

# Class 5: Data Visualization with ggplot

Isaac (PID: A15493599)

## Using ggplot

Load in ggplot Package

```
library(ggplot2)
```

3 Things Necessary for ggplots:

- data: the actual data plotted as a dataframe
- aesthetics: defining axes, how the data is plotted
- geoms: stylistic characteristic of plot

## Scatterplot of Cars

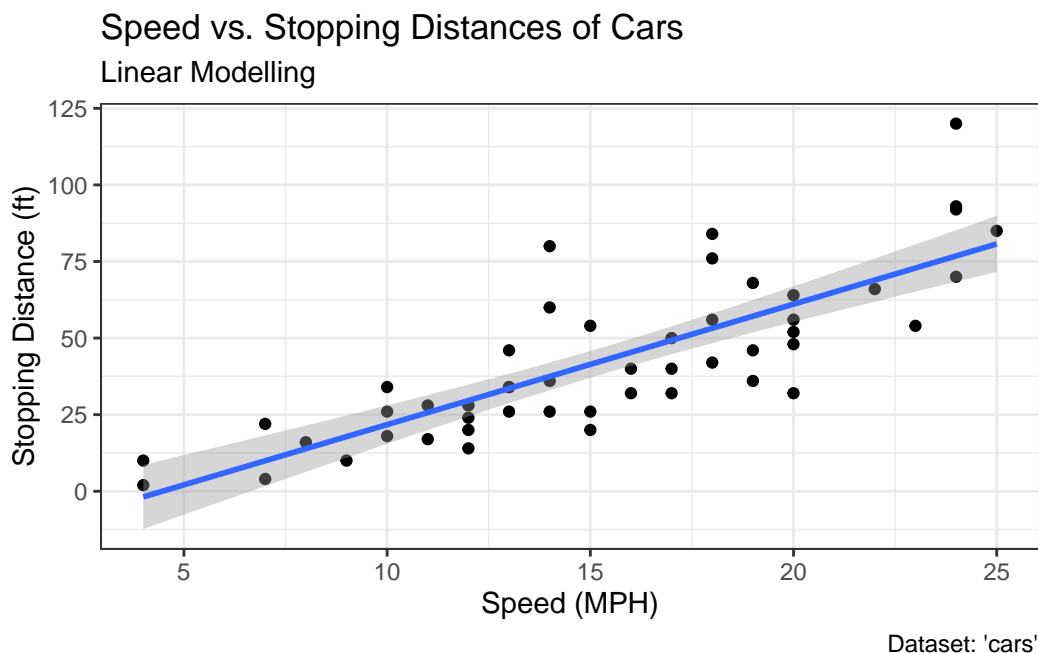
```
head(cars)
```

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

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

```
labs(title="Speed vs. Stopping Distances of Cars",
      x="Speed (MPH)",
      y="Stopping Distance (ft)",
      subtitle = "Linear Modelling",
      caption="Dataset: 'cars'")
```

`geom\_smooth()` using formula = 'y ~ x'



## Using the aes() function in Anti-viral drug trial

```
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

```

Base ggplot of Anti-viral treatment

```

P <- ggplot(genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point(alpha = 0.3)

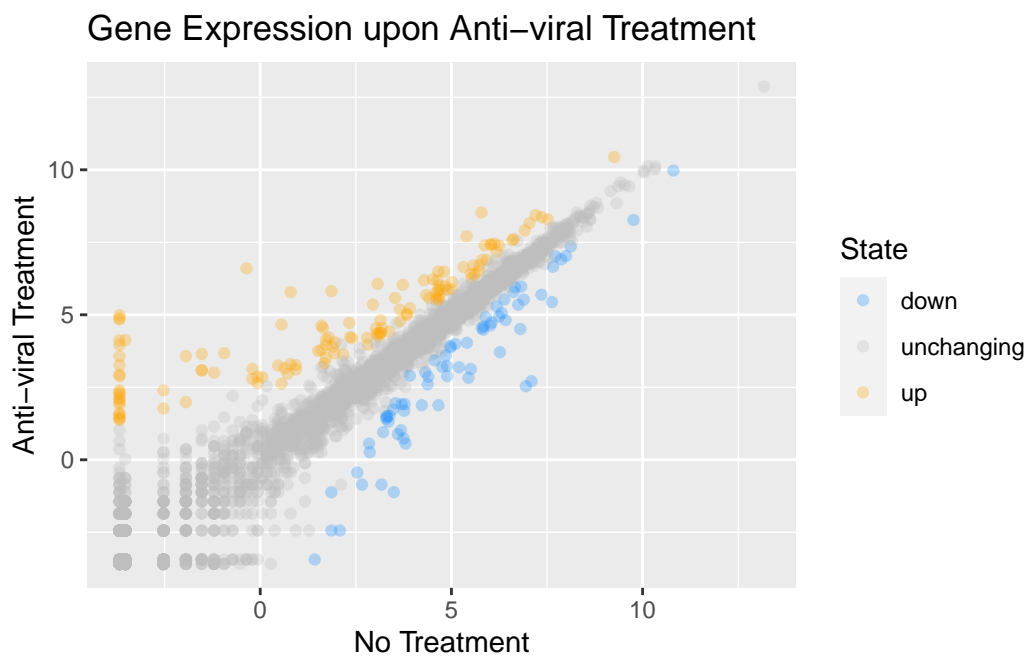
```

Changing Colors and Adding Labels

```

P +
  scale_colour_manual( values=c("dodgerblue1","grey","orange1"))+
  labs(title="Gene Expression upon Anti-viral Treatment",
       x="No Treatment",
       y="Anti-viral Treatment")

```



Install and Load Gapminder and Dplyr

```

library(gapminder)
library(dplyr)

```

Attaching package: 'dplyr'

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

`filter`, `lag`

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

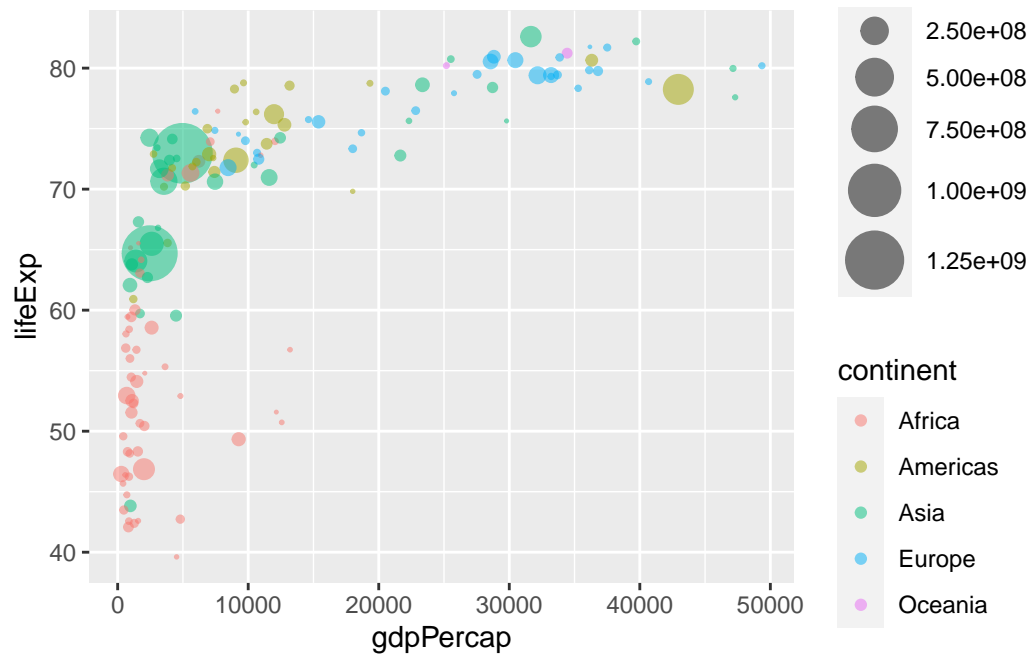
`intersect`, `setdiff`, `setequal`, `union`

Open TSV File from Online

```
url <- "https://raw.githubusercontent.com/jennybc/gapminder/master/inst/extdata/gapminder.  
gapminder <- read.delim(url)  
gapminder_2007 <- gapminder %>% filter(year==2007)
```

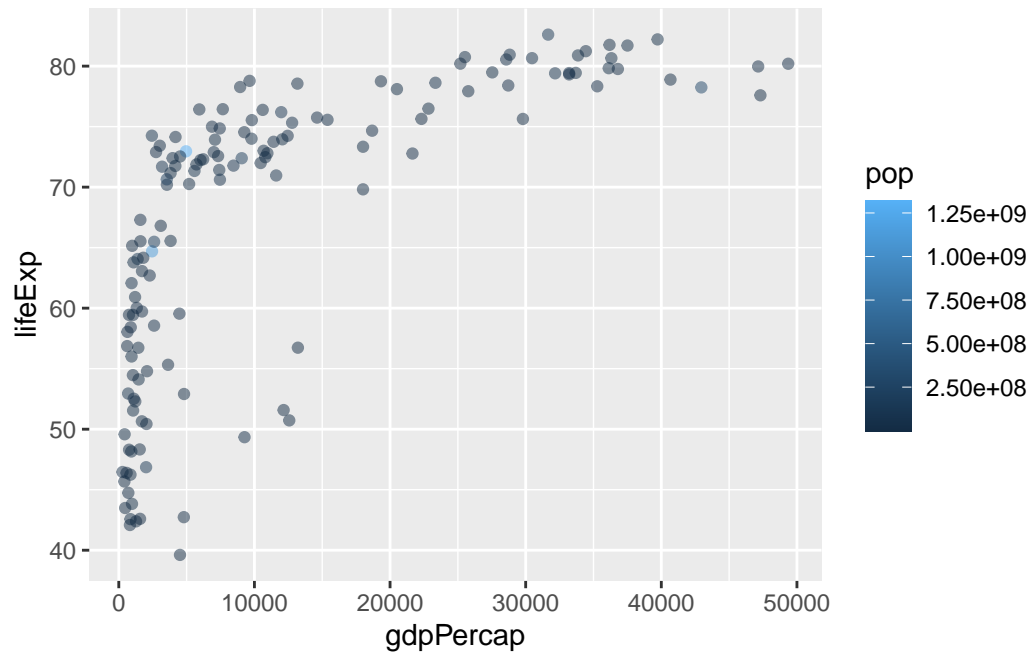
Population by Point Size

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



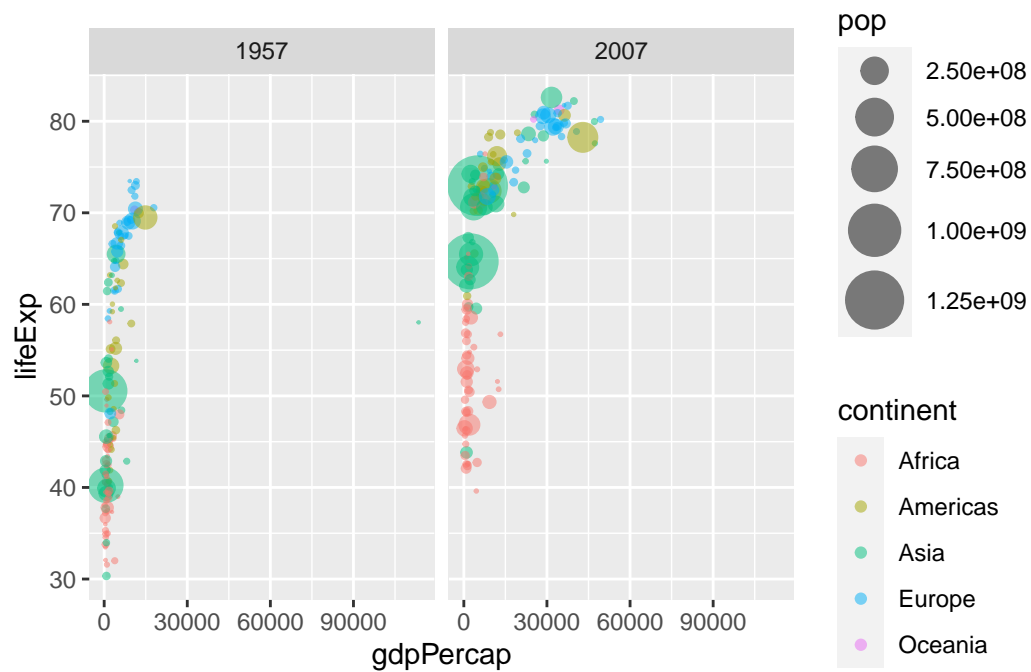
Population by Point Color

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



1957 vs 2007 GDP vs Life Expectancy by Continent

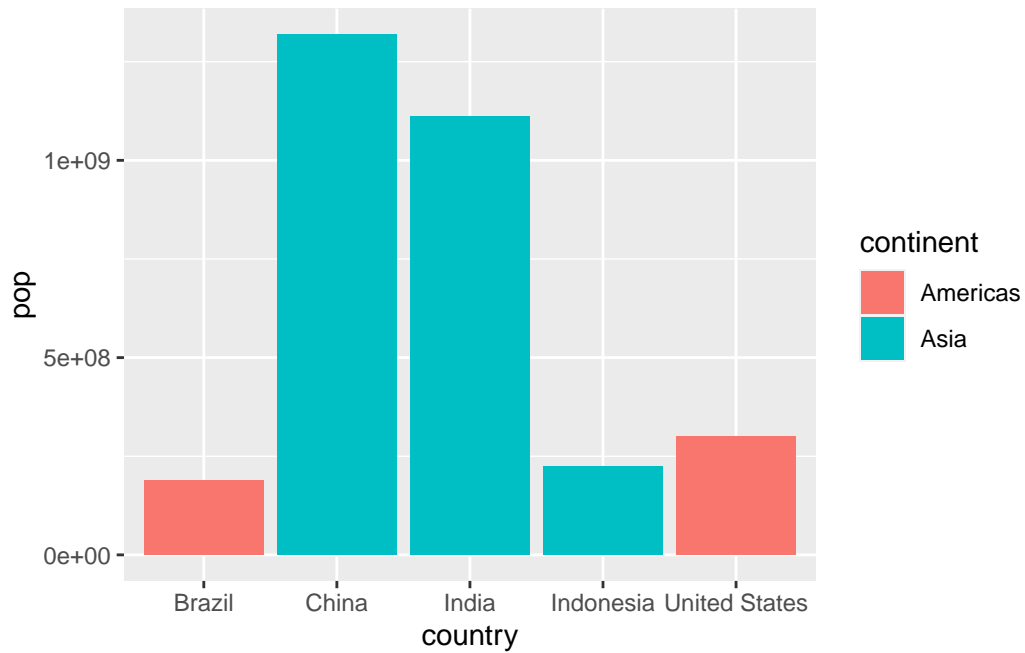
```
gapminder_1957 <- gapminder %>% filter(year==1957 | year==2007)
ggplot(gapminder_1957) +
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point(alpha=0.5) +
  scale_size_area(max_size = 10) +
  facet_wrap(~year)
```



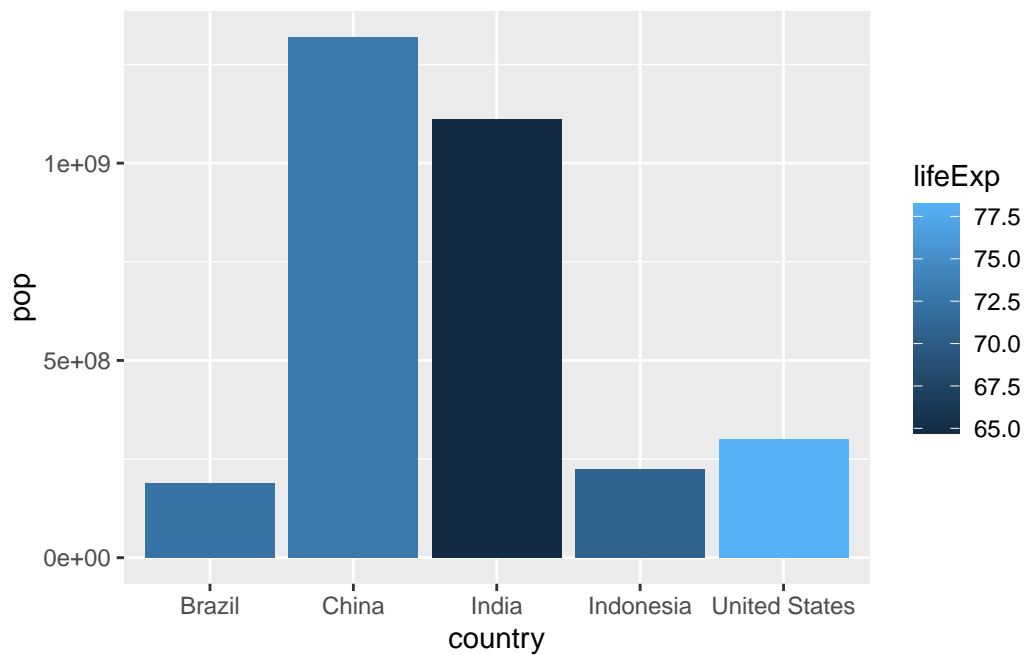
## Bar Charts

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

ggplot(gapminder_top5) +
  geom_col(aes(x = country, y = pop, fill = continent))
```

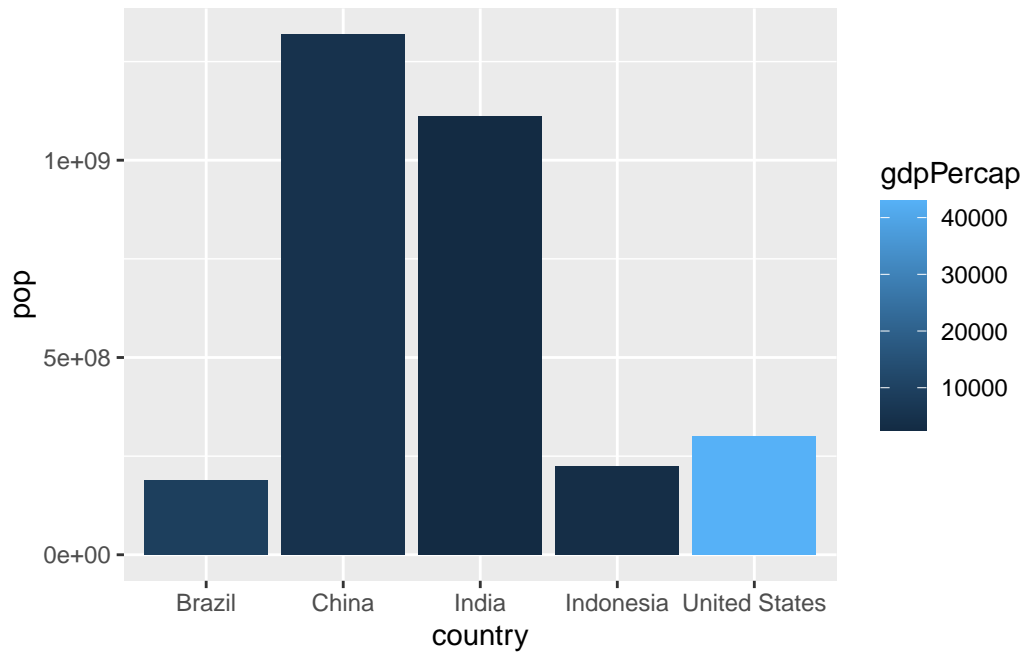


```
ggplot(gapminder_top5) +  
  geom_col(aes(x = country, y = pop, fill = lifeExp))
```

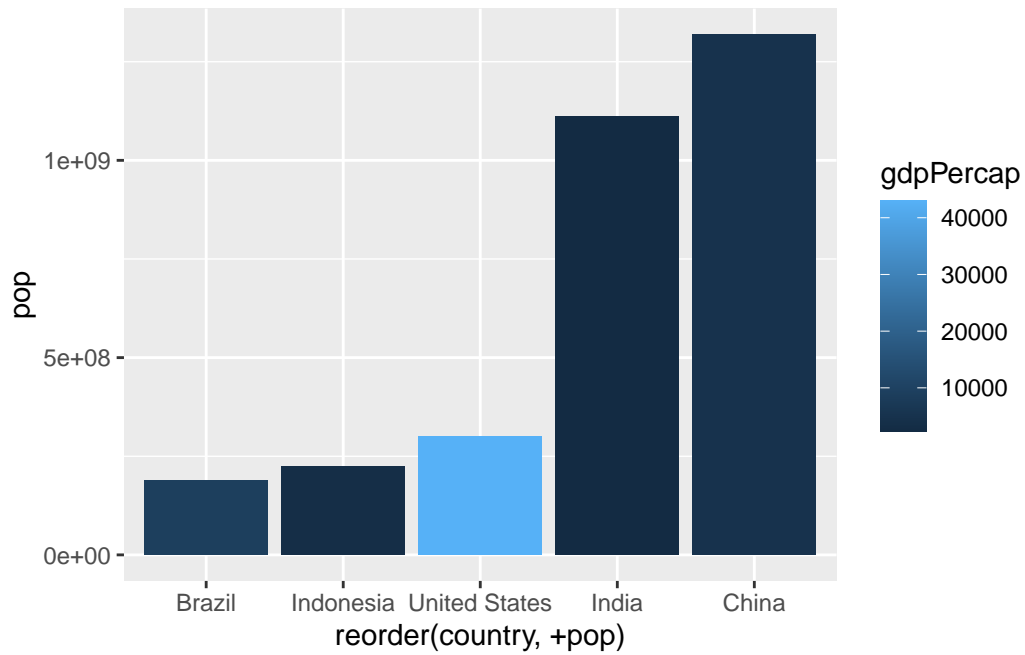




```
ggplot(gapminder_top5) +  
  geom_col(aes(x = country, y = pop, fill = gdpPercap))
```



```
#increasing Pop  
ggplot(gapminder_top5) +  
  aes(x=reorder(country, +pop), y=pop, fill=gdpPercap) +  
  geom_col()
```



```
#decreasing Pop
ggplot(gapminder_top5) +
  aes(x=reorder(country, -pop), y=pop, fill=gdpPercap) +
  geom_col()
```

