

class05: data vis w/ ggplot

Annmarie (PID:16442048)

Graphics systems in R

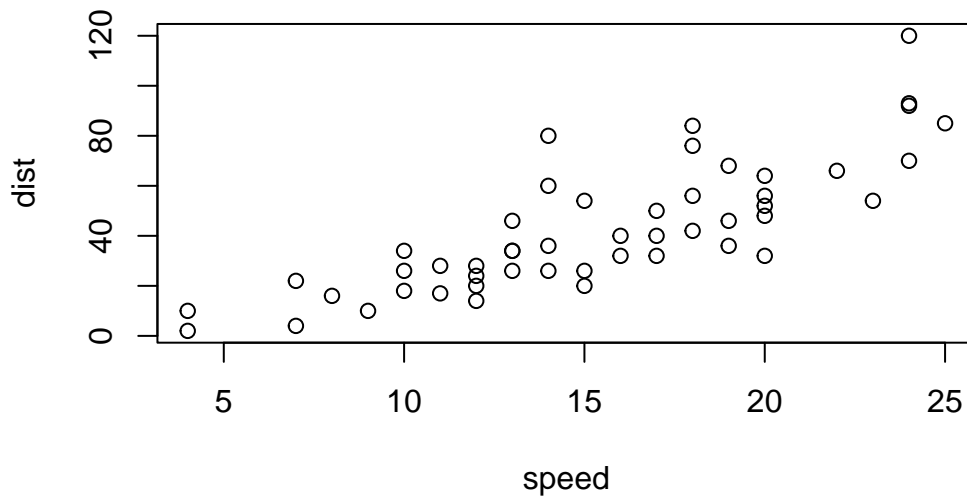
There are many graphics systems in R for making plots and figures.

We have already played a little with “**Base R**” graphics and the `plot()` function.

Today we sill start learning about a popular graphics package called `ggplot2()`.

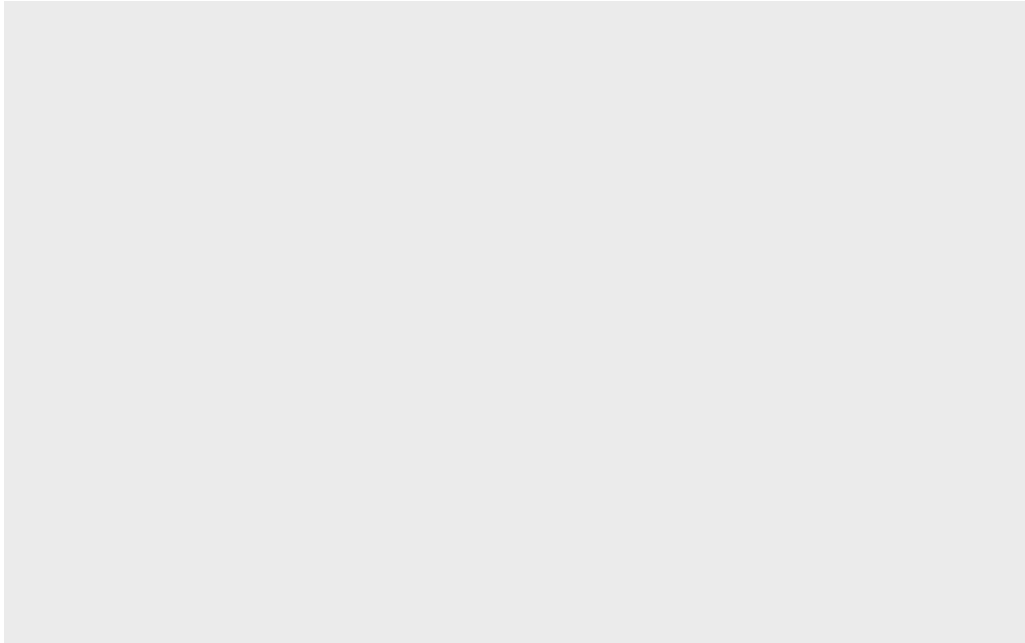
Thus is an add on package - ie. we need to install it. I install it (like I install any package) with the `install.packages()` function.

```
plot(cars)
```



Before I can use the functions from a package I have to load the package from my “library”. We use the `library(ggplot2)` command to look it up

```
library(ggplot2)
ggplot(cars)
```



Every ggplot is made up of at least 3 things: - data (the numbers etc. that will go into your plot) - aes (how the columns of data map to the plot aesthetics) - geoms (how the plot actually looks, points, bars, lines, etc.)

```
ggplot(cars)+
  aes(x=speed, y=dist)+
  geom_point()
```



For simple plots ggplot is more verbose - it takes more to code - than base R plot.

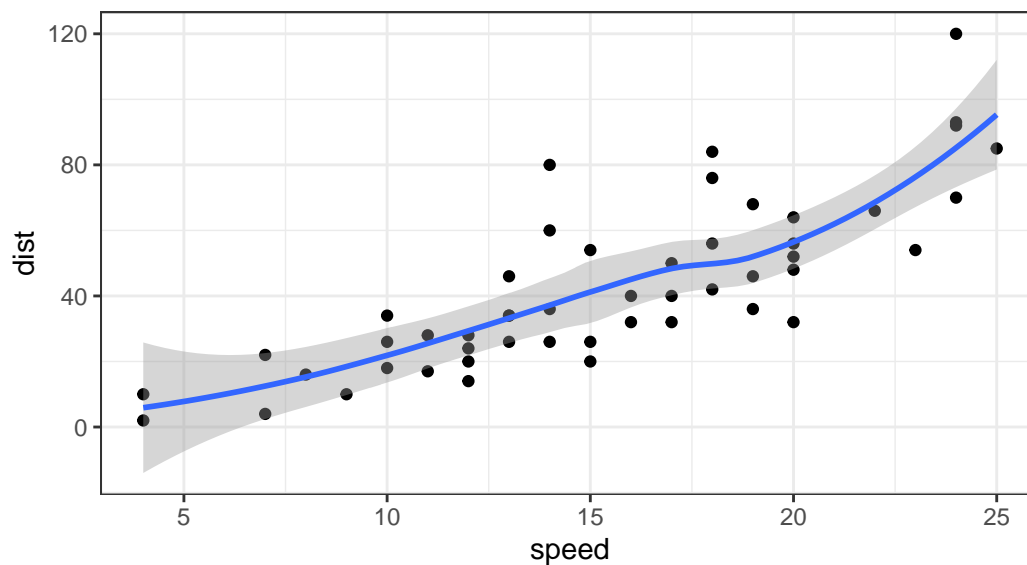
Add some more layers to our ggplot:

```
ggplot(cars)+  
  aes(x=speed, y=dist)+  
  geom_point()+  
  geom_smooth()+  
  labs(title="Stopping distance of old cars",  
        subtitle = "A silly example plot")+  
  theme_bw()
```

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

Stopping distance of old cars

A silly example 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

```
nrow(genes)
```

```
[1] 5196
```

```
colnames(genes)
```

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

```
ncol(genes)
```

```
[1] 4
```

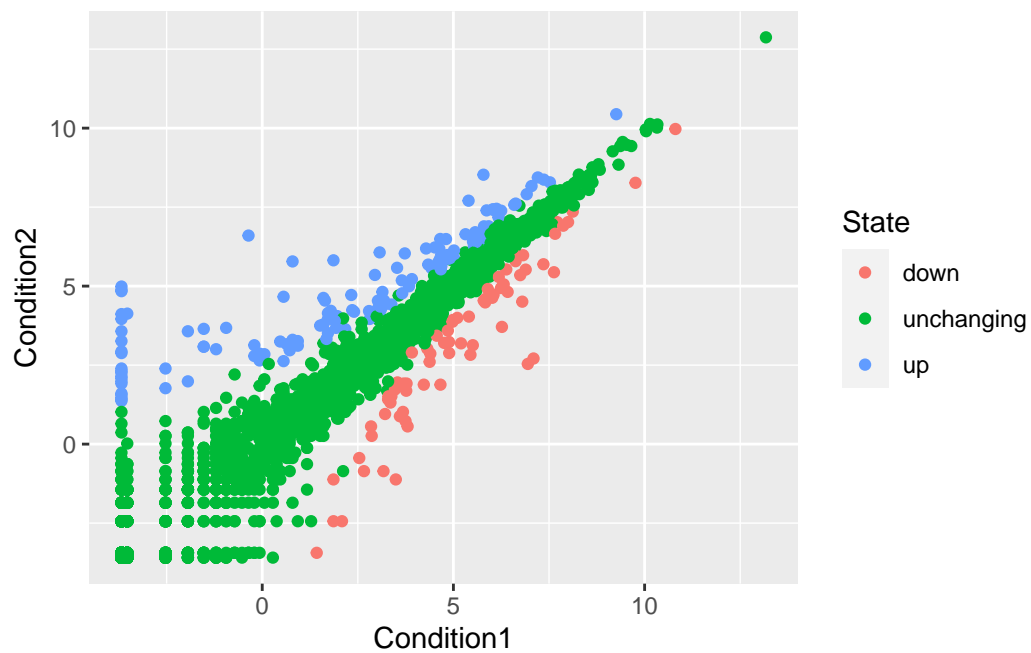
```
table(genes$State)
```

```
down  unchanging      up
   72      4997     127
```

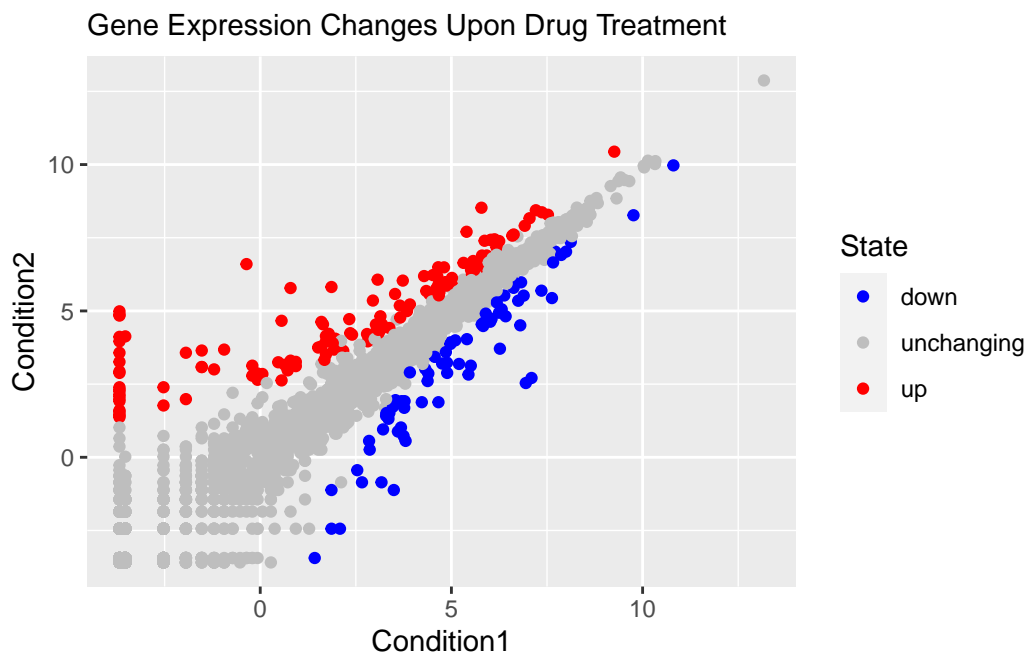
```
round(table(genes$State)/nrow(genes)*100,2)
```

```
down  unchanging      up
 1.39      96.17     2.44
```

```
p <- ggplot(genes)+
  aes(x=Condition1, y=Condition2, col=State)+
  geom_point()
p
```



```
p + scale_colour_manual( values=c("blue","gray","red"))+
  labs(subtitle = "Gene Expression Changes Upon Drug Treatment")
```



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

# install.packages("dplyr") ## un-comment to install if needed
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

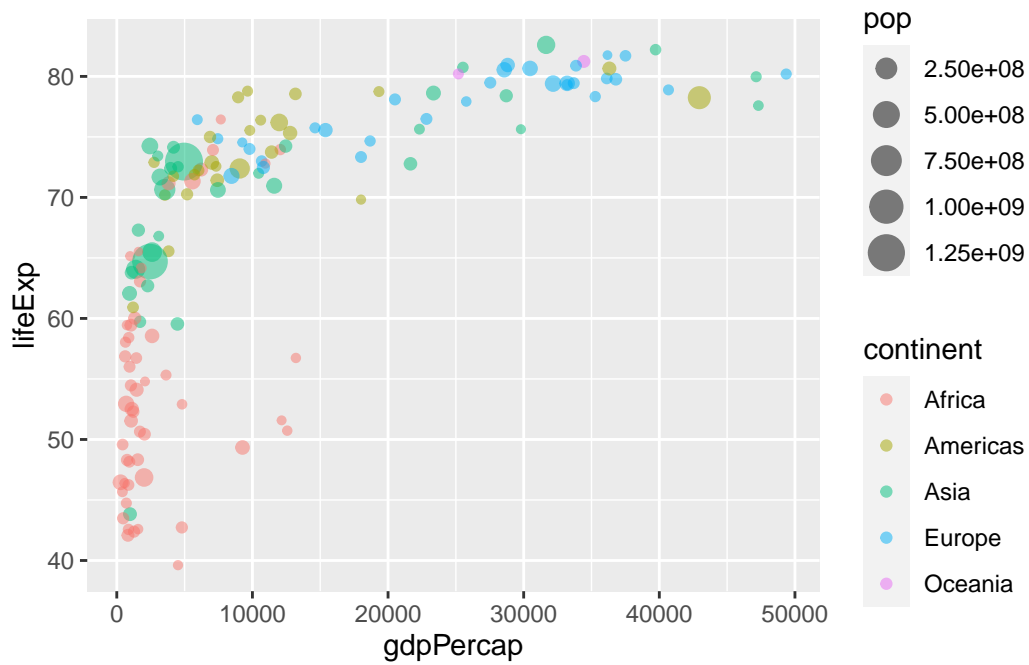
filter, lag

The following objects are masked from 'package:base':

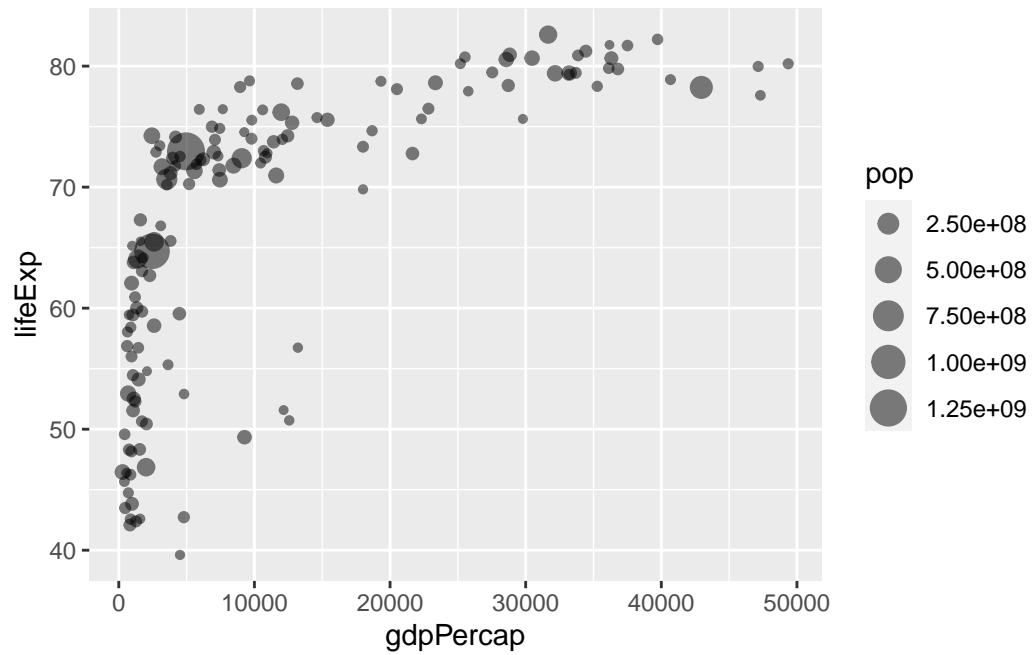
intersect, setdiff, setequal, union

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

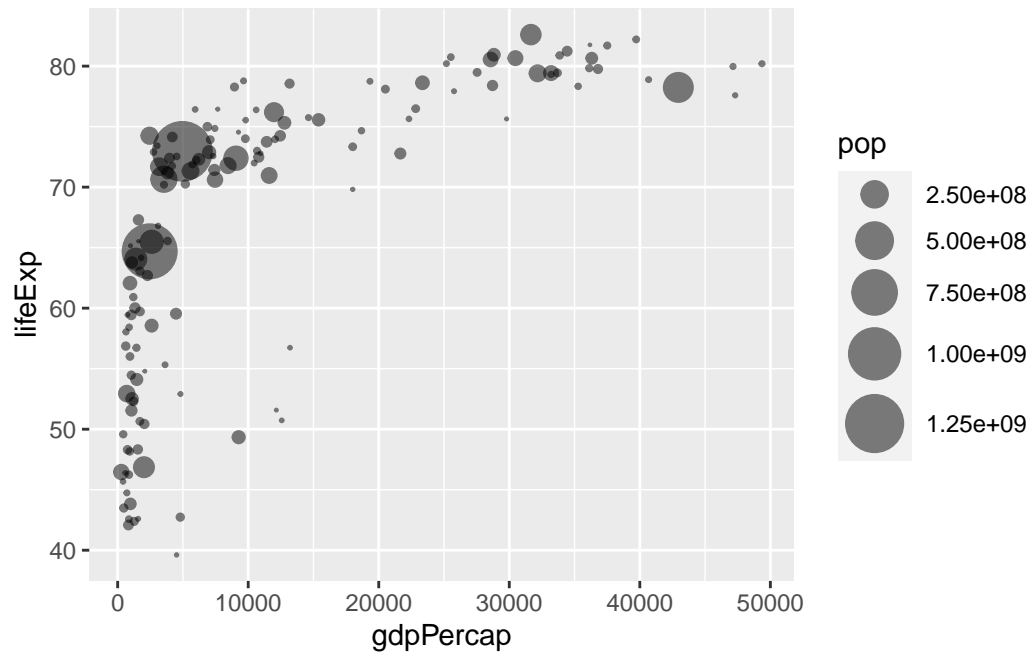
```
ggplot(gapminder_2007)+  
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop)+  
  geom_point(alpha=0.5)
```



```
ggplot(gapminder_2007)+  
  aes(x=gdpPercap, y= lifeExp, size = pop)+  
  geom_point(alpha=0.5)
```



```
ggplot(gapminder_2007)+  
  aes(x=gdpPerCap, y= lifeExp, size = pop)+  
  geom_point(alpha=0.5)+  
  scale_size_area(max_size=10)
```

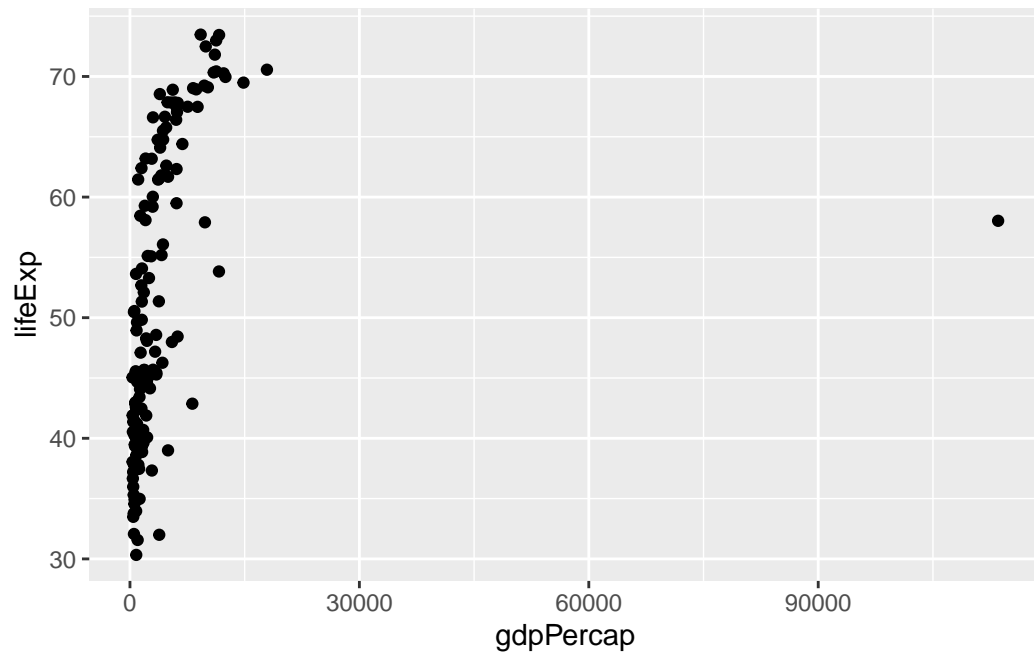



```
gapminder_1957 <- gapminder %>% filter(year==1957)
```

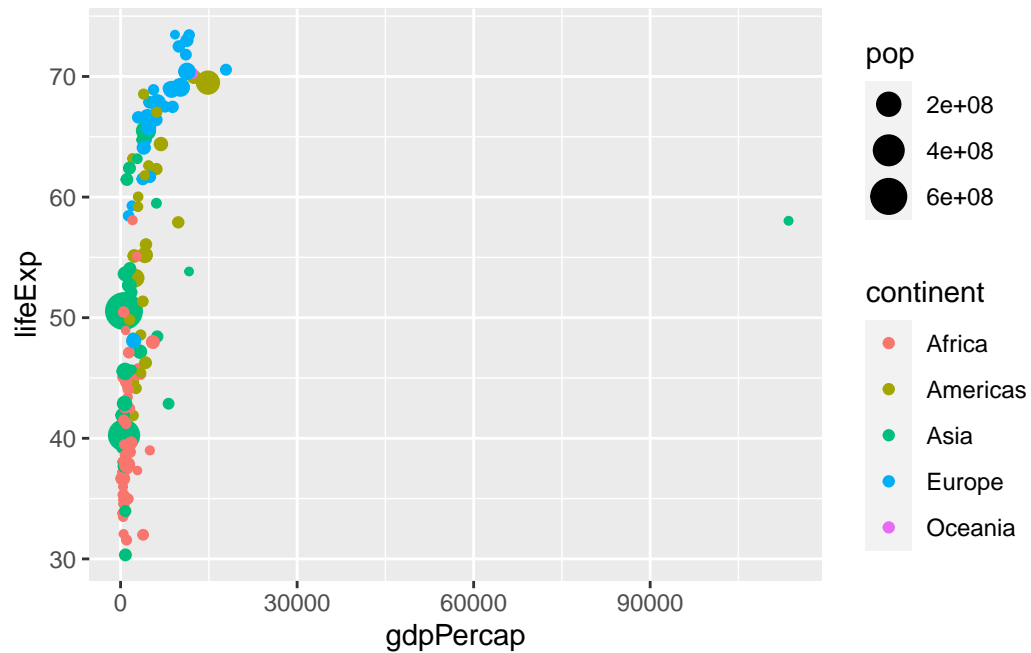
```
colnames(gapminder_1957)
```

```
[1] "country"    "continent"  "year"       "lifeExp"    "pop"        "gdpPercap"
```

```
ggplot(gapminder_1957)+
  aes(x=gdpPercap, y=lifeExp)+
  geom_point()
```



```
ggplot(gapminder_1957)+  
  aes(x=gdpPerCap, y=lifeExp, color=continent, size=pop)+  
  geom_point()
```



```
gapminder_both<- gapminder %>% filter (year==1957 | year==2007)
```

```
ggplot(gapminder_both)+
  geom_point(aes(x = gdpPercap, y = lifeExp, color=continent,
                 size = pop), alpha=0.7) +
  scale_size_area(max_size = 10) +
  facet_wrap(~year)
```

