Data Visualization with R

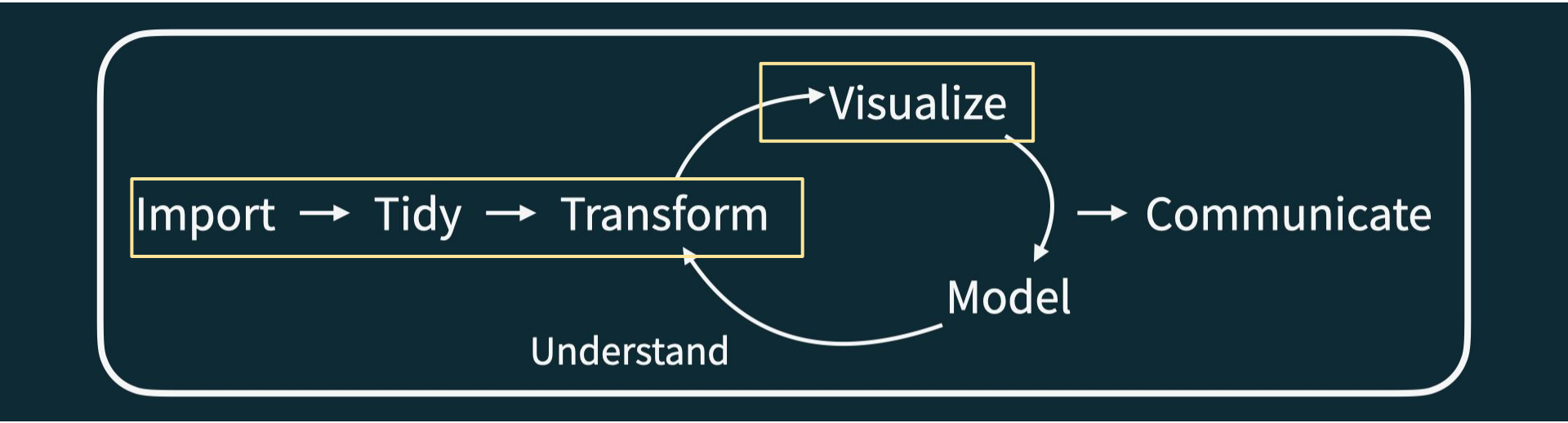
Instructor: Arvind Iyer

www.arvindkiyer.com @Arvind_K_Iyer ayalurarvind@gmail.com

Learning Objectives

- Able to do data formatting and cleaning.
- Learn basics of data visualization with ggplot2.
  - Introduction to grammar of graphics
  - Univariate Graphs (bar plot, pie chart etc)
  - Bivariate Graphs (box plot, violin plot etc)
  - Multivariate Graphs (scatter plot etc)
  - Other Graphs (heatmap etc)
  - Using of Themes
- A hands on workshop with a project at the end. (~70-80% doing)

Data science life cycle

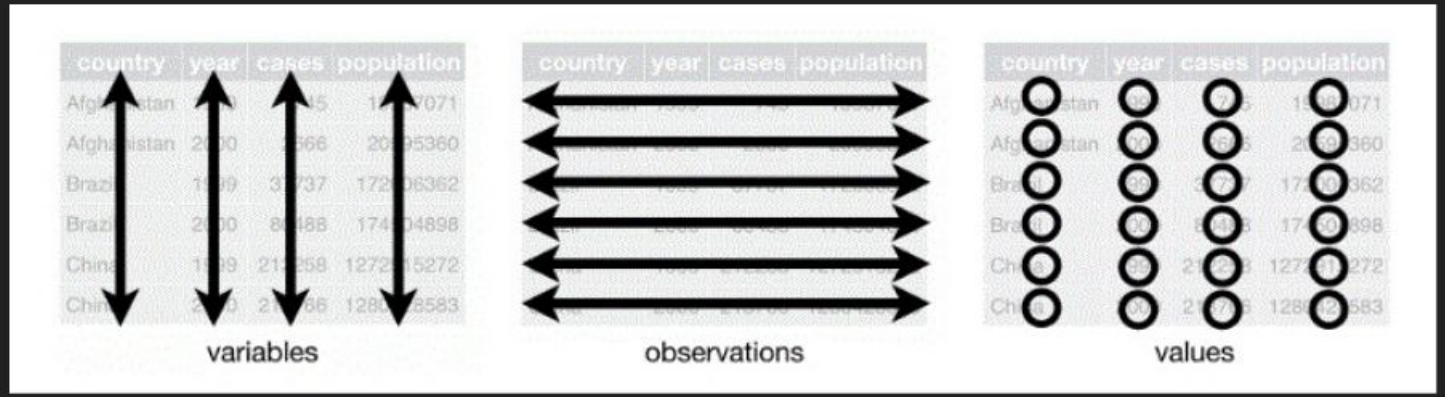



DATA, DATA EVERYWHERE

Data Preparation

Let's read the data in R !!

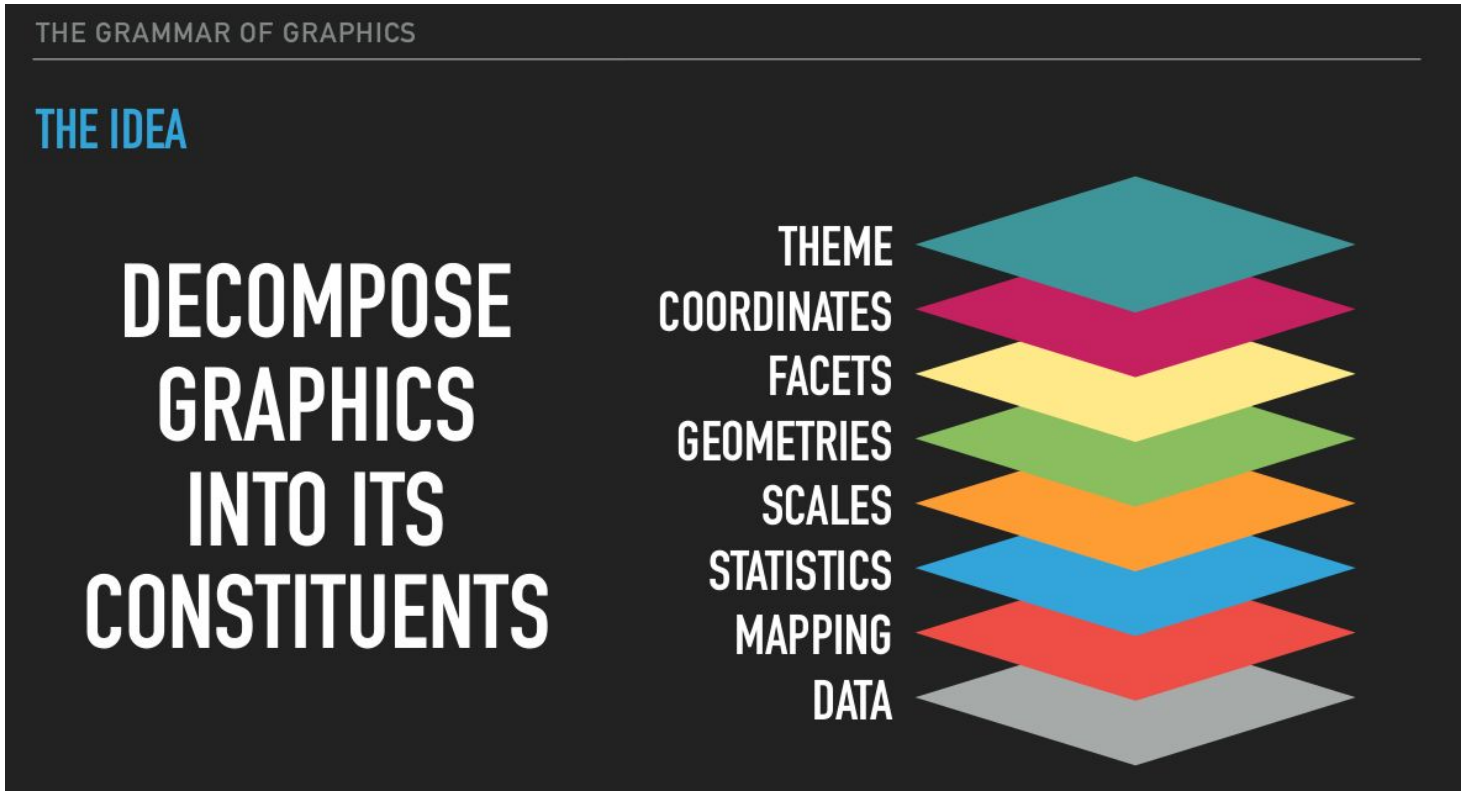
Data Exploration with tidyverse



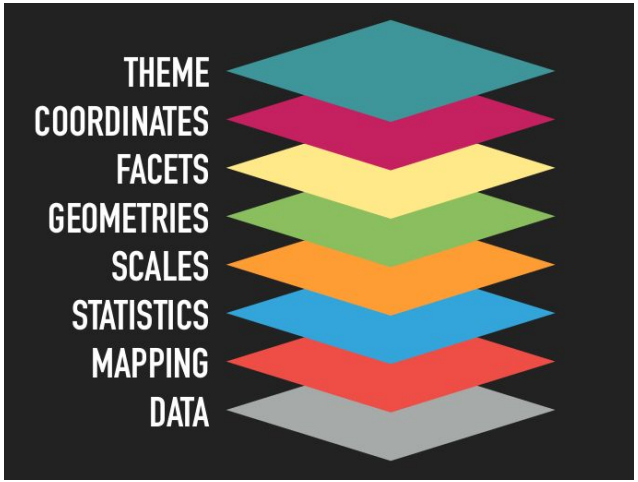
- Each variable in the data set is placed in its own column.
- Each observation is placed in its own row.
- Each value is placed in its own cell.

Data Exploration with tidyverse

Data Visualization



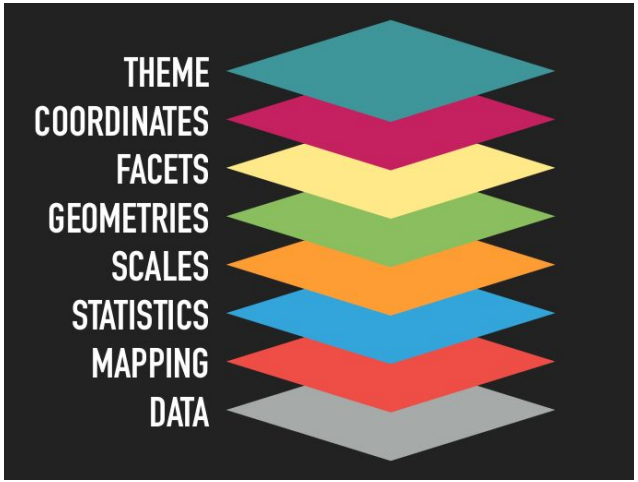
Data Visualization using ggplot2

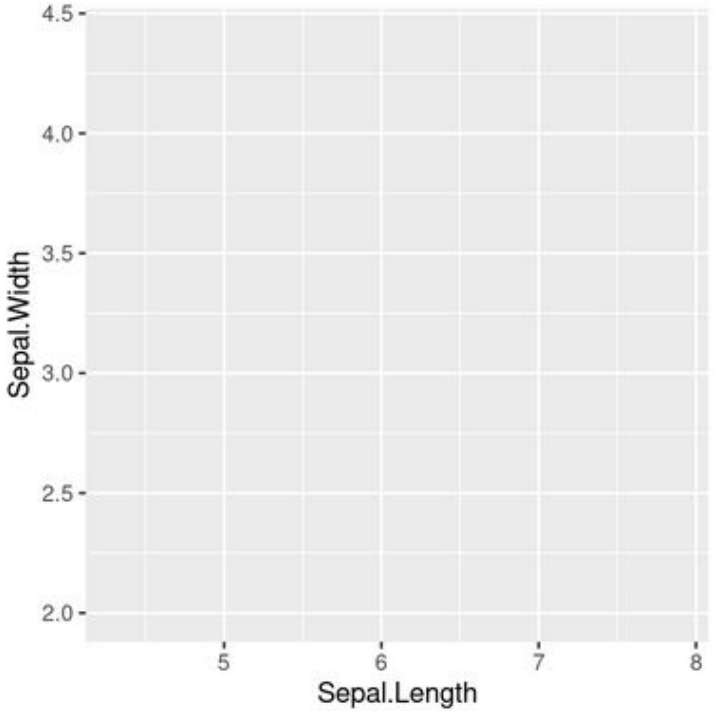



```
ggplot(iris)
```

Figure 2 adds Aesthetics to the plot.

```
ggplot(iris,
      aes(x=Sepal.Length, y=Sepal.Width,
          color=Species))
```





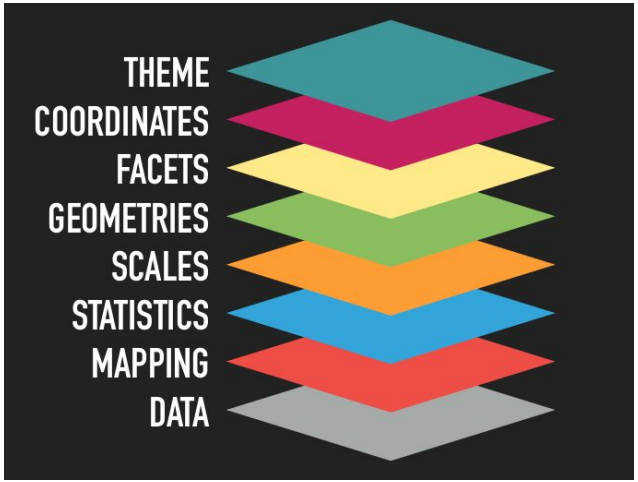
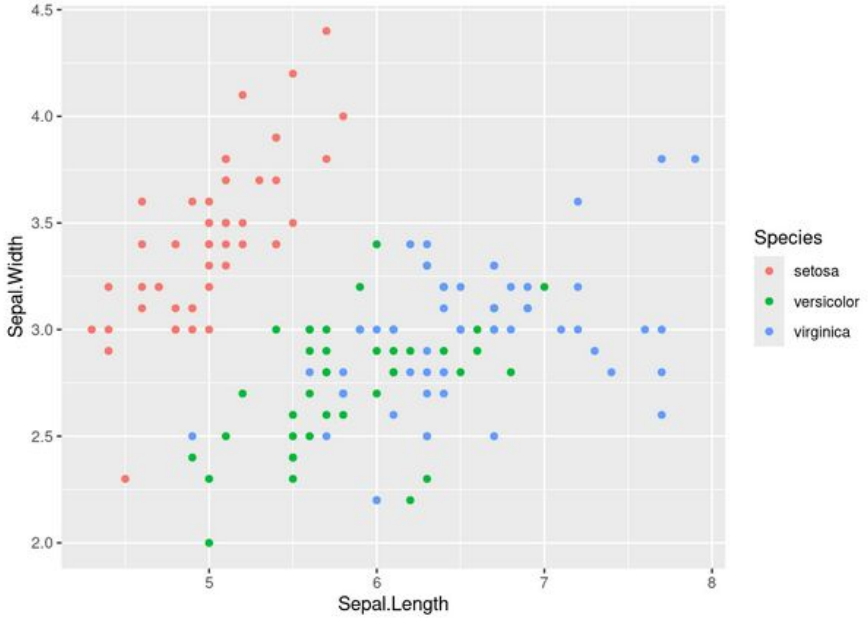


Figure 3 adds Geometries to the plot.

```
ggplot(iris,
       aes(x=Sepal.Length, y=Sepal.Width,
           color=Species))+
  geom_point()
```



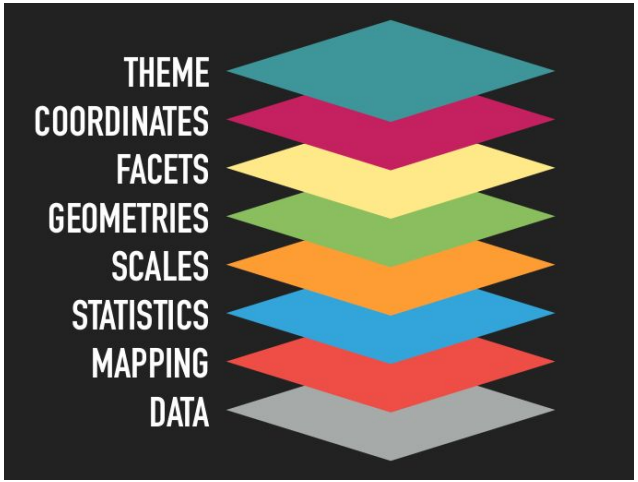
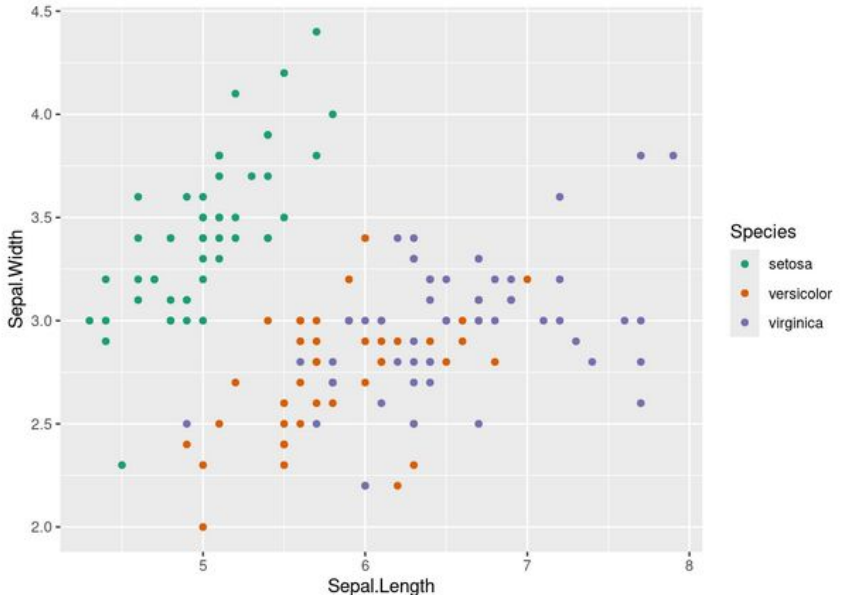


Figure 4 adds Scale to the plot.

```
ggplot(iris,
       aes(x=Sepal.Length, y=Sepal.Width,
           color=Species))+
  geom_point()+
  scale_color_brewer(palette = 'Dark2')
```



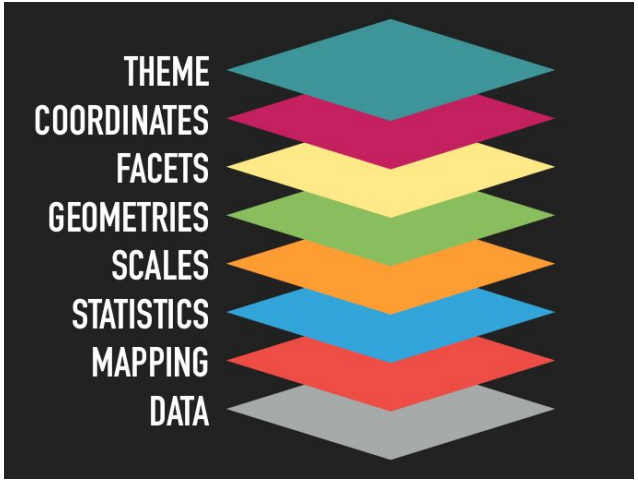
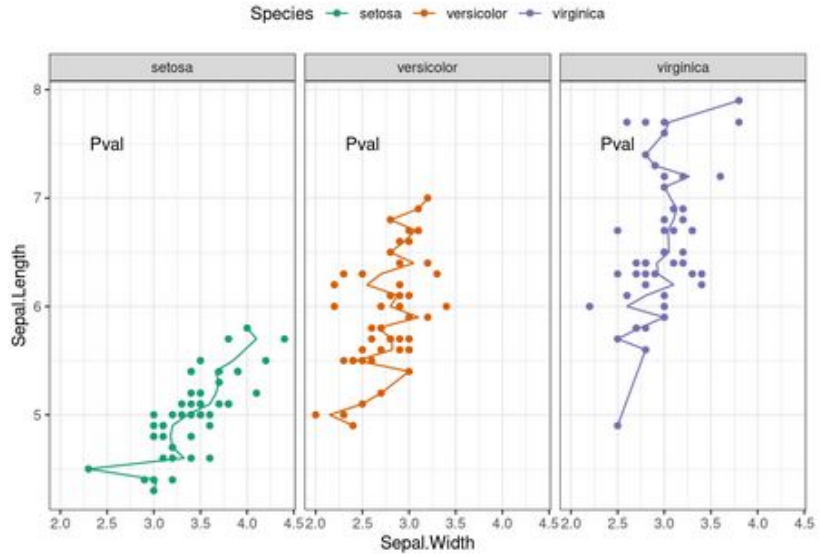


Figure 4: Adding Scales

Figure 4 adding Stats,theme,facets to the plot.

```
ggplot(iris,
       aes(x=Sepal.Length, y=Sepal.Width,
           color=Species)) +
    geom_point() +
    scale_color_brewer(palette="Dark2") +
    stat_summary(fun.y="mean", geom= "line") +
    coord_flip() +
    facet_wrap(~Species) +
    theme_bw() + theme(legend.position="top") +
    annotate("text", x=7.5, y=2.5, label="Pval")
```



References

- Material created have been referred from various sources and book.
  - https://r4ds.hadley.nz/
  - https://rkabacoff.github.io/datavis/
  - https://clauswilke.com/dataviz/
  - https://ggplot2-book.org/

The End.

Things to Do

- Read and try
  - https://clauswilke.com/dataviz/aesthetic-mapping.html
  - https://socviz.co/gettingstarted.html
  - https://r4ds.hadley.nz/data-transform (Chapter-3)
  - https://r4ds.hadley.nz/data-tidy (Chapter-5 section 5.1 and 5.2)
- R reference book
  - https://intro2r.com/

- Complete the homework exercise.