**ggplot2**

**Grammar of Graphics**

2 principles

1. Graphics = distinct layers of grammatical elements
2. Meaningful plots through aesthetic mapping

|  |  |
| --- | --- |
|  |  |

ggplot(mtcars, aes(x = wt, y = mpg)) + geom\_point()

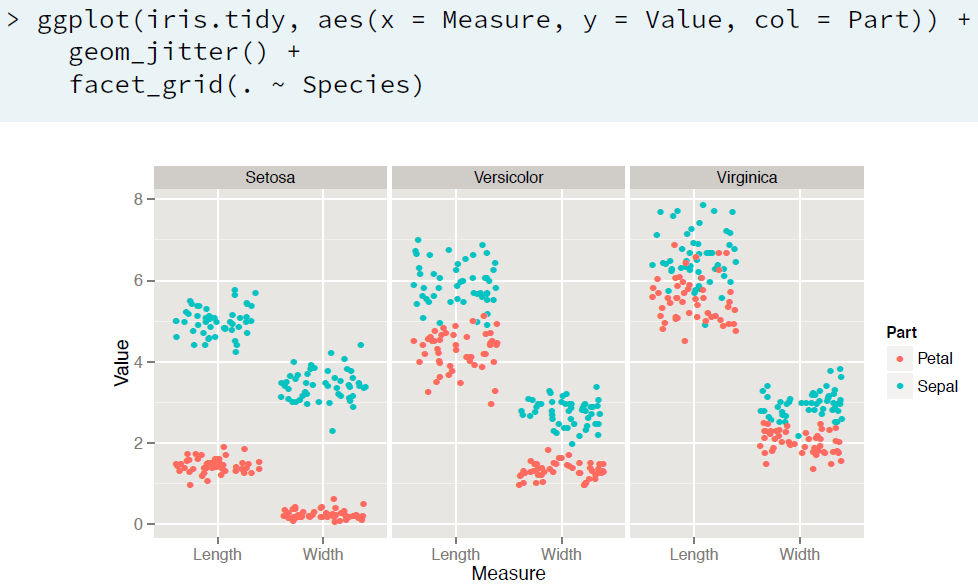
ggplot(mtcars, aes(x = wt, y = mpg, **color** = disp)) + geom\_point()

ggplot(mtcars, aes(x = wt, y = mpg, **size** = disp)) + geom\_point()

|  |  |
| --- | --- |
|  |  |

**Data**

iris.wide (pg15) & iris.tidy (pg23) & facet\_grid()



**Aesthetics Layer**

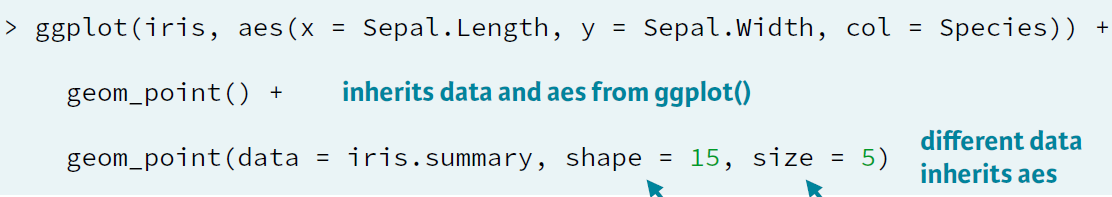
* Column can be mapped onto visible ***aesthetic***
* Aesthetics in aes(), ***attributes*** in geom\_(col=”red”)
* aes() can also be called in geom\_(), but done usually when you want to include multiple data sources
* ggplot(mtcars, aes(x=wt, y=mpg, fill=cyl, col=am)) + geom\_point(**shape**=21, **size**=4, **alpha**=0.6)
  + aes has to be associated with columns
  + attributes are given along with geom\_\*() and don’t have columns associated with them
* ggplot(mtcars, aes(x = wt, y = mpg, fill = cyl, **label**=rownames(mtcars))) + geom\_**text**(color='red')
* **Modifying Aesthetics**
  + geom\_bar(**postion**=”< stack, fill, dodge, … >”)
  + **scale\_\*** functions
    - scale\_x/y\_continuous/discrete(“title”, limits, breaks, expand, ….)
    - ex: scale\_x\_continuous(“x-axis”, limits=c(0,10), breaks=seq(0,10,2))
  + **labs**(x,y,col, …)

|  |  |
| --- | --- |
| Aesthetics for Continuous Variables | Aesthetics for Categorical Variables |

**Geometry Layer**

**Scatter Plots**: geom\_point()

* aes() inside geom\_\*() can be used to use different aesthetics for different layer of geom\_\*() . Same goes for different data for different geom\_\*()
* aes() inside geom\_\*() is same as aes() in ggplot()



* ggplot(iris, aes(x = Sepal.Length, y = Sepal.Width, col = Species)) + geom\_**jitter**(**shape**=1, **alpha**=0.6)
  + to visualize the density, use jitter along with alpha & shape(hollow shapes preferred)

**Bar Plots:**

* **Histogram**: geom\_histogram() : x-axis : continuous variables
  + ggplot(df, aes(x=x1)) + geom\_histogram(**binwidth**=0.1)
  + ggplot(df, aes(x=x1)) + geom\_histogram(aes(y=**..density..**), binwidth=0.1)
  + ggplot(df, aes(x=x1, fill=cat\_var)) + geom\_histogram(binwidth=0.1, **position**=”stack/fill/dodge”)

|  |  |  |
| --- | --- | --- |
|  |  |  |

* **Bar Plot**: geom\_bar() : x-axis = categorical variables
  + ggplot(df, aes(x=cat\_var)) + geom\_bar(stat=”bin”)
  + Custom Color Palettes

blues <- brewer.pal(9, "Blues")

blue\_range <- colorRampPalette(blues)

ggplot(Vocab, aes(x = education, fill = vocabulary)) +

geom\_bar(position = "fill") +

scale\_**fill\_manual**(**values**=blue\_range(11))

* + Overlapping bar plots

posn\_d <- **position\_dodge**(**width**=0.2)

ggplot(mtcars, aes(x = cyl, fill = am)) + geom\_bar(position=posn\_d)

* **Line Plots:** geom\_line()
  + Plotting different categories

ggplot(df, aes(x=Year, y=Capture, **linetype**=Species)) + geom\_line()

* + Proportional Trends

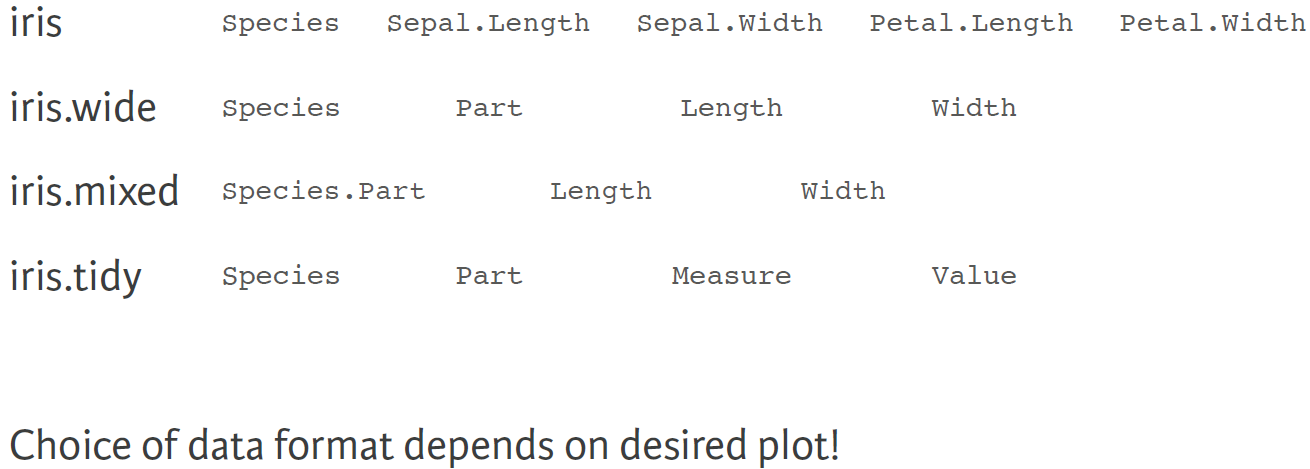
ggplot(df, aes(x=Year, y=Capture, **fill**=Species)) + geom\_**area**(position=”fill”)

* + ggplot(economics, aes(x=date, y=unemploy/pop)) + **geom\_rect**(data=recess, aes(xmin=begin, xmax=end, ymin=-Inf, ymax=+Inf), **inherit.aes=FALSE**, fill="red", alpha=0.2) + geom\_line()

**qplot**

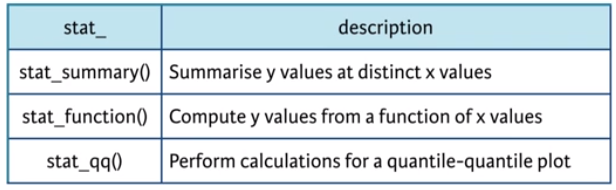
* Quick and dirty way for plotting, not very flexible, doesn’t follow grammar of graphics
* qqplot( x, y, data, shape/size/col, postion, jitter, alpha=I(value) )

**Wrap-Up**



**Statistics Layer**

* Two types: called within a geom and called independtly



* **stat\_smooth**(method=”loess/lm/…”, se, aes(group=1, col=”text”), span)
  + ex: ggplot(Vocab, aes(x = education, y = vocabulary, col = year, group = factor(year))) +

stat\_smooth(method = "lm", se = FALSE, alpha = 0.6, size = 2) +

scale\_color\_gradientn(colors = brewer.pal(9, "YlOrRd"))

* **stat\_summary**(geom=”***errorbar***”, fun.data=mean\_sdl, fun.args=list(…), witdh) # adds error bars
* **stat\_summary**(geom=”***point***”, fun.y=mean) # adds point for the mean values
* **stat\_summary**(geom="***linerange***", fun.data =<custom\_function>, position=posn.d, size=3)
* **stat\_function**(fun=dnorm, colour, arg=list(…))

­

**Coordinate Layer**

* zooming-in : + coord\_cartesian(xlim=c(4.5, 5.5))
* aspect ratio : + coord\_fixed(0.055)

**Facets Layer**

* … + facet\_grid(row ~ column)

**Themes Layer**



* … + theme(plot.background=element\_rect(color=”black”, size=3))
* … + theme(panel.grid=element\_line(color=”red”))
* … + theme(axis.text=element\_text(color=”red”))
* my.theme <- theme( … )
* theme\_update( … )
* theme\_set( … )