Data Visualization ggplot2 vs Seaborn

Yuqian Wang
Jackie Feng

BASICS

ggplot2: is a R package based on the grammar of graphics, build every graph from the same components: a data set, a coordinate system, and geoms

```
install.packages("ggplot2")
library("ggplot2")
```

```
ggplot (data = <DATA> ) +
<GEOM_FUNCTION> (mapping = aes( <MAPPINGS> ), stat = <STAT> ,
                position = <POSITION> ) +
<COORDINATE_FUNCTION> +
<FACET_FUNCTION> +
<SCALE_FUNCTION> +
<THEME_FUNCTION>
```
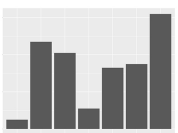
Seaborn: is a library for making statistical graphics in Python. It builds on top of matplotlib and integrates closely with pandas data structures.

```
import seaborn as sns
import matplotlib.pyplot as plt
```

```
sns. <FUNCTION>
```
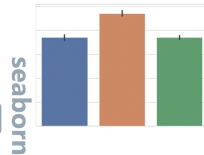
ONE CATEGORICAL VARIABLE

Bar Chart

gglot
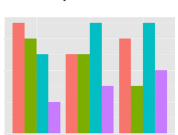```
d <- ggplot(df, aes(x))
d + geom_bar()
```
seaborn
```
sns.barplot()
```

## TWO CATEGORICAL VARIABLES

### Grouped Bar Chart

gglot
```
d <- ggplot(df, aes(x))
d + geom_bar(fill=z)
```
seaborn
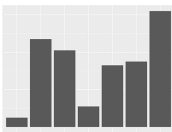```
sns.barplot(hue=z)
```
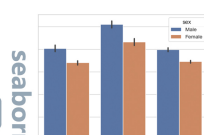
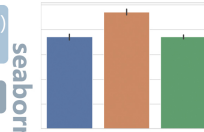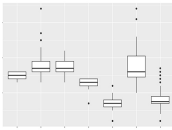## ONE CATEGORICAL & ONE CONTINUOUS VARIABLE

Bar Chart

gglot
```
d <- ggplot(df, aes(x,y))
d + geom_col()
```
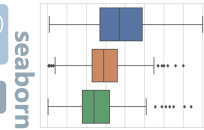seaborn
```
sns.barplot()
```

Box Plot

gglot
```
d <- ggplot(df, aes(x,y))
d + geom_boxplot()
```
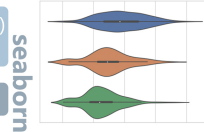seaborn
```
sns.boxplot()
```
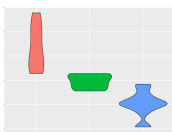
Violin Plot

gglot
```
d <- ggplot(df, aes(x,y))
d + geom_violin()
```
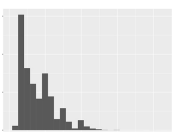seaborn
```
sns.violinplot()
```
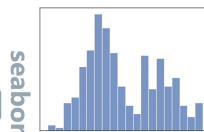
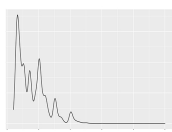## ONE CONTINUOUS VARIABLE

Histogram

gglot
```
d <- ggplot(df, aes(x))
d + geom_histogram()
```
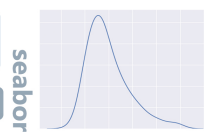seaborn
```
sns.histplot()
```

Density Plot

gglot
```
d <- ggplot(df, aes(x))
d + geom_density()
```
seaborn
```
sns.kdeplot()
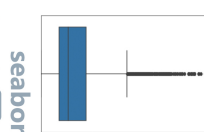```
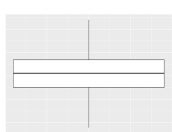
### Box Plot
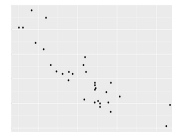
gglot
```
d <- ggplot(df, aes(x))
d + geom_boxplot()
```
seaborn
```
sns.boxplot()
```
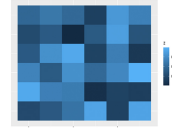
## TWO CONTINUOUS VARIABLES

### Scatter Plot

gglot
```
d <- ggplot(df, aes(x,y))
d + geom_point()
```
seaborn
```
sns.scatterplot()
```

### Heatmap

gglot
```
d <- ggplot(df, aes(x))
d + geom_tile()
```
seaborn
```
sns.heatmap()
```

## AESTHETICS

Colour related aesthetics: colour, fill, and alpha
```
aes_colour_fill_alpha colour color fill
```

Aesthetics: grouping
```
aes_group_order group
```

Differentiation related aesthetics: linetype, size, shape
```
aes_linetype_size_shape linetype size shape
```

Position related aesthetics: x, y, xmin, xmax, ymin, ymax, xend, yend
```
aes_position x y xmin xmax ymin ymax xend yend
```

Setting themes
```
sns.set_style("whitegrid") (darkgrid, whitegrid, dark, white, ticks)
```

Choosing color palettes
```
color_palette()
```

## FACET

Facets divide a plot into subplots based on the values of one or more discrete variables.
```
facet_grid (rows=vars(x)/cols=vars(x))
```
seaborn
```
sns.FacetGrid(col = x, row = y)
```