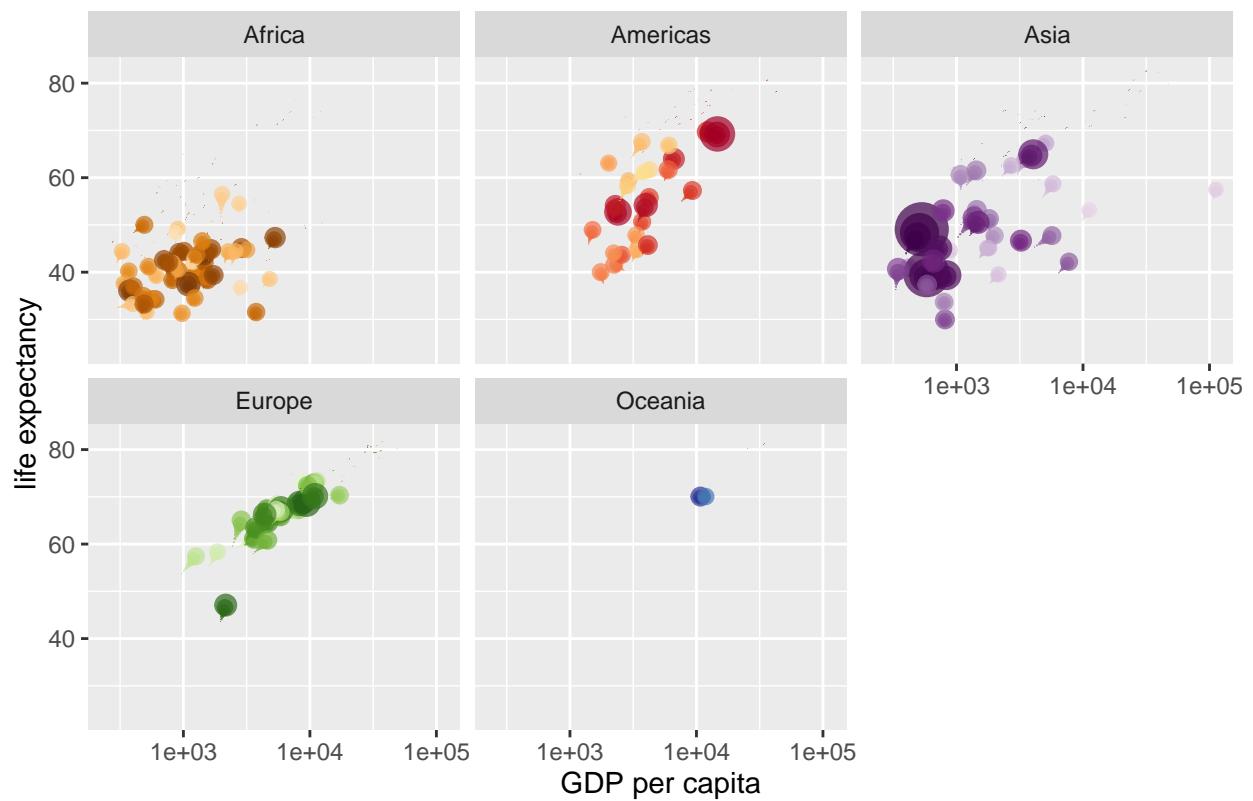
Year: 1955



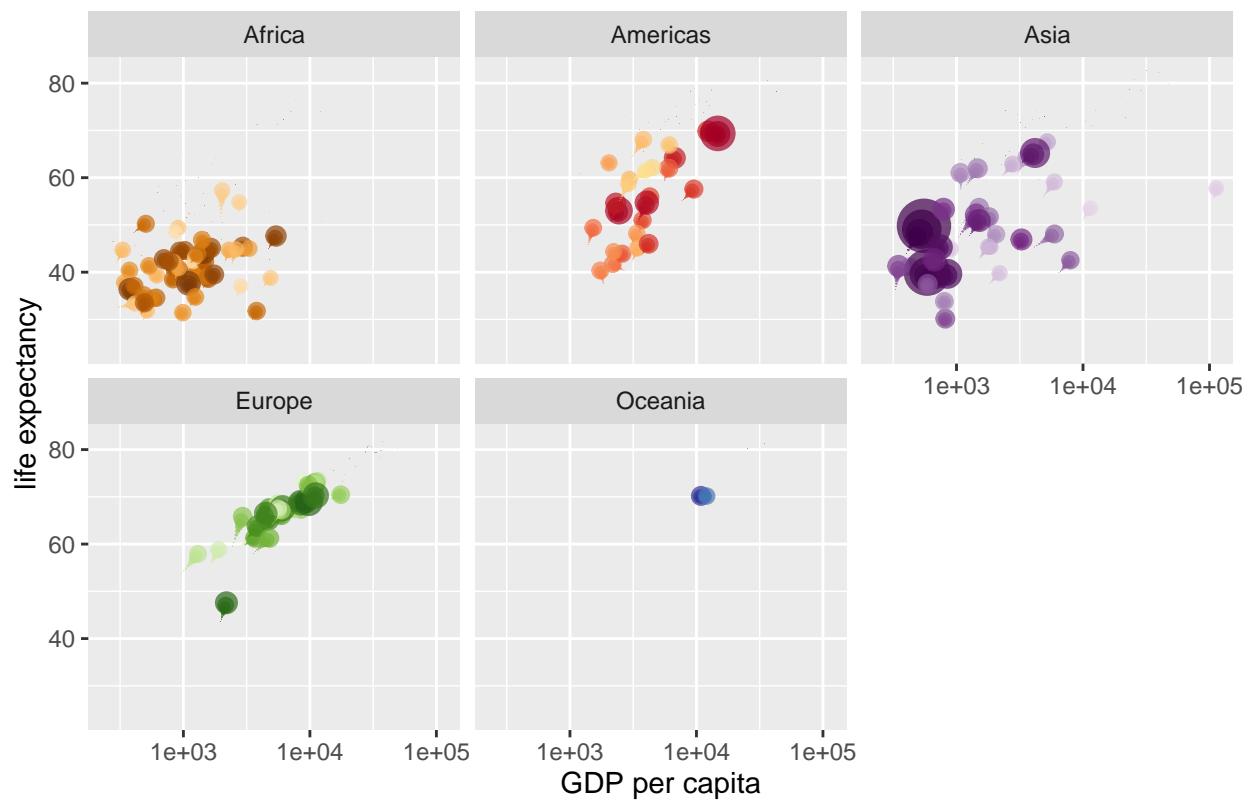
Year: 1955



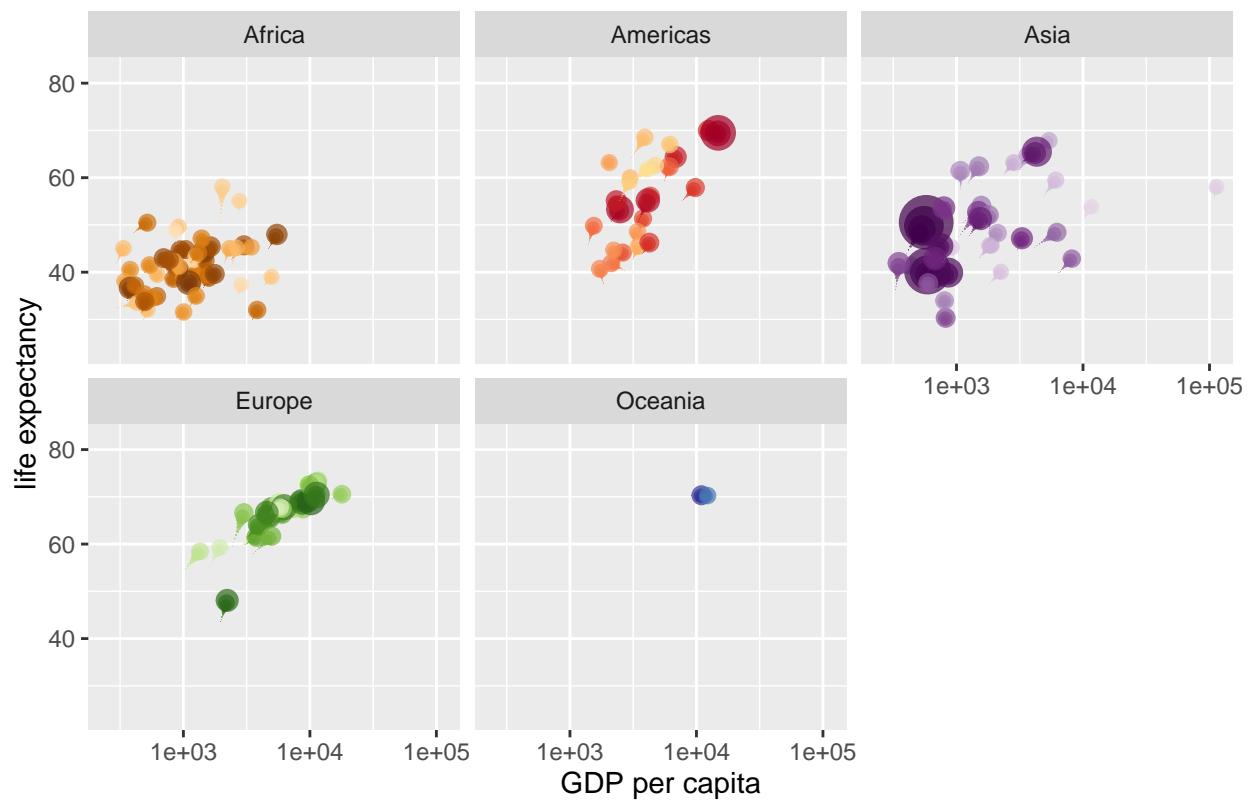
Year: 1956



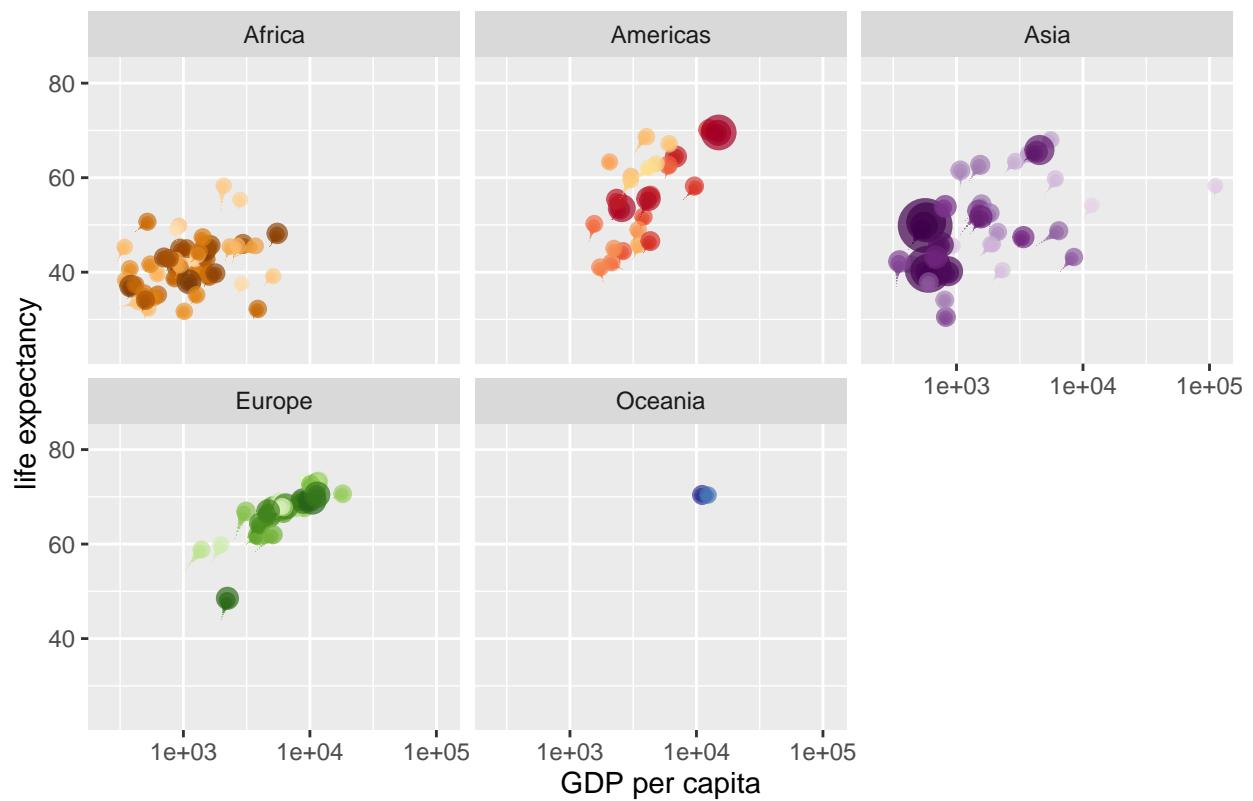
Year: 1956



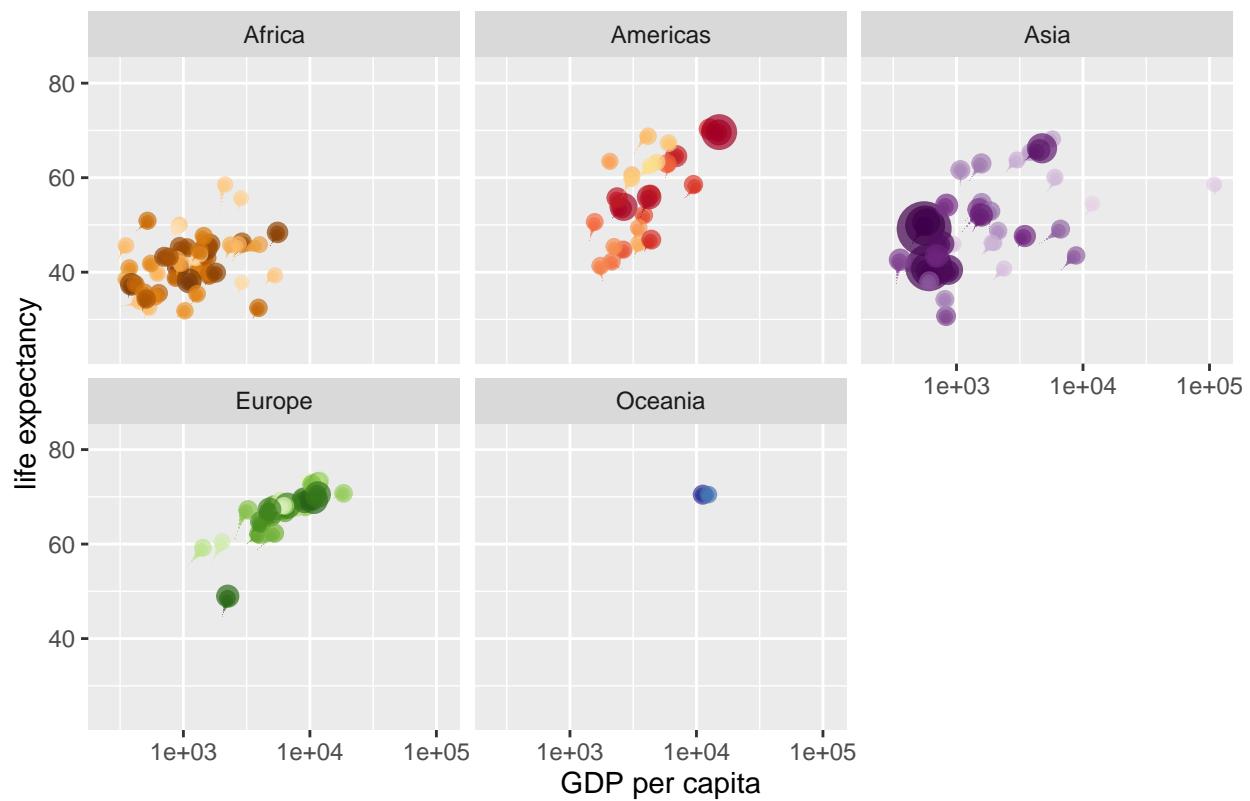
Year: 1957



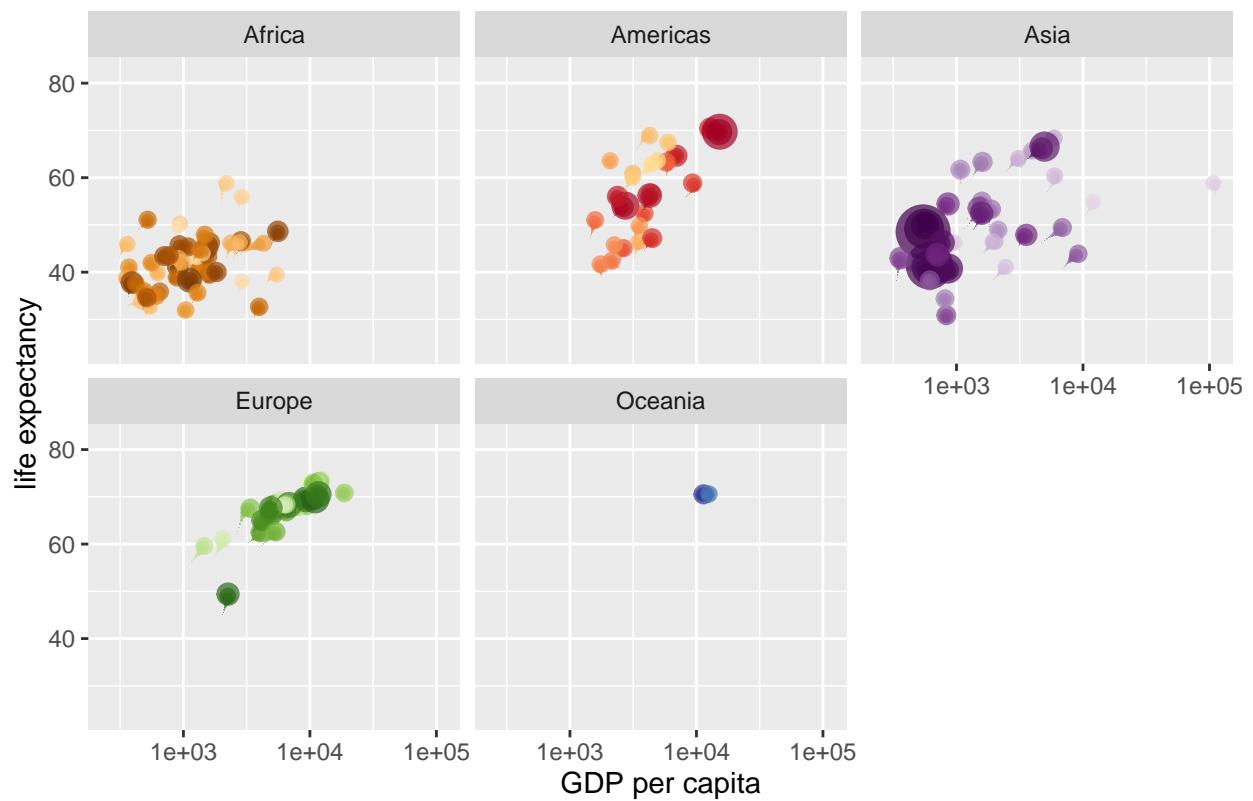
Year: 1958



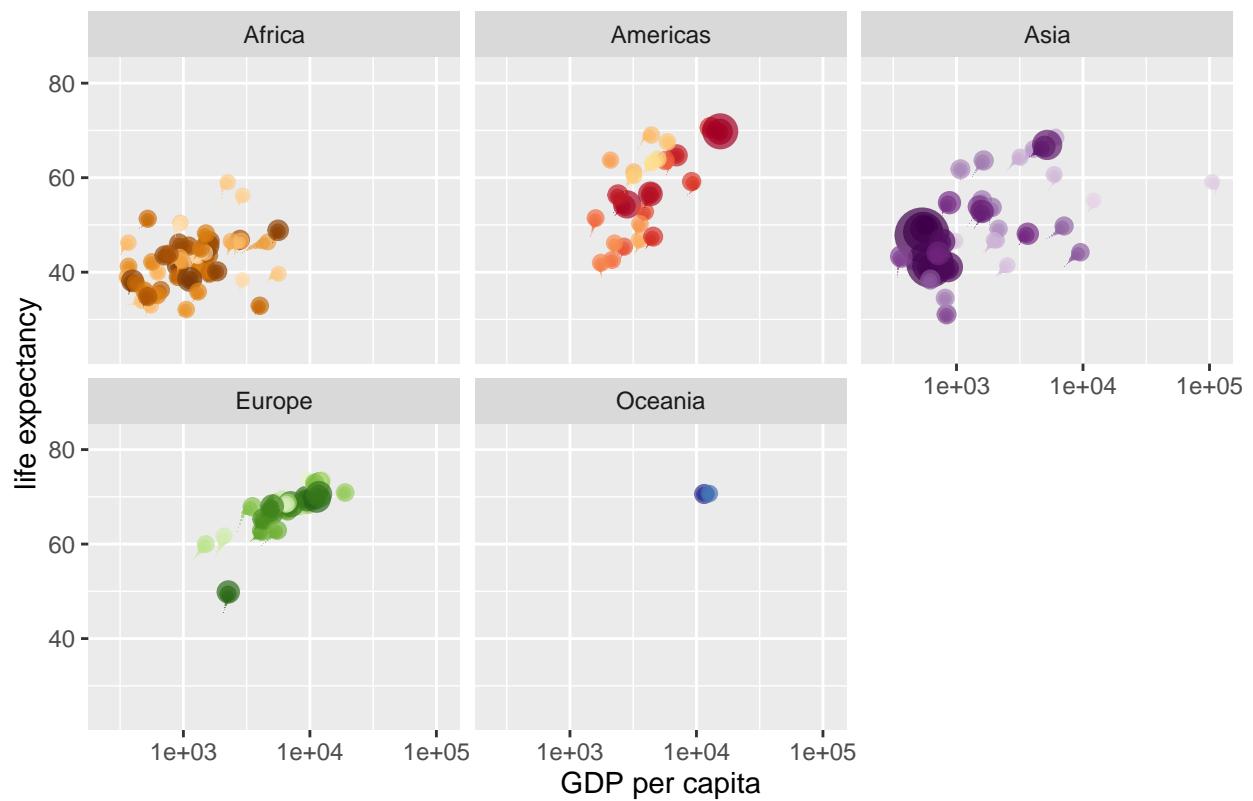
Year: 1958



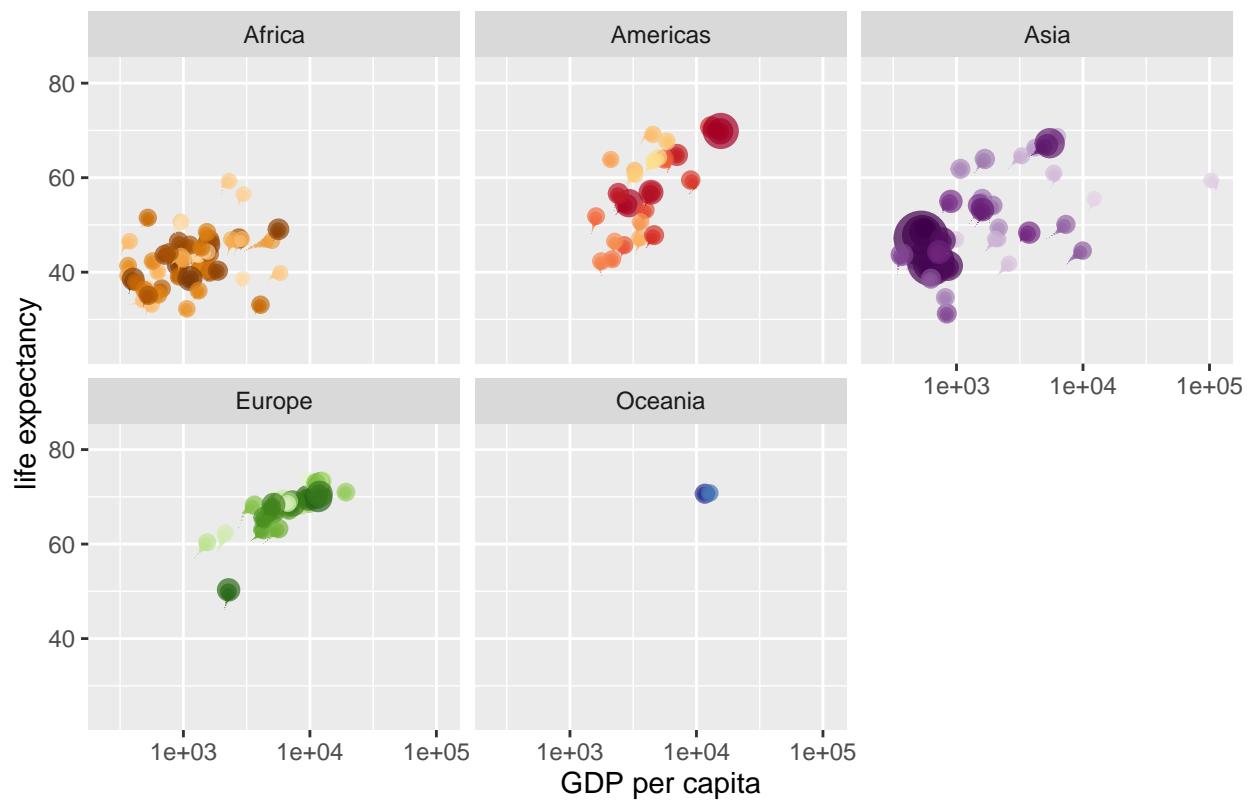
Year: 1959



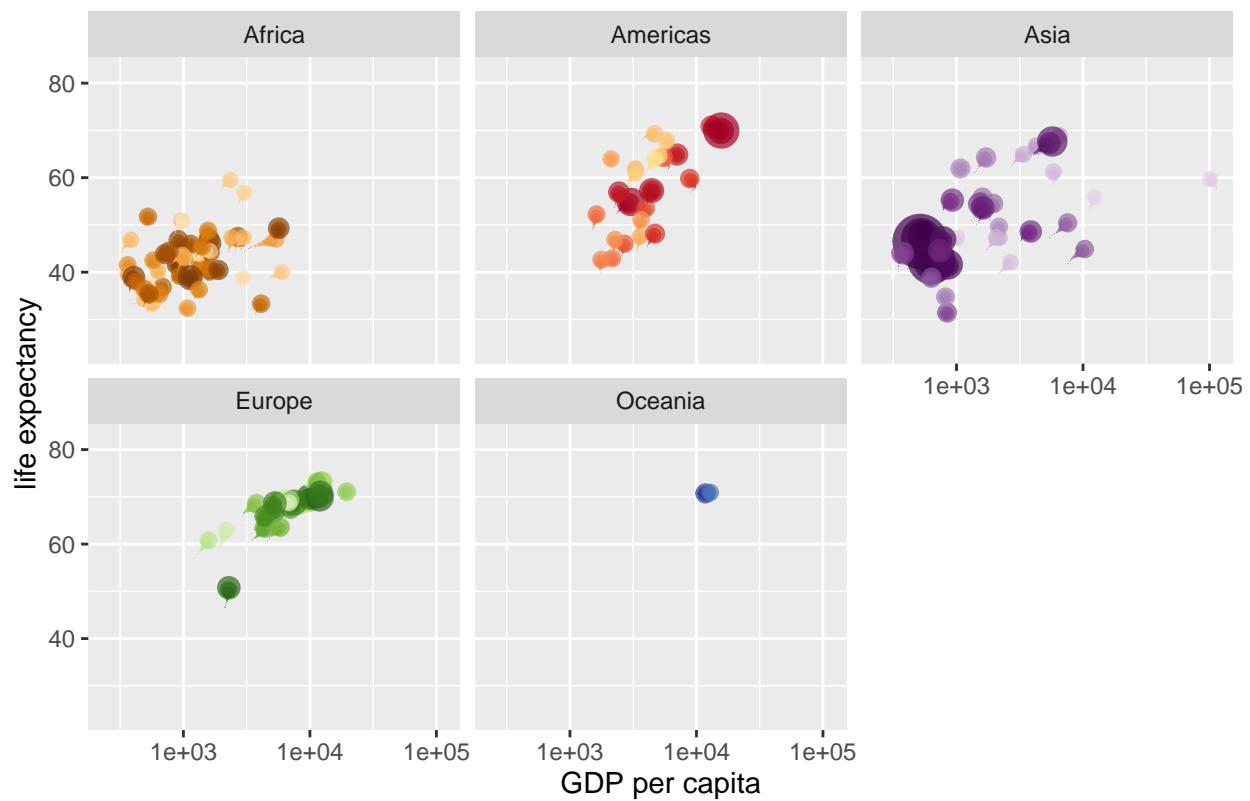
Year: 1959



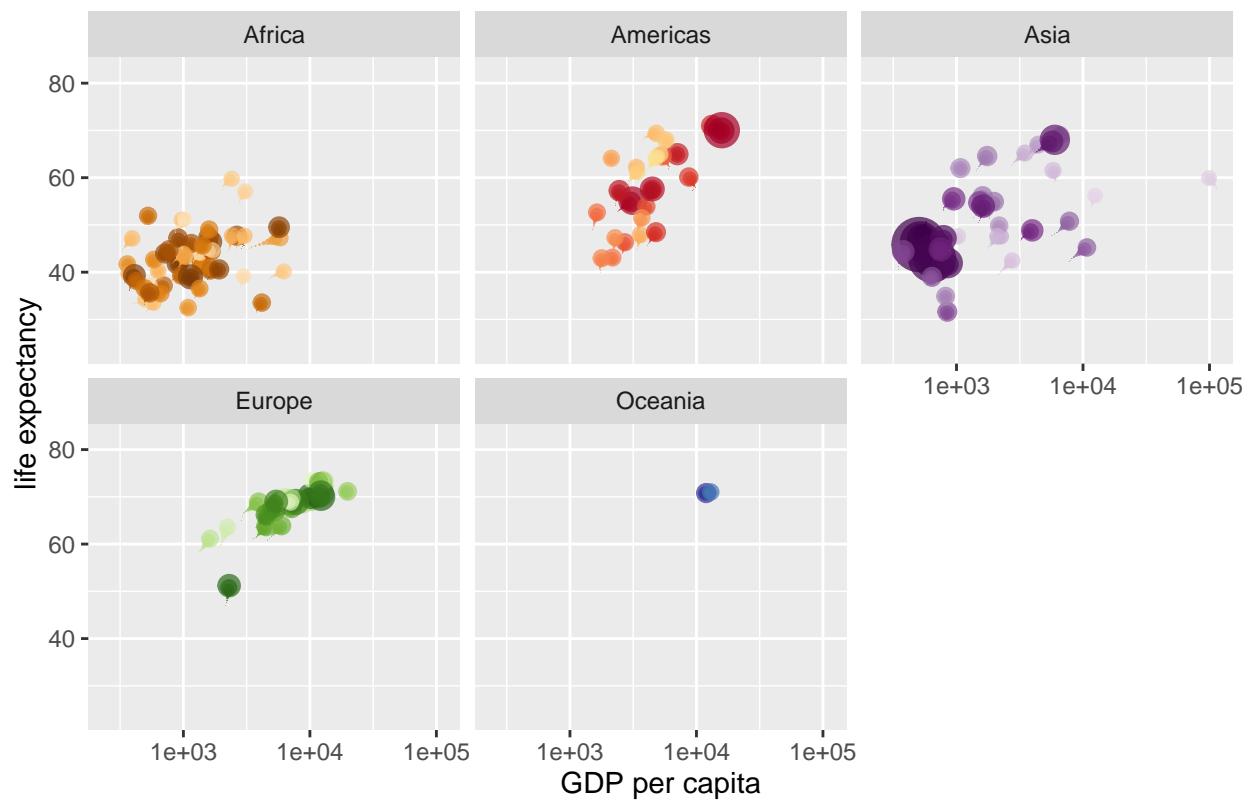
Year: 1960



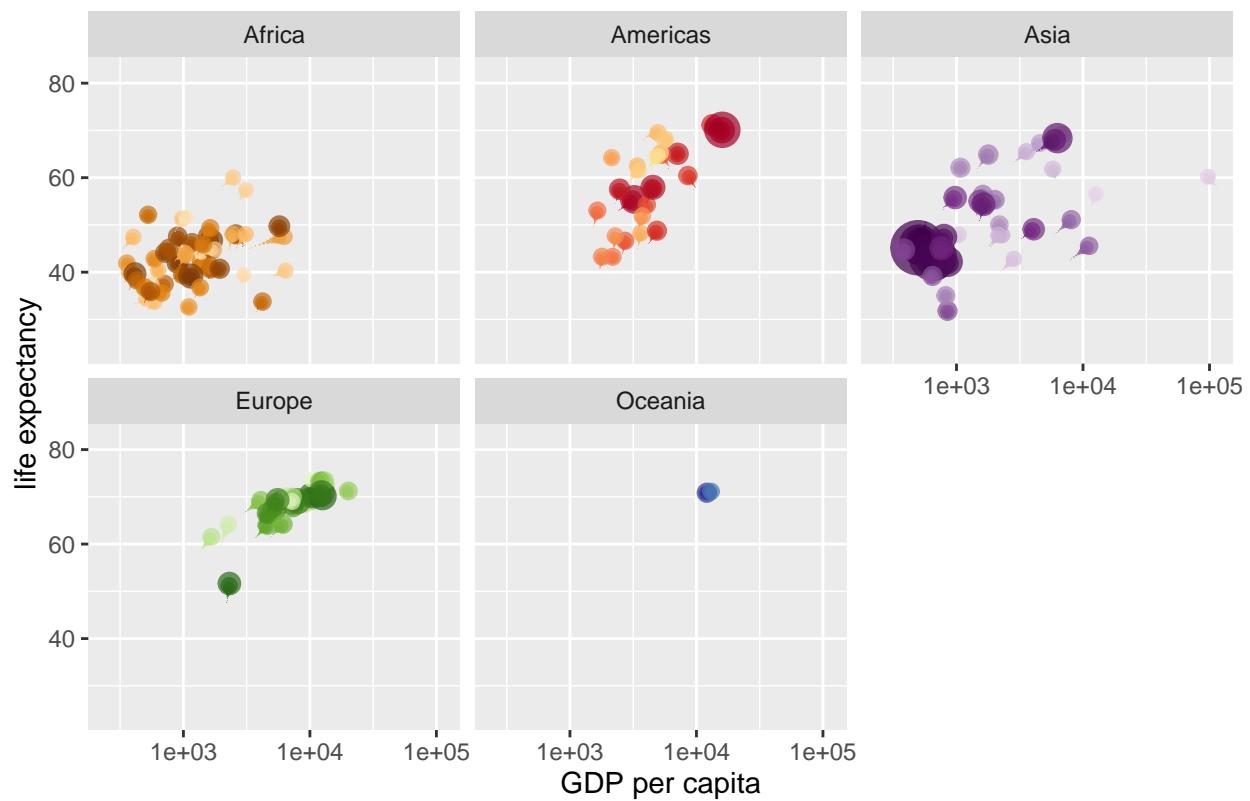
Year: 1960



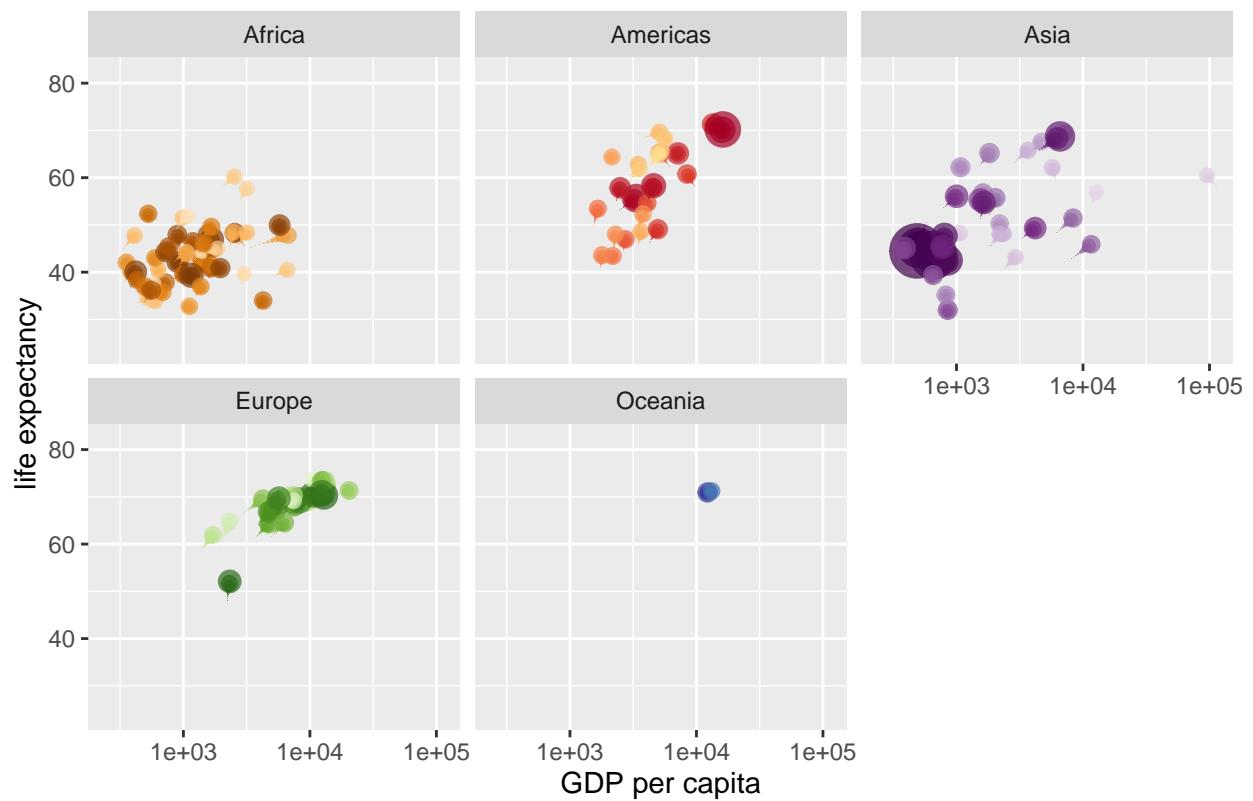
Year: 1961



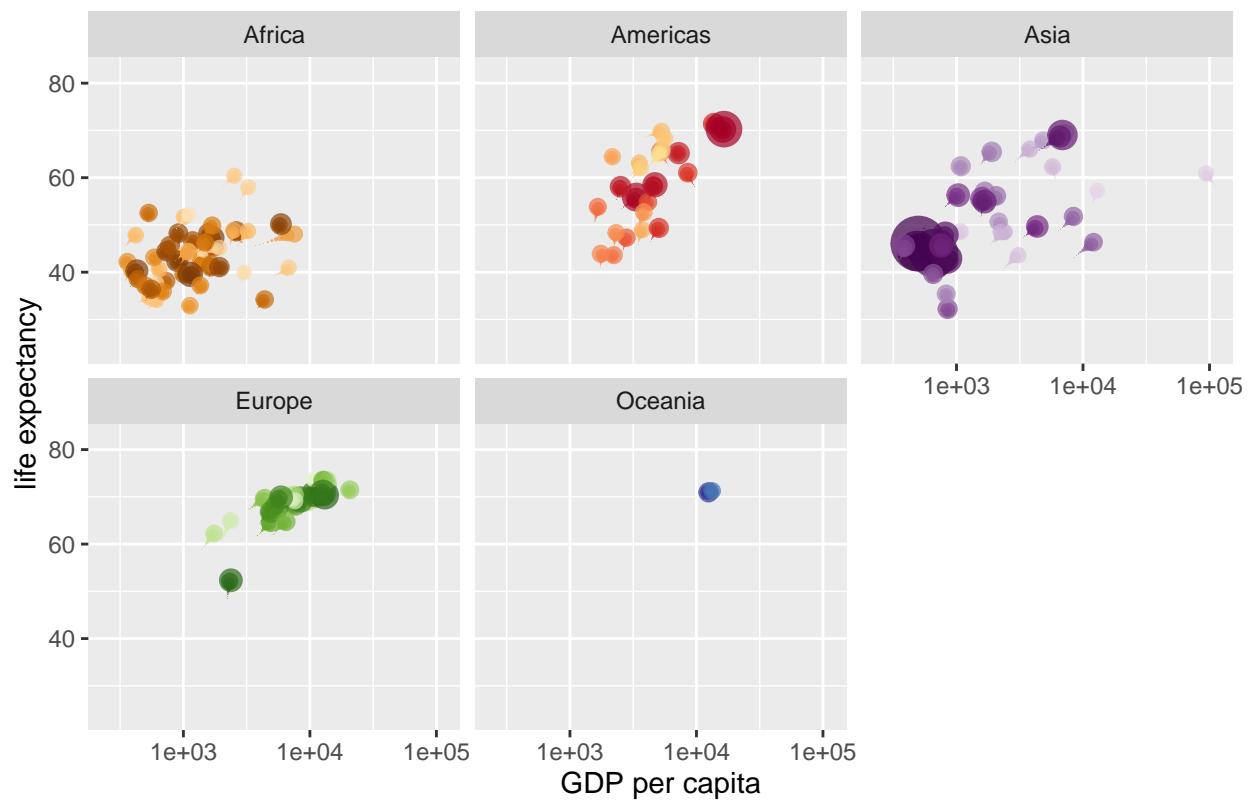
Year: 1961



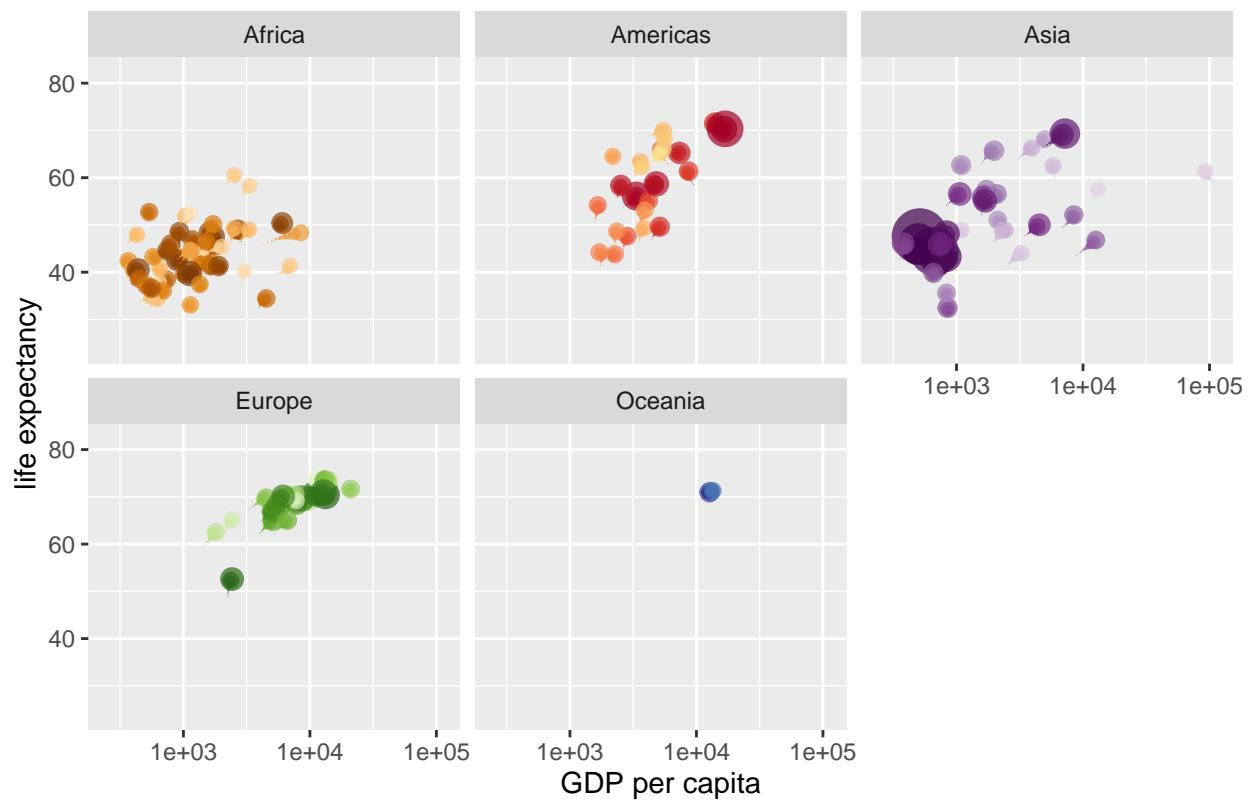
Year: 1962



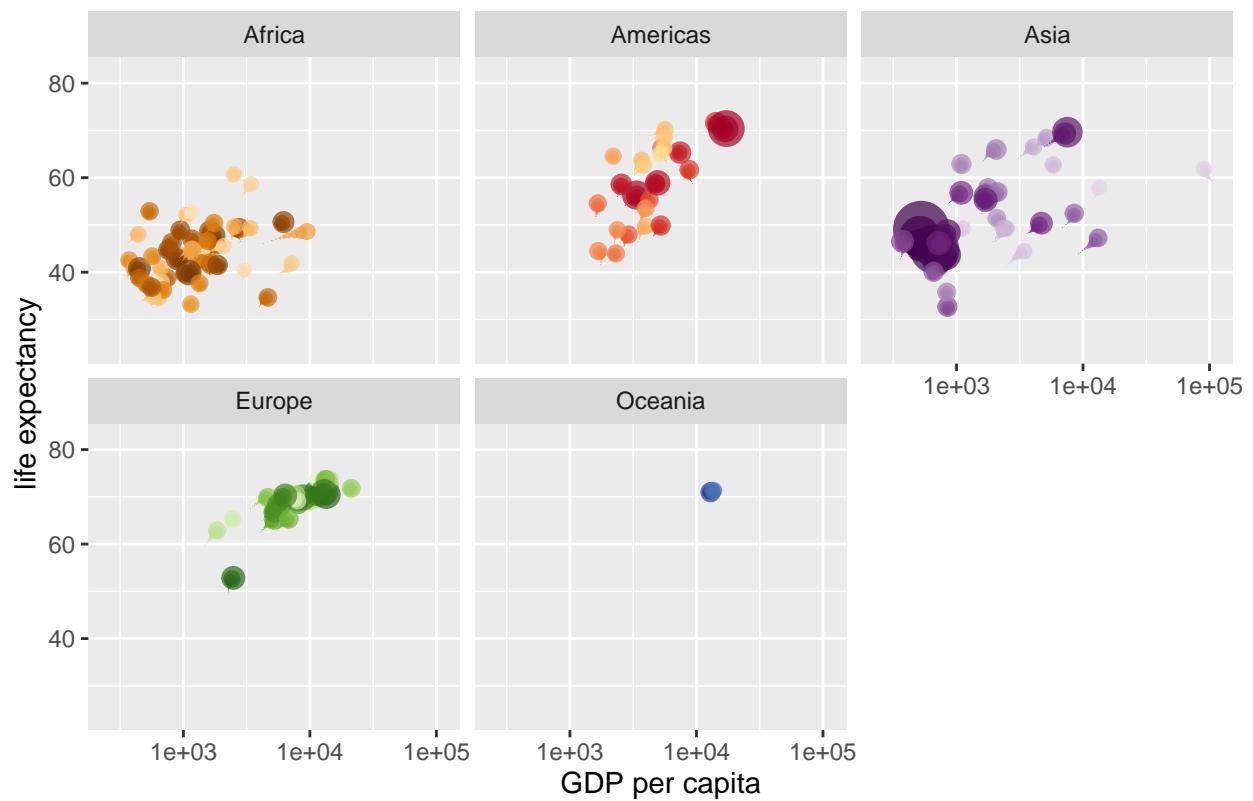
Year: 1963



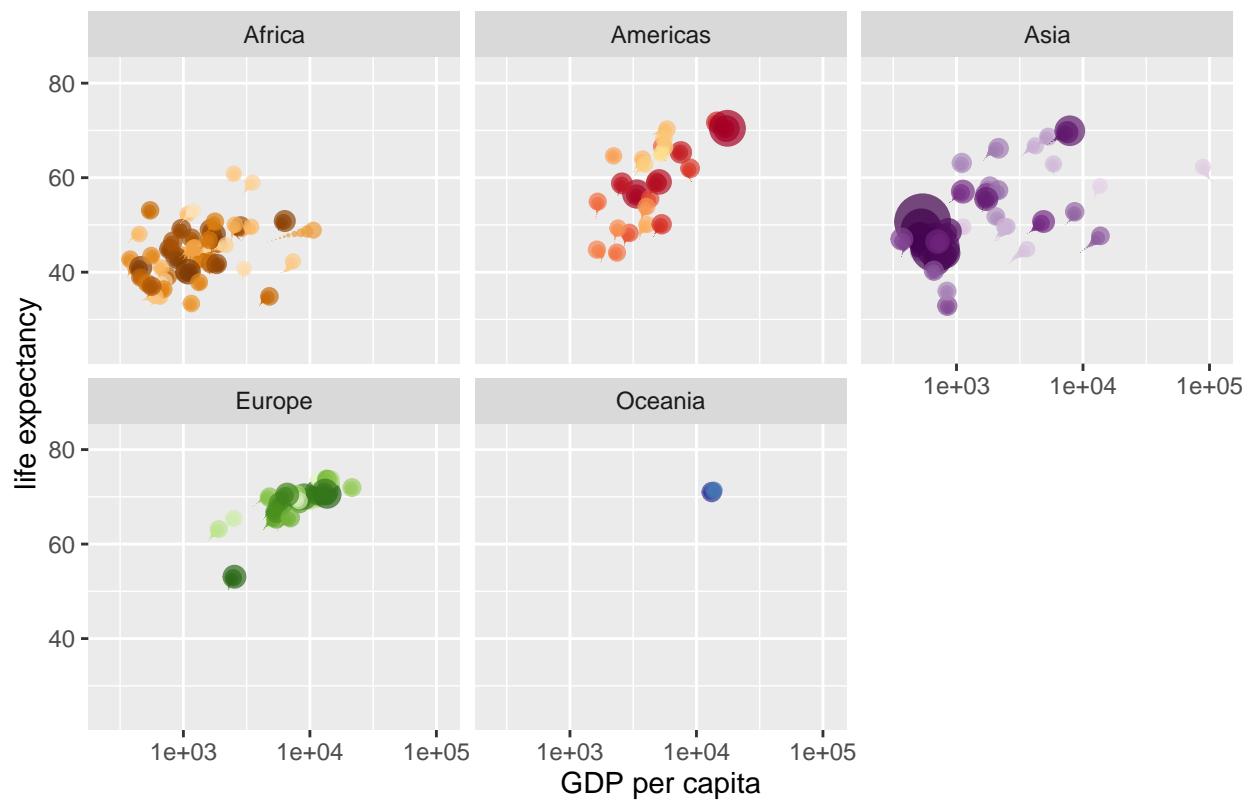
Year: 1963



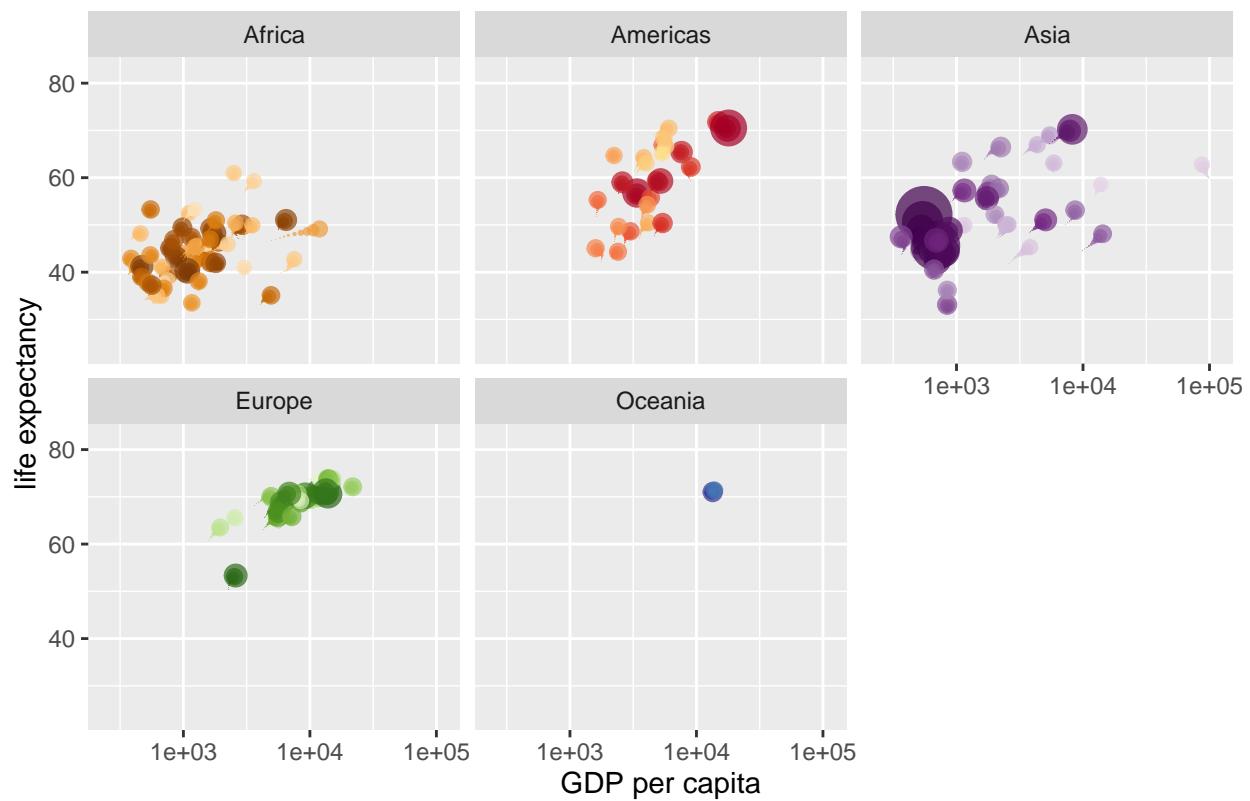
Year: 1964



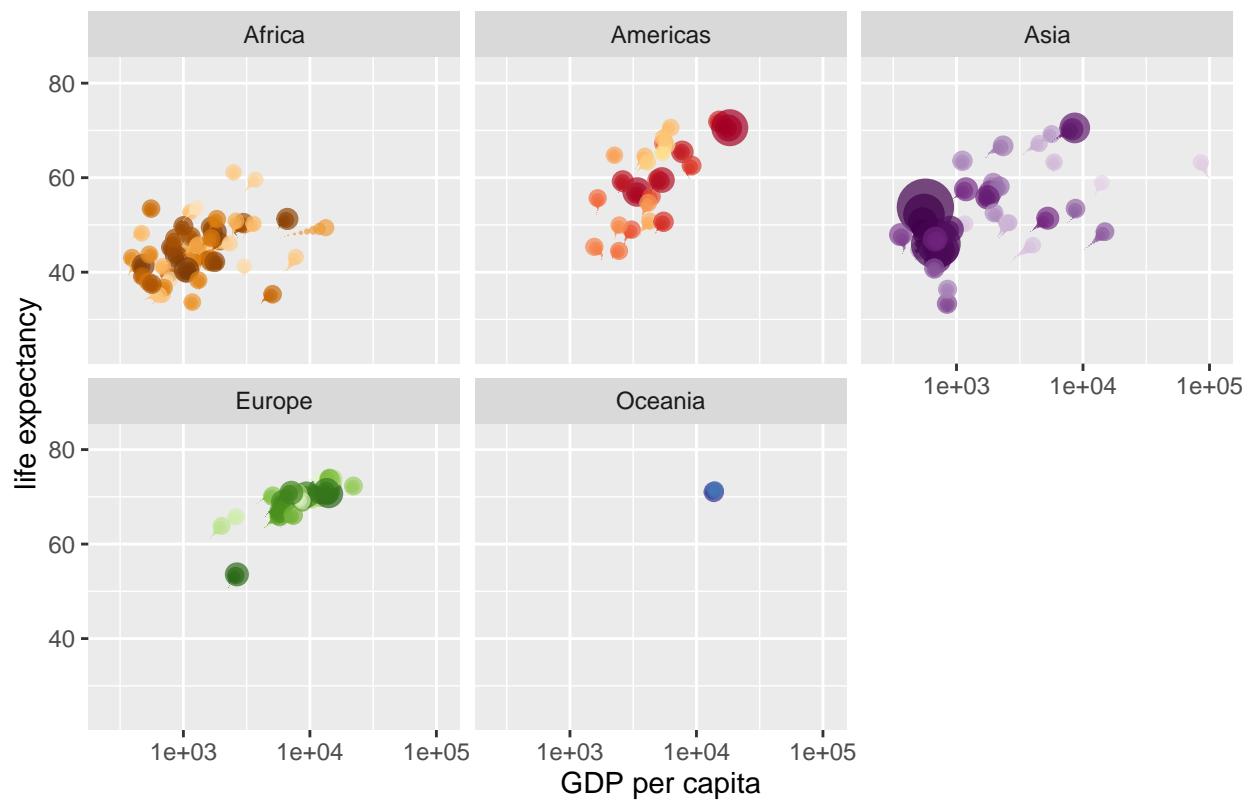
Year: 1964



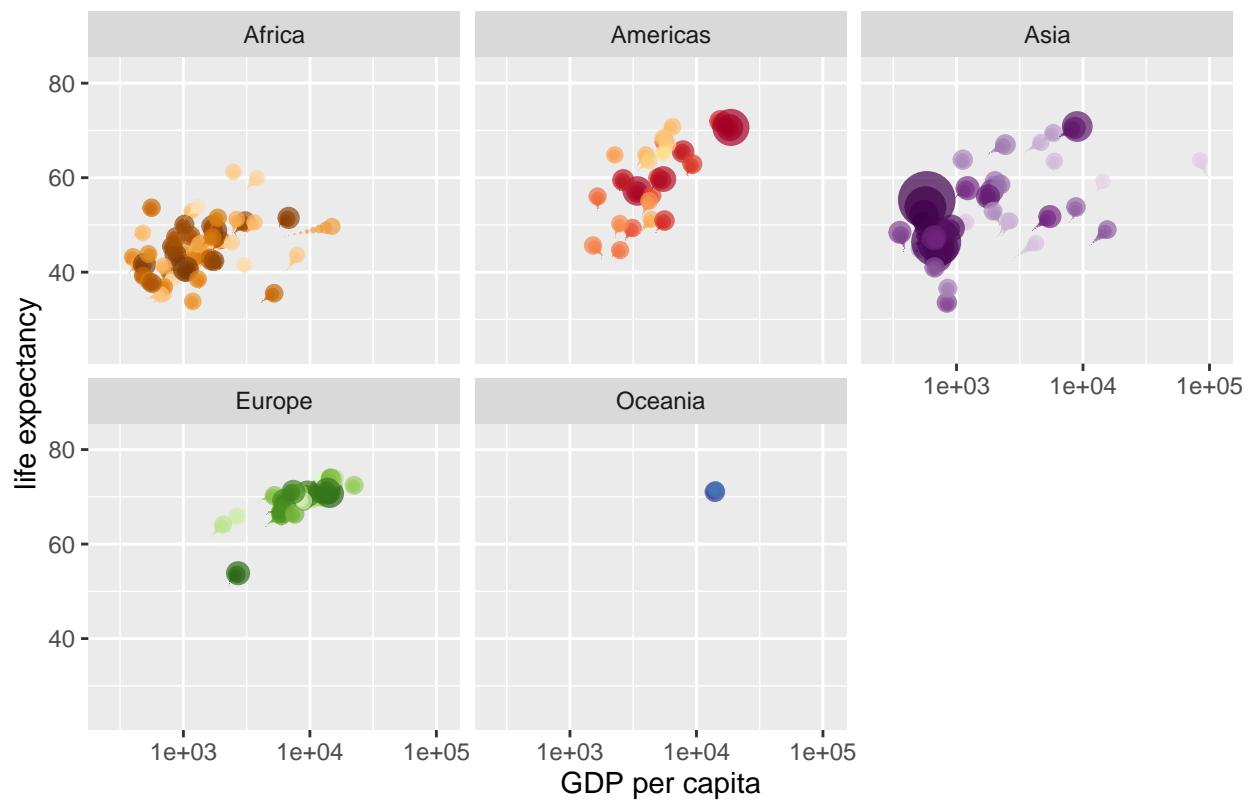
Year: 1965



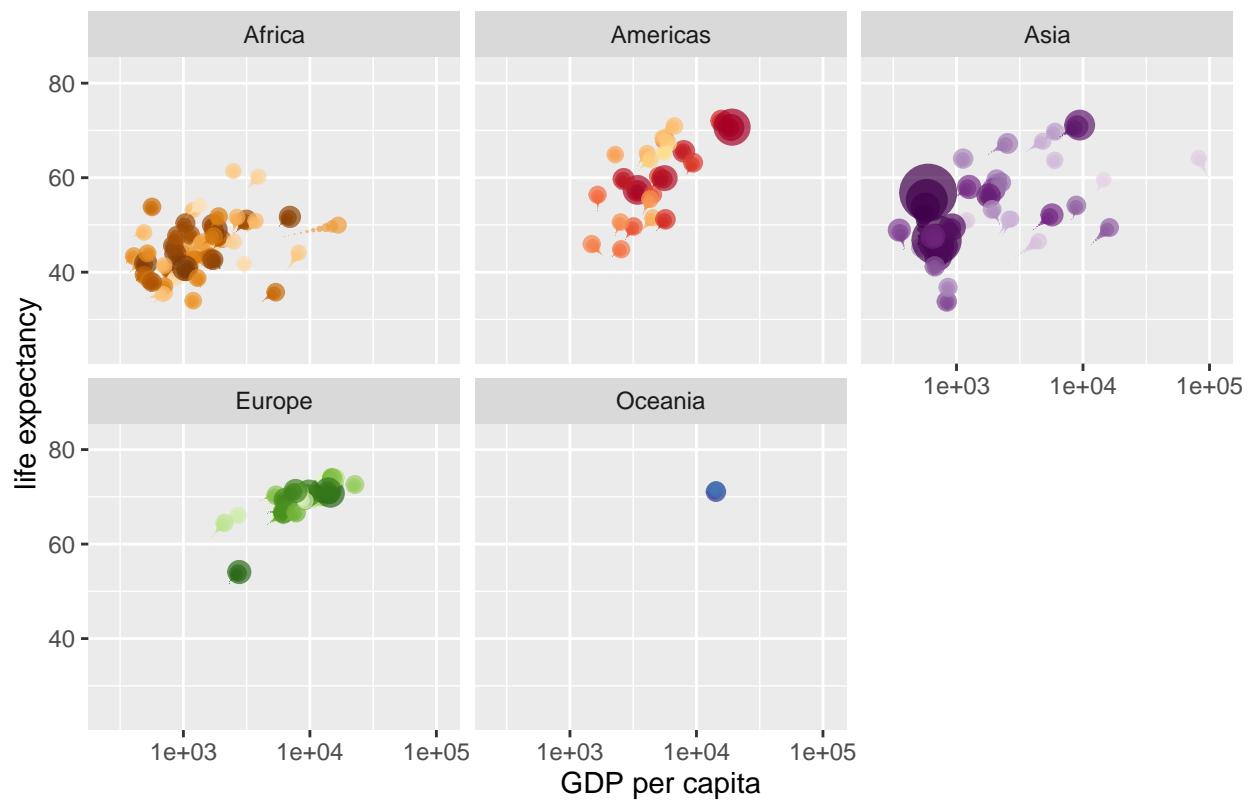
Year: 1965



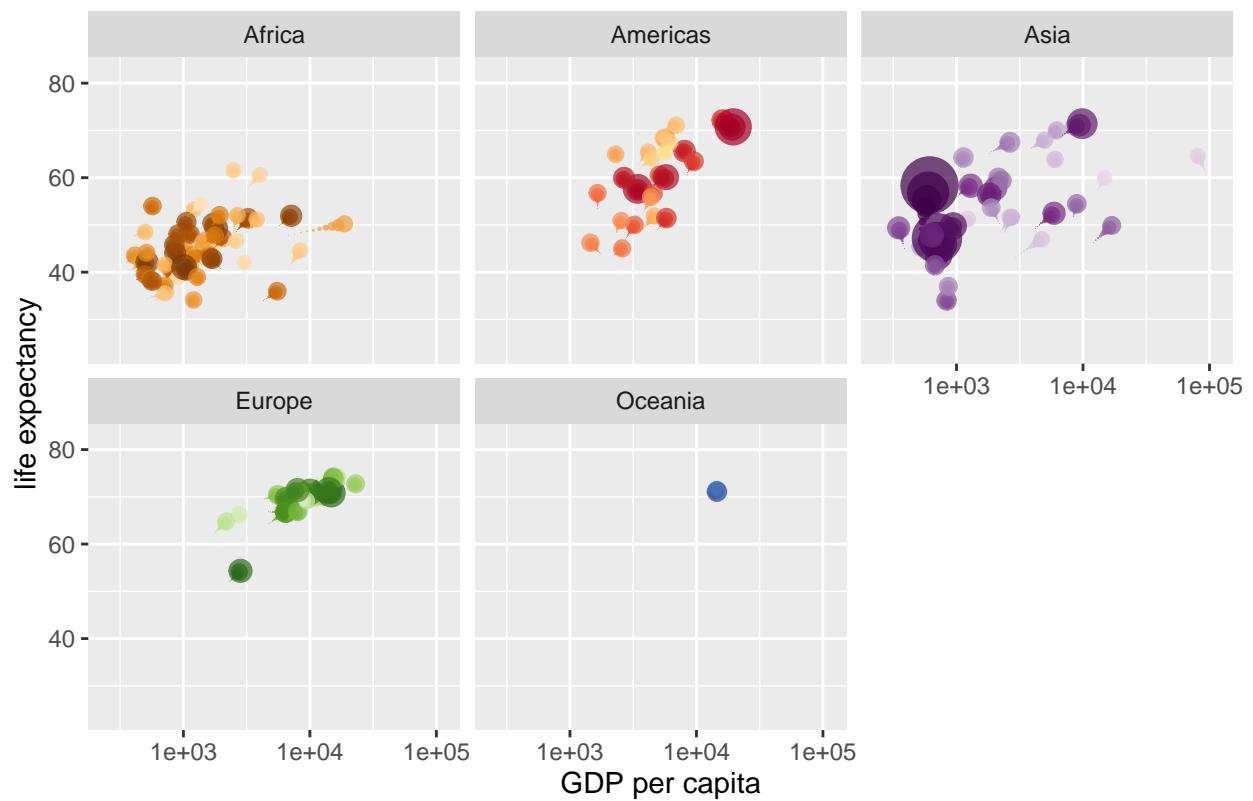
Year: 1966



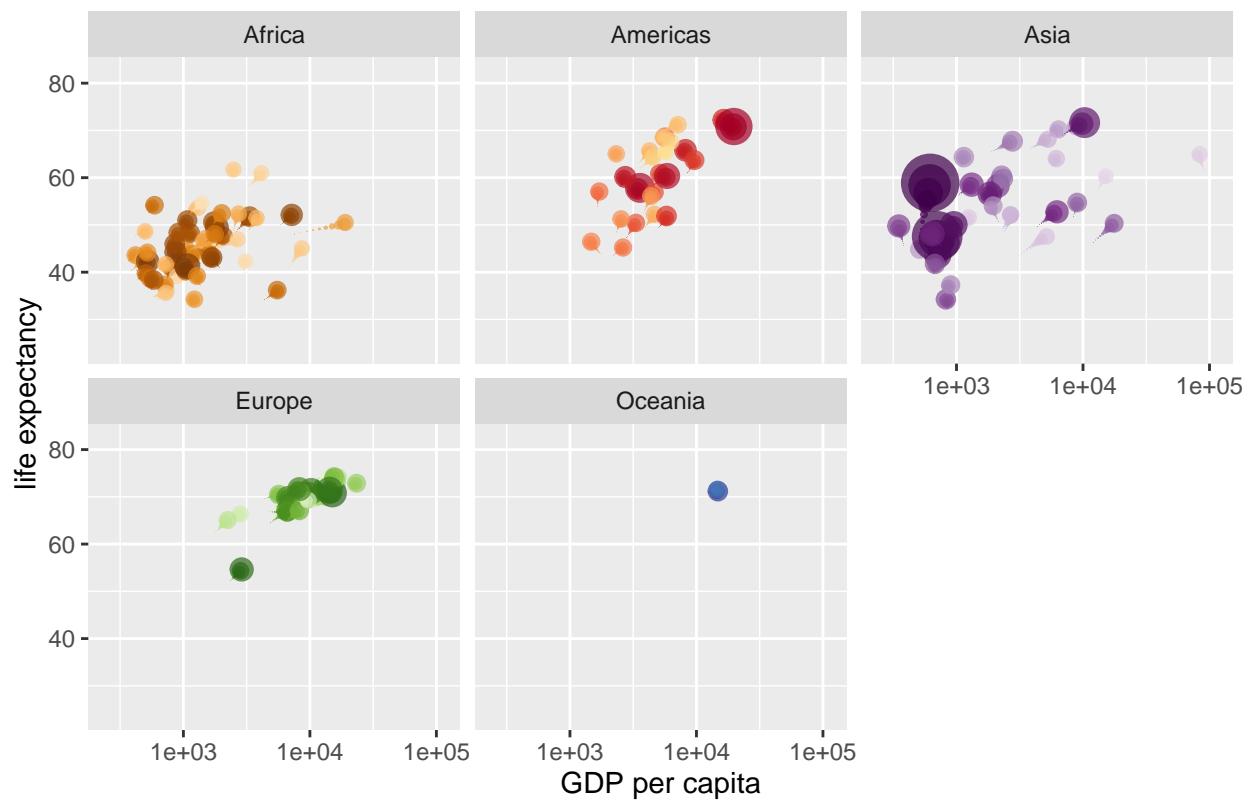
Year: 1966



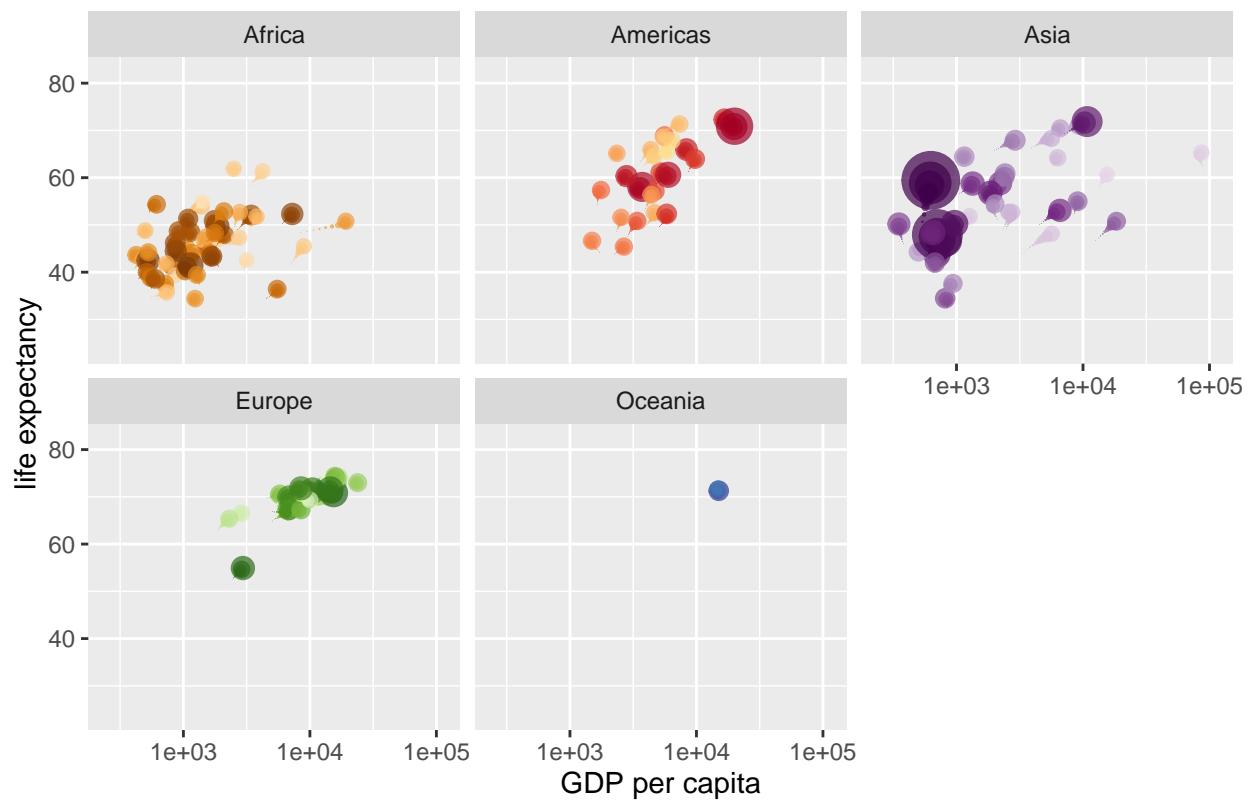
Year: 1967



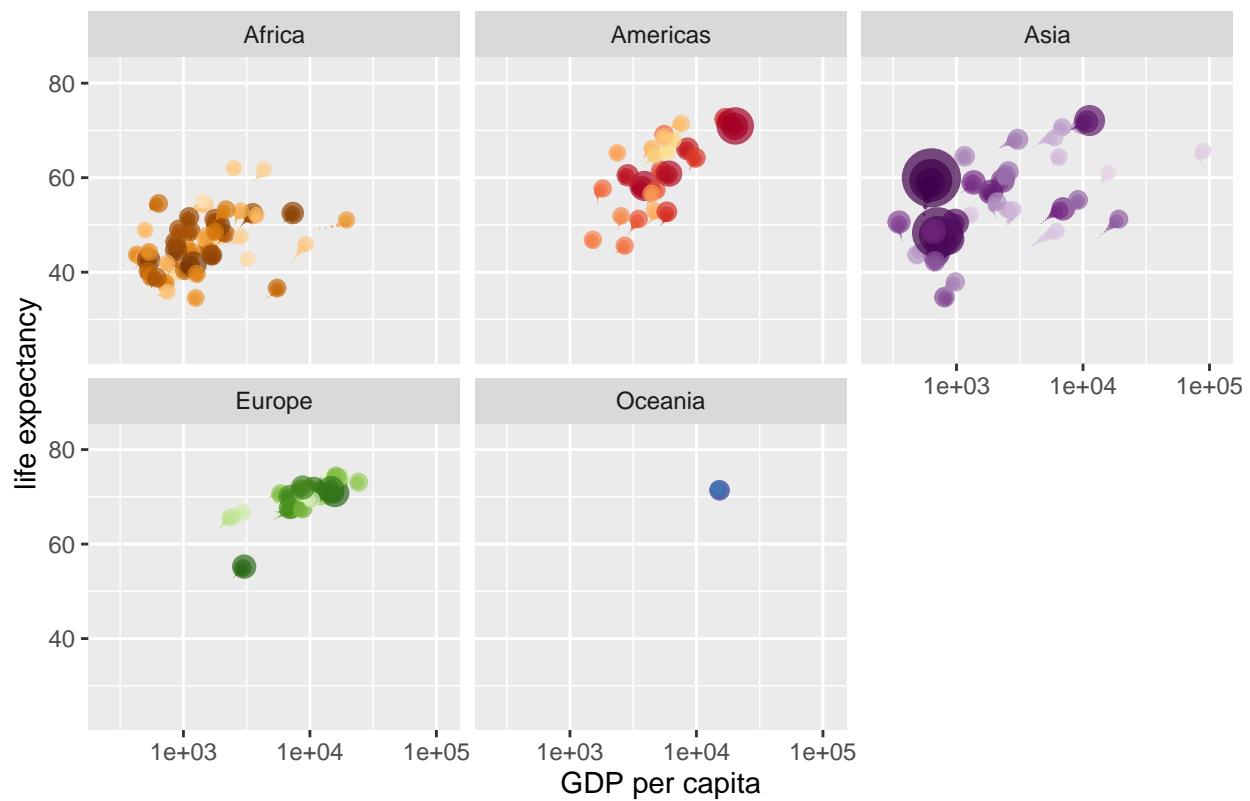
Year: 1968



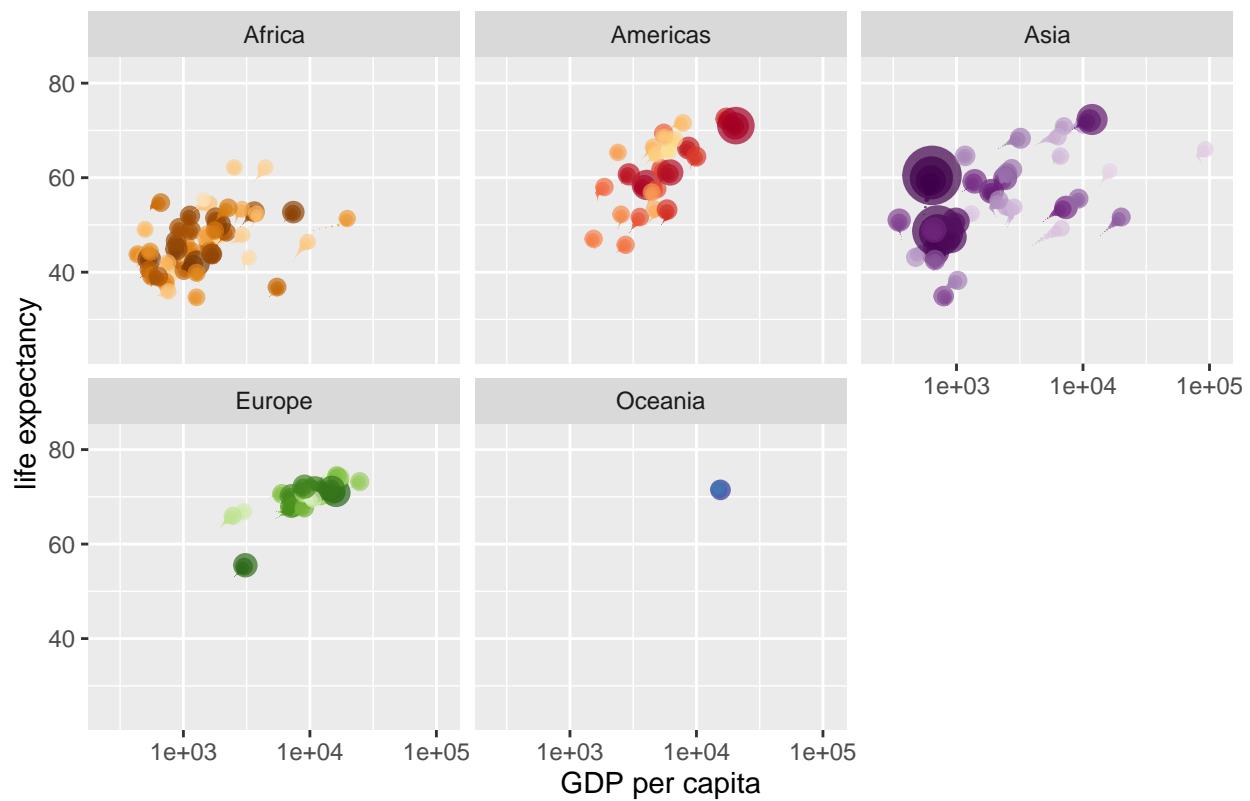
Year: 1968



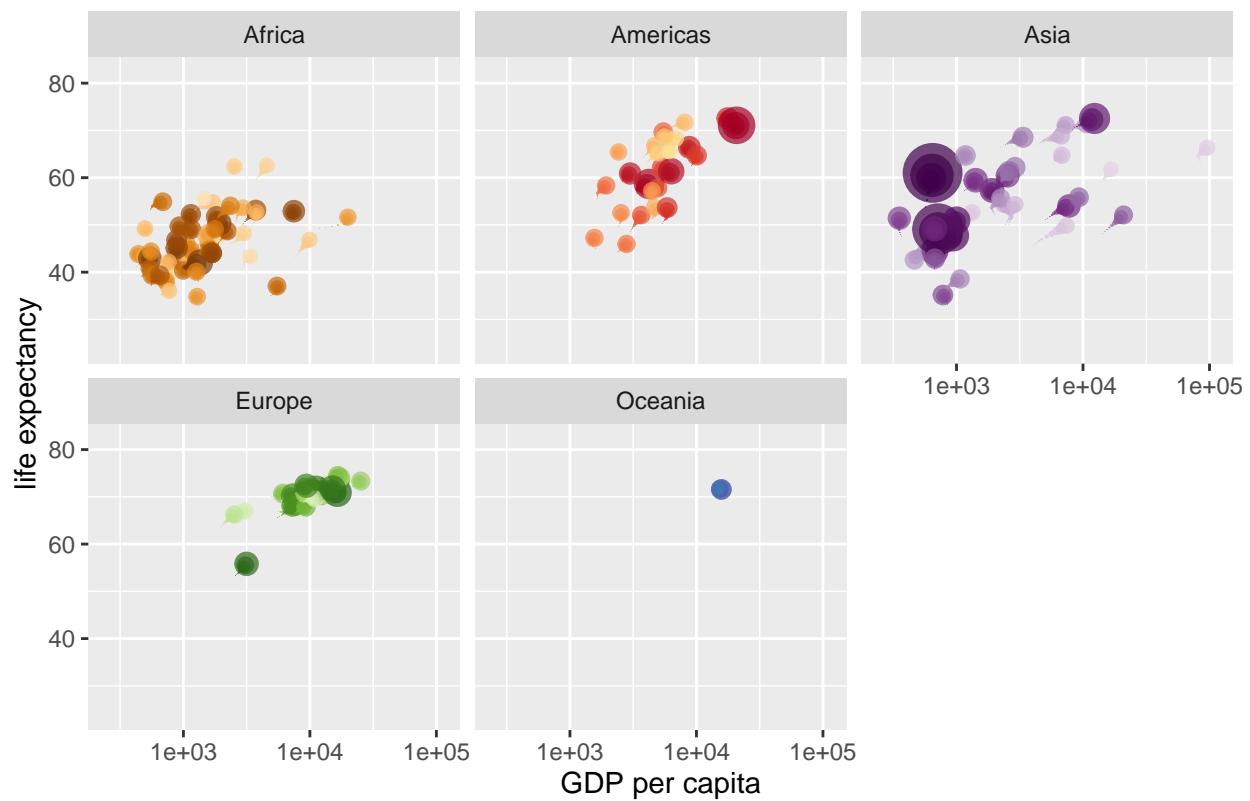
Year: 1969



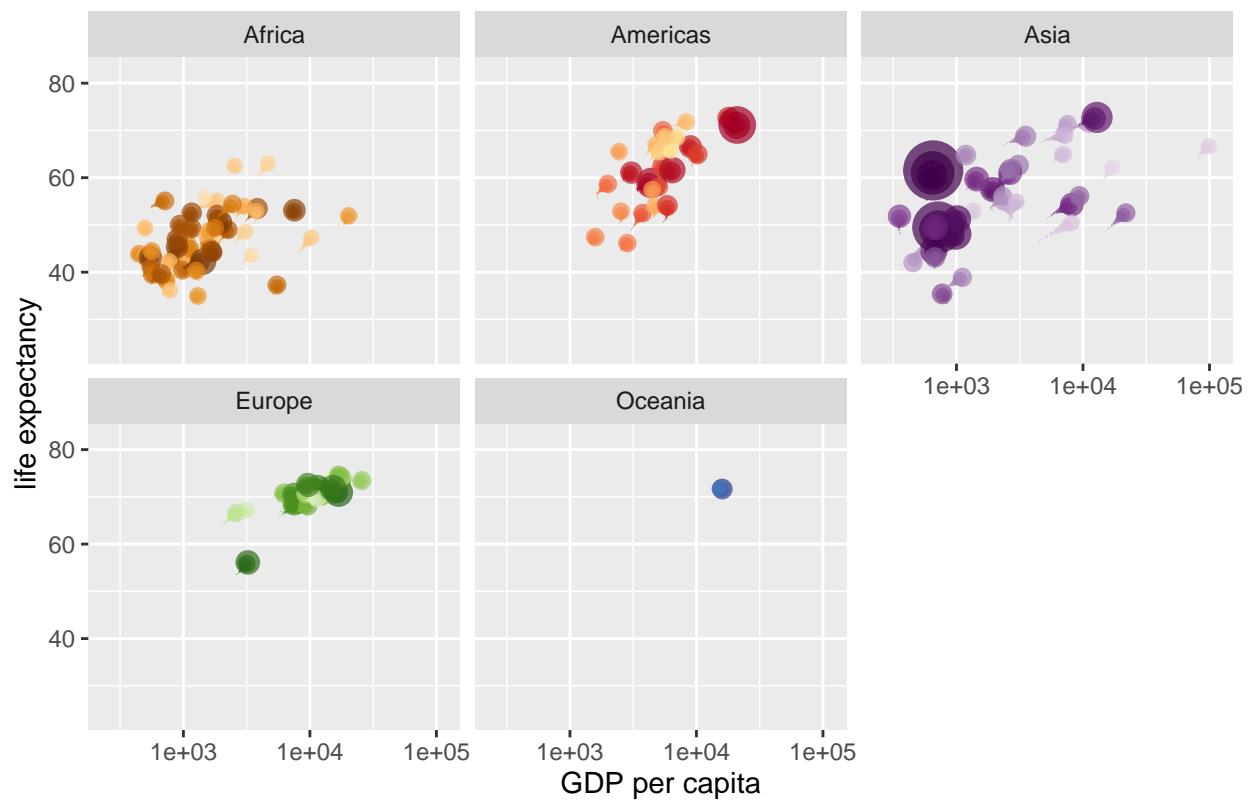
Year: 1969



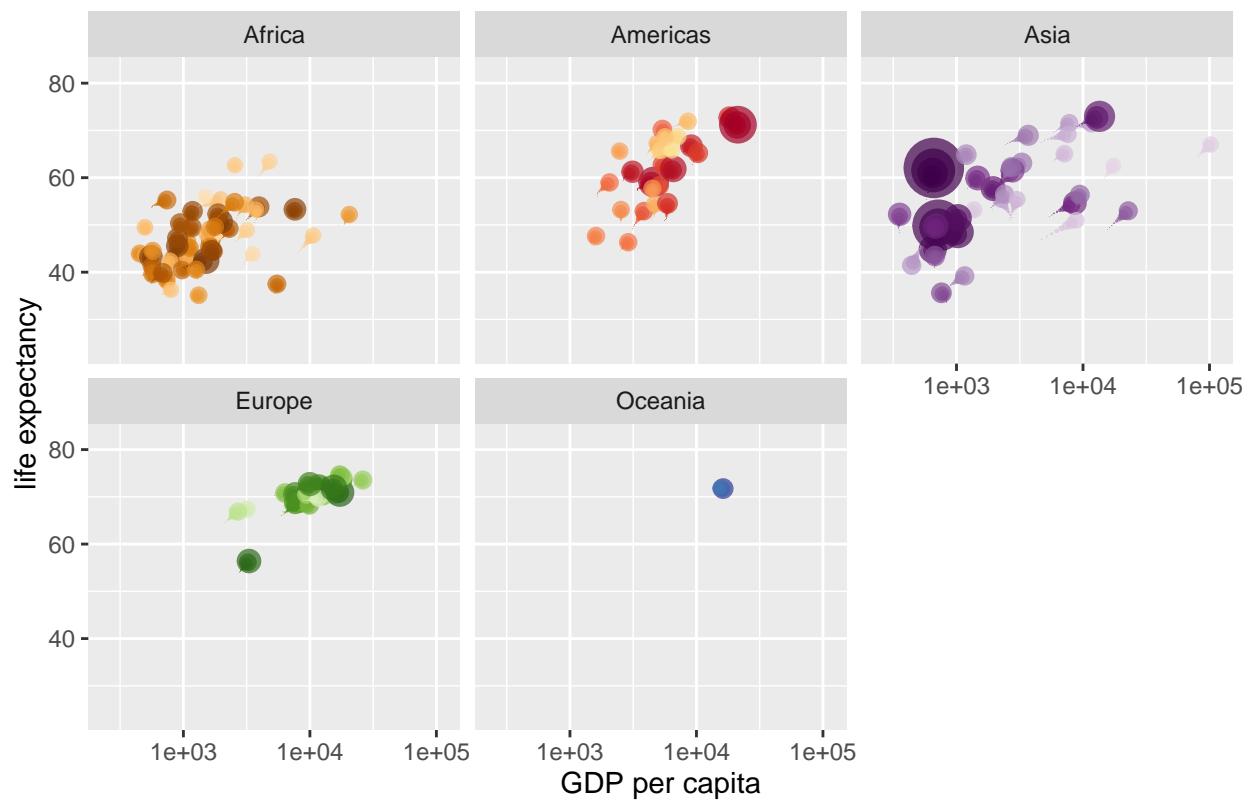
Year: 1970



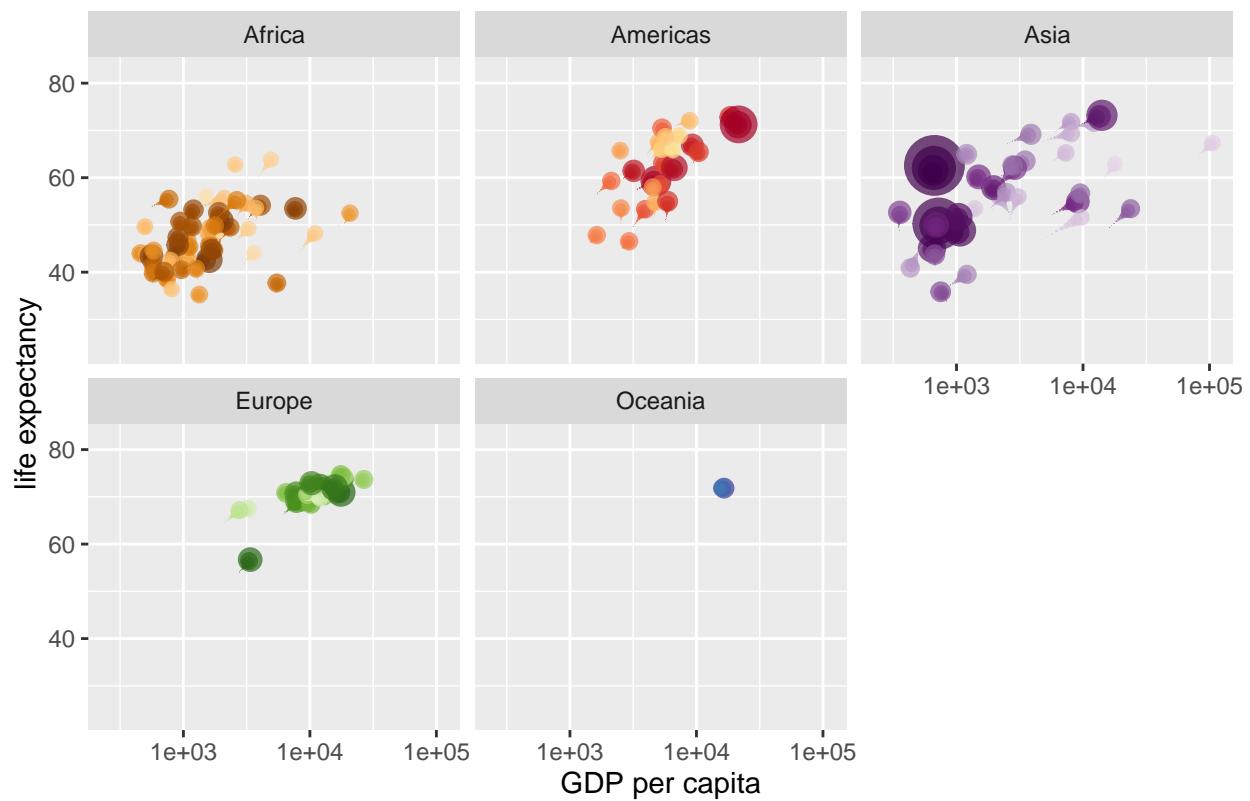
Year: 1970



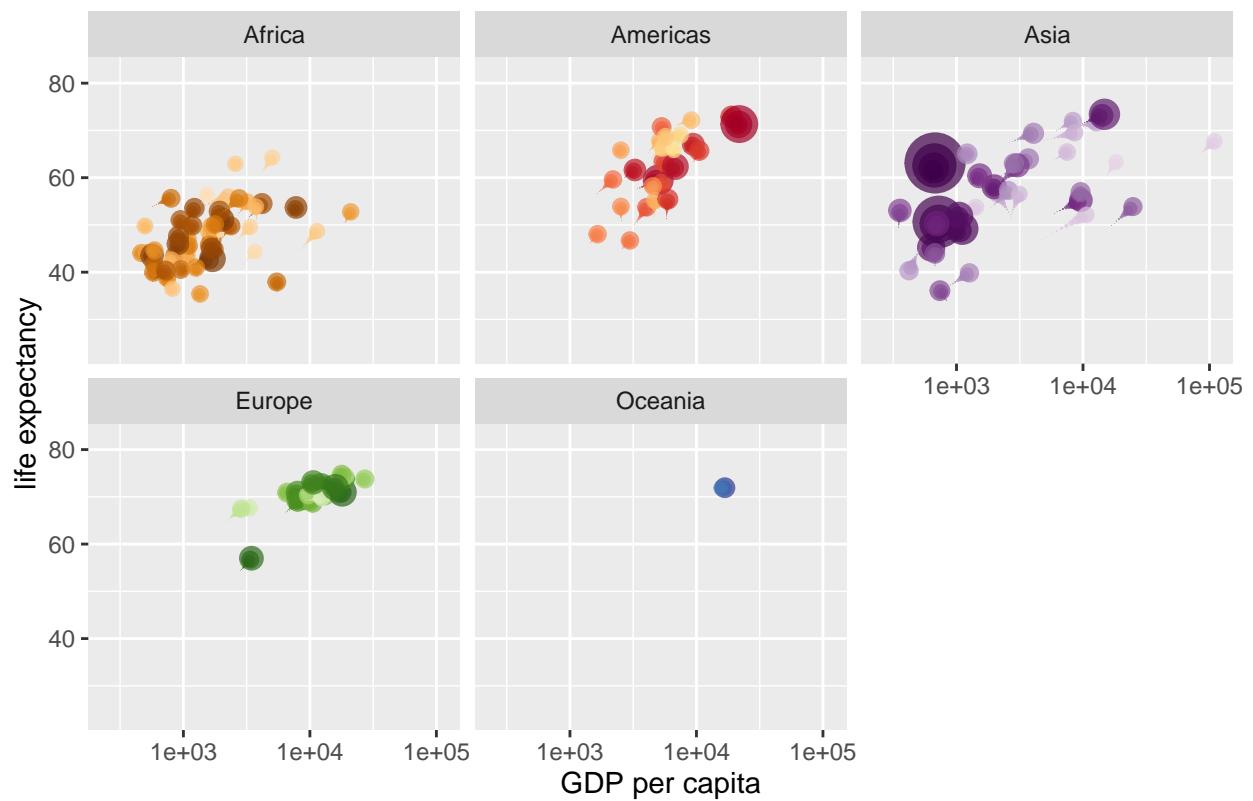
Year: 1971



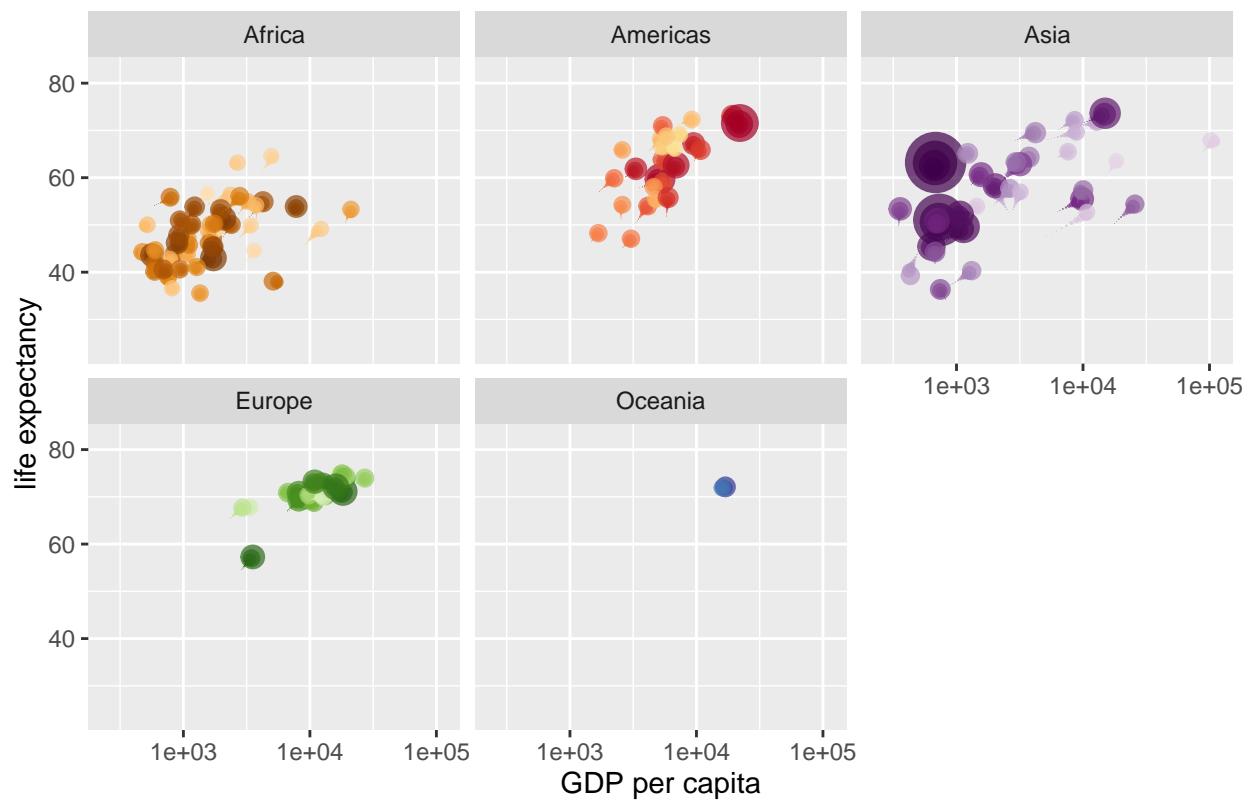
Year: 1971



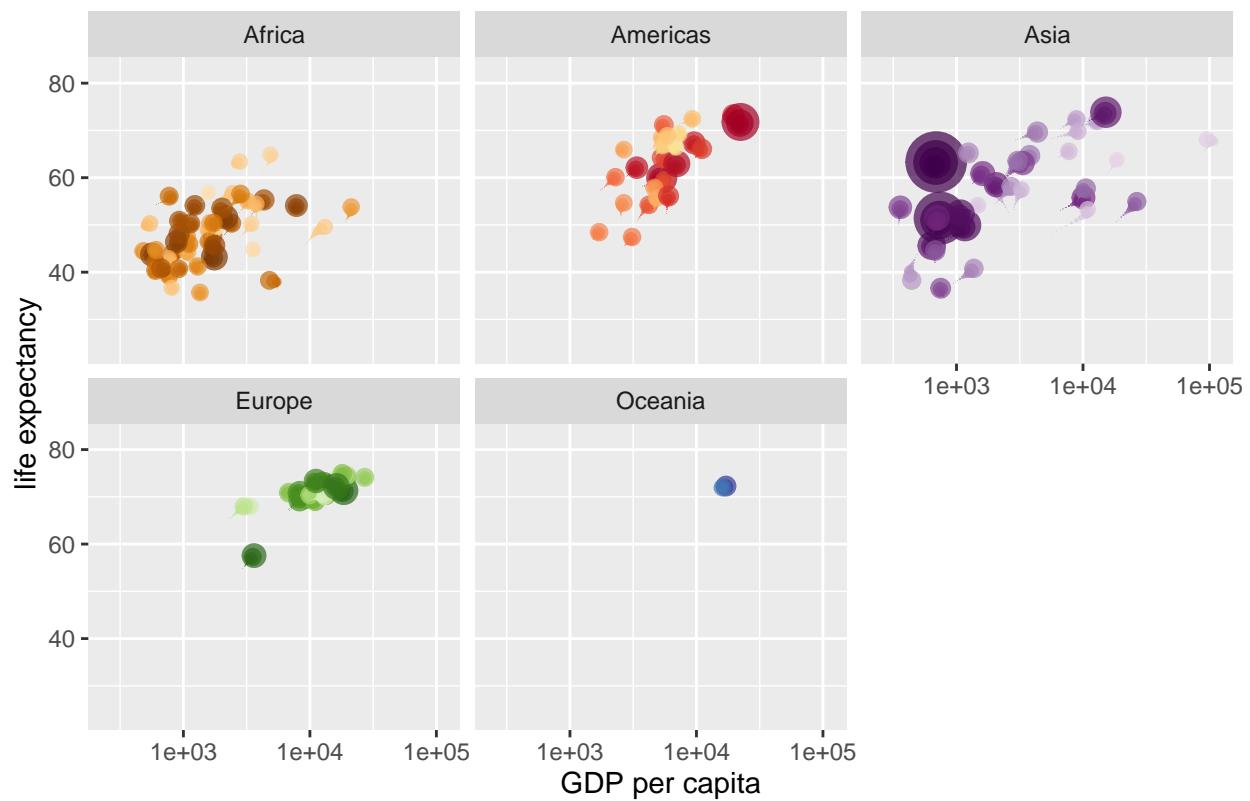
Year: 1972



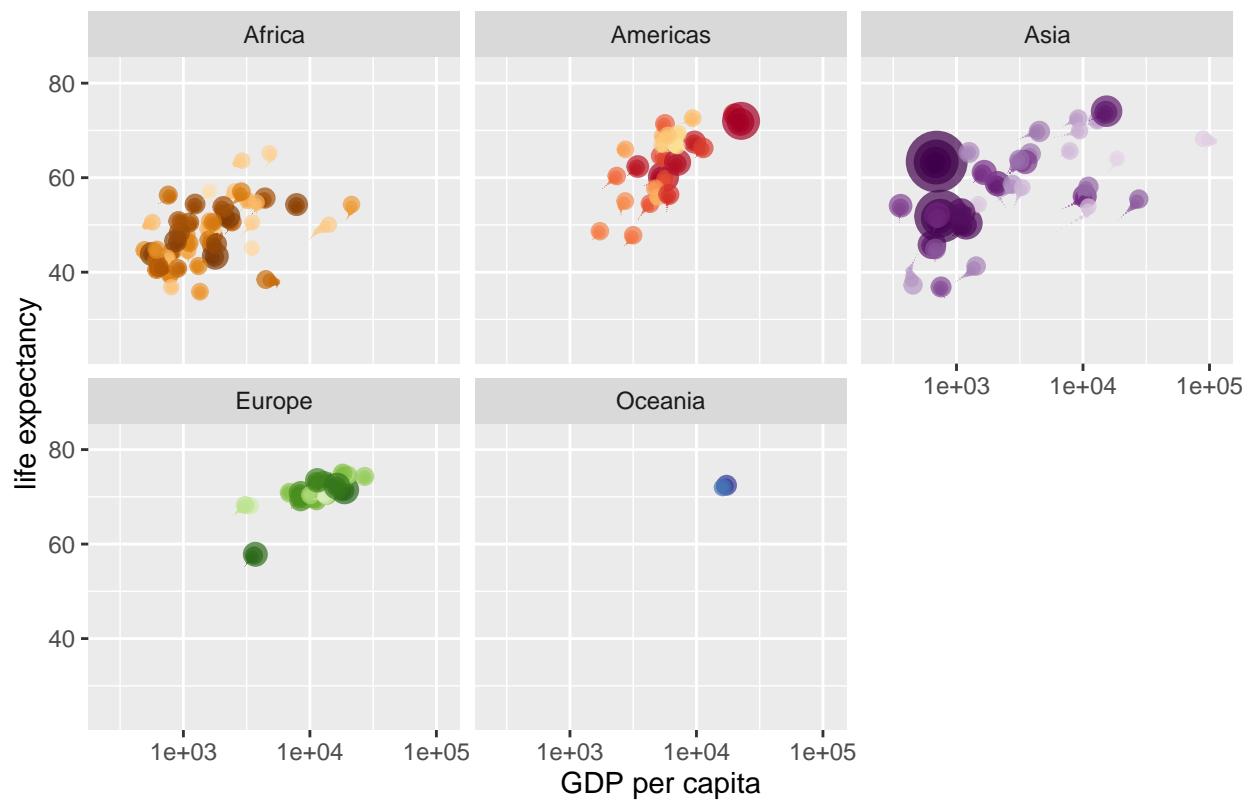
Year: 1973



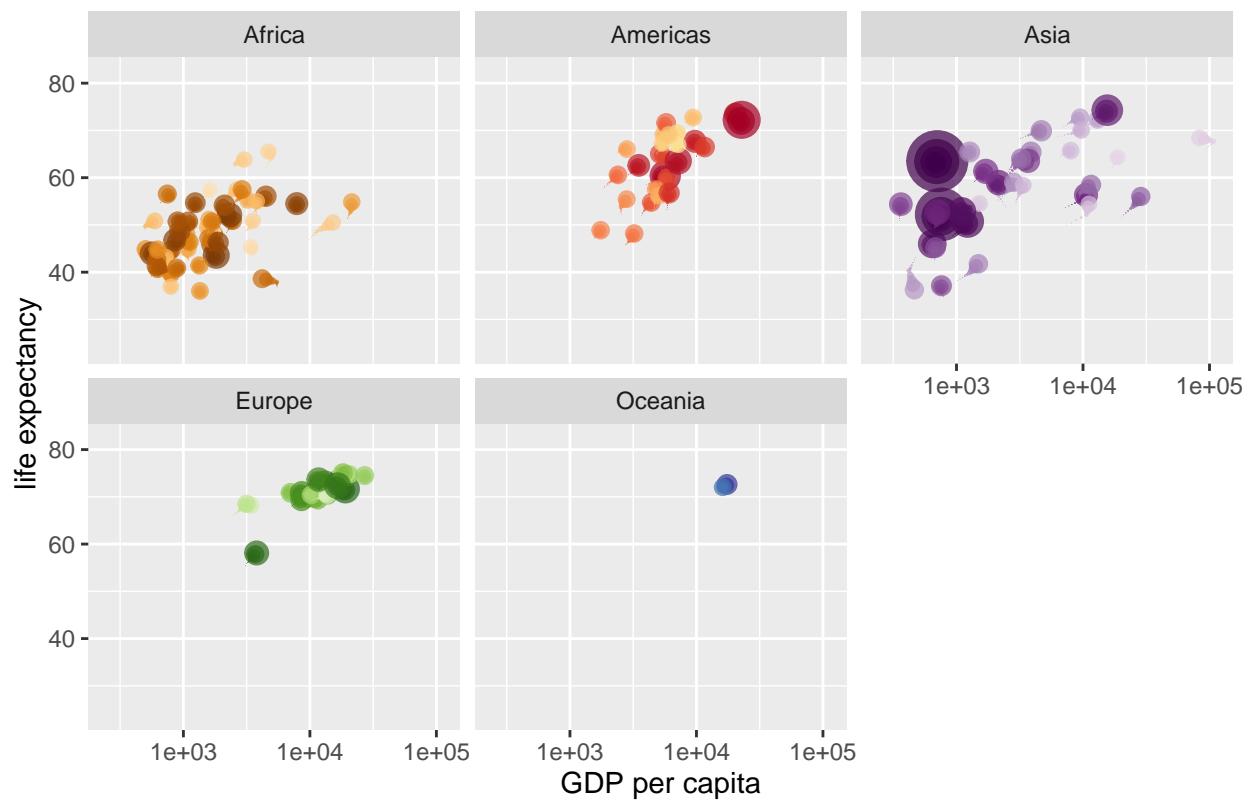
Year: 1973



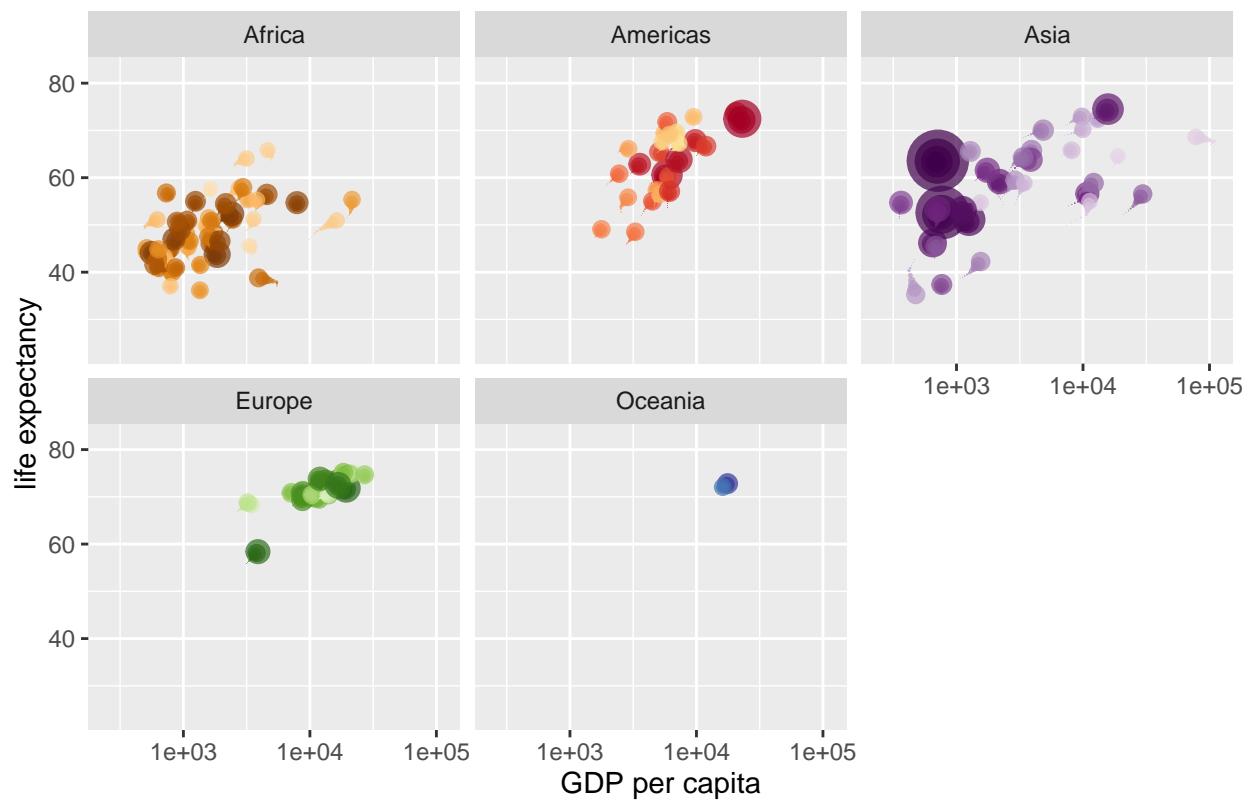
Year: 1974



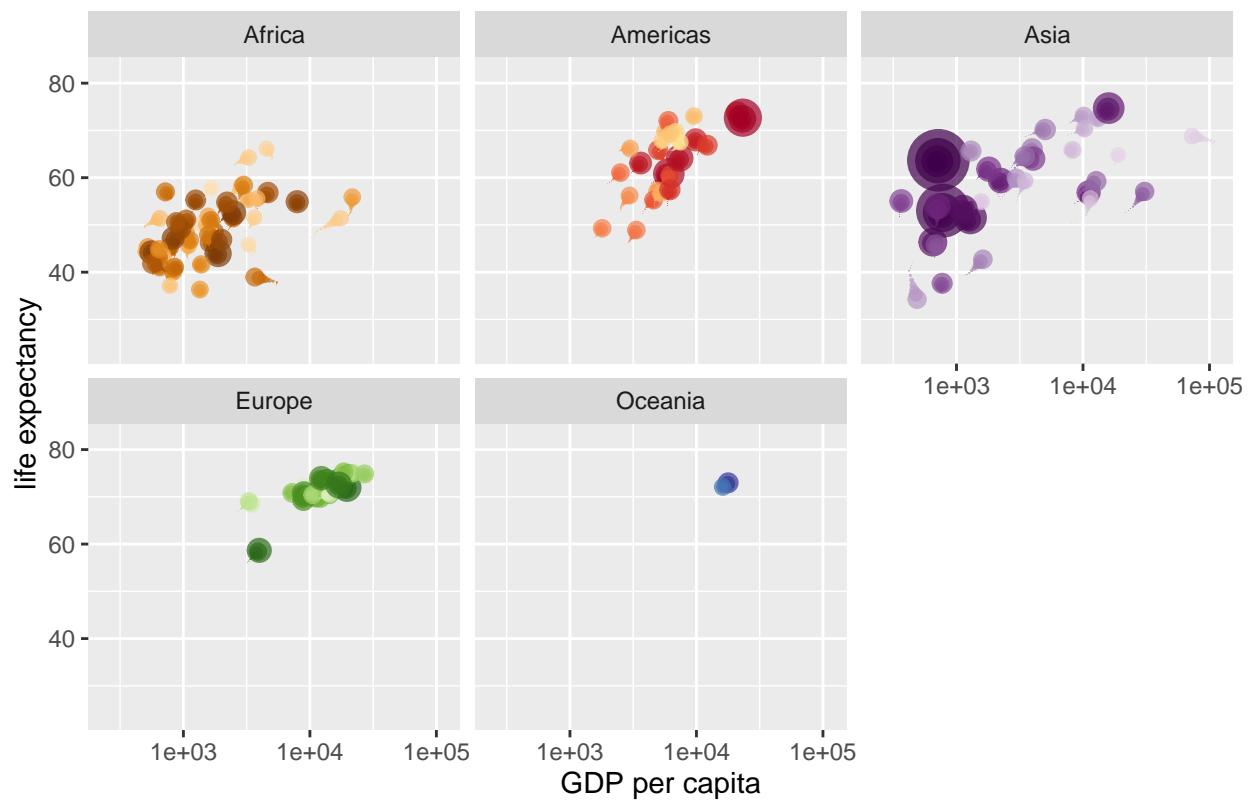
Year: 1974



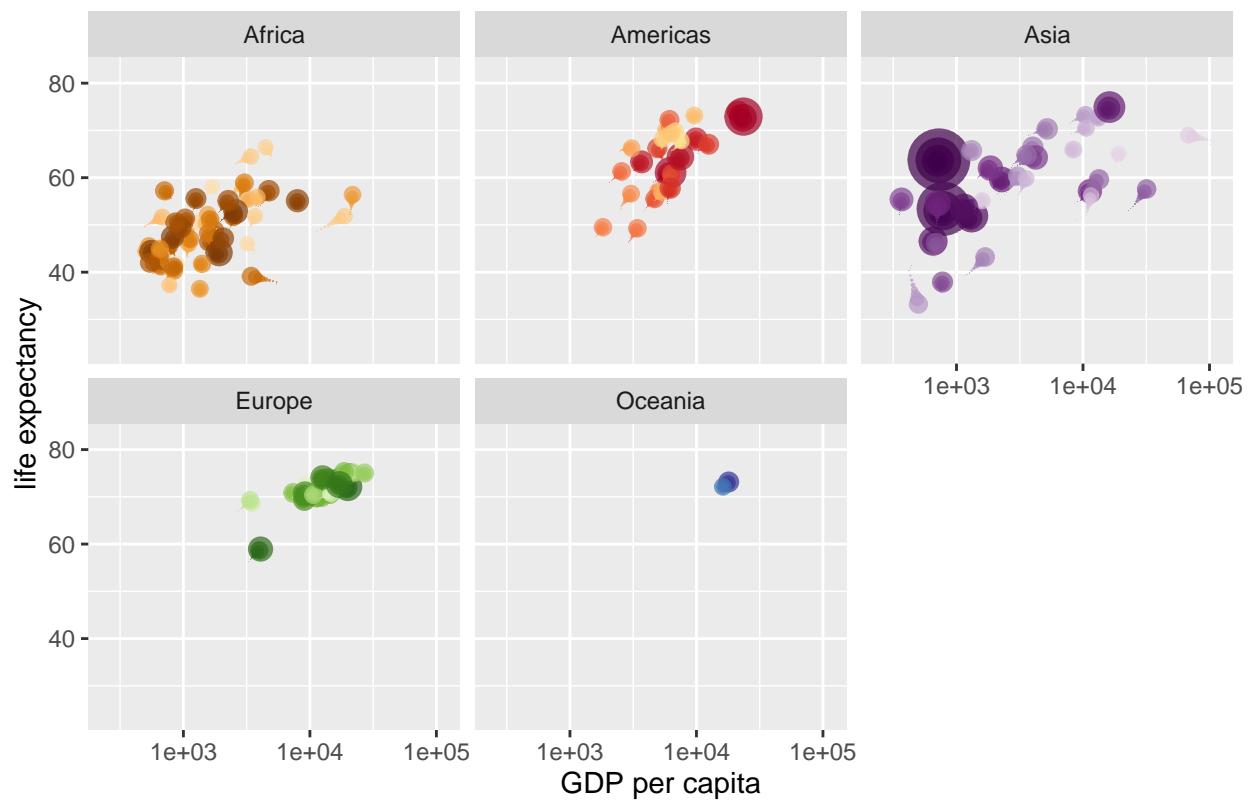
Year: 1975



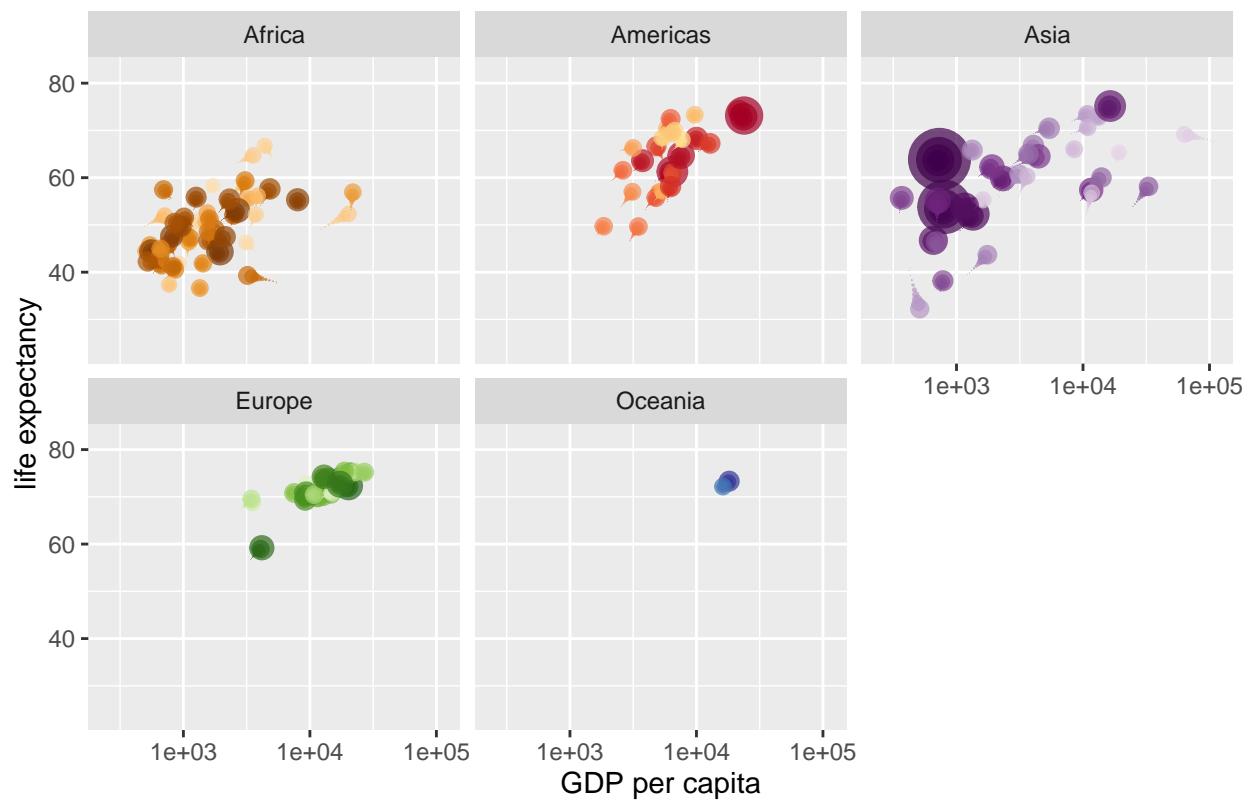
Year: 1975



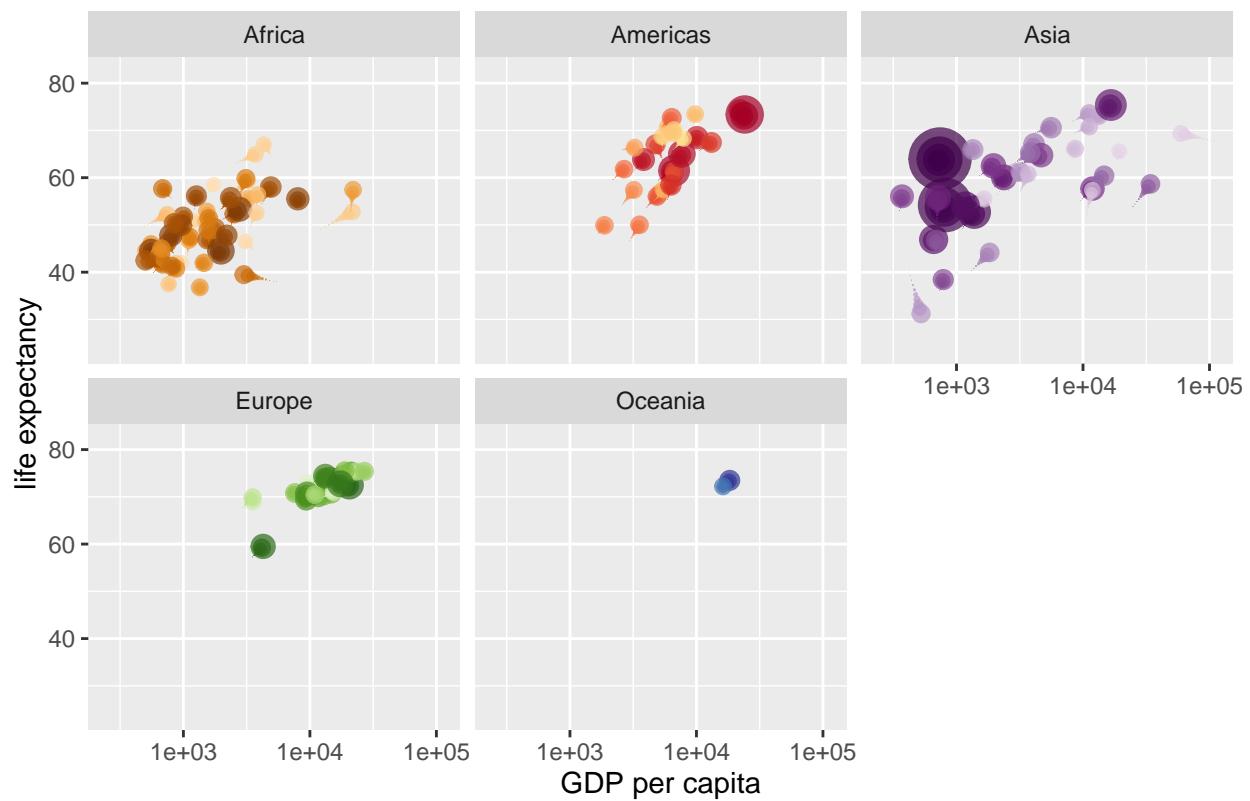
Year: 1976



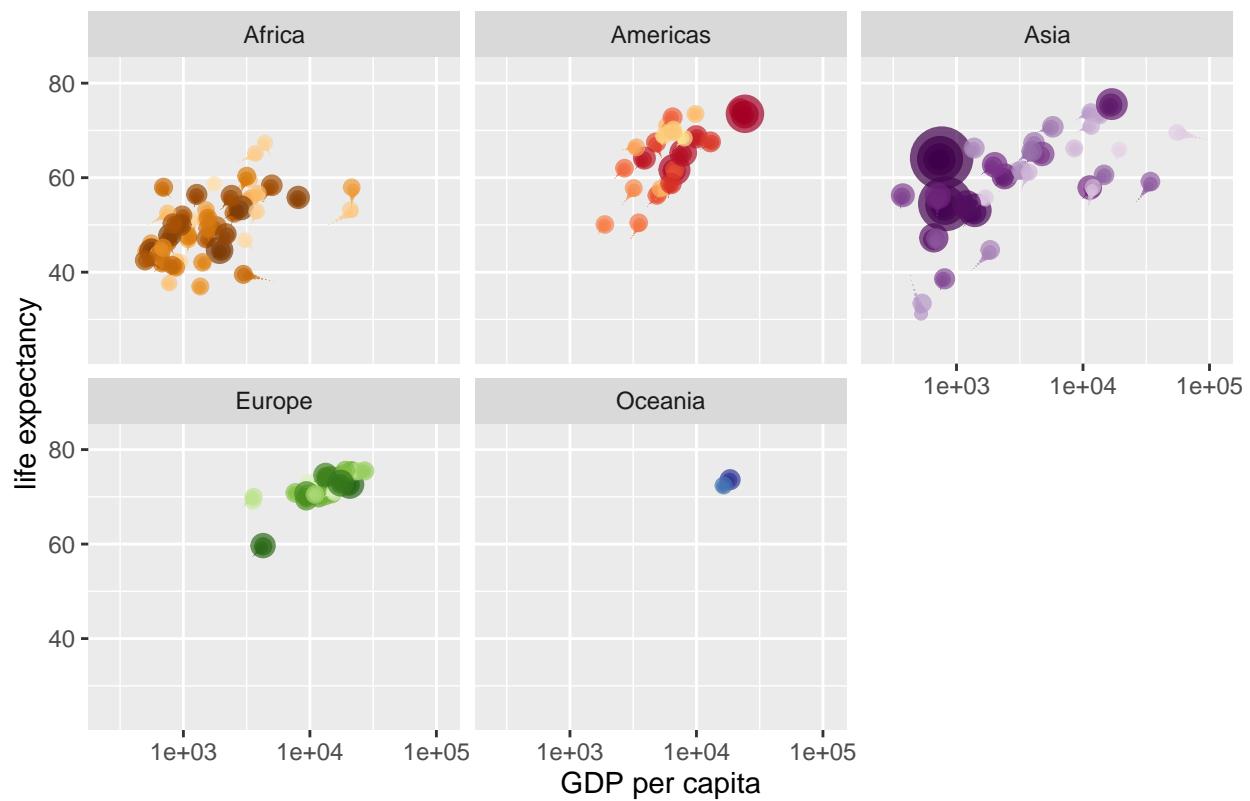
Year: 1976



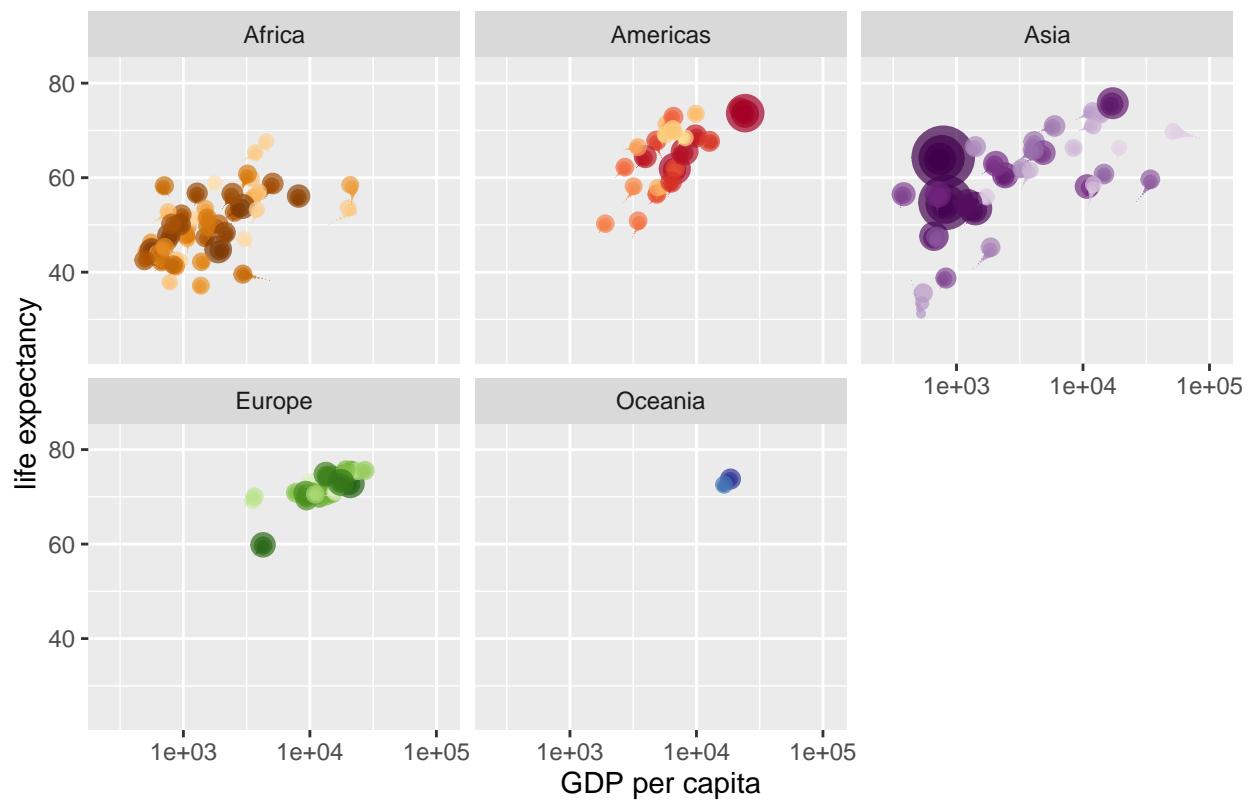
Year: 1977



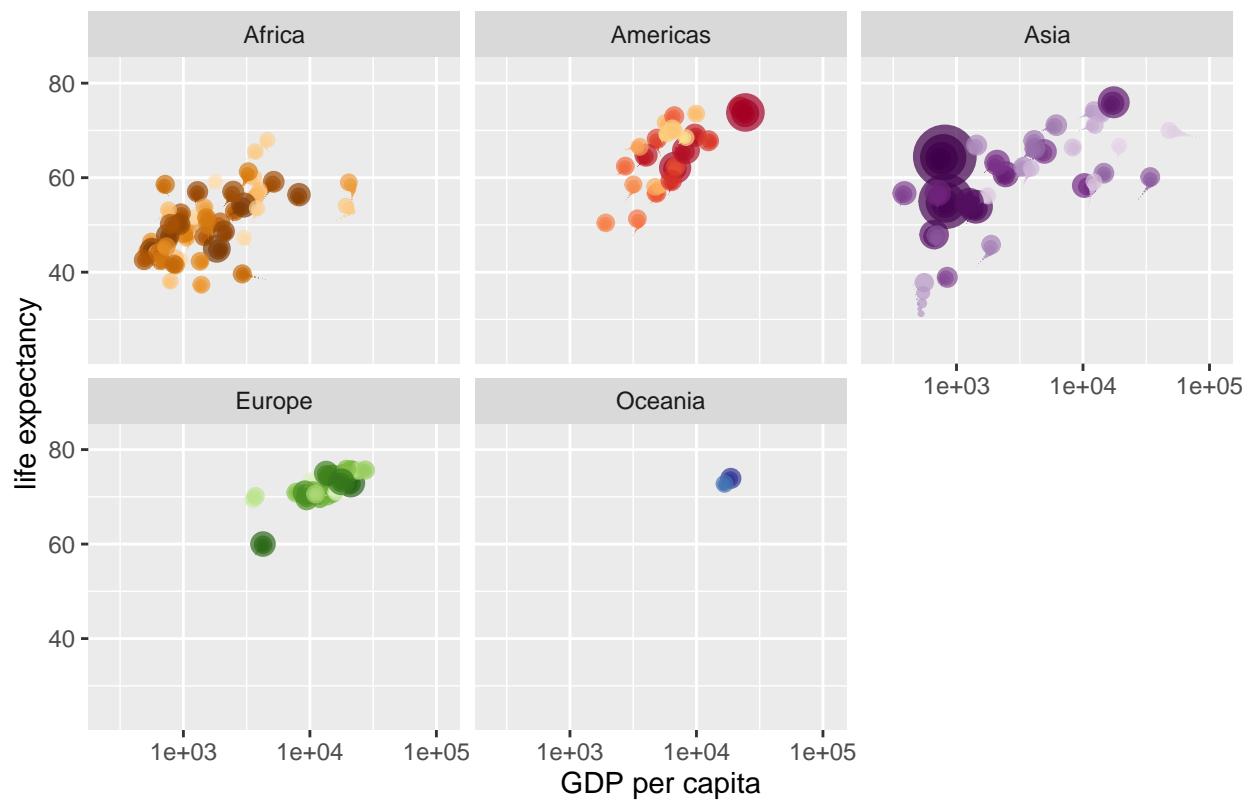
Year: 1978



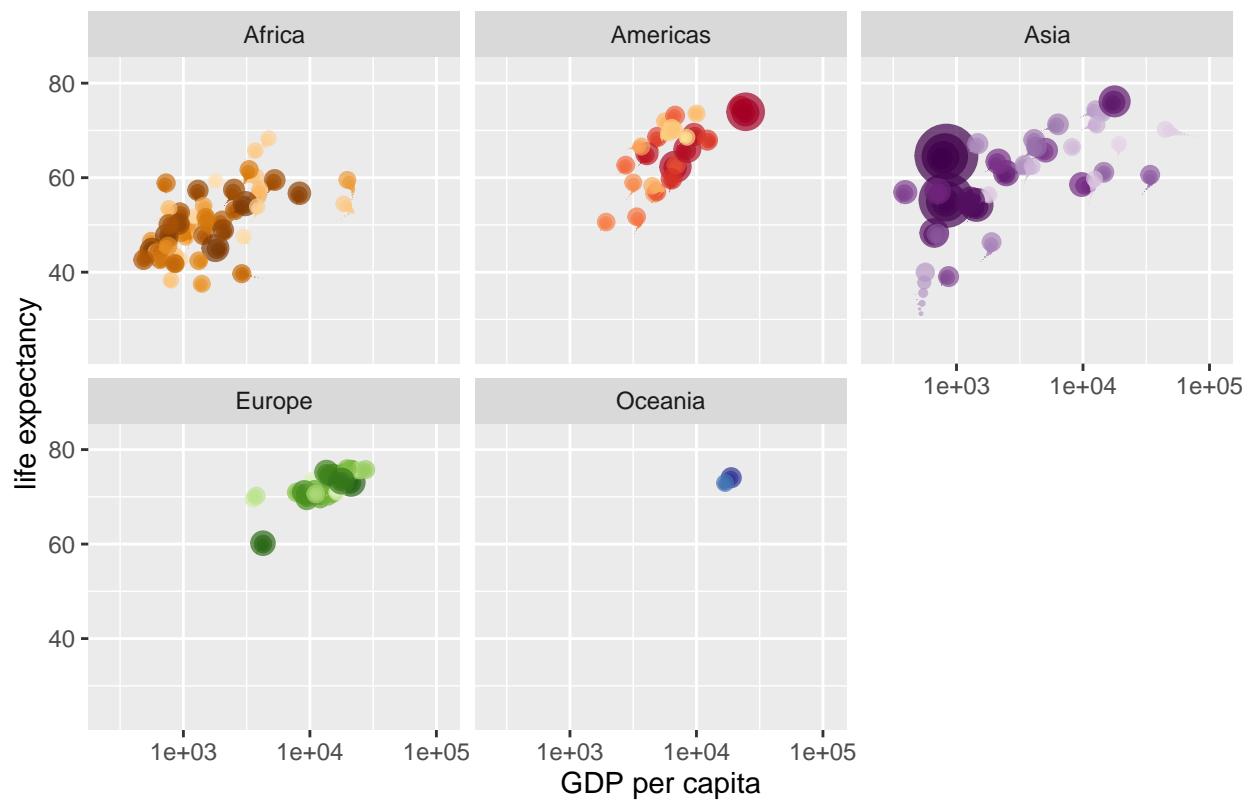
Year: 1978



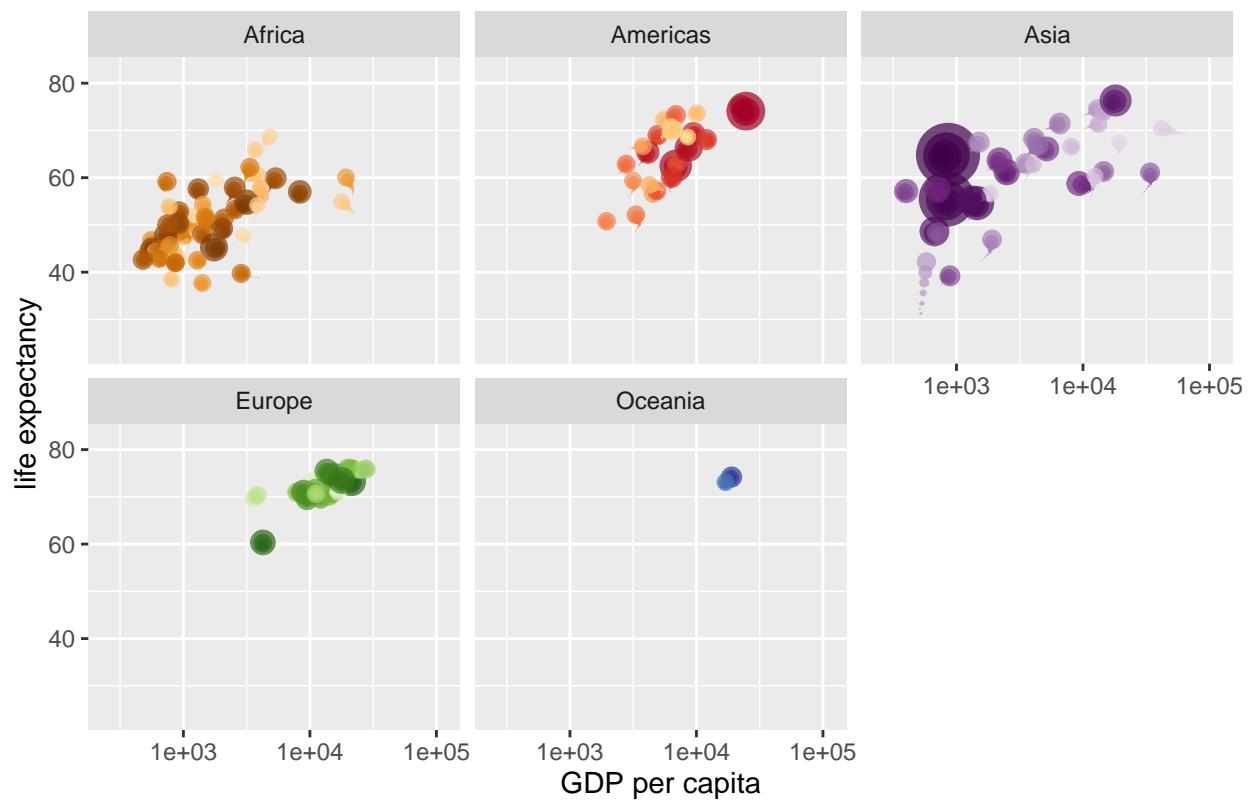
Year: 1979



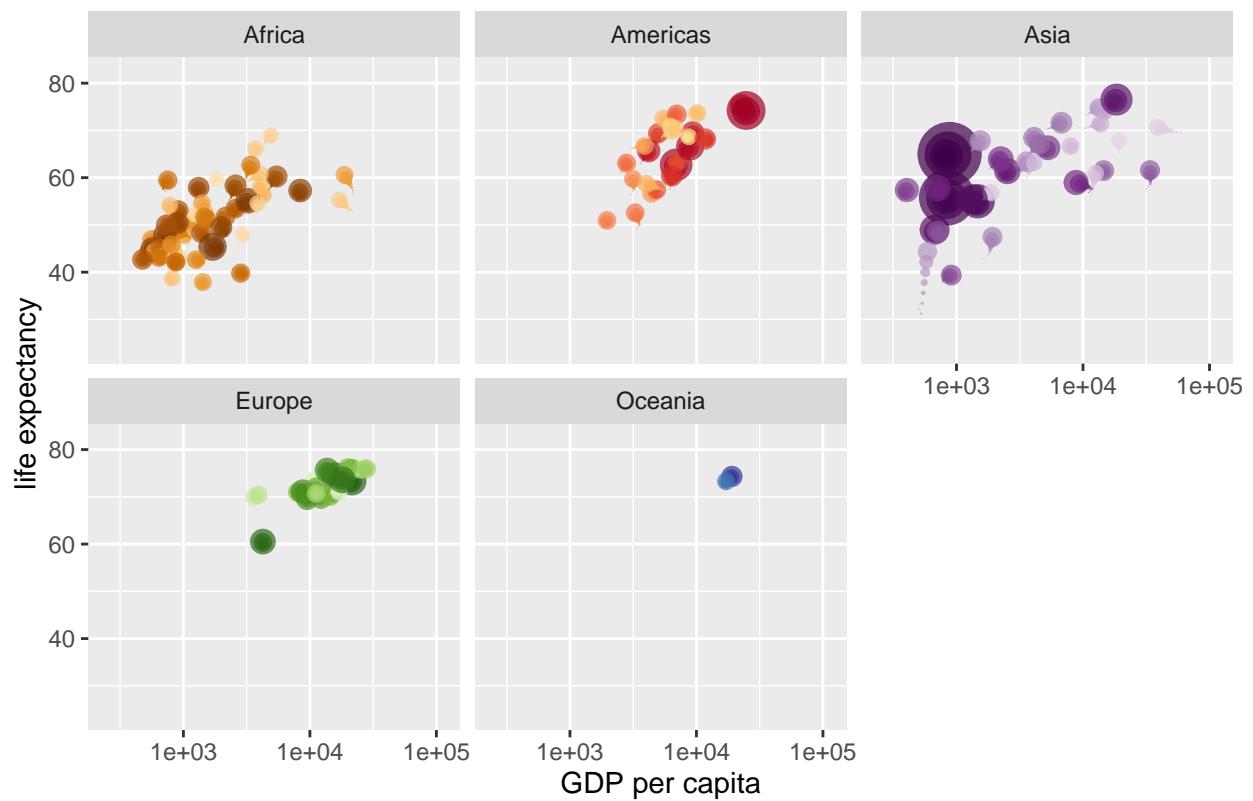
Year: 1979



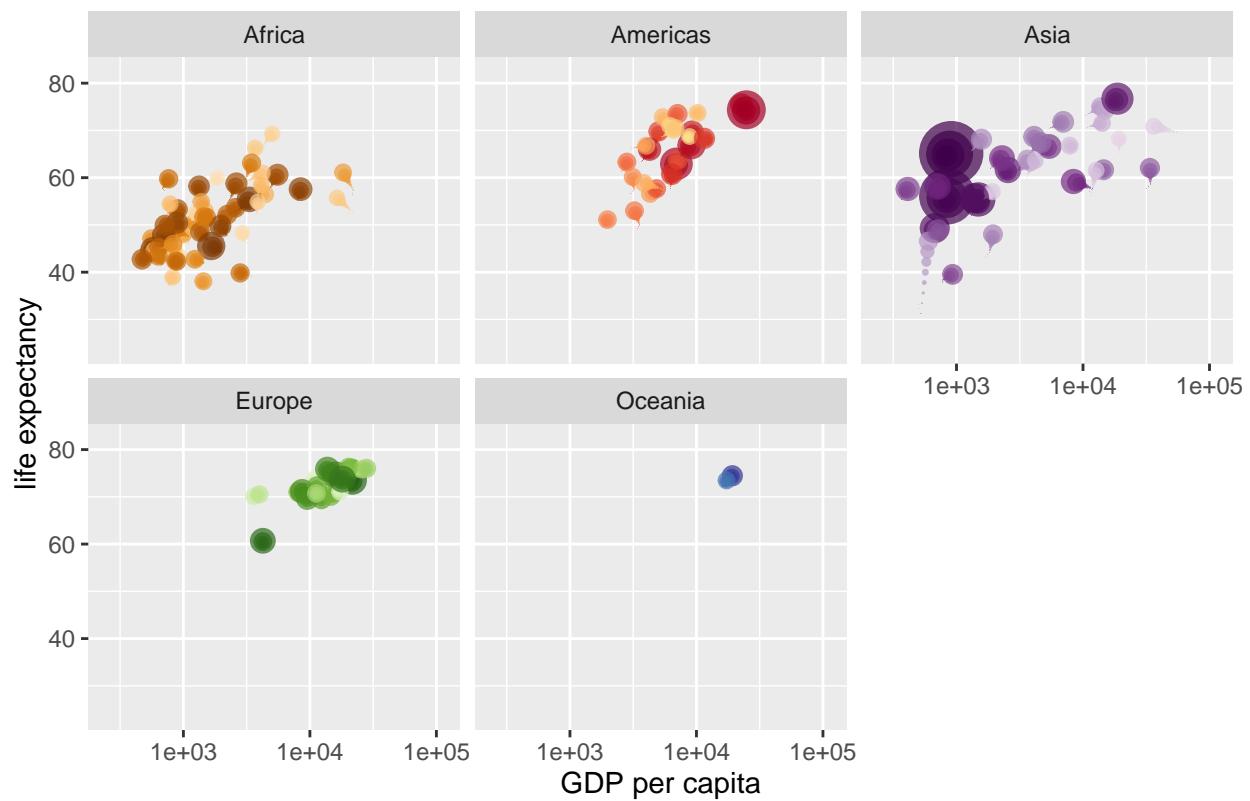
Year: 1980



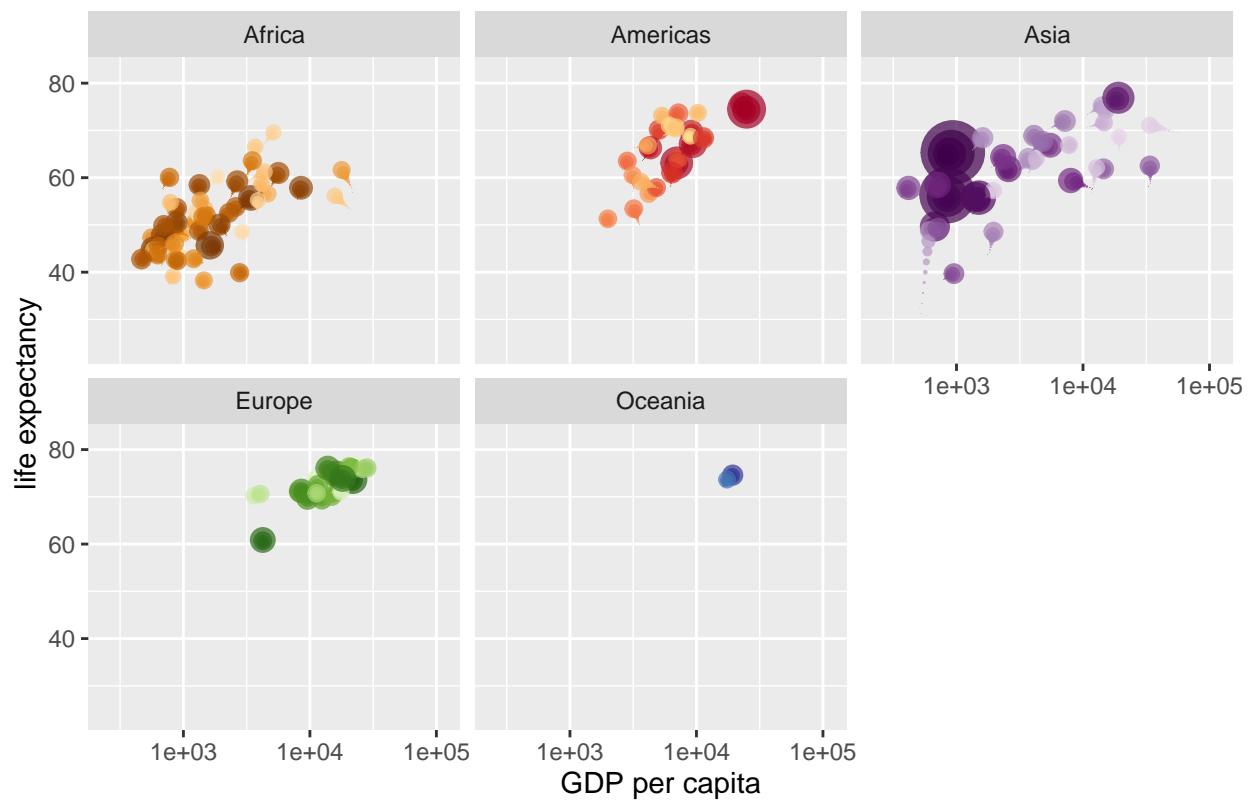
Year: 1980



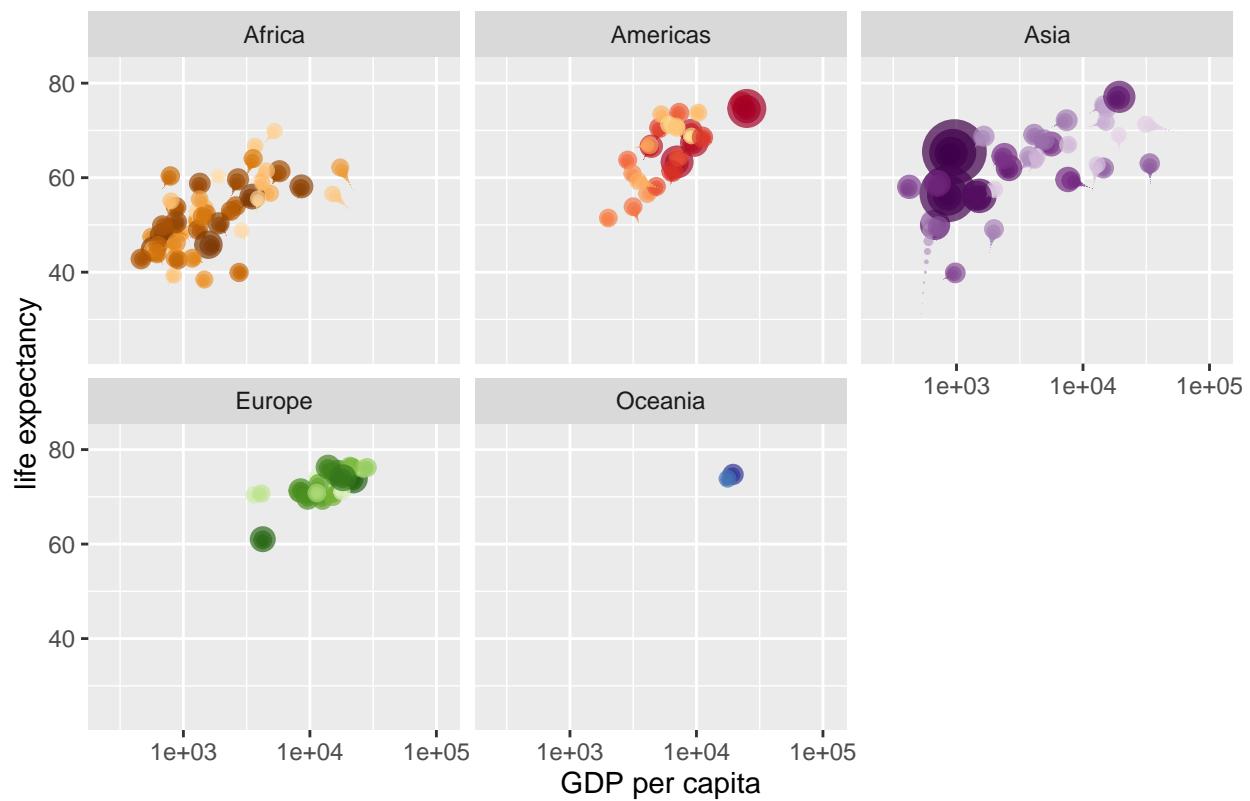
Year: 1981



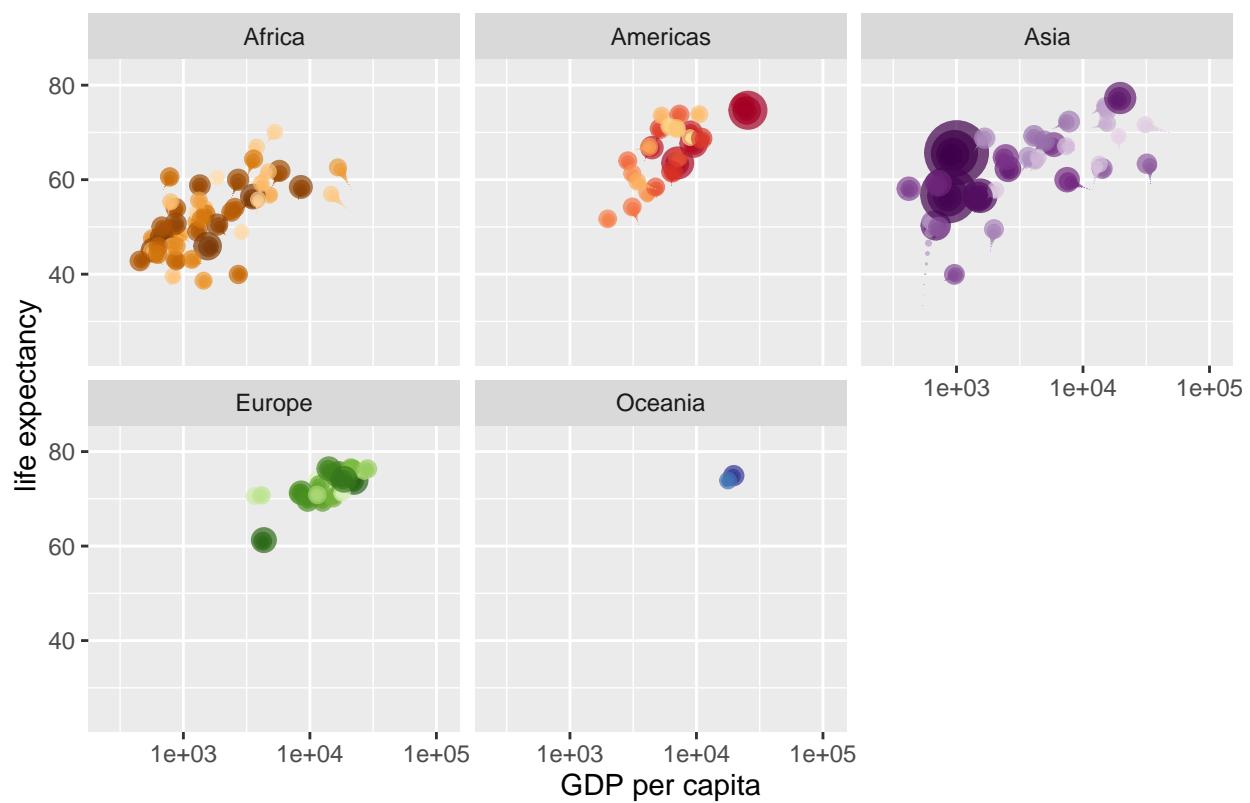
Year: 1981



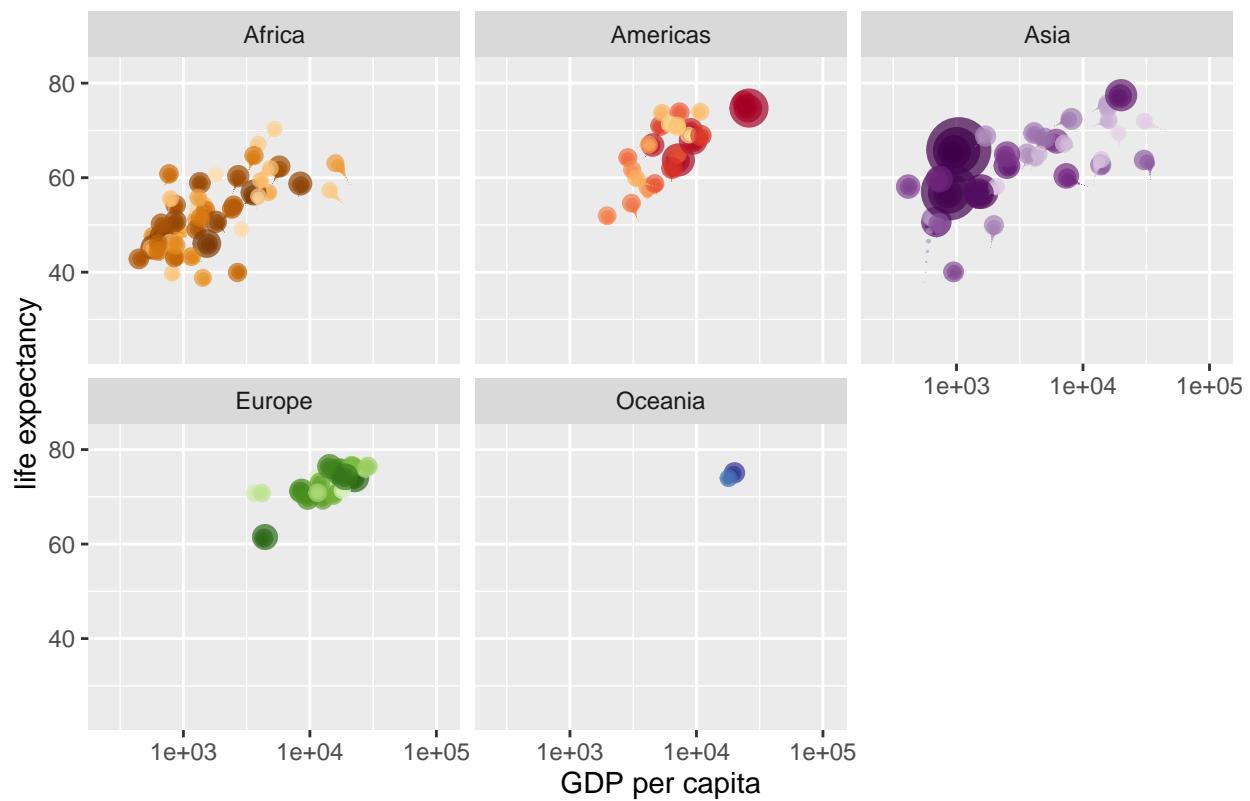
Year: 1982



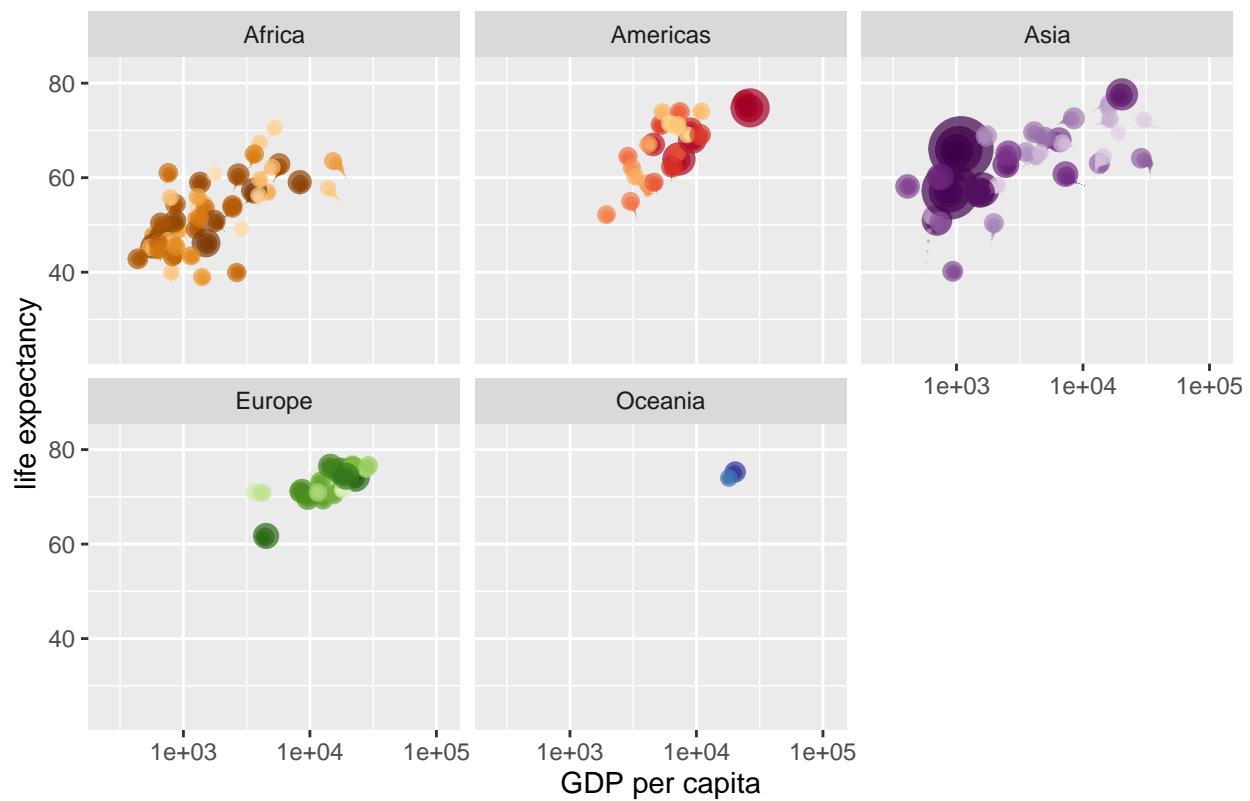
Year: 1983



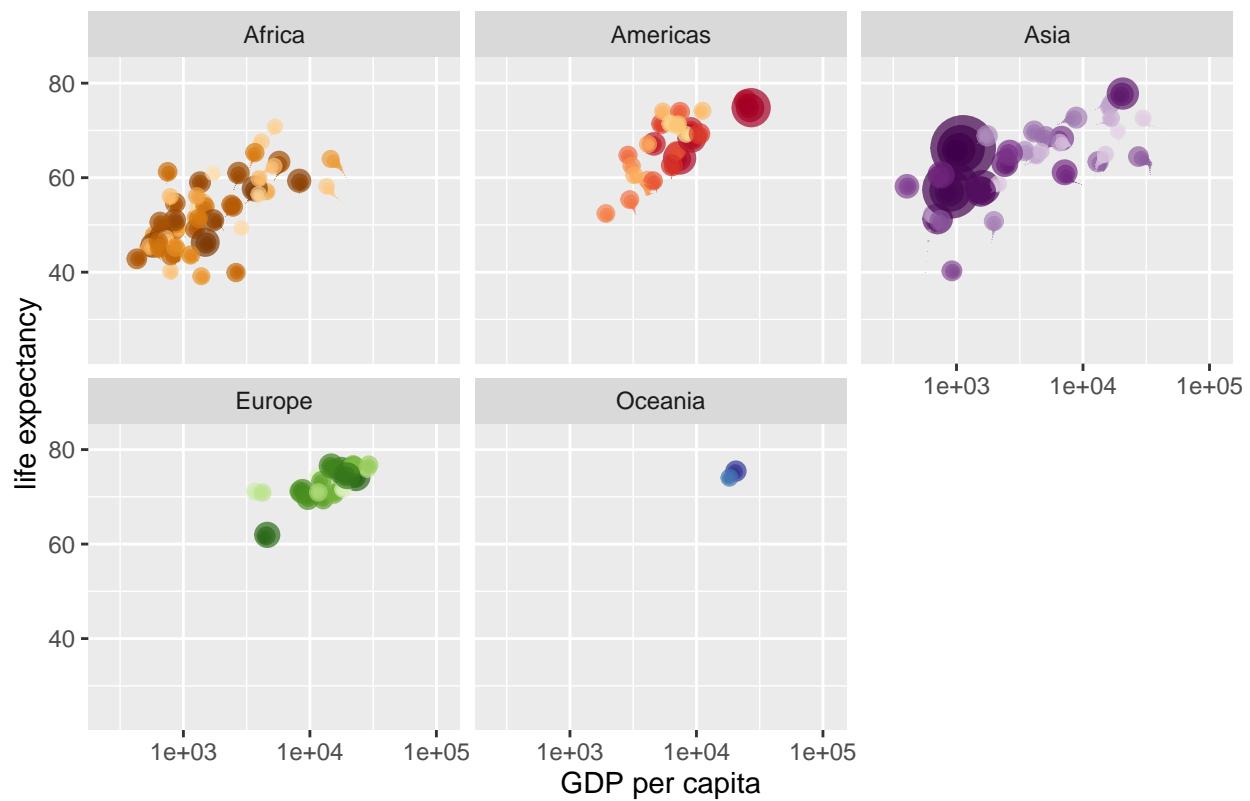
Year: 1983



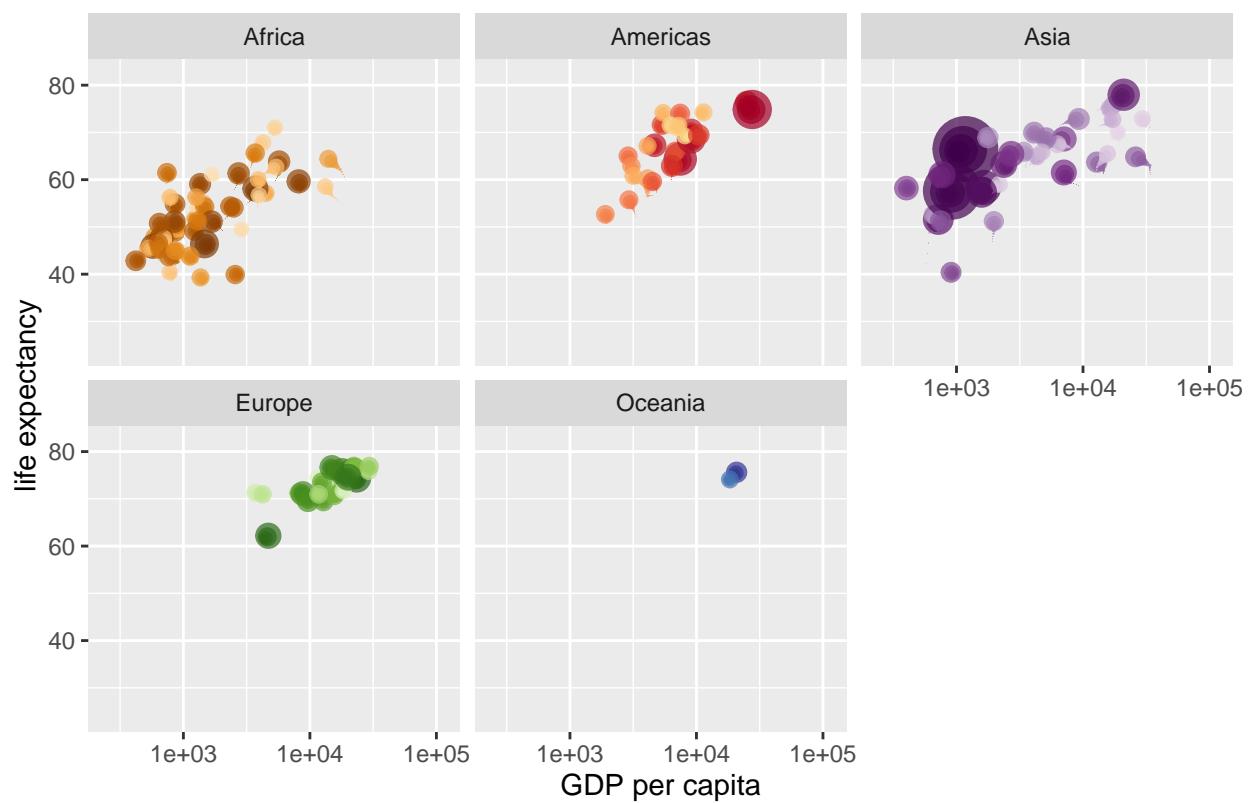
Year: 1984



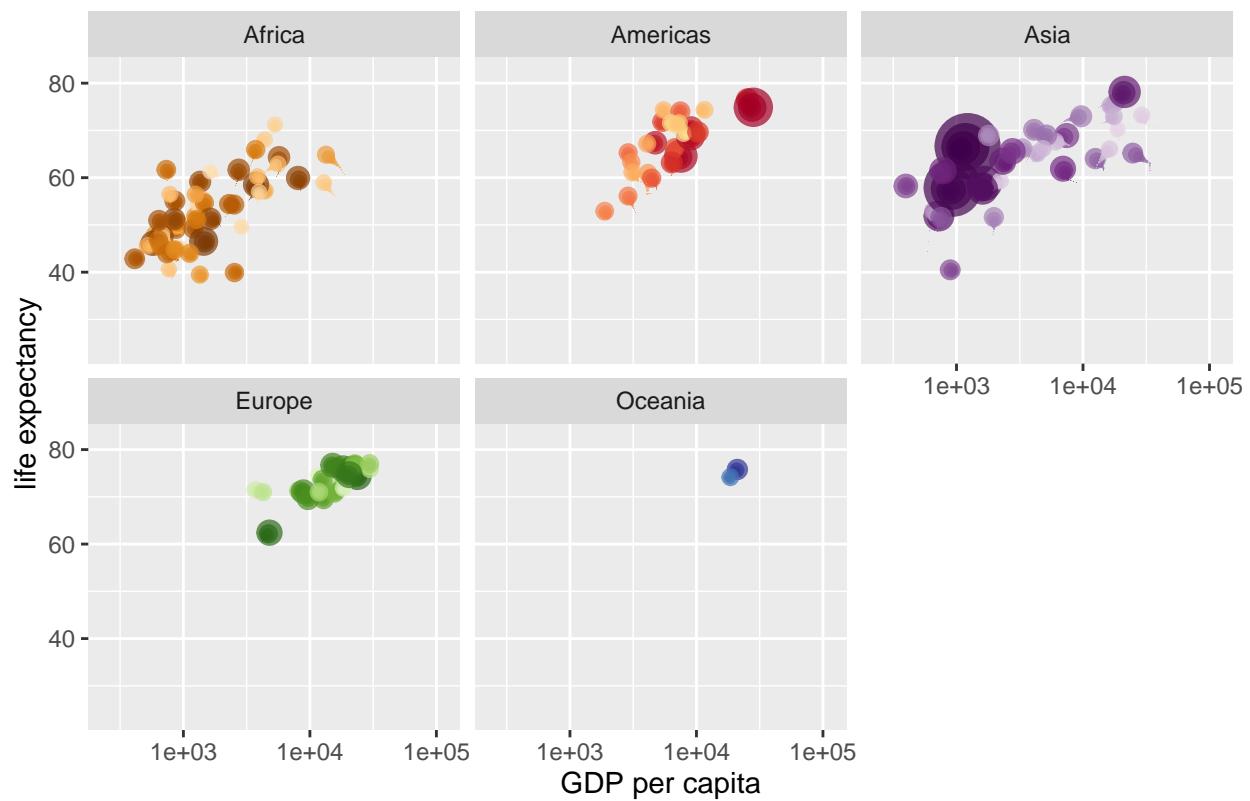
Year: 1984



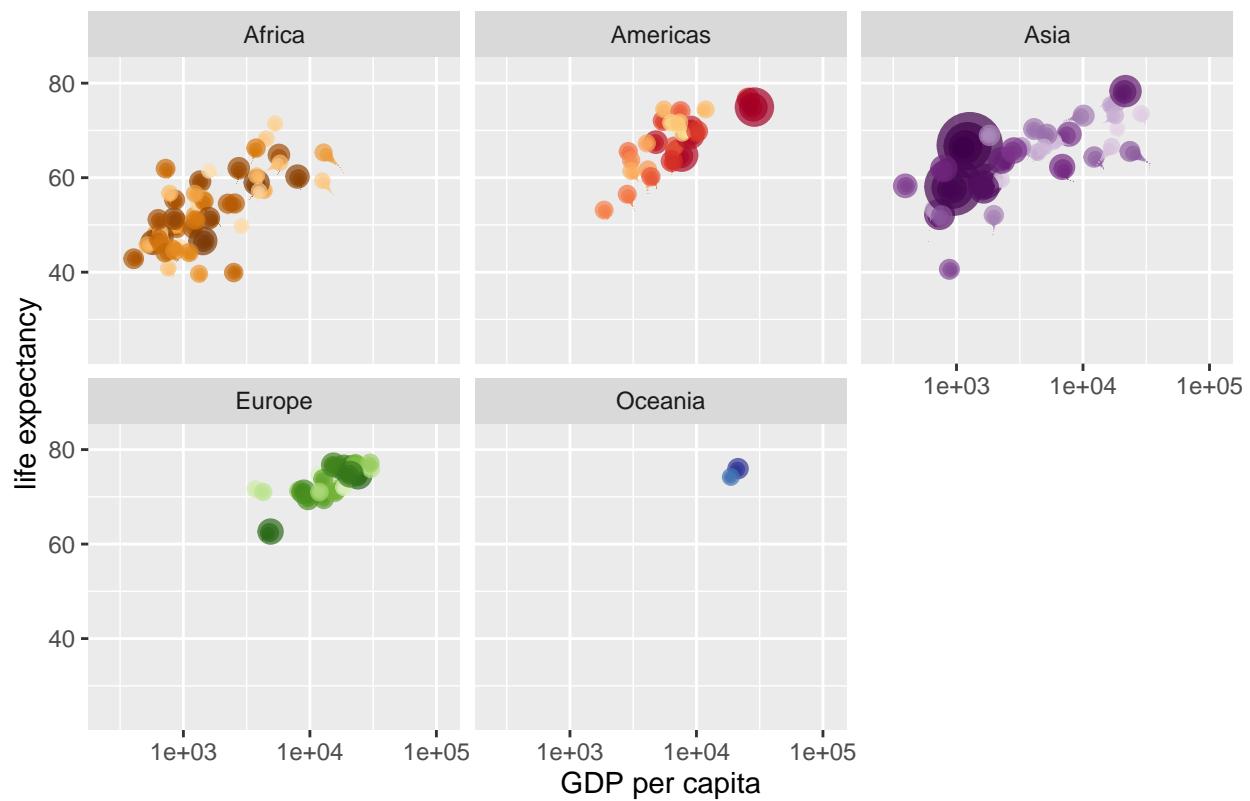
Year: 1985



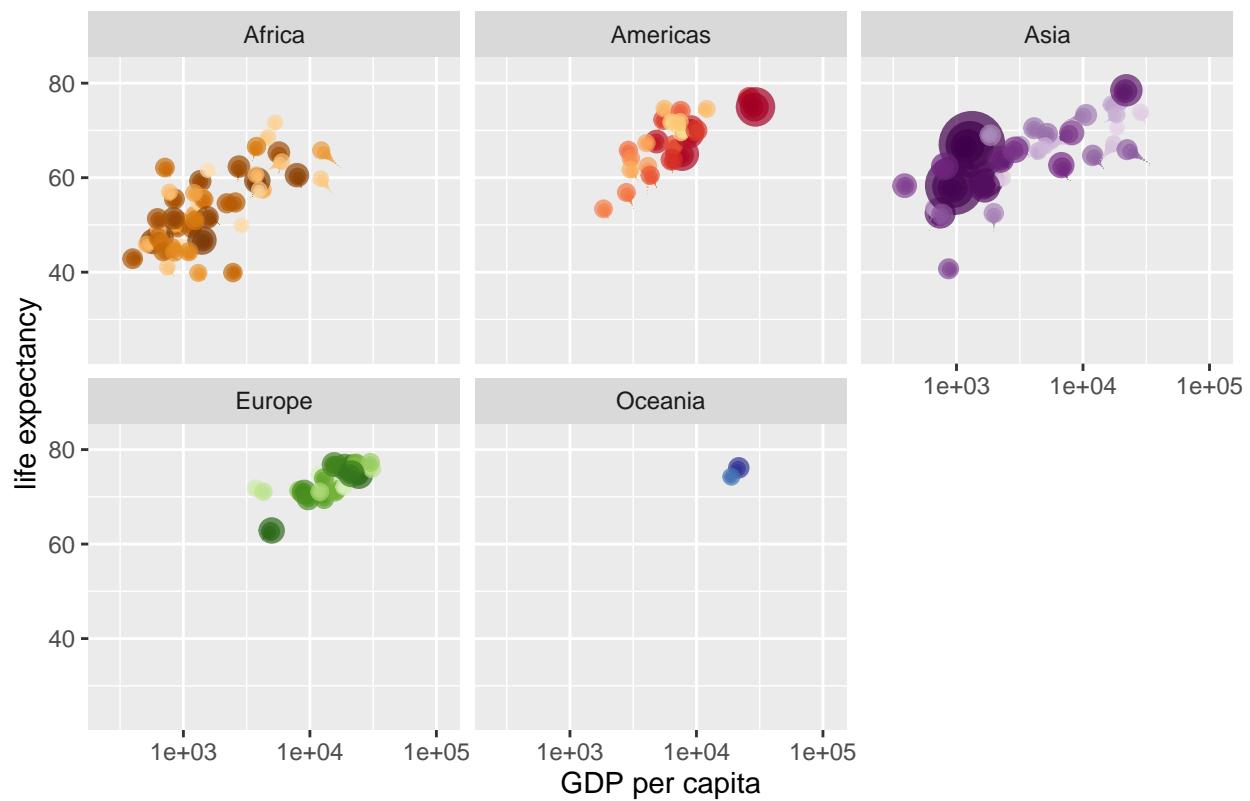
Year: 1985



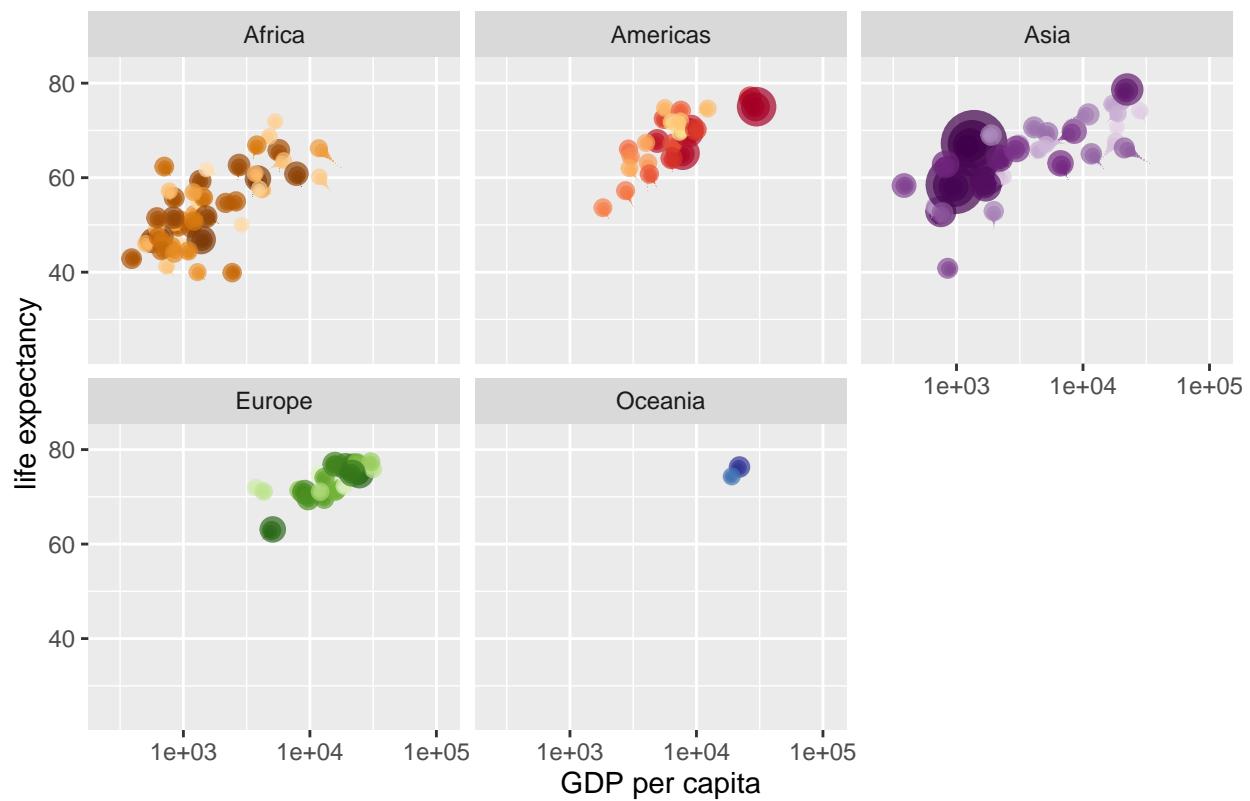
Year: 1986



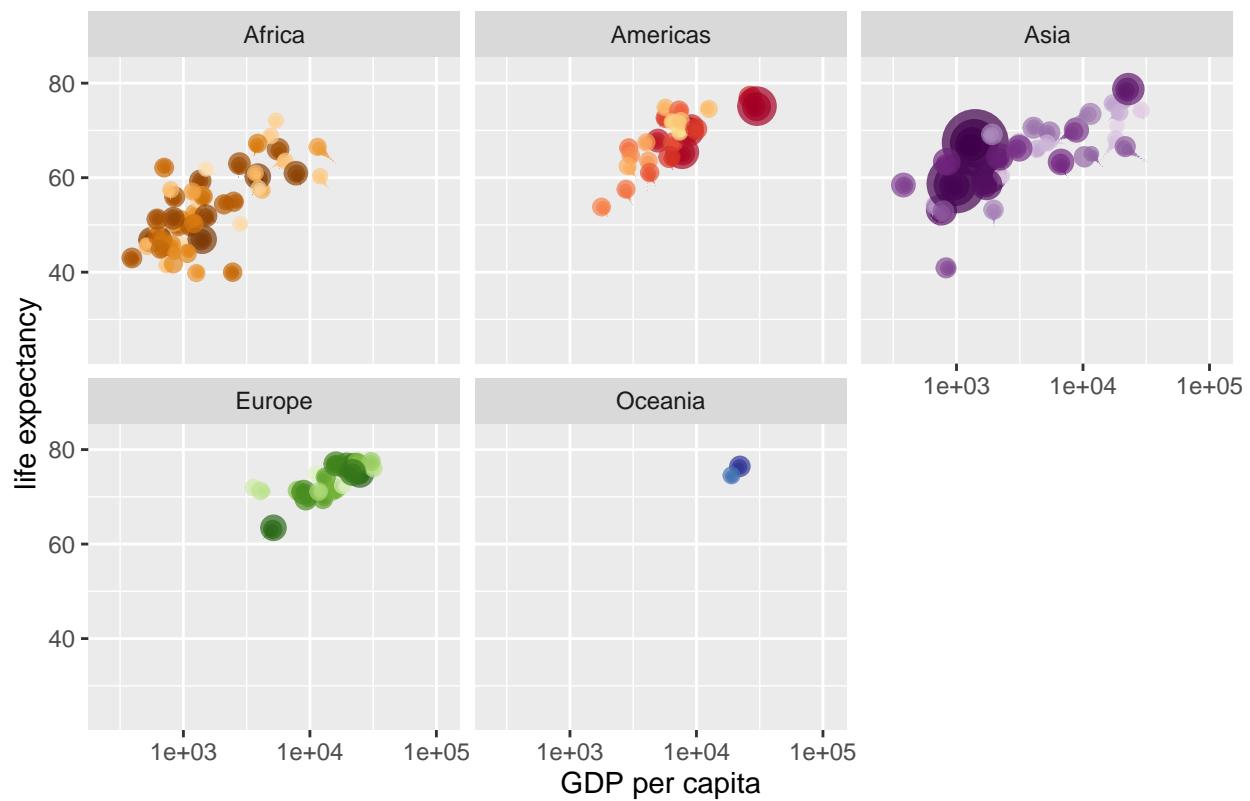
Year: 1986



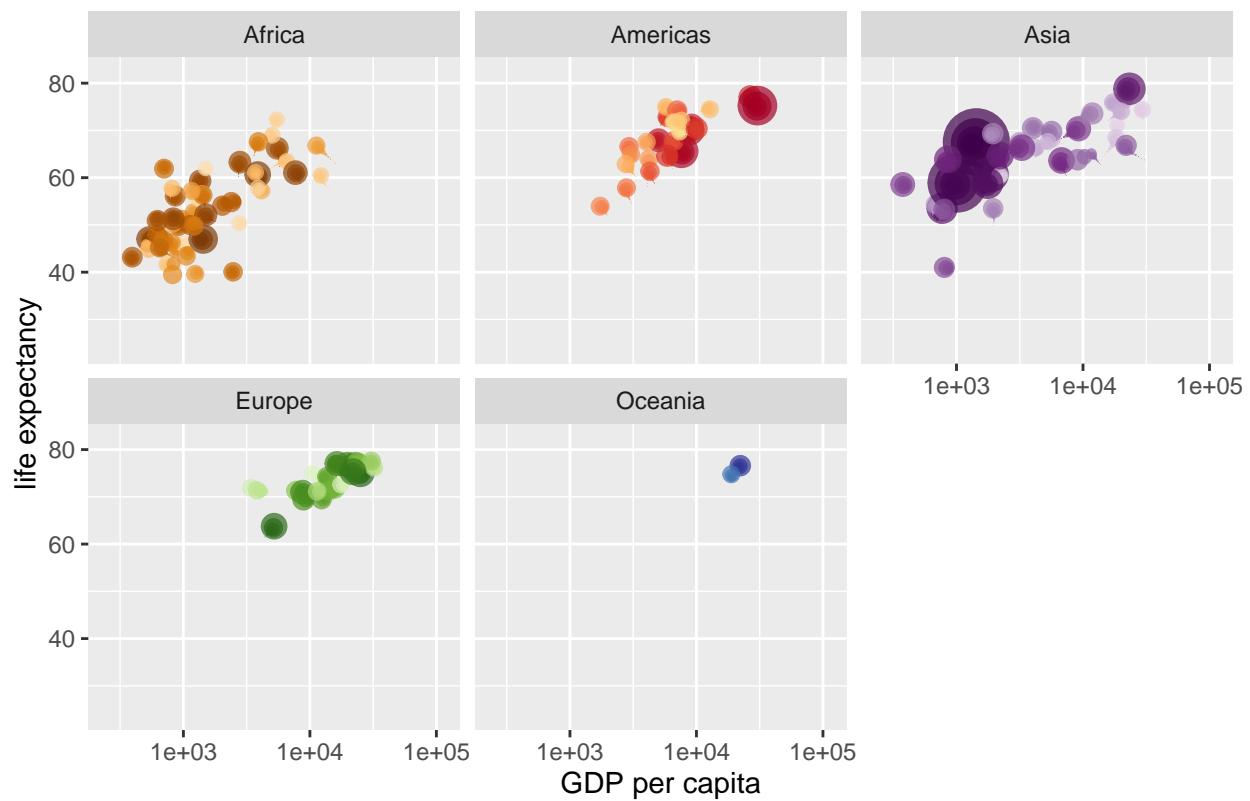
Year: 1987



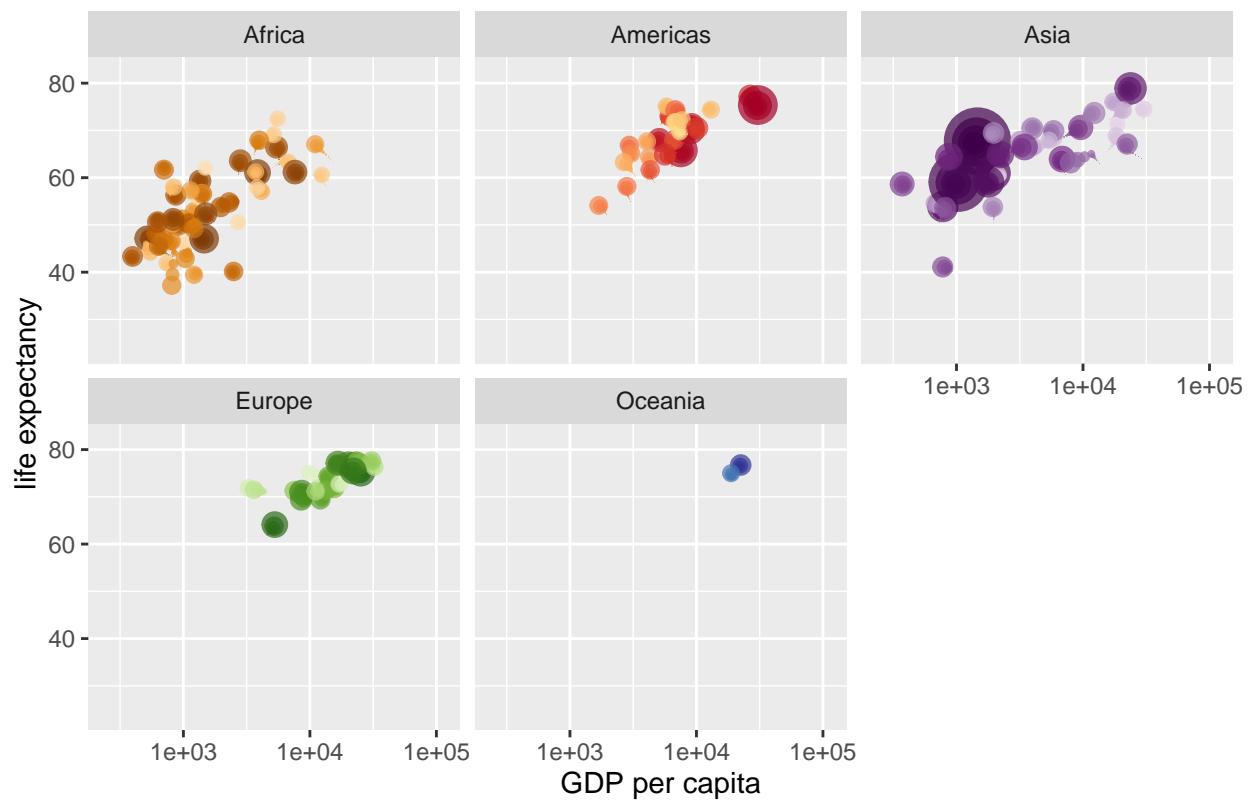
Year: 1988



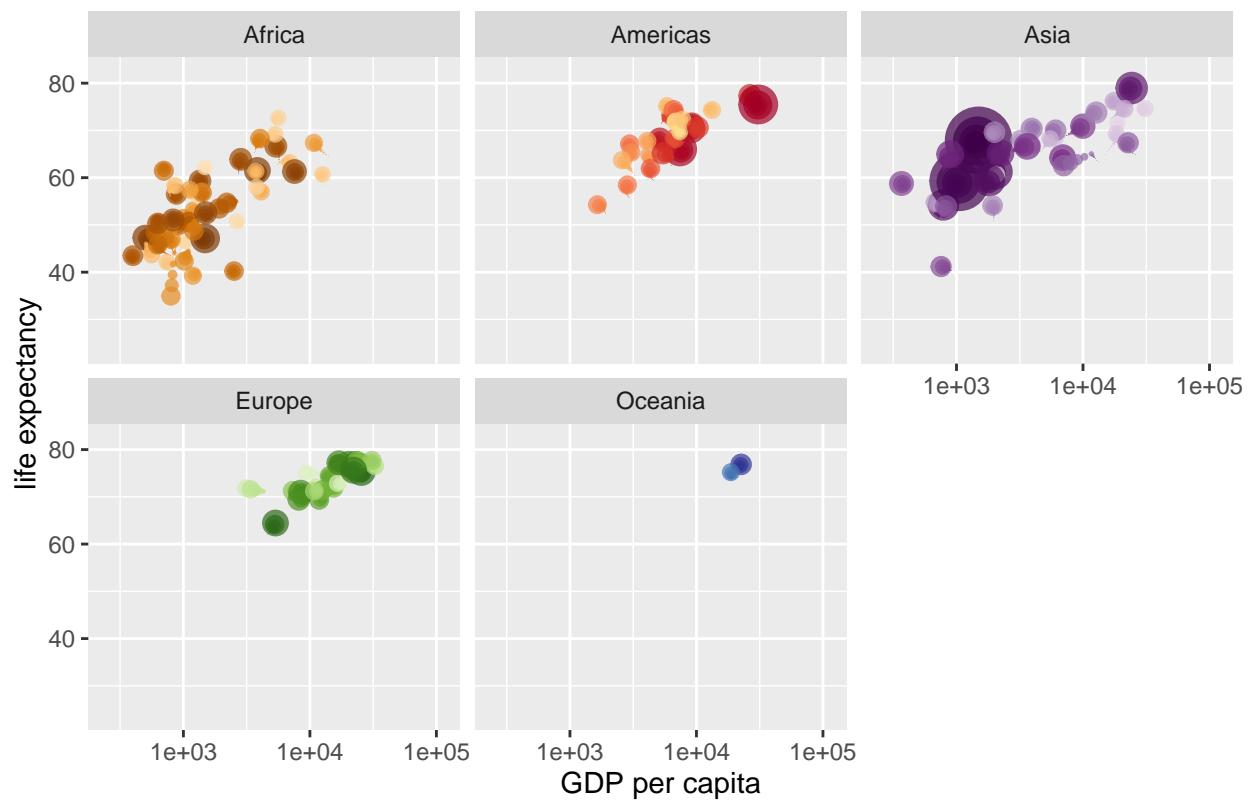
Year: 1988



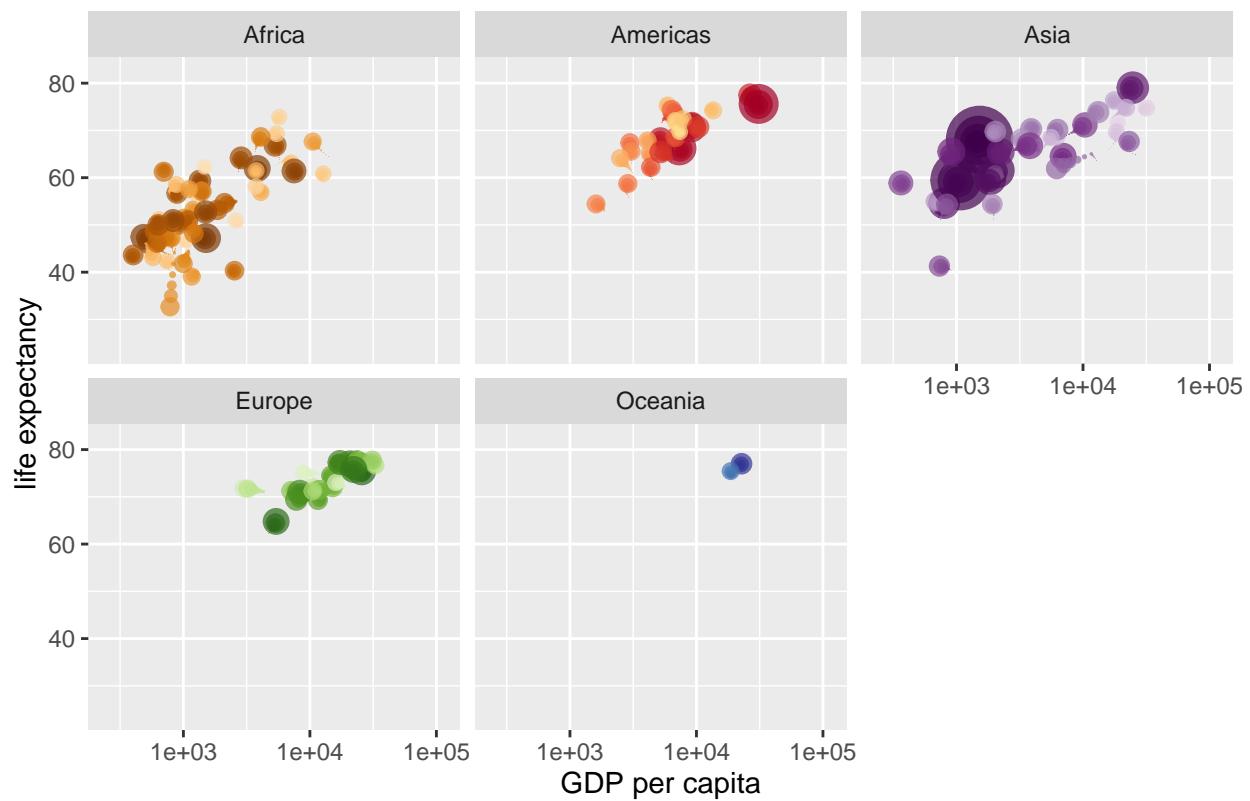
Year: 1989



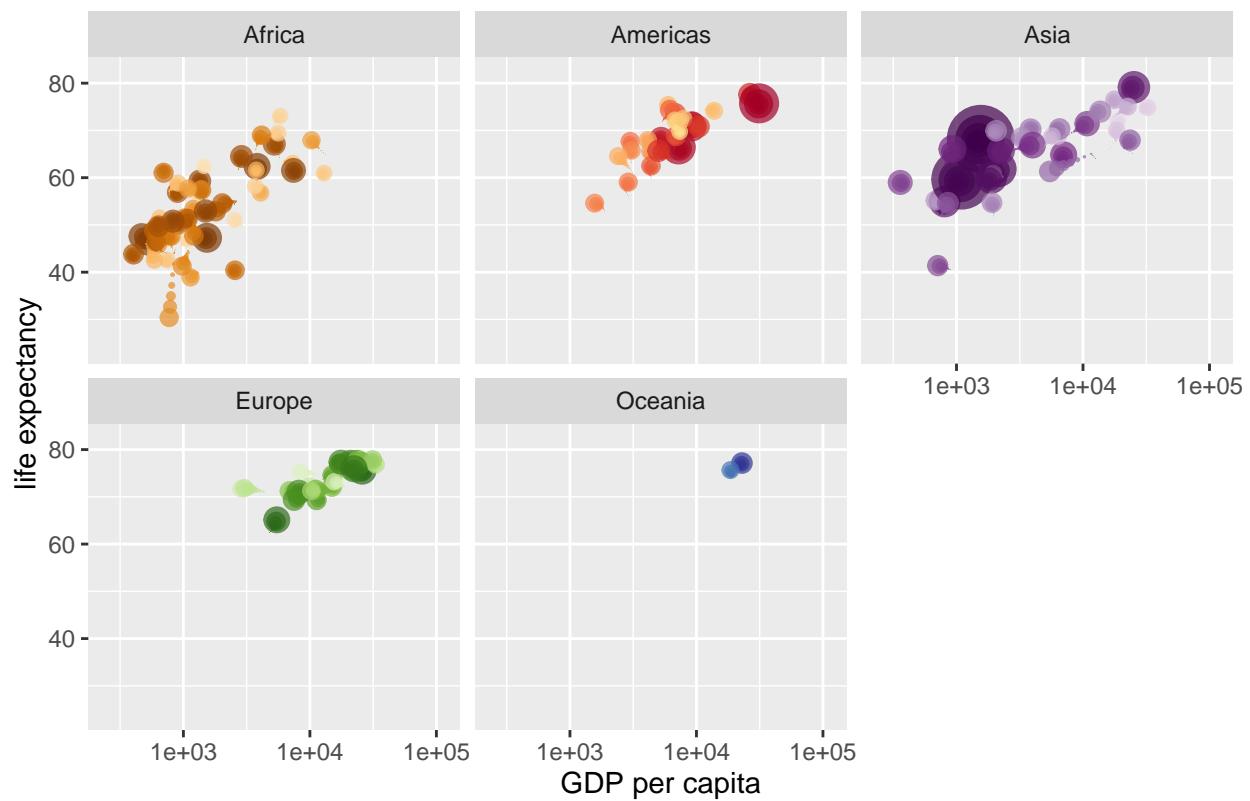
Year: 1989



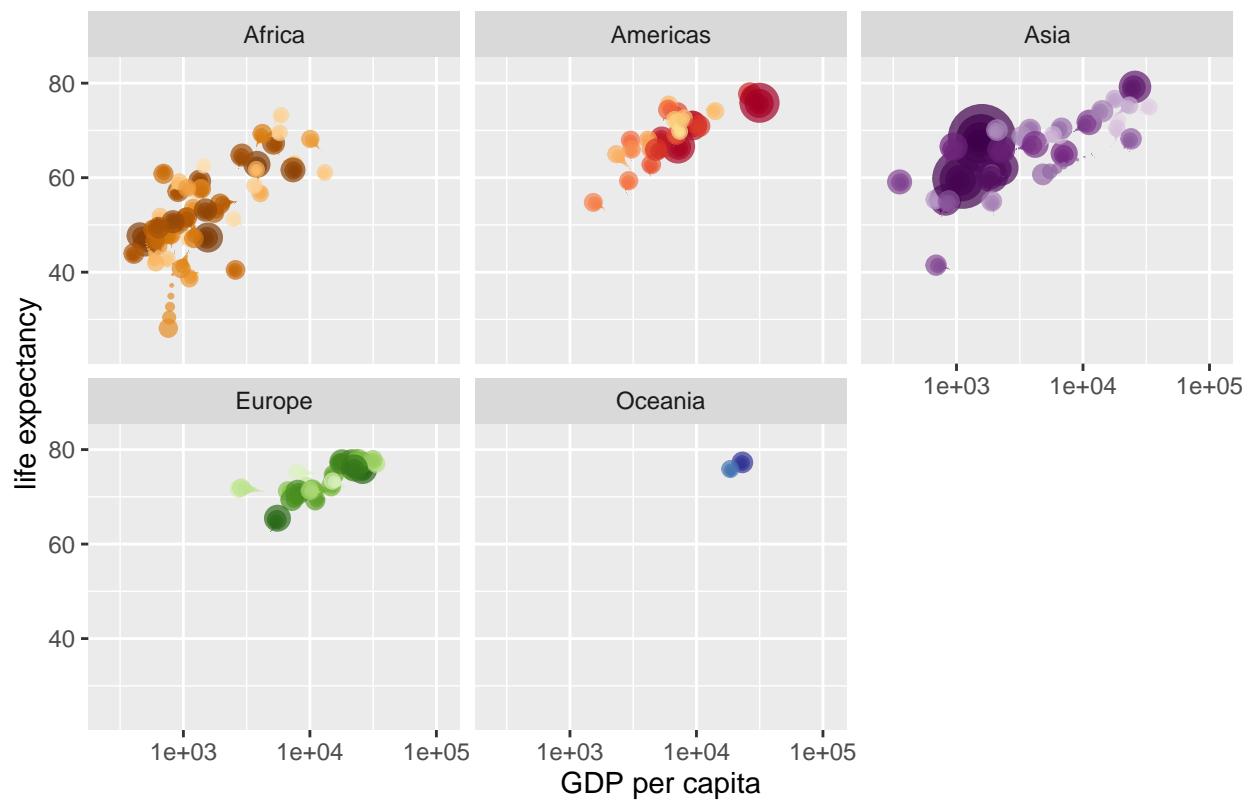
Year: 1990



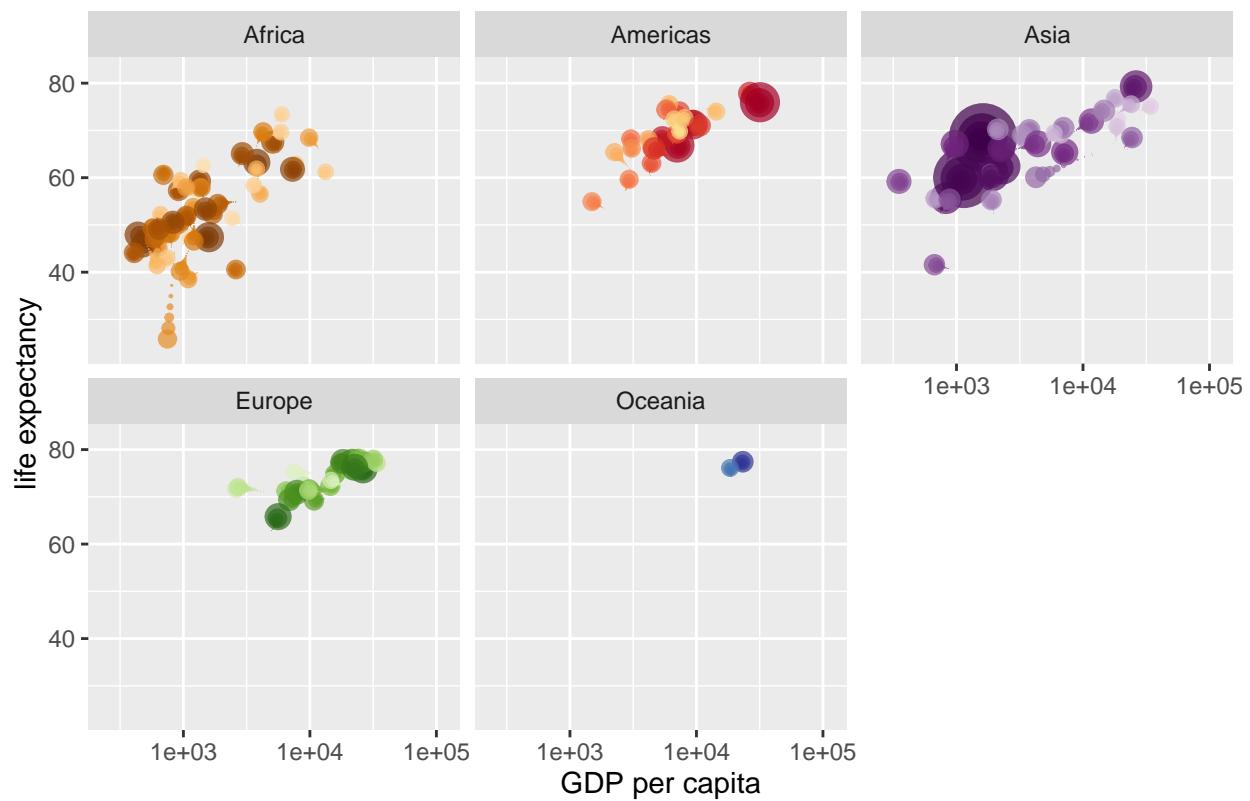
Year: 1990



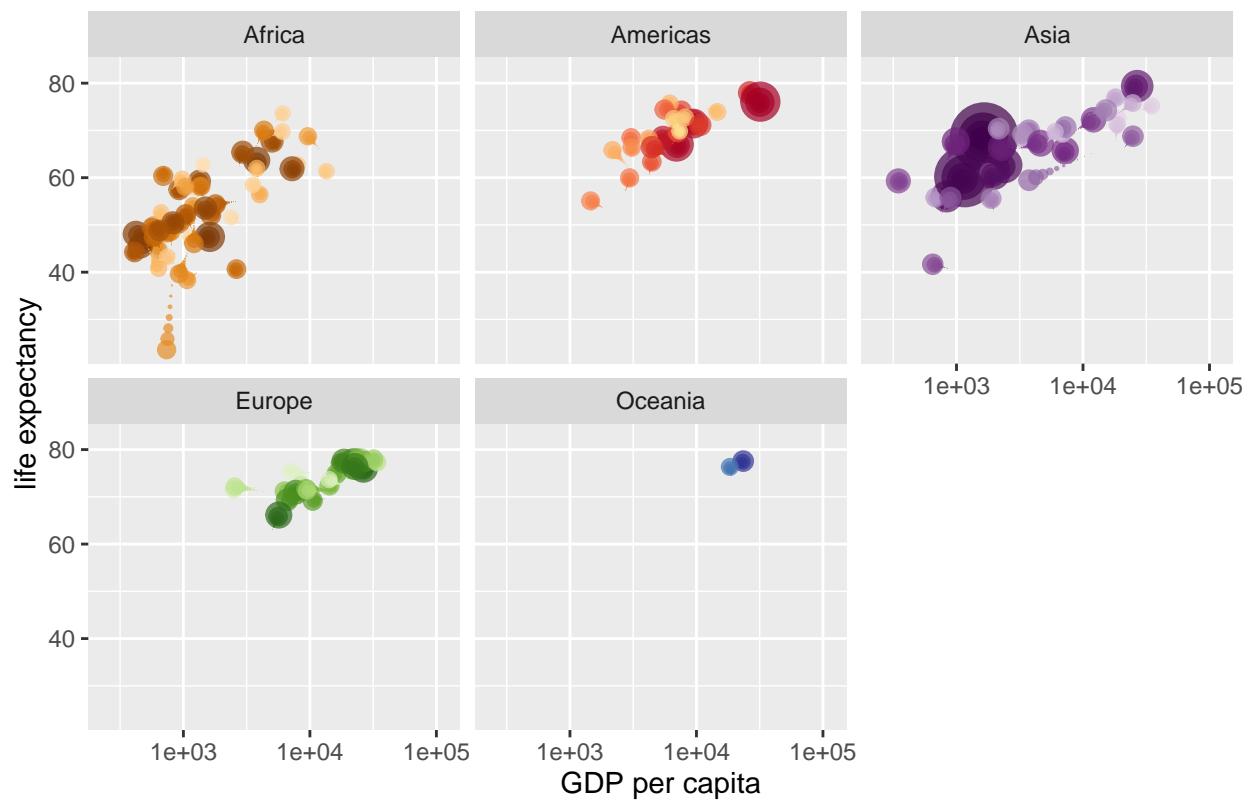
Year: 1991



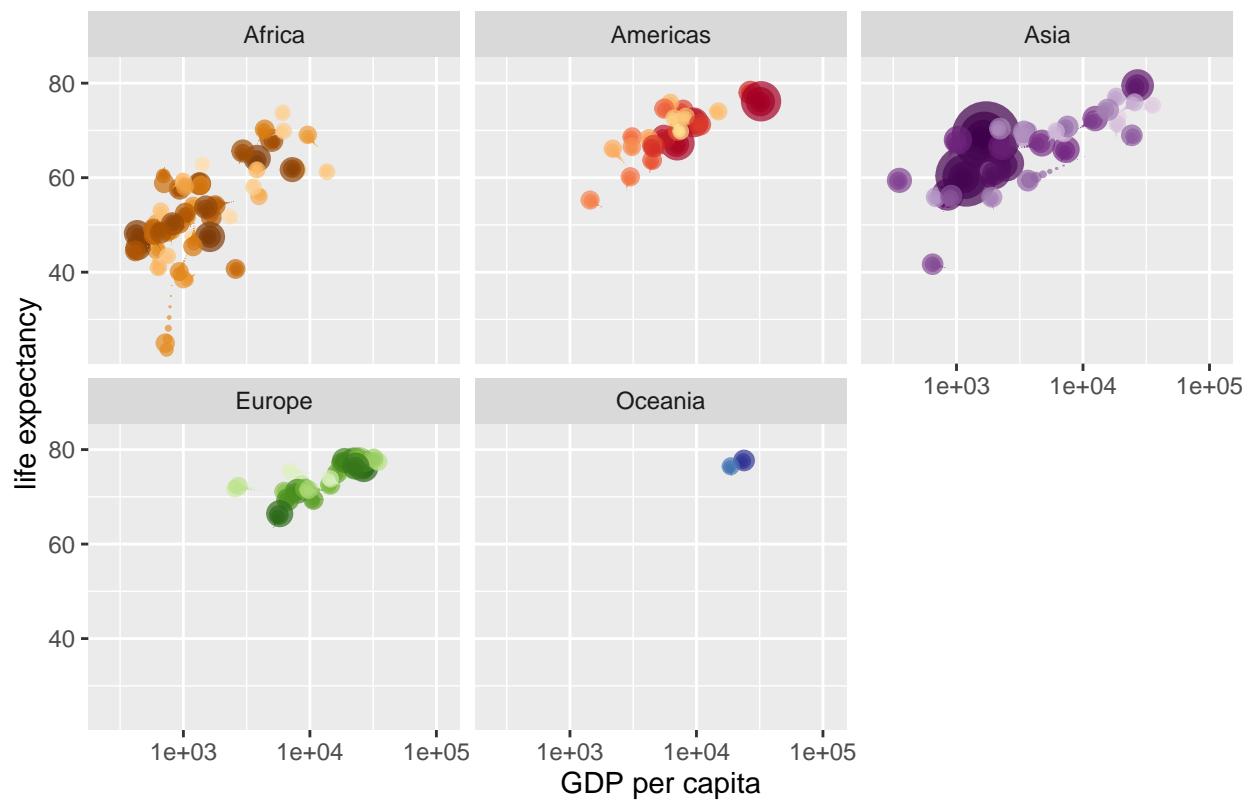
Year: 1991



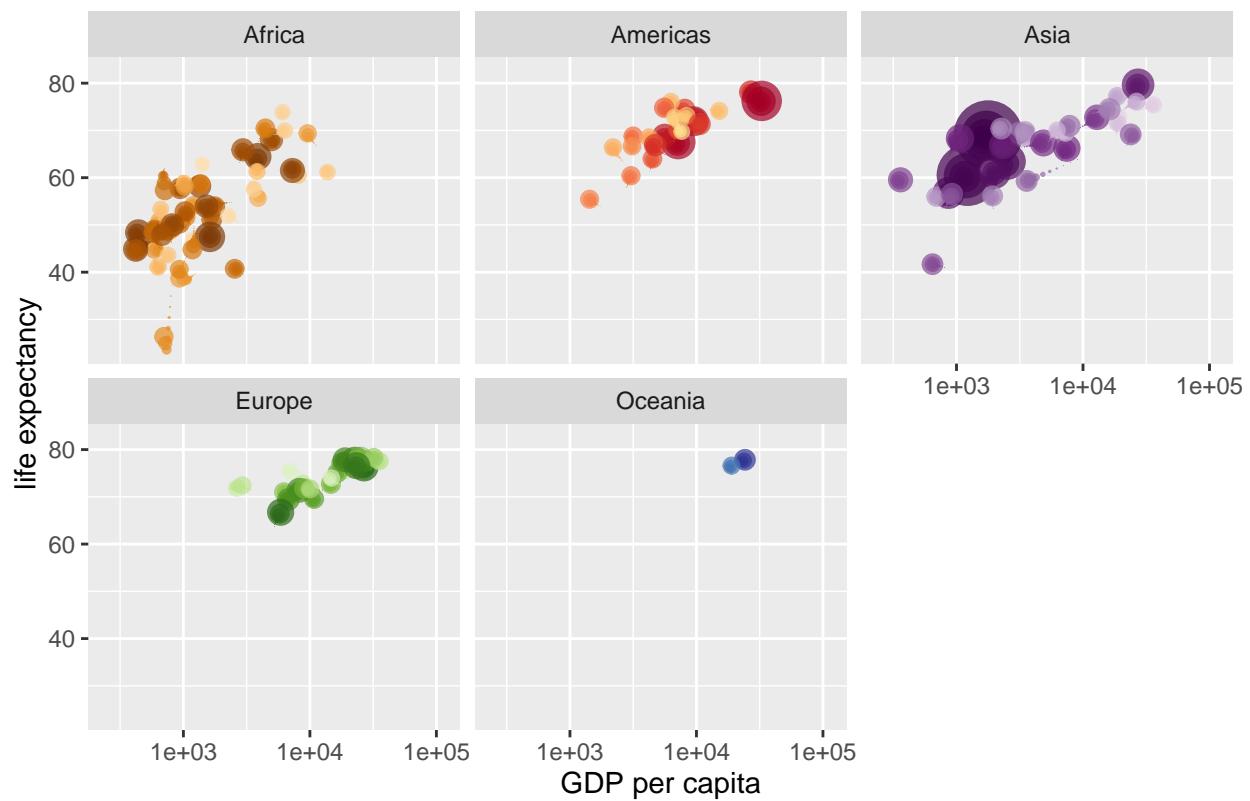
Year: 1992



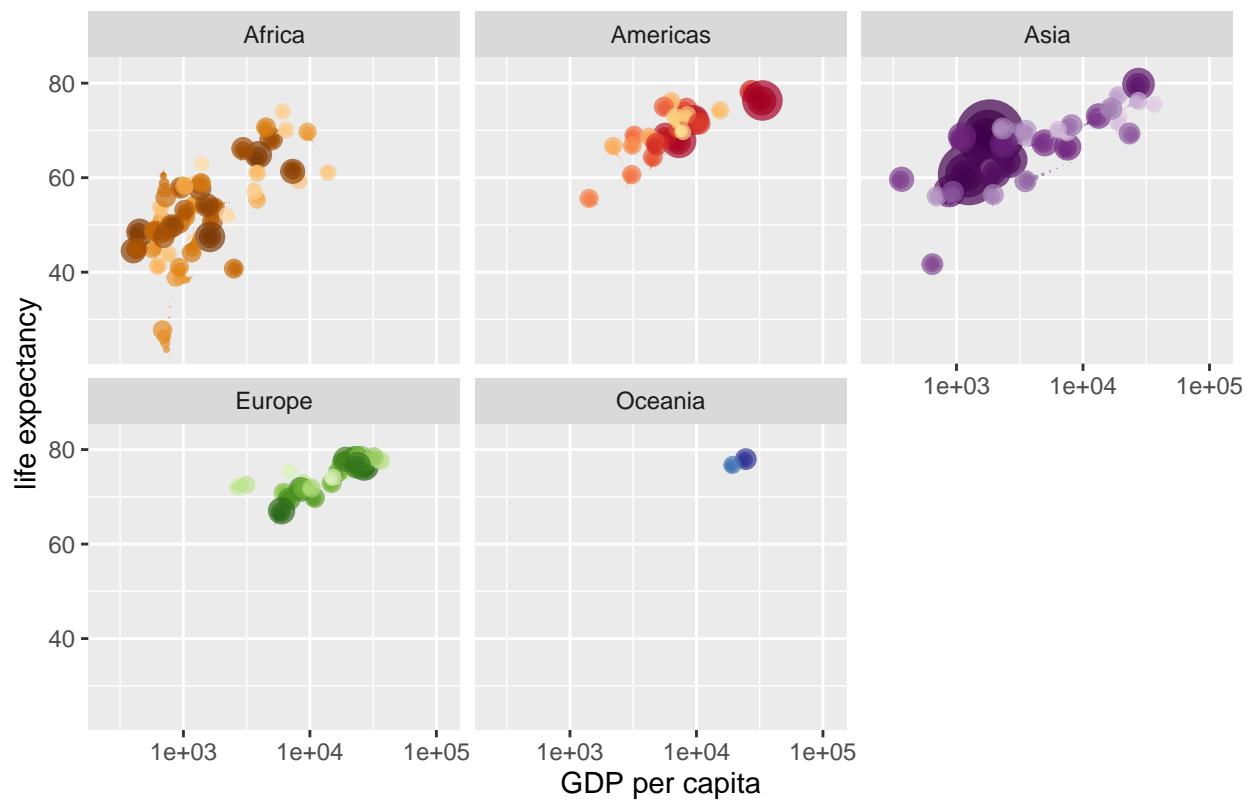
Year: 1993



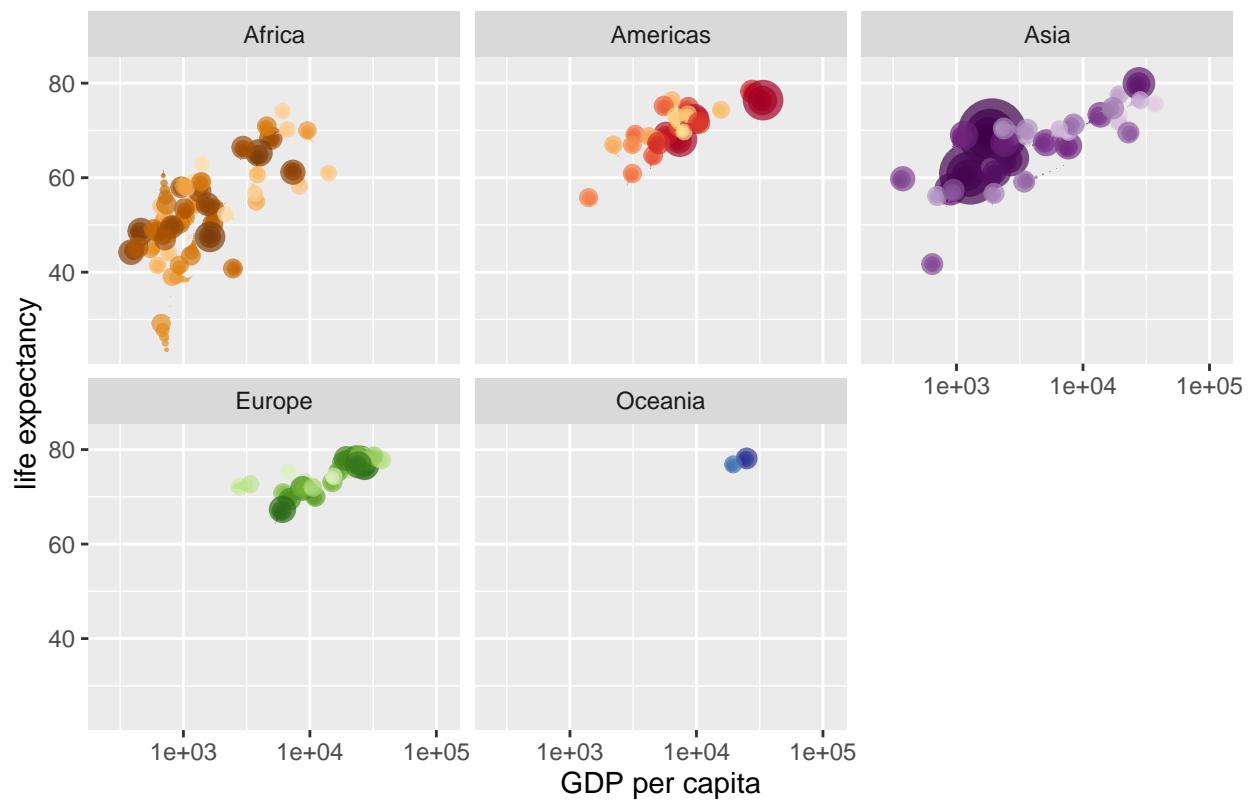
Year: 1993



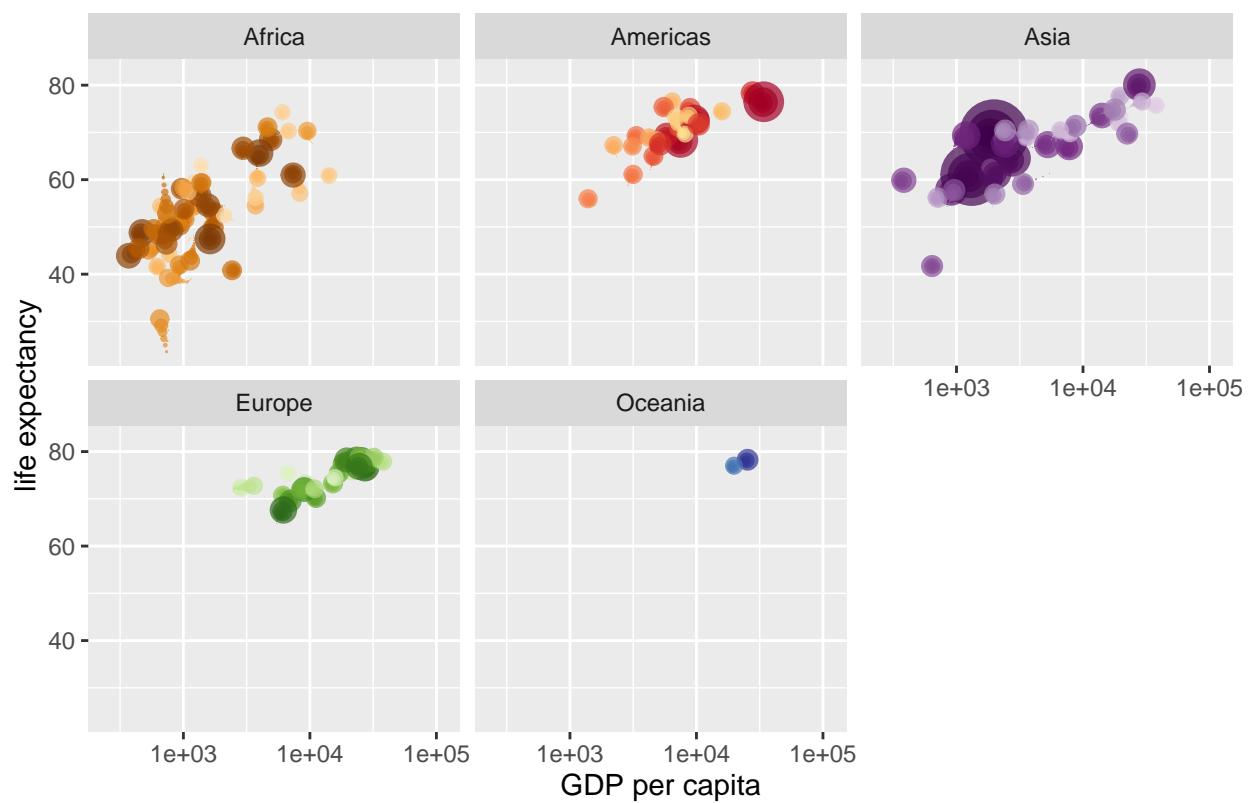
Year: 1994



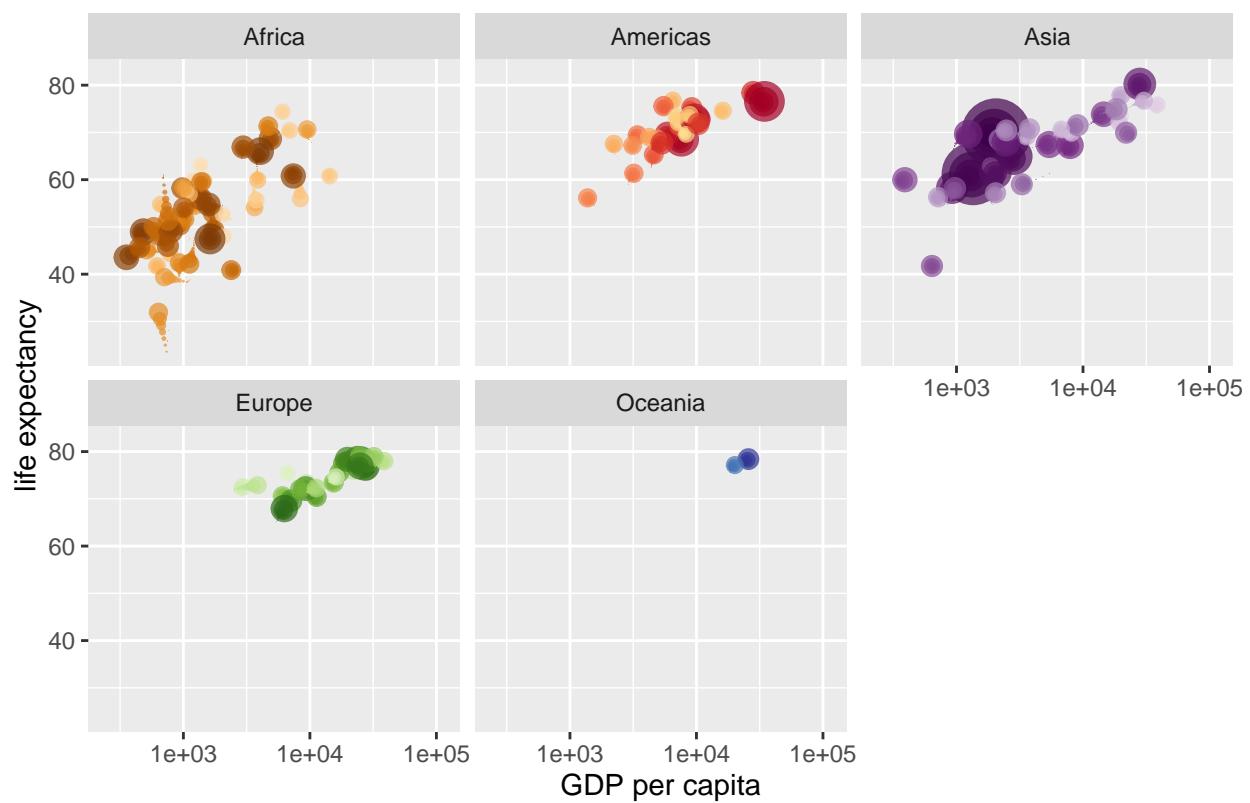
Year: 1994



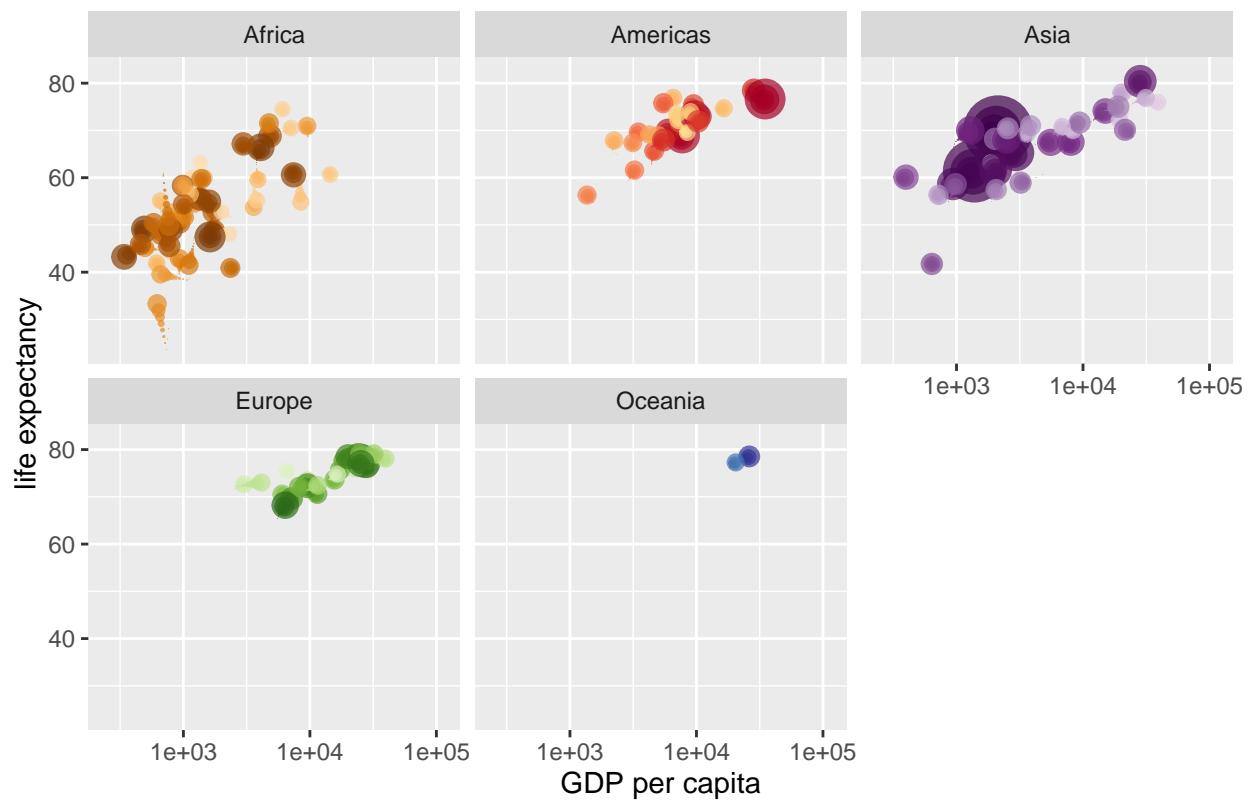
Year: 1995



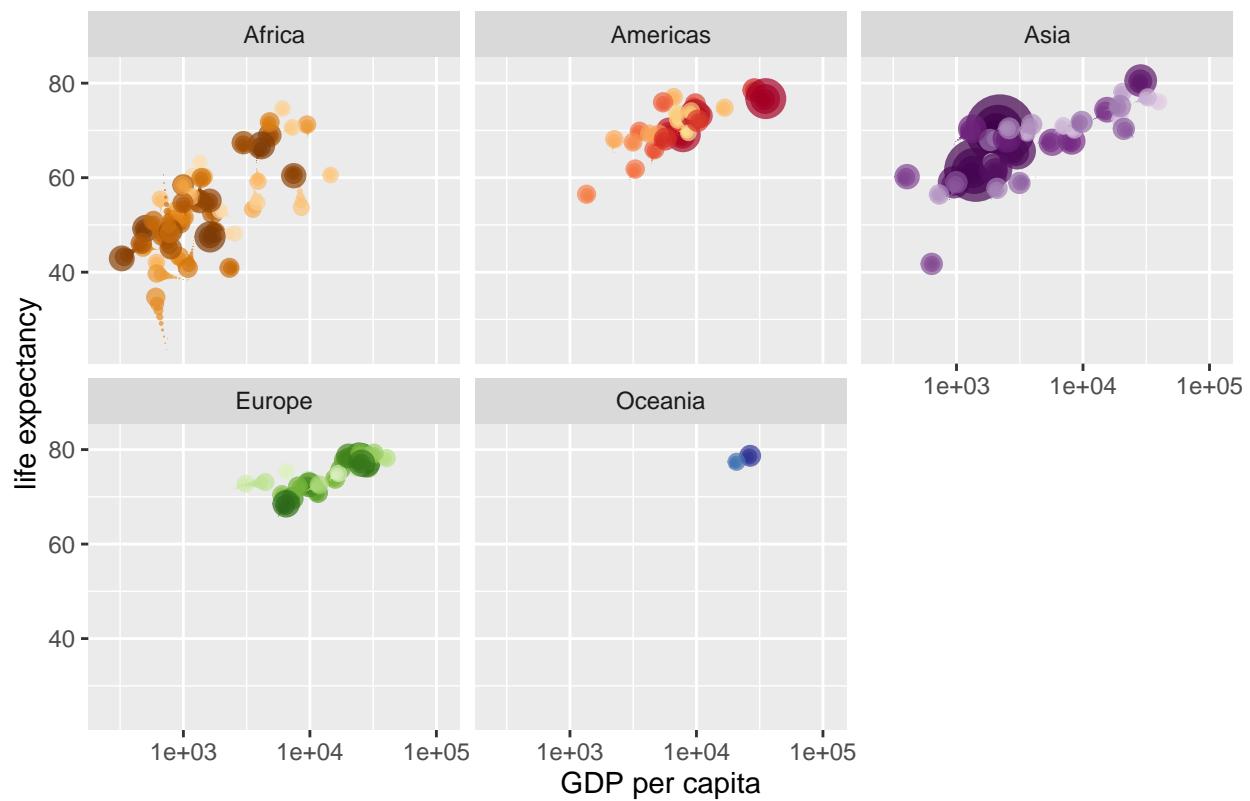
Year: 1995



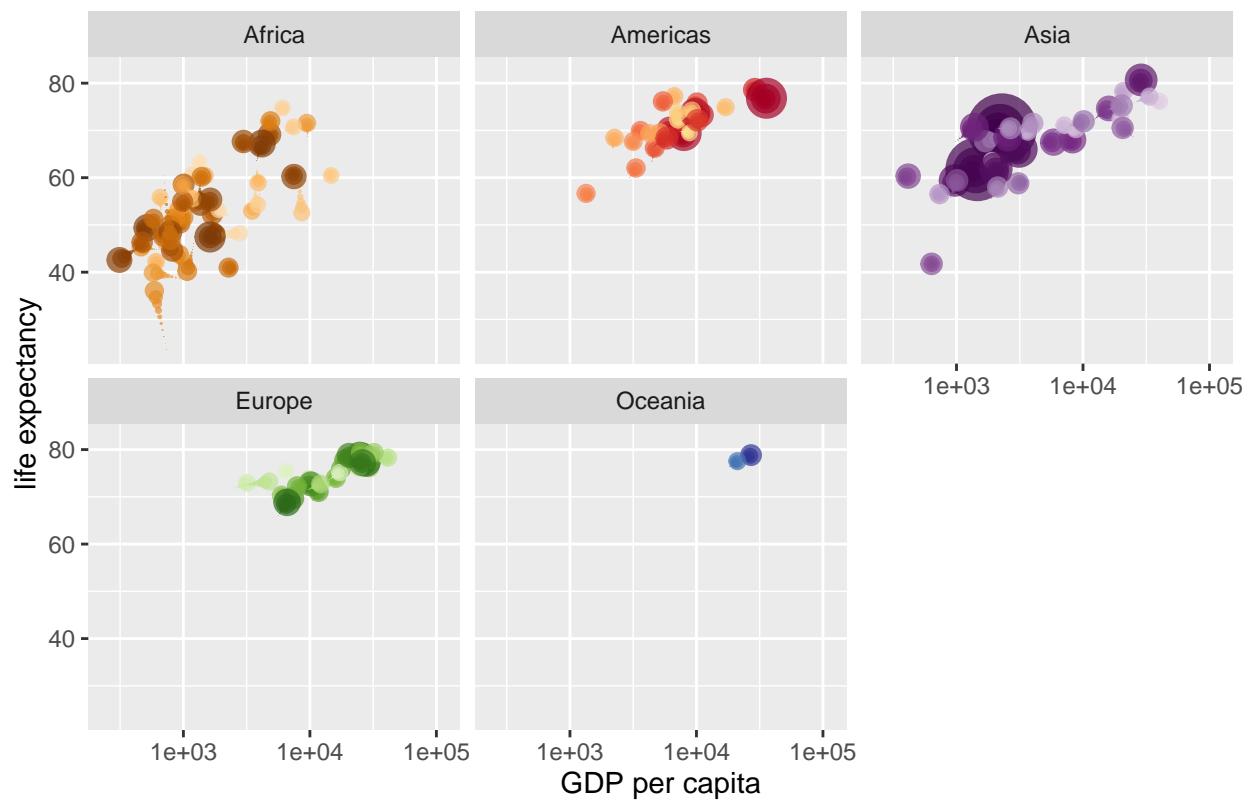
Year: 1996



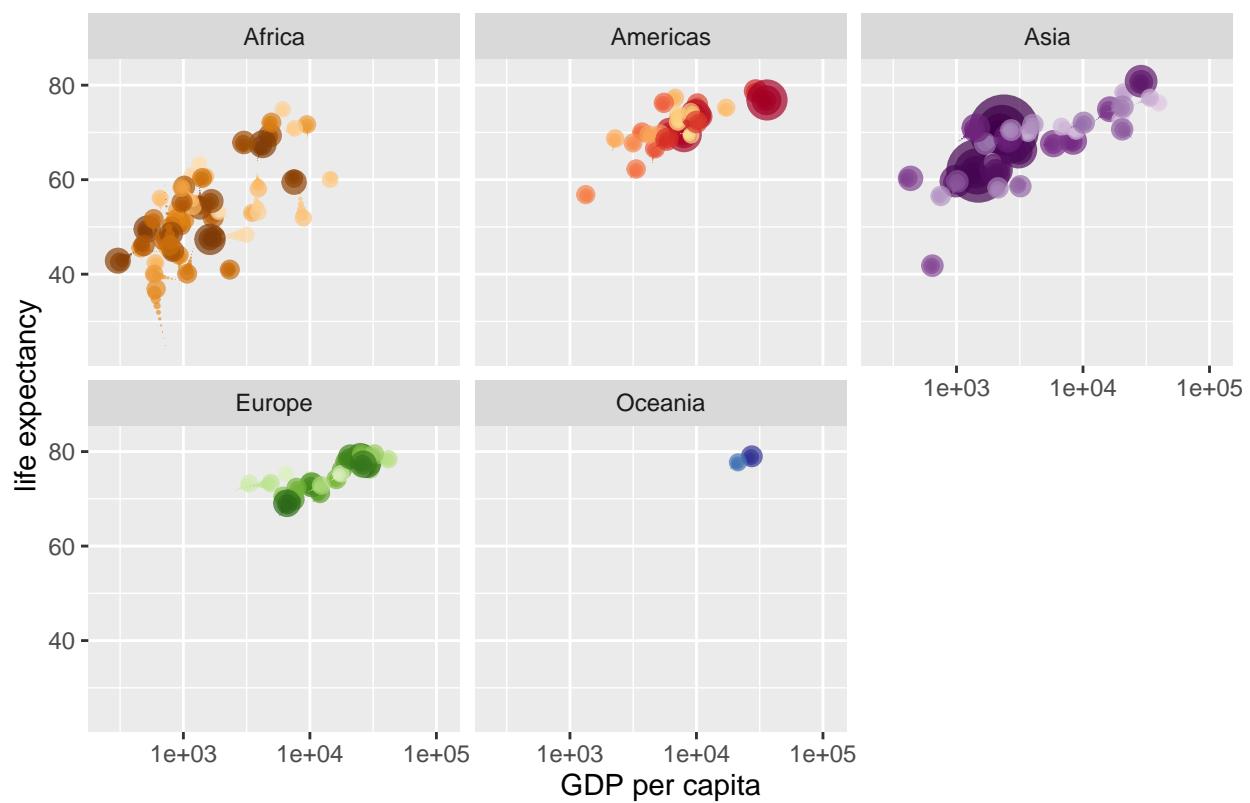
Year: 1996



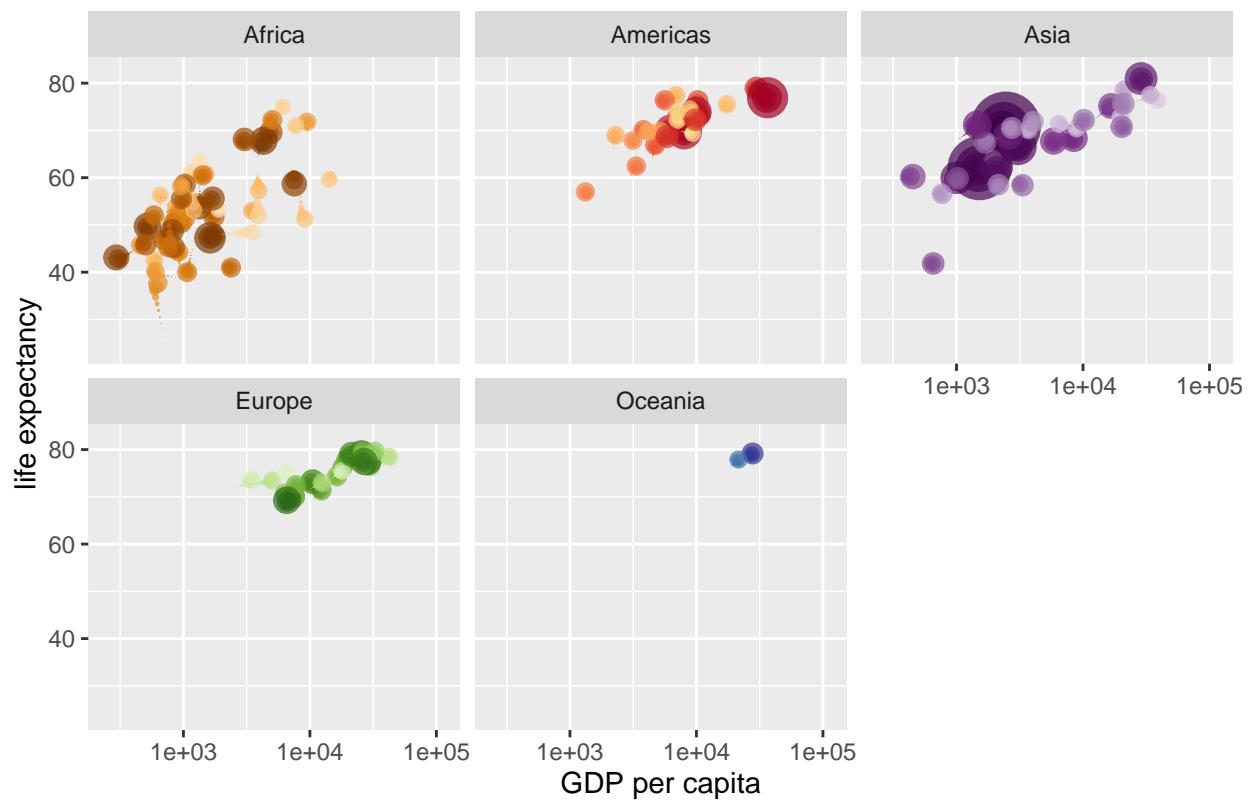
Year: 1997



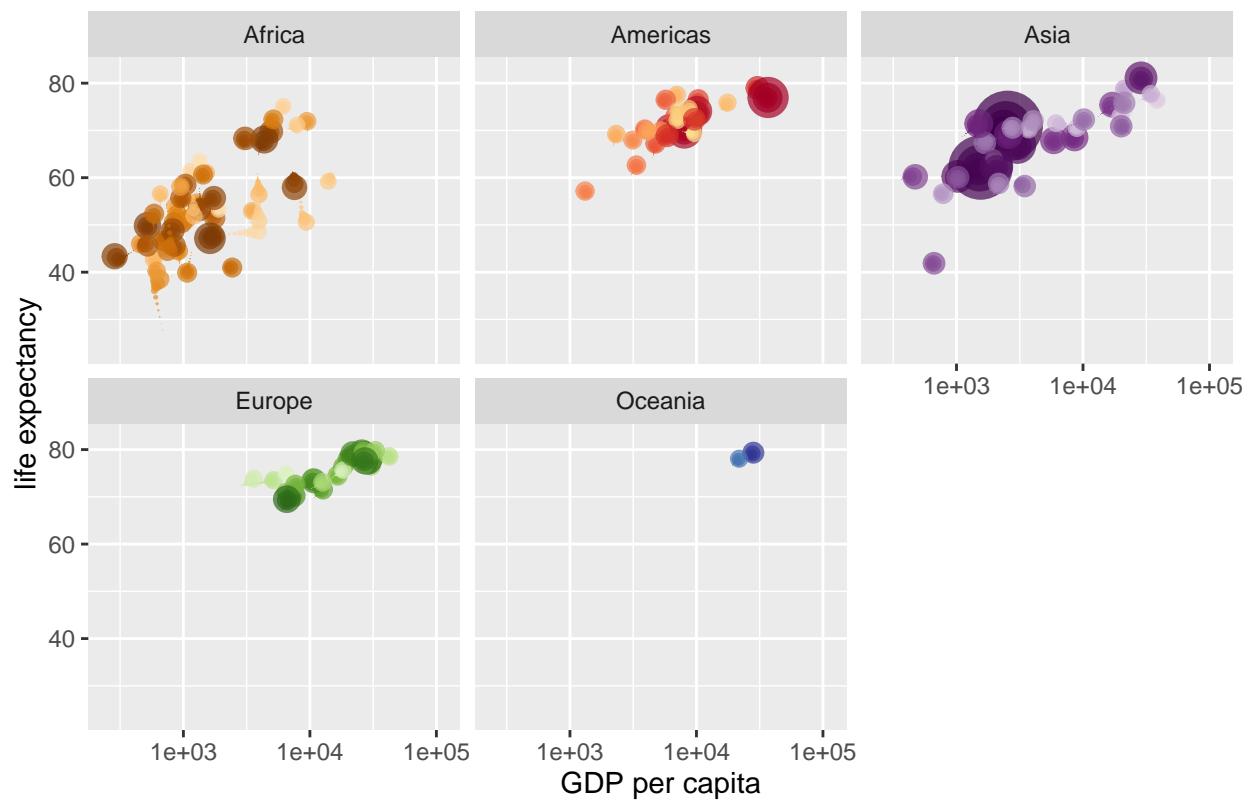
Year: 1998



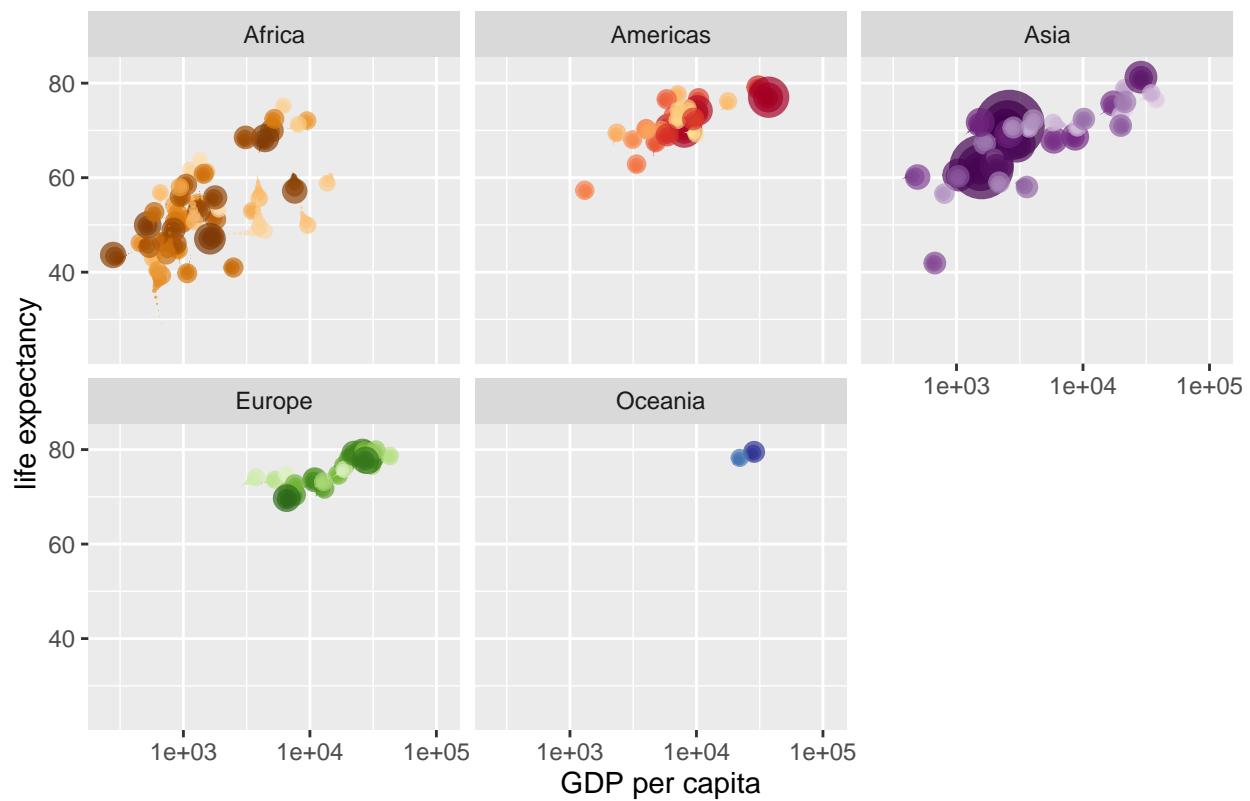
Year: 1998



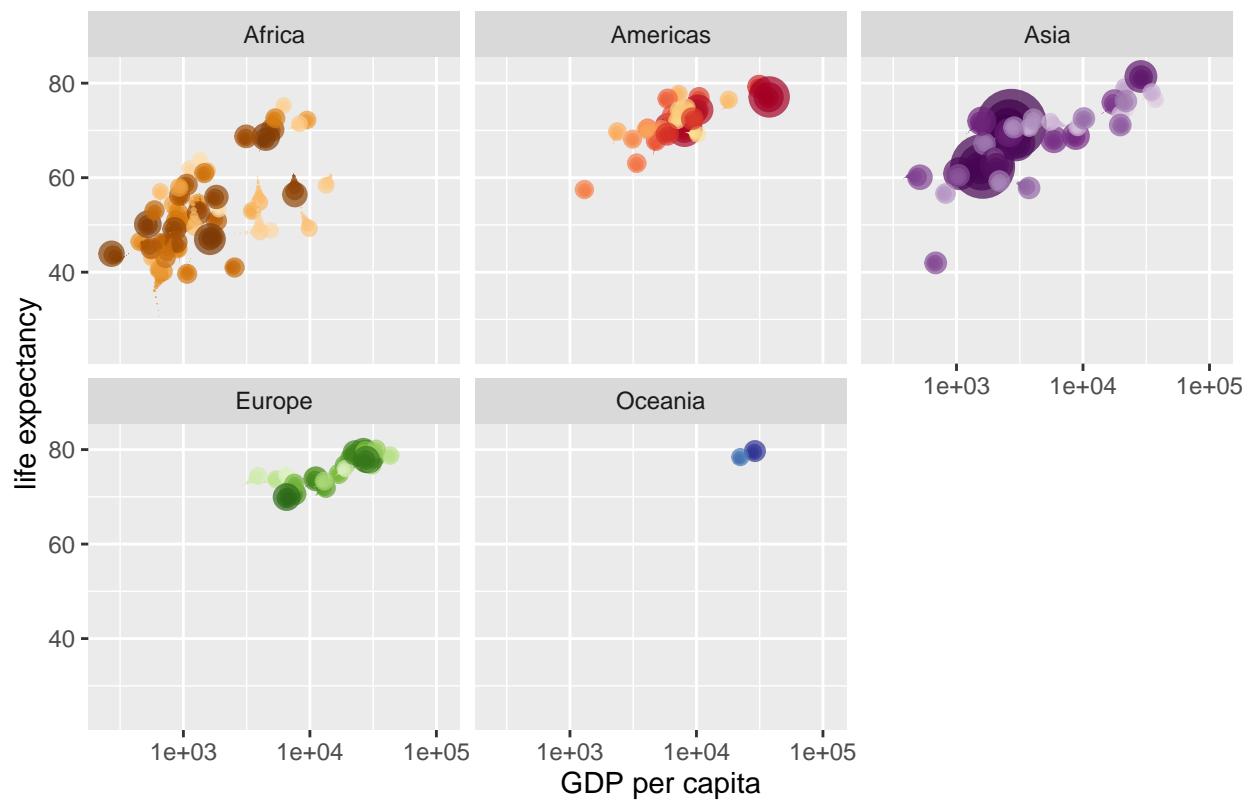
Year: 1999



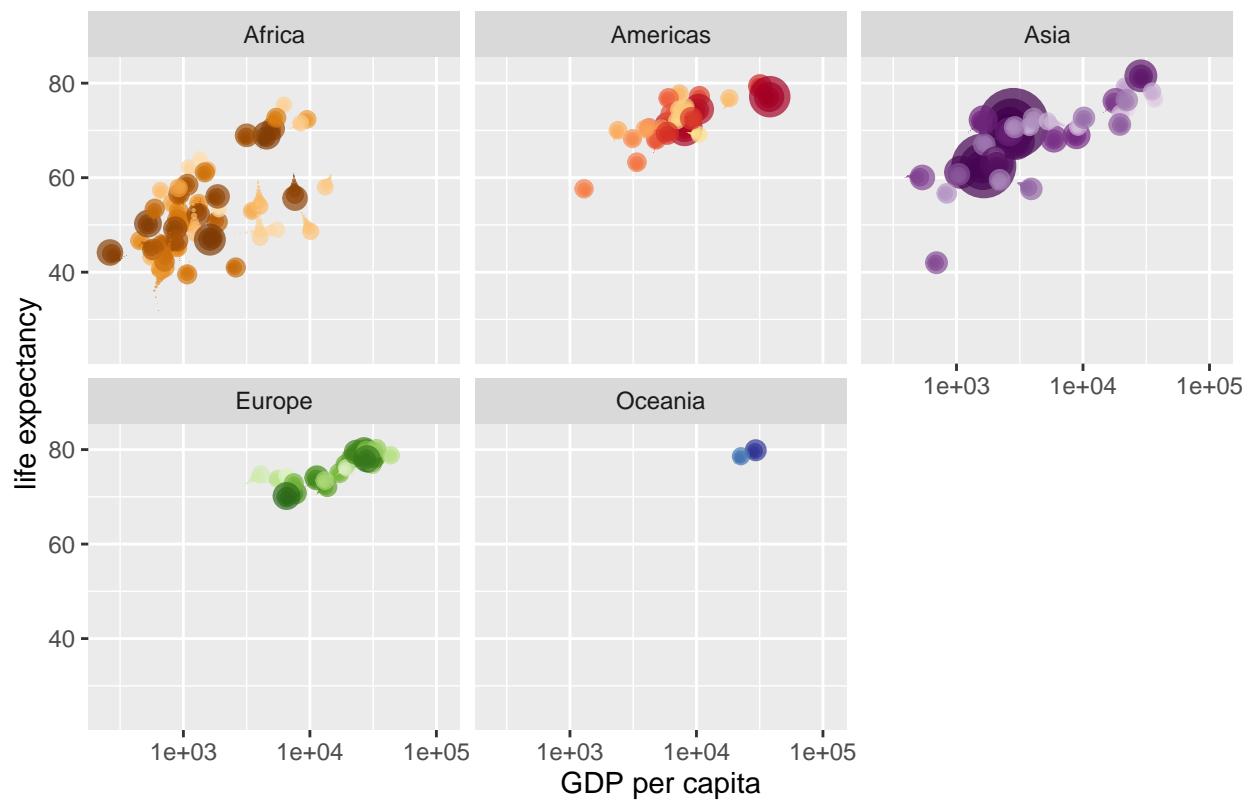
Year: 1999



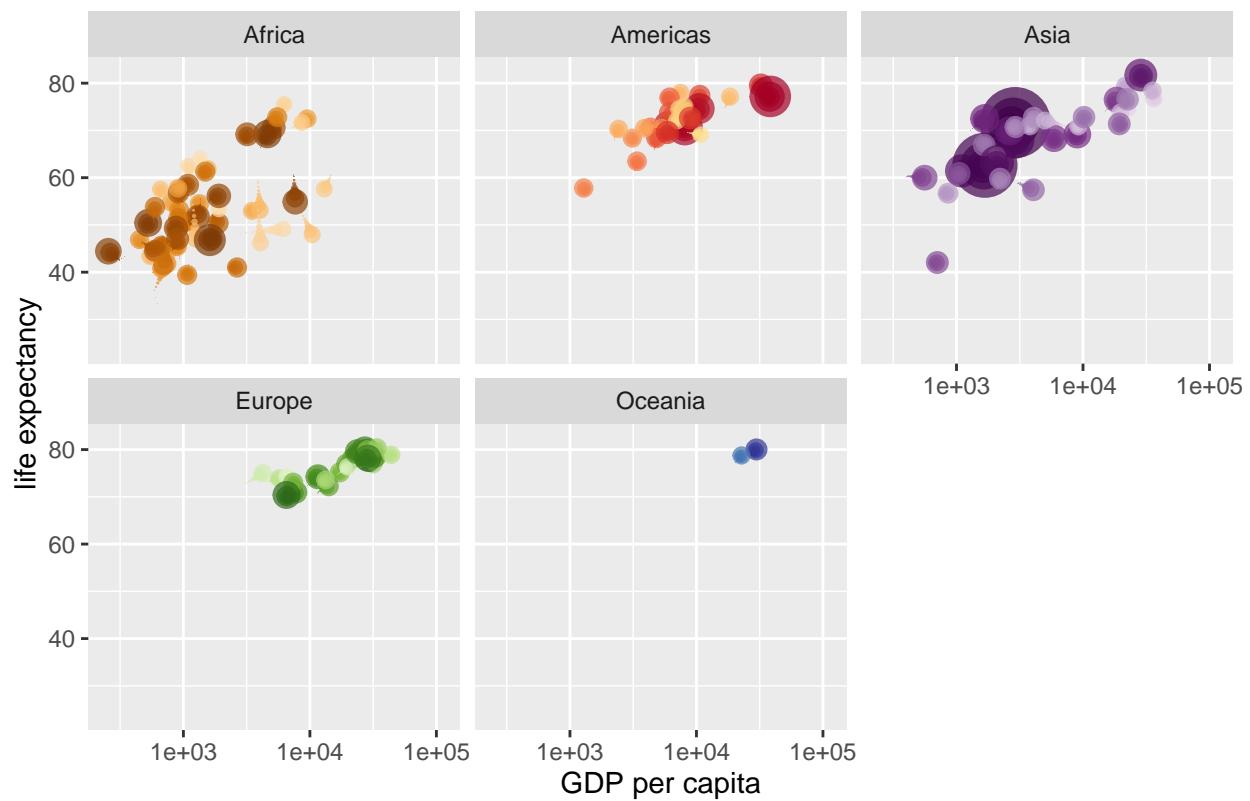
Year: 2000



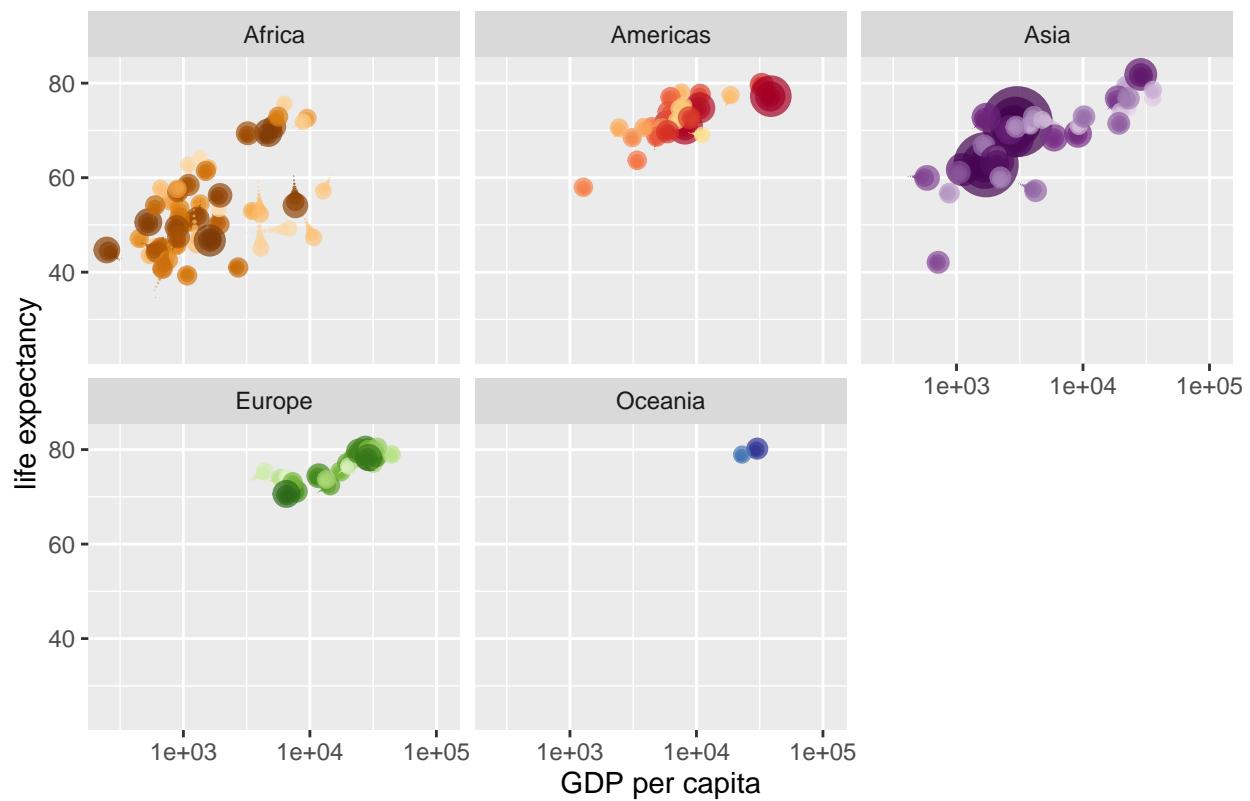
Year: 2000



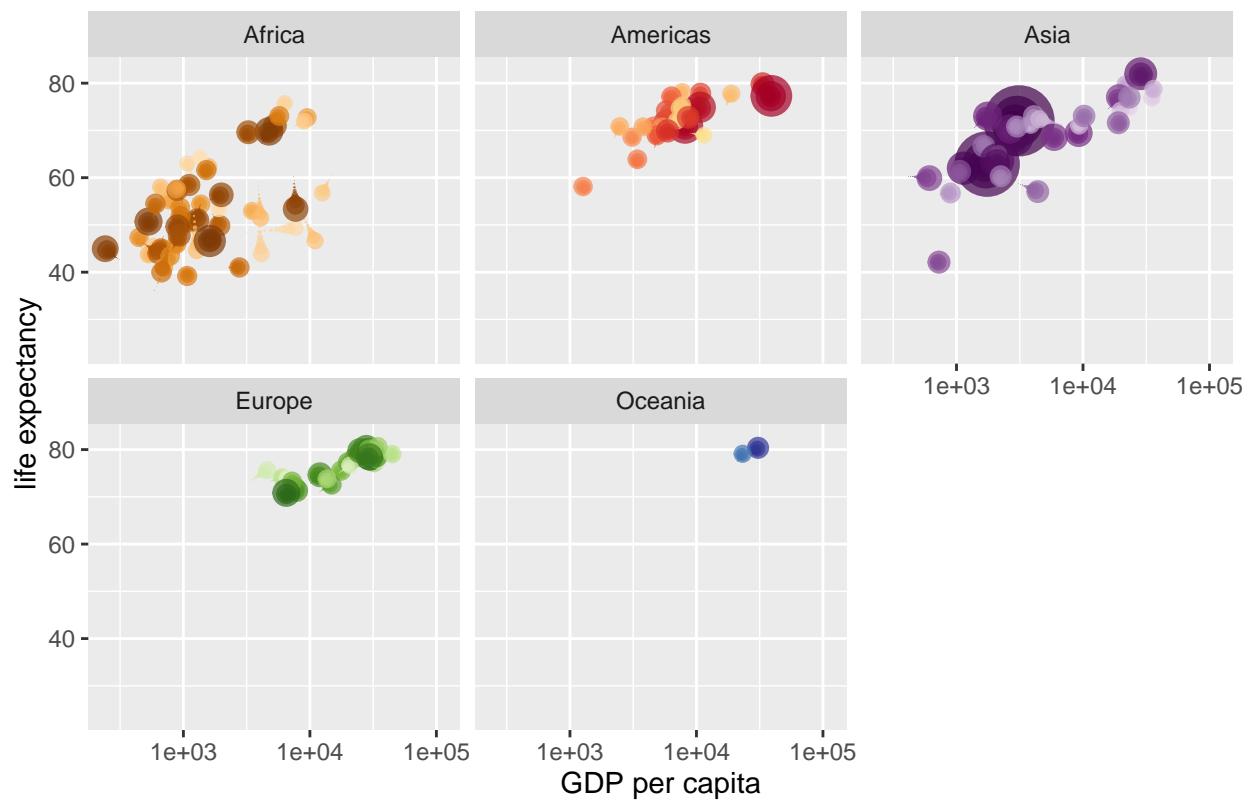
Year: 2001



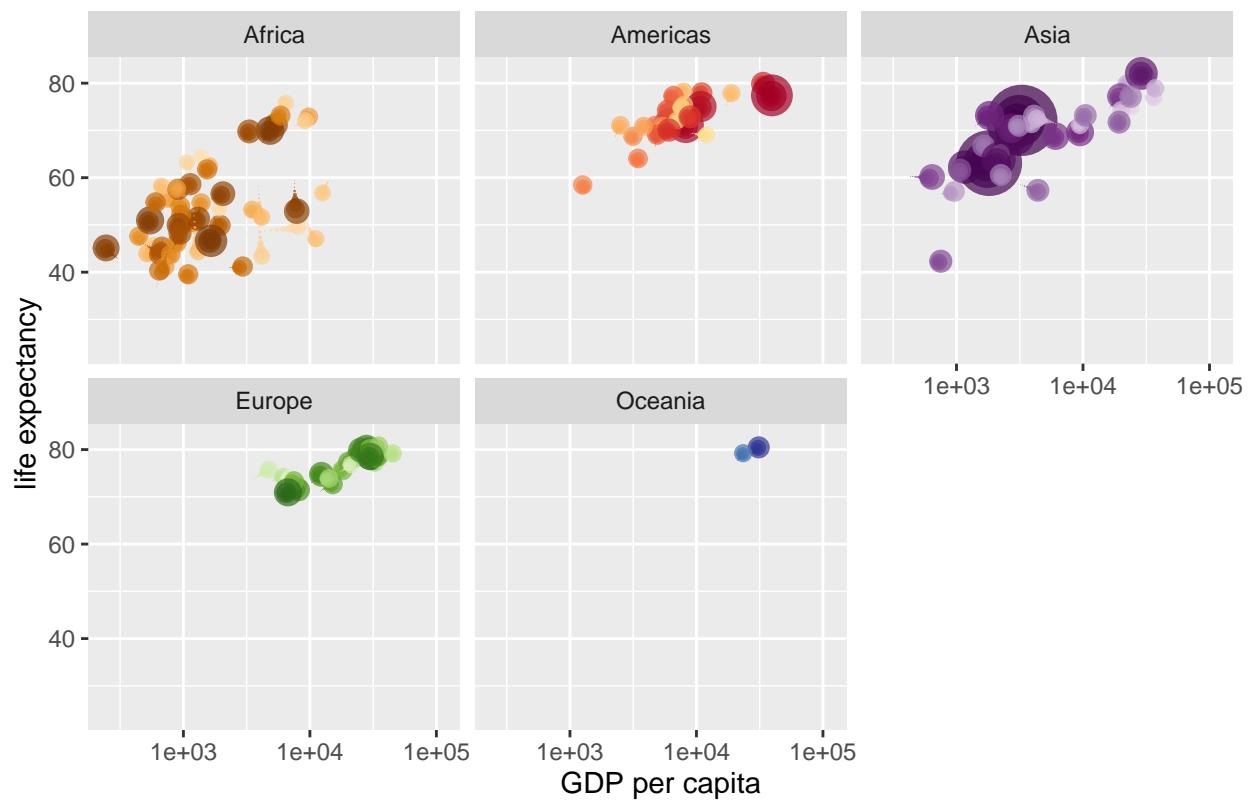
Year: 2001



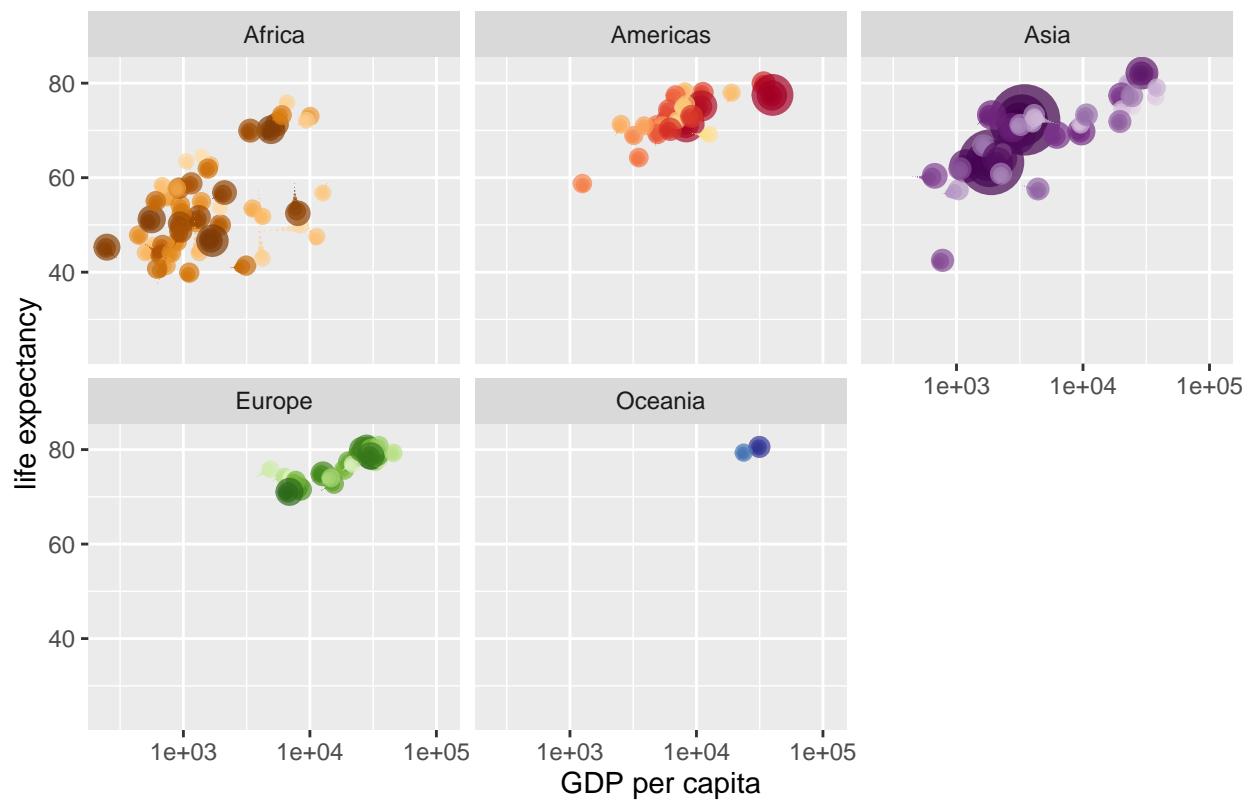
Year: 2002



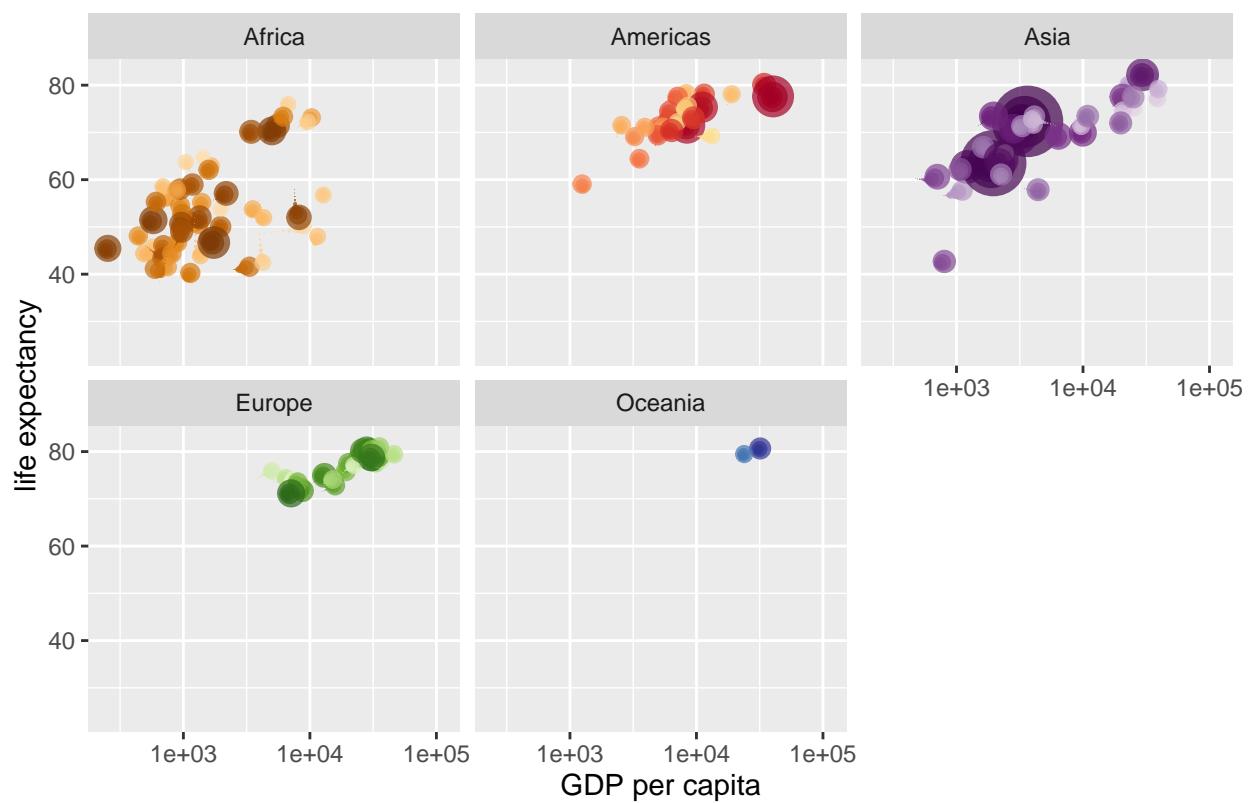
Year: 2003



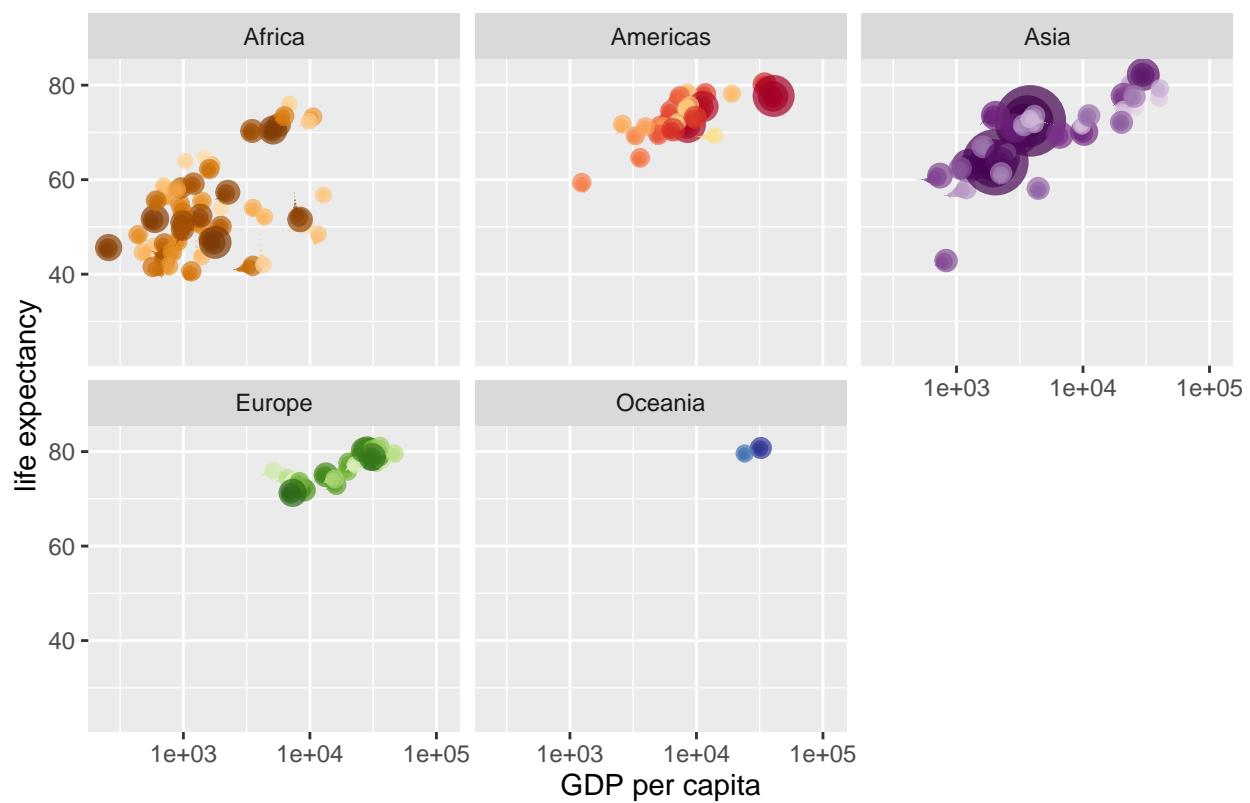
Year: 2003



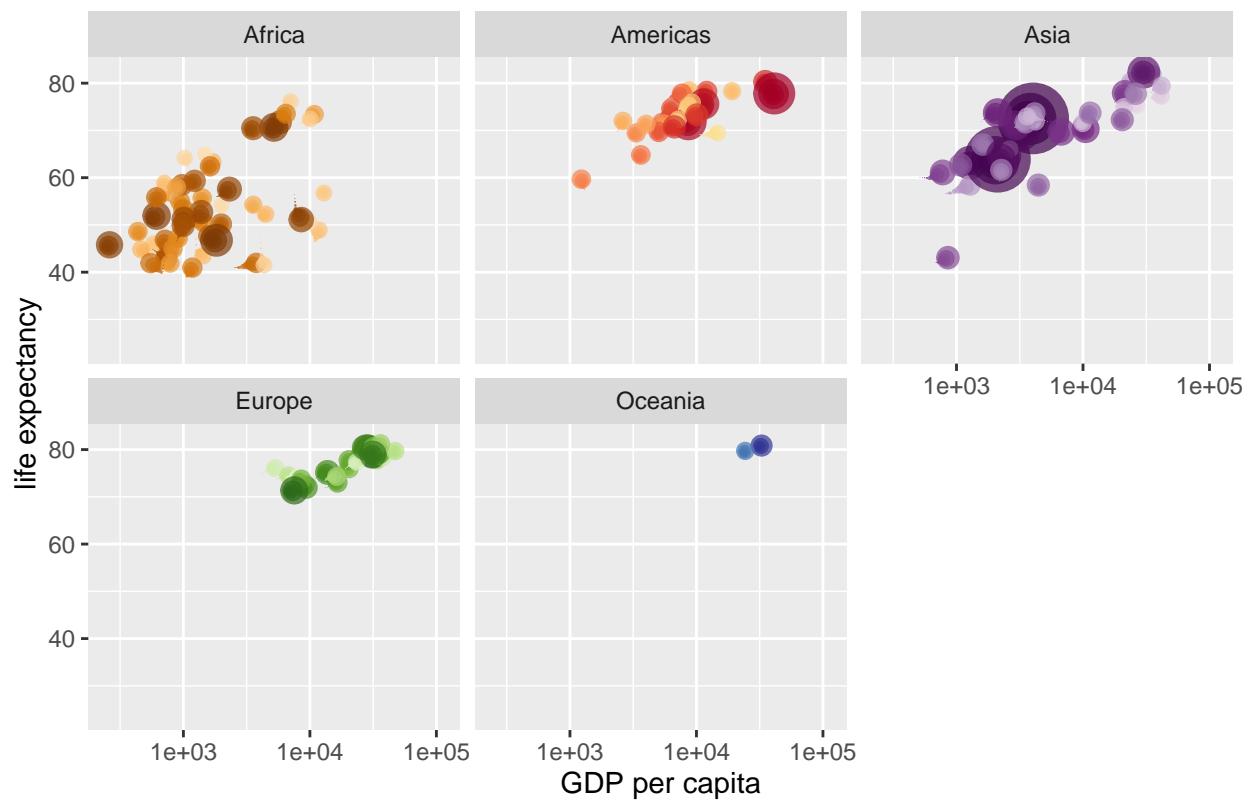
Year: 2004



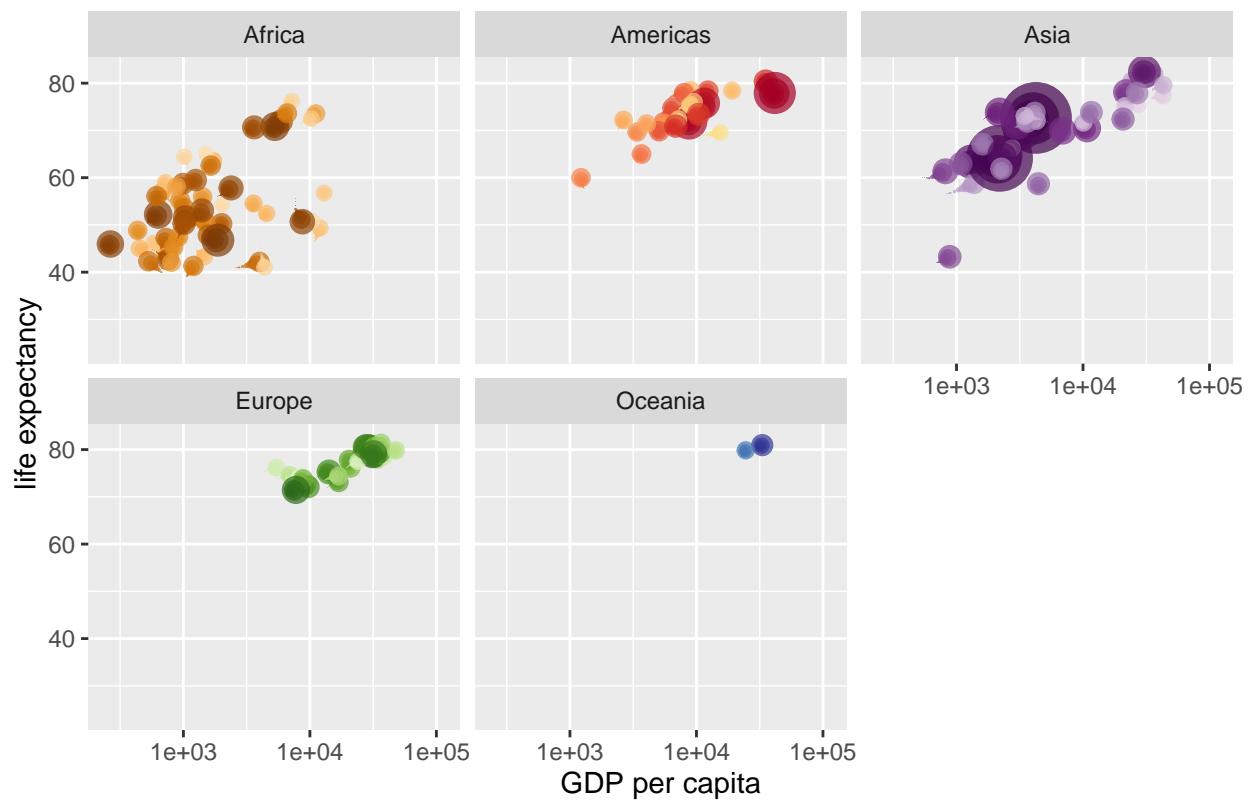
Year: 2004



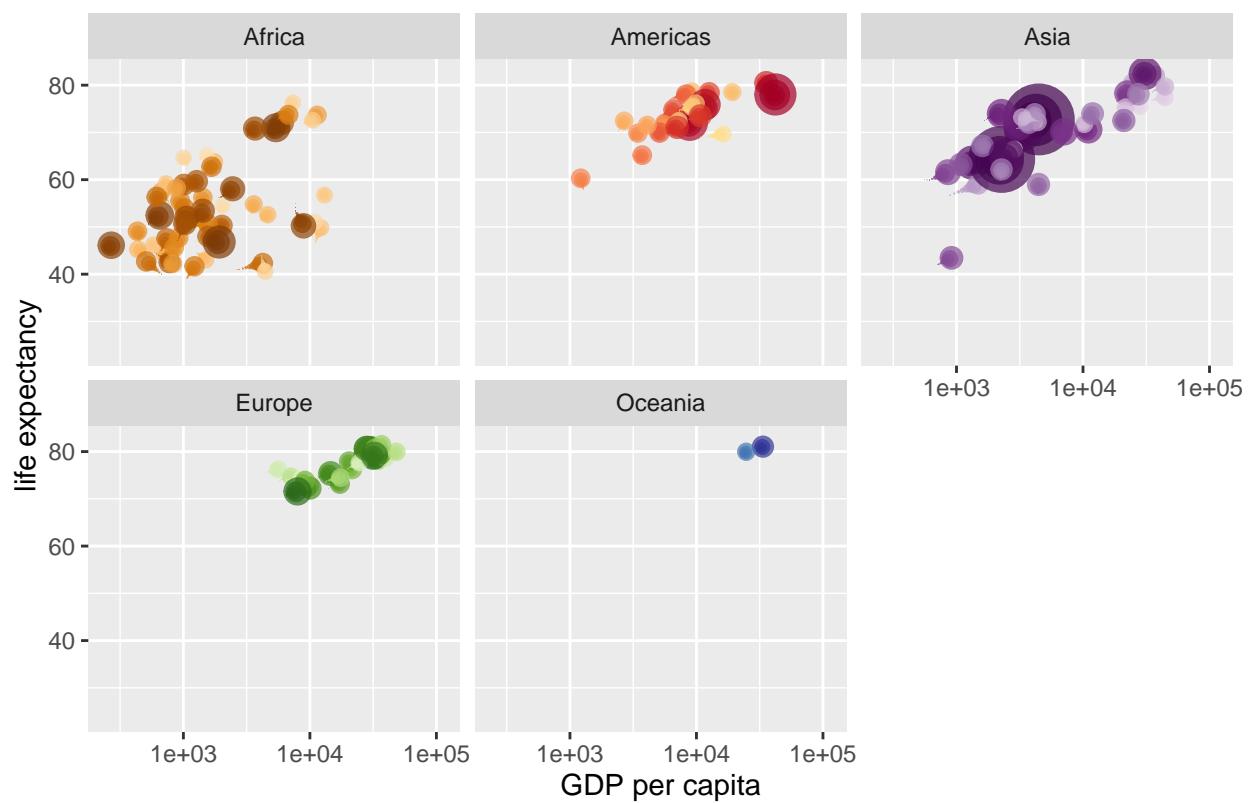
Year: 2005



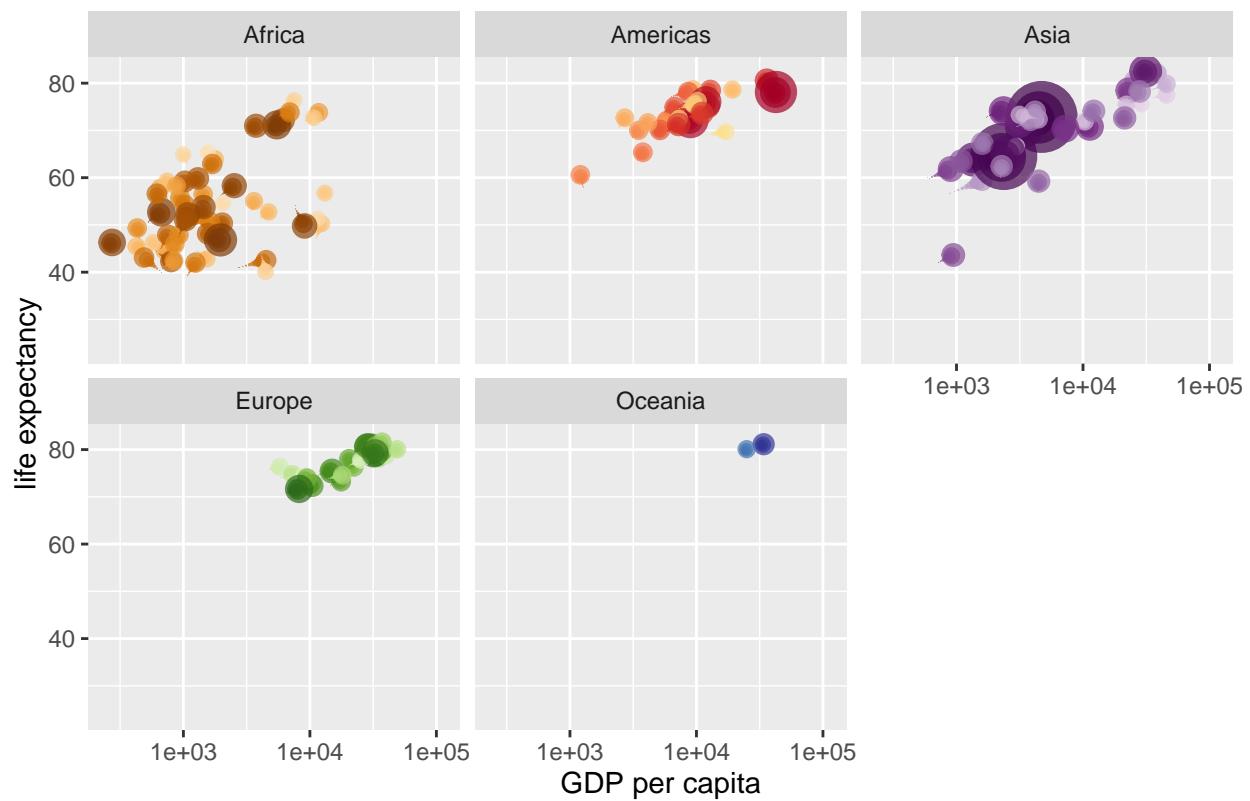
Year: 2005



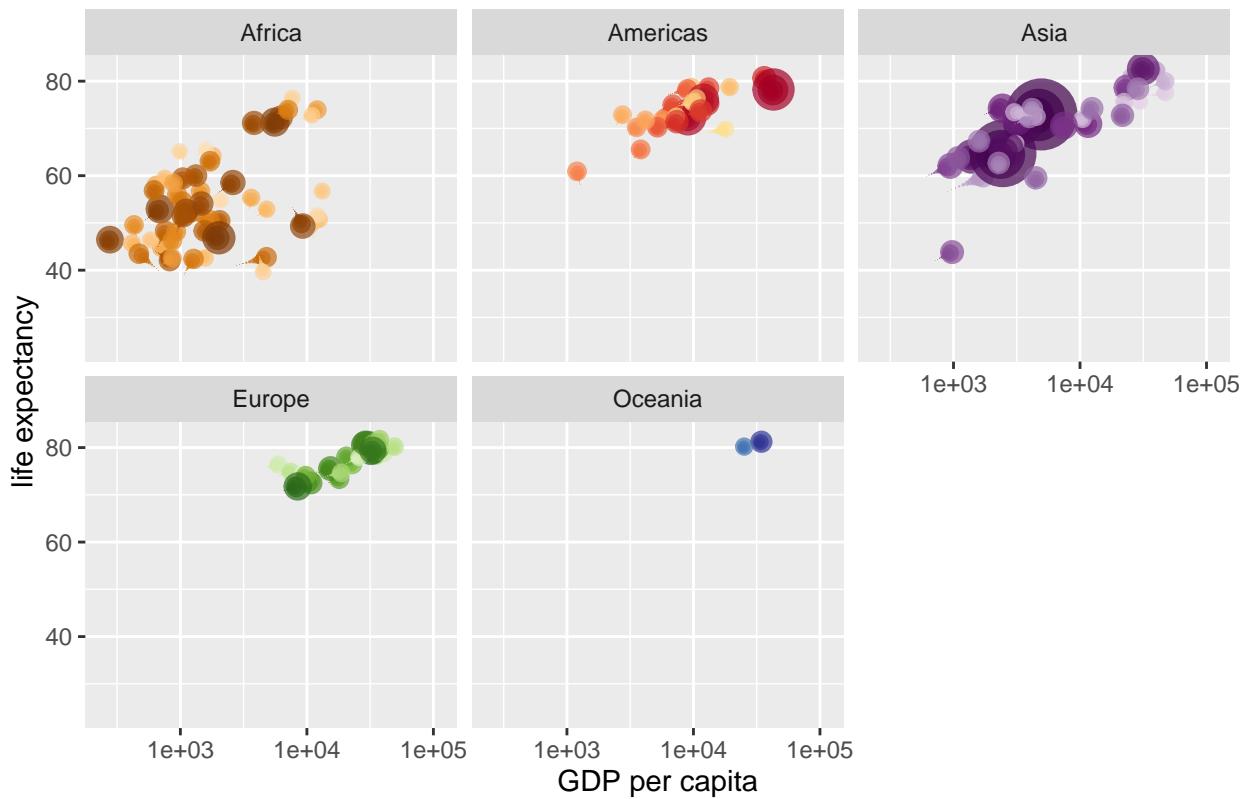
Year: 2006



Year: 2006



Year: 2007



```
#combining graphs
```

```
library(patchwork)

#plots
p1 <- ggplot(mtcars) + geom_point(aes(mpg, disp))
p2 <- ggplot(mtcars) + geom_boxplot(aes(gear, disp, group = gear))
p3 <- ggplot(mtcars) + geom_smooth(aes(disp, qsec))
p4 <- ggplot(mtcars) + geom_bar(aes(carb))

# patchwork combines them
(p1 | p2 | p3) /
  p4
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

