

class 05 data visualization with ggplot2

Trinity Lee A16639698

Using GGLOT

The ggplot2 package does not already come installed with R

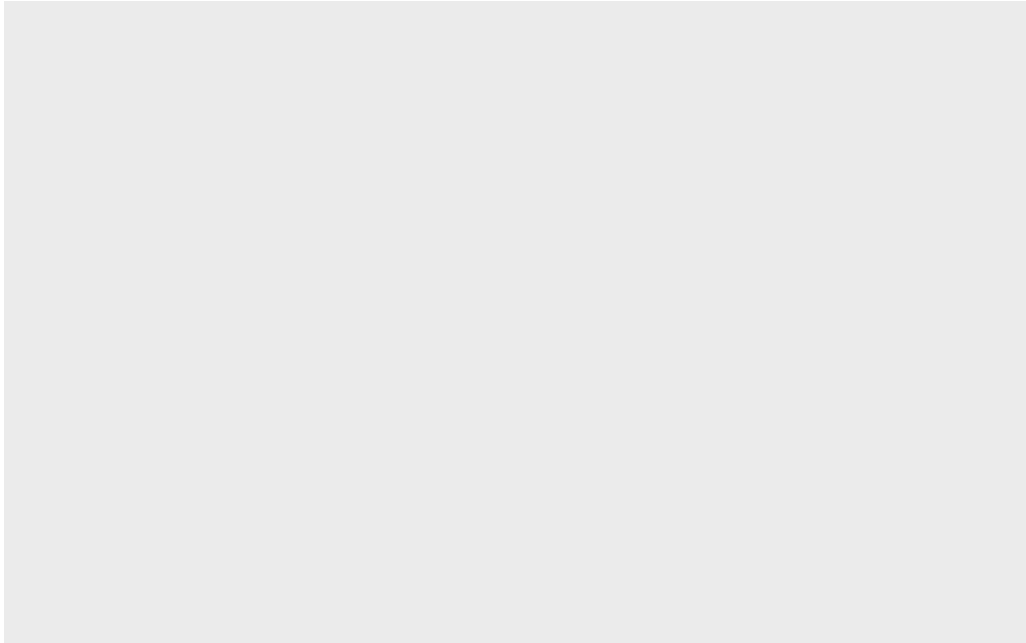
We have to use `install.packages()` function to install ggplot2

```
head(cars)
```

	speed	dist
1	4	2
2	4	10
3	7	4
4	7	22
5	8	16
6	9	10

To use ggplot I need to load it up before I can call any of the functions in the package. I do this with the `library()` function.

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



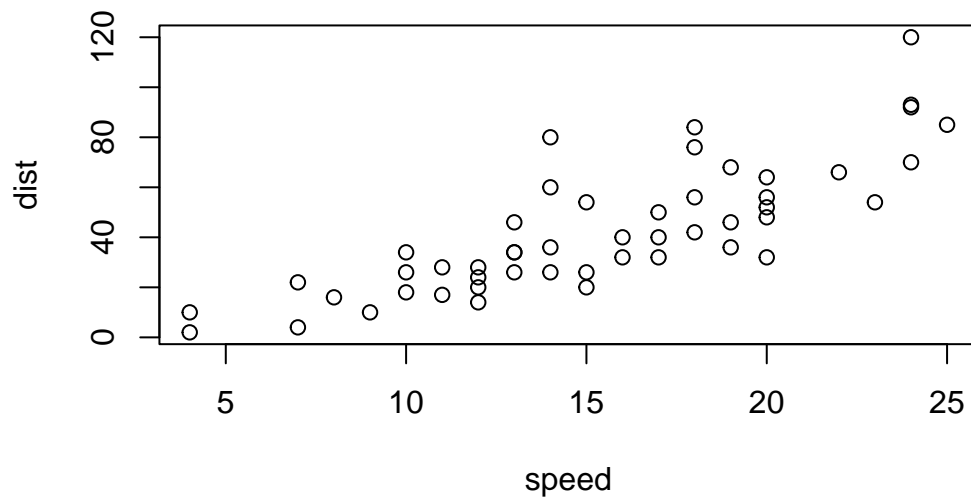
All ggplot figures have at least three things: - data (the stuff we want to plot) - aesthetic mapping (aes vales) - geoms

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



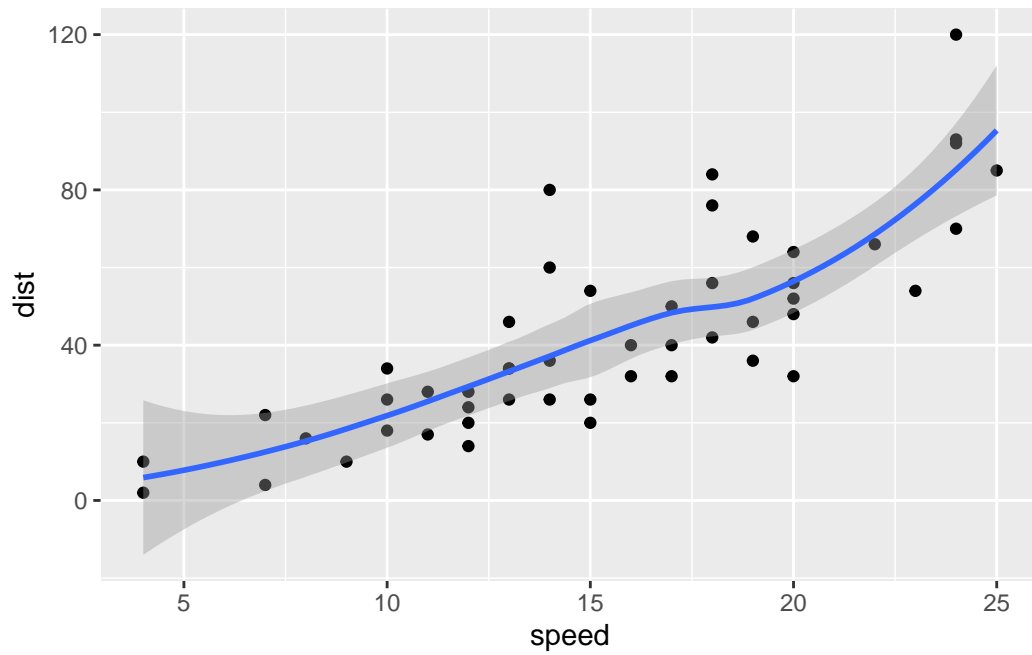
ggplot is not the only graphing system in R there are lots of others. There is even “base R” graphics.

```
plot(cars)
```



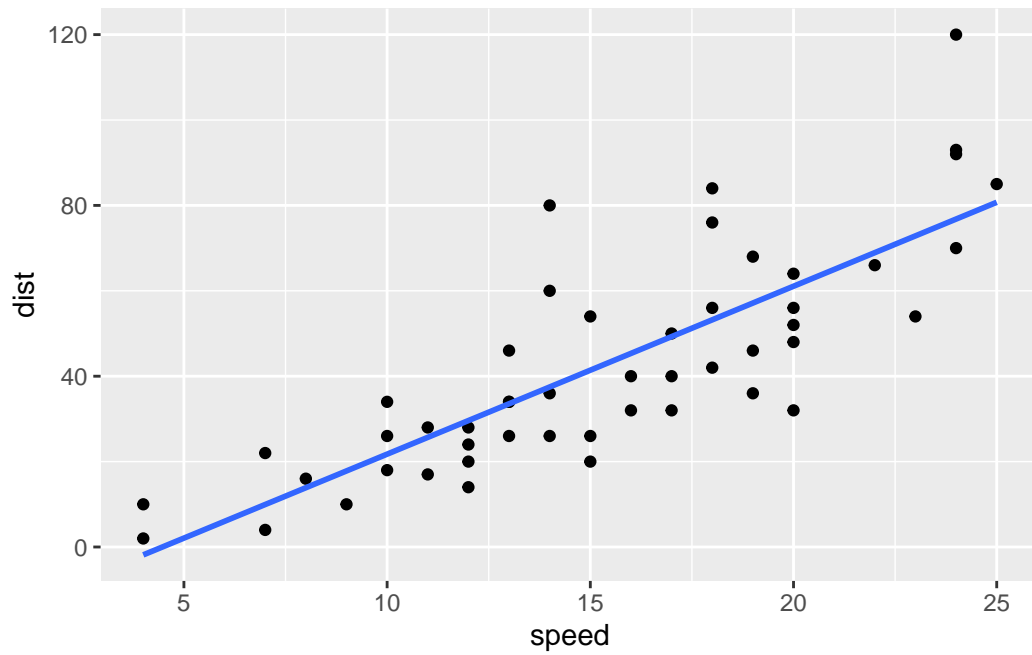
```
ggplot(cars)+aes(x=speed,y=dist) + geom_point() + geom_smooth()
```

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



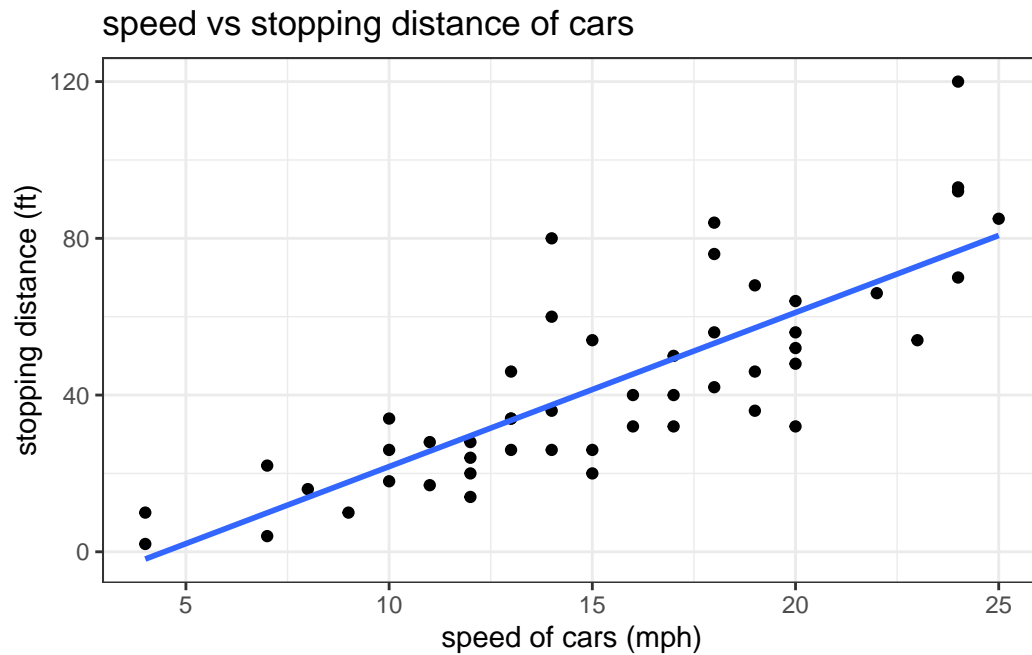
```
ggplot(cars)+aes(x=speed,y=dist) + geom_point() + geom_smooth(method="lm",se=FALSE)
```

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



```
ggplot(cars)+aes(x=speed,y=dist) + geom_point() + geom_smooth(method="lm",se=FALSE)+labs(t
```

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



```
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
```

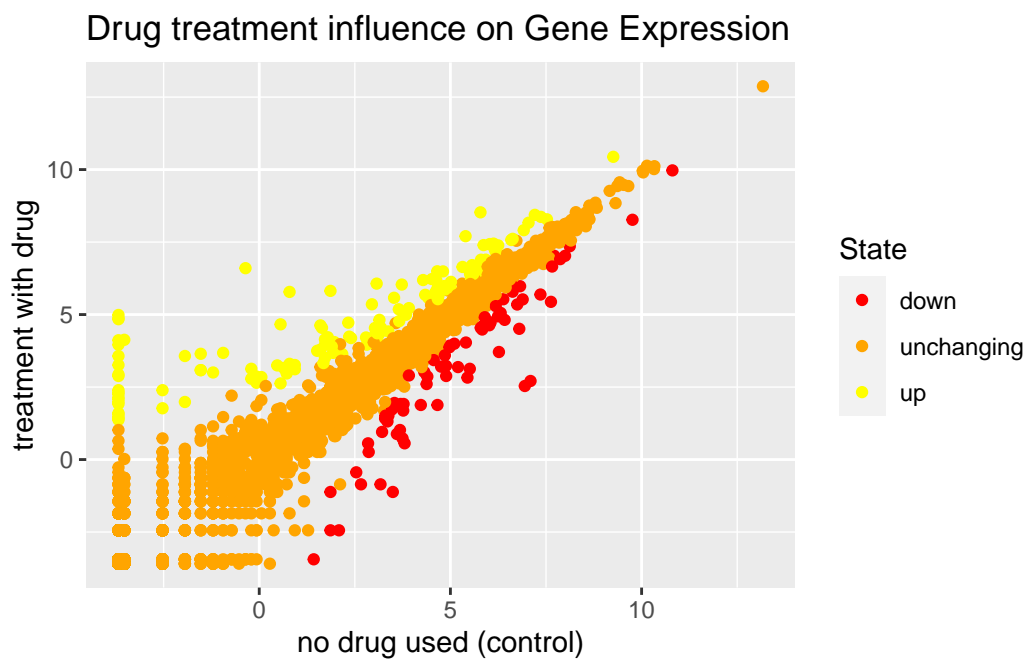
```
table(genes$State)
```

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

```
table(genes$State)/nrow(genes)*100
```

```
down  unchanging      up  
1.385681  96.170131  2.444188
```

```
p<-ggplot(genes)+aes(x=Condition1, y=Condition2, col=State)+geom_point()  
p+scale_color_manual (values=c("red","orange","yellow"))+labs(title="Drug treatment influence
```




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

Attaching package: 'dplyr'

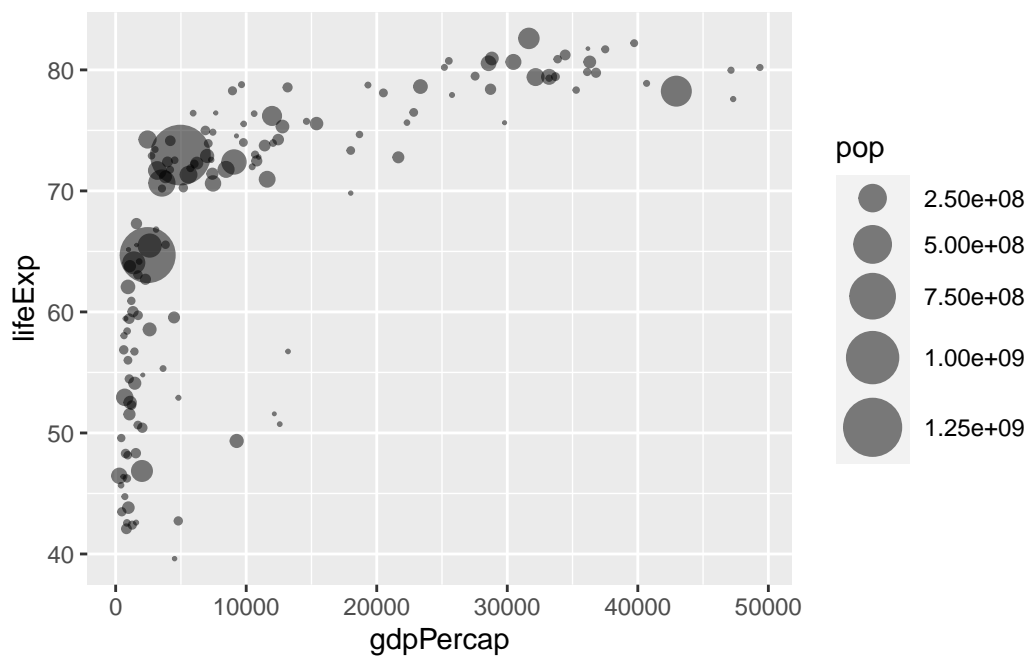
The following objects are masked from 'package:stats':

filter, lag

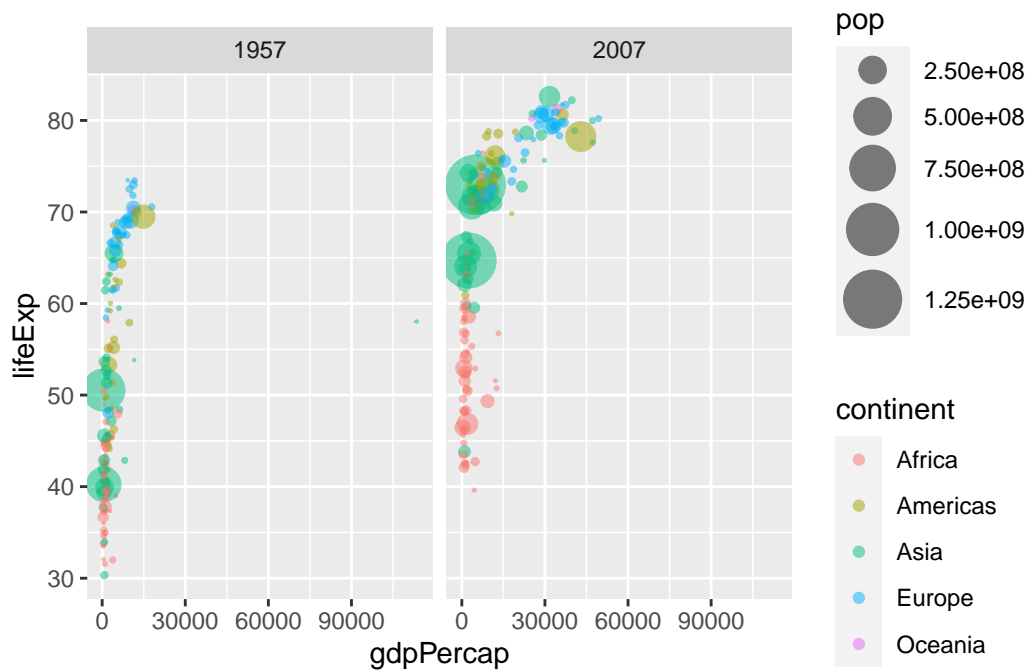
The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(gapminder)
gapminder_2007 <- gapminder %>% filter(year==2007)
ggplot(gapminder_2007)+aes(x=gdpPercap, y=lifeExp, size=pop)+geom_point(alpha=0.5)+scale_s
```



```
gapminder_1957<-gapminder %>% filter(year==1957|year==2007)
ggplot(gapminder_1957)+aes(x=gdpPercap, y=lifeExp, size=pop,color=continent)+geom_point(al
```



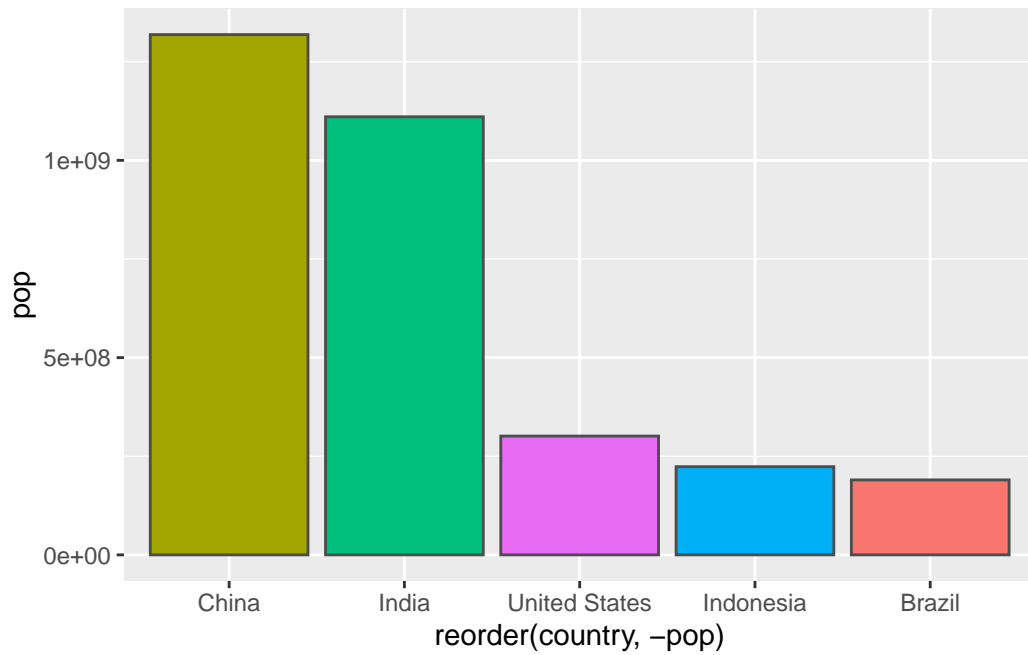
```
gapminder_top5 <- gapminder %>%
  filter(year==2007) %>%
  arrange(desc(pop)) %>%
  top_n(5, pop)
```

```
gapminder_top5
```

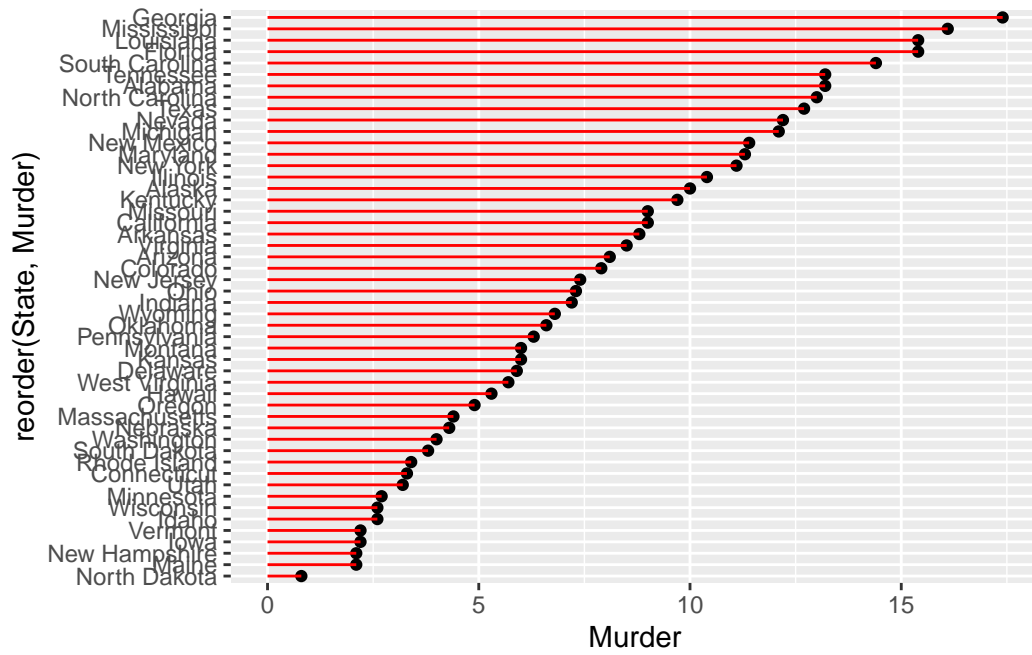
A tibble: 5 x 6

	country	continent	year	lifeExp	pop	gdpPercap
	<fct>	<fct>	<int>	<dbl>	<int>	<dbl>
1	China	Asia	2007	73.0	1318683096	4959.
2	India	Asia	2007	64.7	1110396331	2452.
3	United States	Americas	2007	78.2	301139947	42952.
4	Indonesia	Asia	2007	70.6	223547000	3541.
5	Brazil	Americas	2007	72.4	190010647	9066.

```
ggplot(gapminder_top5)+aes(x=reorder(country,-pop),y=pop,fill=country)+ geom_col(col="gray
```



```
USArrests$State <- rownames(USArrests)
ggplot(USArrests) +
  aes(x=reorder(State,Murder), y=Murder) +
  geom_point() +geom_segment(aes(x=State,xend=State,y=0,yend=Murder),color="red")+coord_fl
```



““