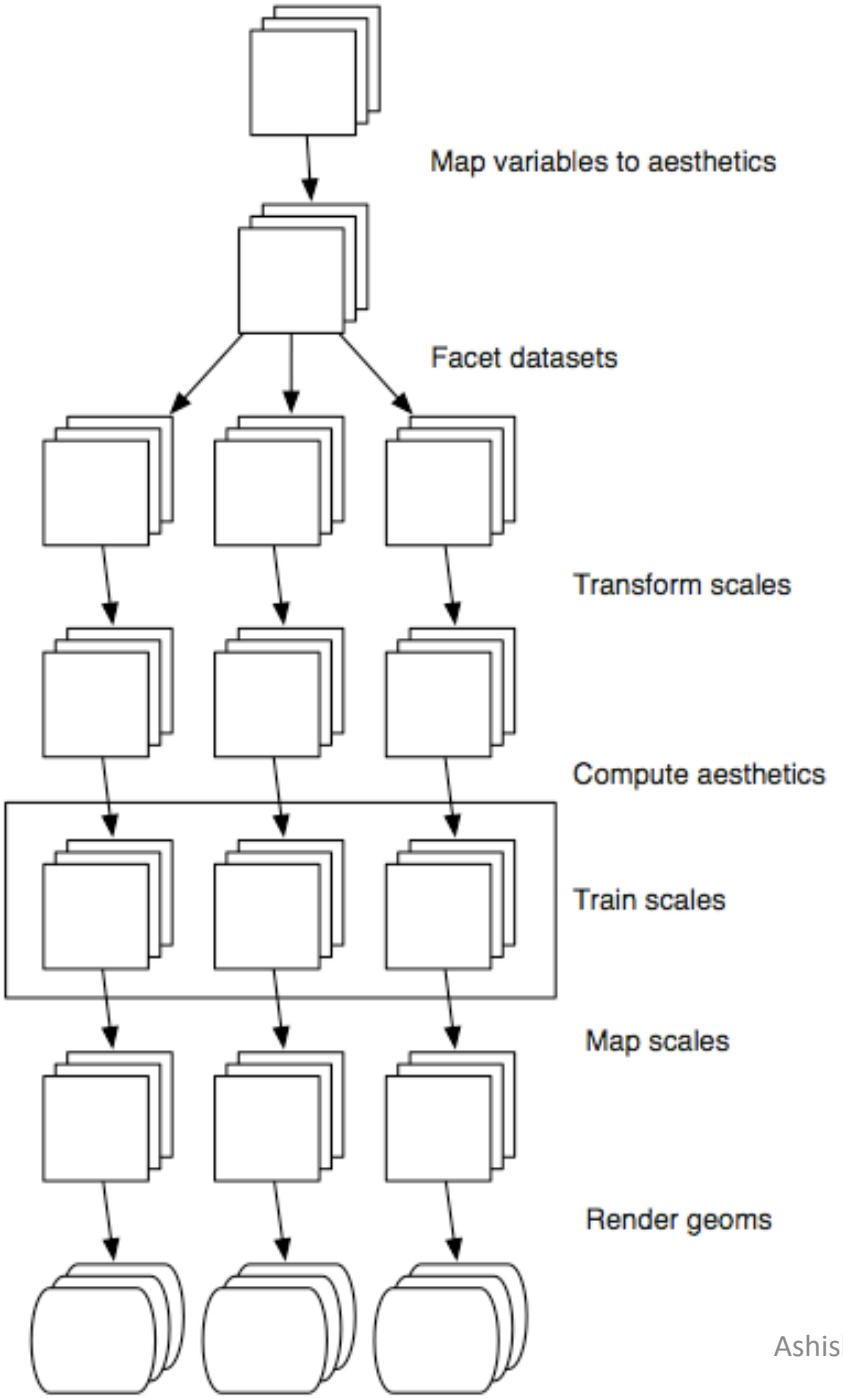
Ashish Mahabal
California Institute of Technology

R - VII

Sponsored by the Jet Propulsion Laboratory and by the Keck Institute for Space Studies, California Institute of Technology



Anatomy of a plot

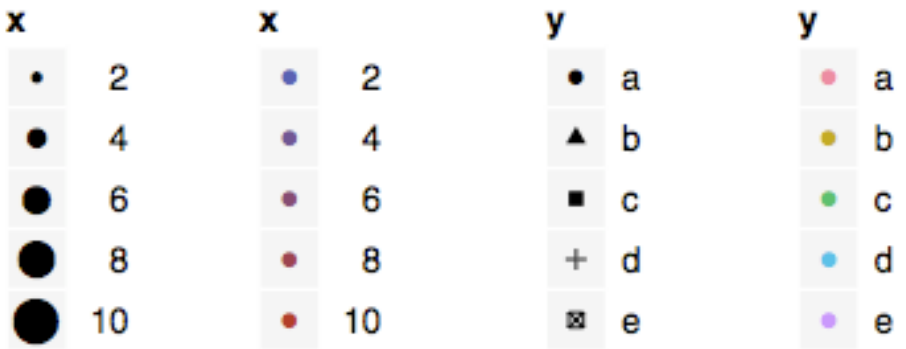
- Overall:
  - dataset + mappings
  - layers
  - scale per mapping
  - coordinate system
  - faceting specification
- Per layer:
  - data+mapping; geom; stat; pos

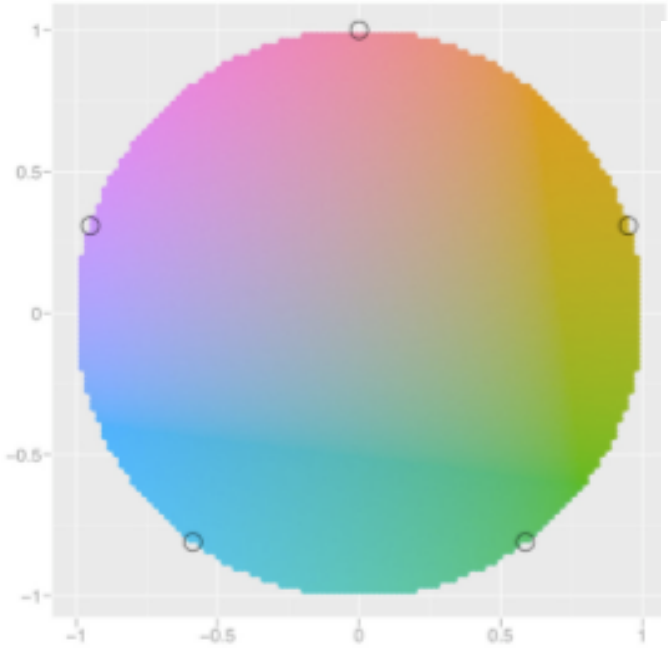
Fig 3.7 ggplot2, Hadley

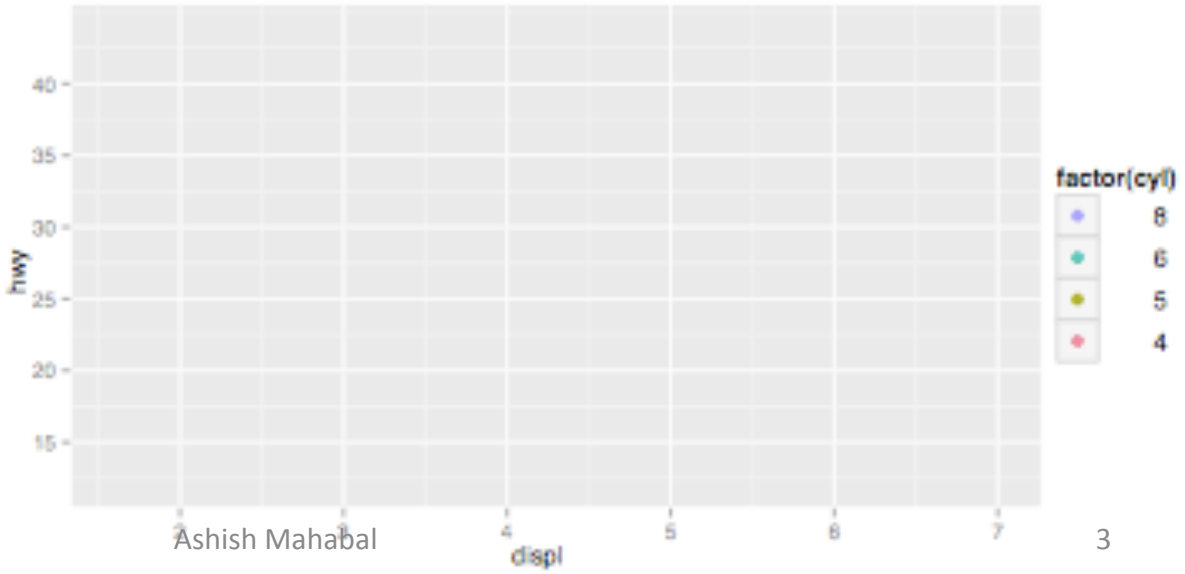
What is a plot (made of)?

- A ***data***set
- Set of ***mapping***s from variables to aesthetics
- At least one layer, each made of
  - A geometric object (***geom***)
  - A statistical transformation (***stat***)
  - A position adjustment
  - Perhaps a dataset and aesthetic mappings
- One ***scale*** per mapping
- A ***coord***inate system
- ***Faceting*** specification

Example of scale
continuous and discrete variables







```
p <- ggplot(data = crts_6obj,
aes(amplitude,std), colour = factor(object))
```

Things to do with an object

- print() – automatically unless in a loop
- ggsave() – to write out an image
- summary() – summary about the makeup
- save() – to save the object

summary(p)

- data: name, ra, dec, amplitude, beyond1std, fpr_mid20, fpr_mid35, fpr_mid50, fpr_mid65, fpr_mid80, linear_trend, max_slope, med_abs_dev, med_buf_range_per, pair_slope_trend, percent_amplitude, pdfp, skew, kurtosis, std, ls, rcorbor, magratio, data_num, object, dmjd [1619x26]
- mapping: x = amplitude, y = std
- faceting: facet_null()

summary(p + geom_point()) additionally gives:

- geom_point: na.rm = FALSE
- stat_identity:
- position_identity: (width = NULL, height = NULL)

saving/loading plot/object

- > # Save plot object to disk
- > save(p, file = "plot.rdata")
- > # Load from disk
- > load("plot.rdata")
- > # Save png to disk
- > ggsave("plot.png", width = 5, height = 5)

Adding layers

- p <- ggplot(diamonds, aes(carat, price, colour = cut))
- p <- p + layer(geom = "point") #stat/pos defaults
- layer(geom, geom_params, stat, stat_params, data, mapping, position)
- p <- ggplot(diamonds, aes(x = carat))
- p <- p + layer(
  - geom = "bar",
  - geom_params = list(fill = "steelblue"),
  - stat = "bin",
  - stat_params = list(binwidth = 2))
- p

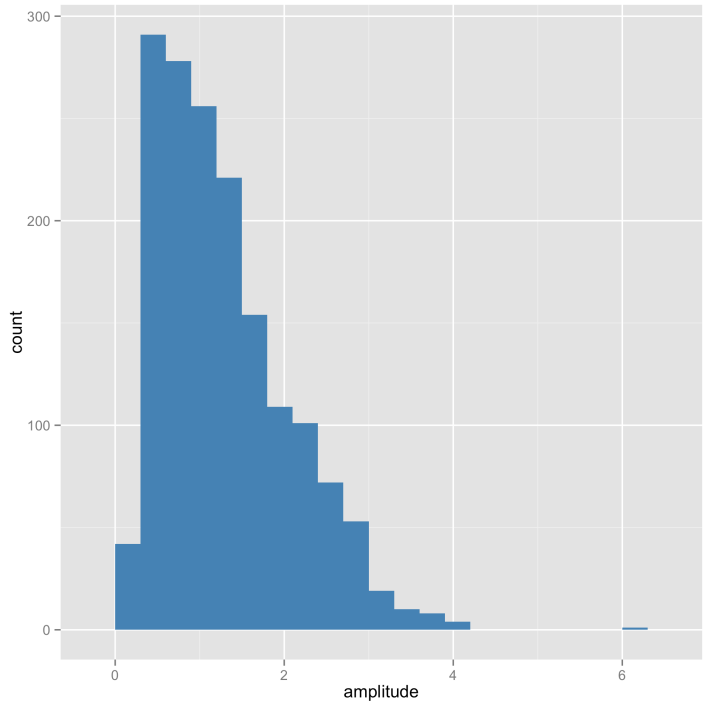
Every geom associated with a (default) stat

- geom_histogram(binwidth = 0.3, fill = "steelblue")

And every stat with a (default) geom

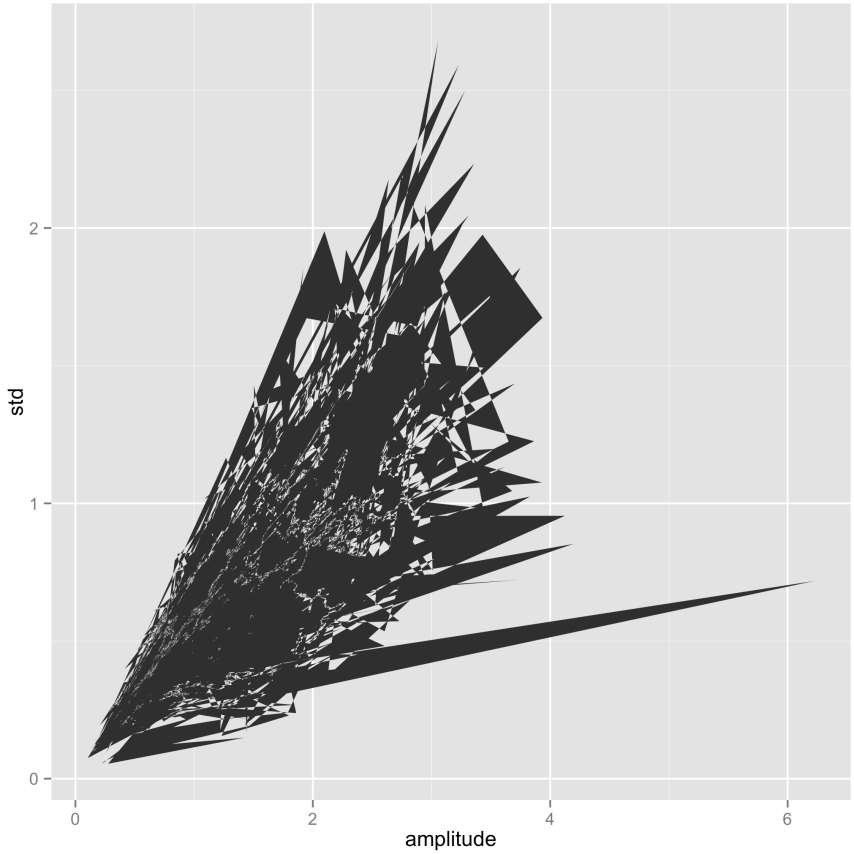
stat: bin

geom: bar



Some geoms (29)

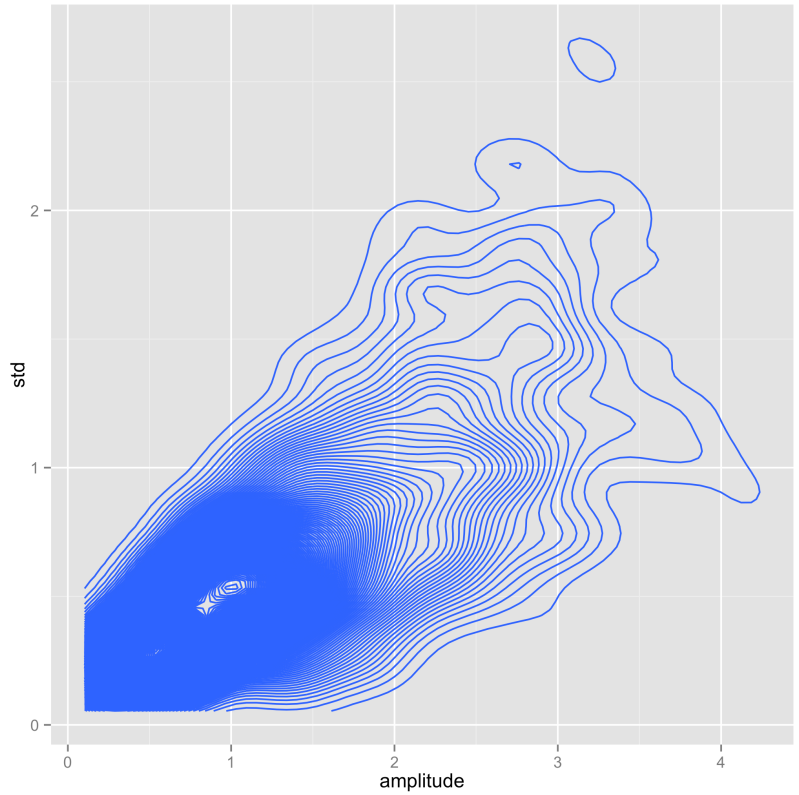
- bar
- boxplot
- contour
- line
- point
- step
- text



Some stats (15)

- bin
- boxplot
- contour
- density2d
- identity
- quantile
- unique

```
p2 <- ggplot(data = crts_6class, aes(amplitude,std))
p + layer(geom = "density2d",geom_params = list(fill =
"steelblue"),stat = "density2d",stat_params =
list(binwidth = 0.01))
```



Sundry points

```
R> ggplot(data, aes(x, y)) + geom_point()

# Aesthetics
R> ggplot(data, aes(x, y)) + geom_point(colour = "red")

# Layers
R> ggplot(data, aes(x, y)) + geom_point(method = "lm") +
    geom_smooth(method = "lm")

# Scales and axes
ggplot(data, aes(x, y)) + geom_point() + scale_x_log10() + scale_y_log10()

# Plot options
R> ggplot(data, aes(x, y)) + geom_point() + opts(title = "title", aspect.ratio = 1)
```

Homework

- Apply to the astronomy dataset various graphing techniques learnt
- http://astro.caltech.edu/~aam/datasets/crts6_obj.txt
- Determine the most discriminating columns for the six different classes (AGN, blazars, CV, Flares, RR Lyrae, SN)

- amplitude, beyond1std, fpr_mid20, fpr_mid35, fpr_mid50, fpr_mid65, fpr_mid80, linear_trend, max_slope, med_abs_dev, med_buf_range_per, pair_slope_trend, percent_amplitude, pdfp, skew, kurtosis, std, ls, rcorbor, magratio

Use plot(), qplot() and ggplot2() methods