

## OBJECTIVE

\* Learn Exploratory Data Analysis with R

\* Create static and interactive graphs

\* Tips and tricks to do it quickly and rapidly

## INTENT

"... information display should be documentary, comparative, causal and explanatory, quantified, multivariate, exploratory." – Edward Tufte

## REQUIREMENTS

Install Software:  
R and RStudio

Install Packages:  
ggplot2, ggmap, GGally, ggvis, shiny, manipulate, dplyr, hexbin

## AUDIENCE

\* Business Analyst, Data Scientist, and anybody doing data munging, exploration and visualisation

\* Have (some) prior experience in data analysis and viz

\* Familiarity with R & RStudio

## FACILITATOR

Amit Kapoor

@amitkaps

amit@narrativeviz.com

<http://amitkaps.com>

## 0. INTRODUCTION

The Grammar of Graphics

*data*: data to visualise

*aes*: aesthetic mapping

*geom*: geometric object

*stat*: statistical transformation

*scales*: scale mapping

*coord*: coordinate system

*facet*: subset & small multiples

## 1. PORTRAIT

Distribution Representation

*About* – What is the portrait of each variable?

*Visualisations*

\* *Categorical* – Bar, Column, Stacked, Polar

\* *Continuous* – Histogram, Frequency Polygon

*Interaction* – Selection

*Tips + Tricks* – Coordinate changes, Scale transform, Binwidth

## 2. COMPARISON

Comparative Representation

*About* – How does the variable compare with others?

*Visualisations*

\* *Categorical vs. Categorical* – Stacked, Position, Mosaic

\* *Categorical vs. Continuous* – Histogram with colour, Frequency Polygon, Box Plot, Points

\* *Continuous vs. Continuous* – Scatterplot

*Interaction* – Navigation, Annotation

*Tips + Tricks* – Position, Alpha, Jitter, Binning

## STATIC & INTERACTIVE EXPLORATORY DATA ANALYSIS IN R

## 6. FLOW

Relationship, Heirarchy

*About* – How does the variable related to others?

NOT COVERED IN THIS SESSION

## 5. TIMELINE

Positions across Time

*About* – When is the variable occurring across time?

NOT COVERED IN THIS SESSION

## 4. MAP

Position across Space

*About* – Where does the variable show on a map?

*Visualisations*

\* *Scatter with lon/lat*

\* *Geocode and Mapping*

\* *Simple plotting with maps*

*Interaction* – Navigation, Pan & Zoom

*Tips + Tricks* – Scatter, Simple Maps

## 3. MULTI VARIABLE

Deduction, Prediction

*About* – Why are the two or more variables correlated?

*Visualisations*

\* *Multiple Aesthetics* – Colour, Shape, Size

\* *Facetting* – Grid, Wrap

\* *Matrix*

\* *Parallel Coordinates*

\* *Projections & Tours*

*Interaction* – Linking, Brushing, Tours

*Tips + Tricks* – Facetting, Sorting, Filtering