MODELING THE PAST - PREDICTING THE FUTURE

An interactive workshop using
Python for machine learning

notebooks you'll need:
bit.ly/TFNWI

@GrimmScientist
allen@GrimmScience.com

WHAT?

MACHINES AND SNAKES

(particularly python)

- http://core0.staticworld.net/images/article/2013/06/shutterstock_42883318_friendly_robot-100040734-medium.jpg

SET UP ENVIRONMENT

- We'll be using python's sklearn and the ipython notebook.
- An easy way to do it:
  - Continuum Analytic's Anaconda distribution
- An easier way to do it:
  - Find a friend

(you're going to want Graphviz, too)

A WALK TO THE COFFEE SHOP

- How humans do learning….

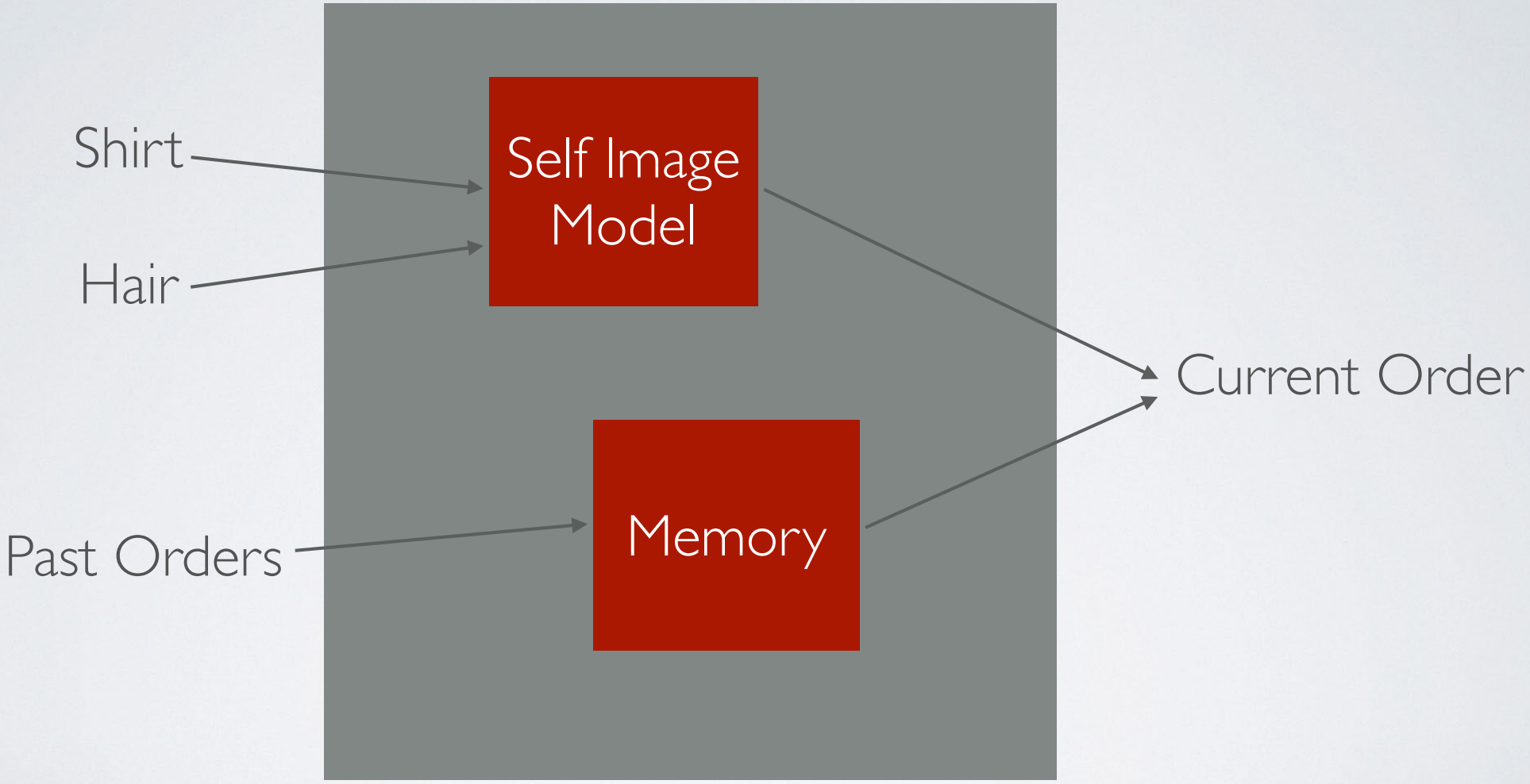
A WALK TO THE COFFEE SHOP

- Person in suit gets a mocha
- Teenager gets a latte
- Horde of college kids gets the house coffee
- Your regular gets the espresso (as always)

A WALK TO THE COFFEE SHOP

- What did your mind do in that situation?
  - variable selection
  - correlations / data modelling
  - predictions
  - action from observation (rather than reaction)

THE BLACK (COFFEE) BOX



MACHINE LEARNING

- The systematic extraction of information from data.
- A set of methods that attempt to reduce a dataset into a model that maintains true structure and ignores the noise.

LET'S PLAY

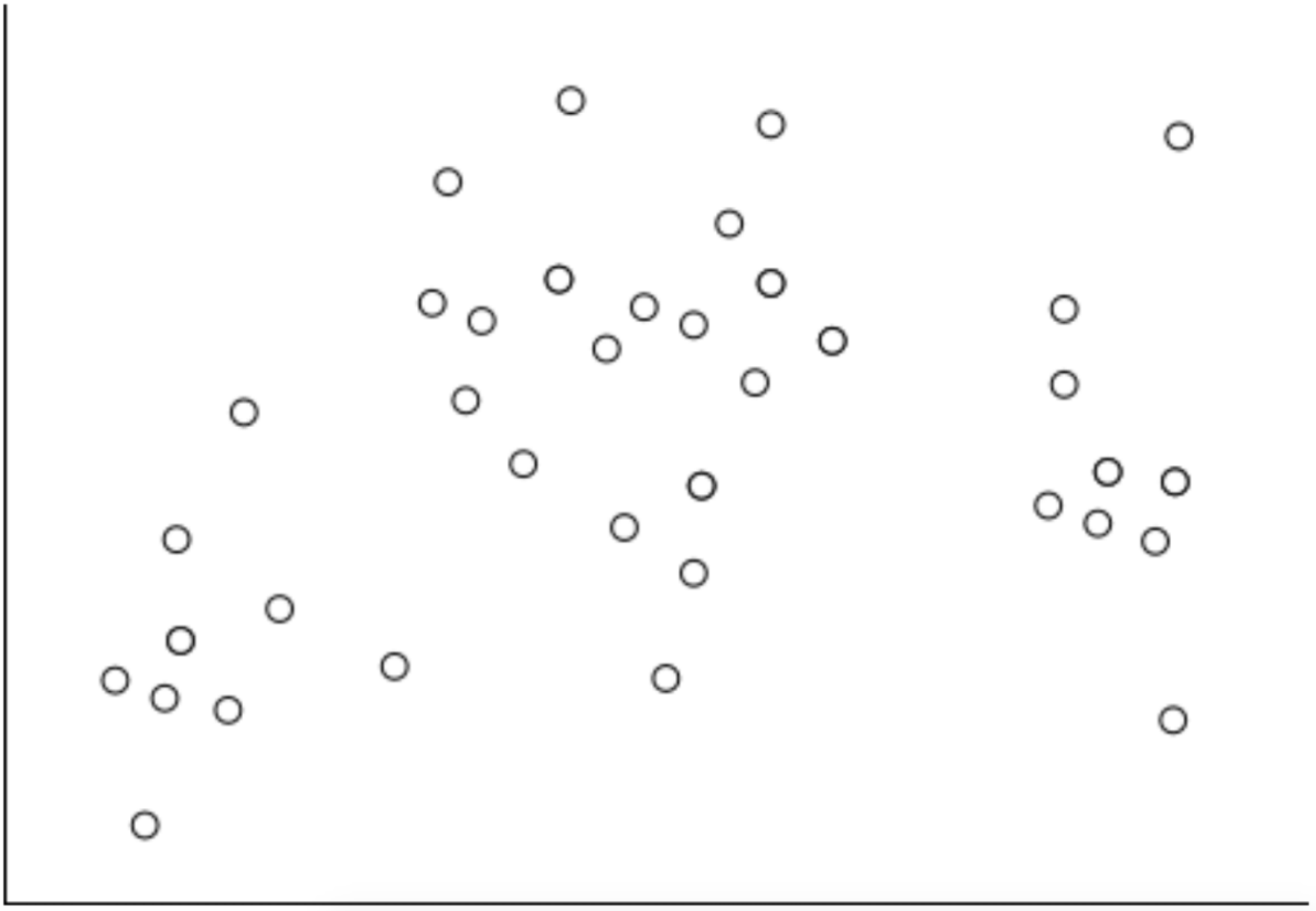
- Fisher's Irises
- Manually Made
- Digits
- Hack Oregon's Campaign Finance

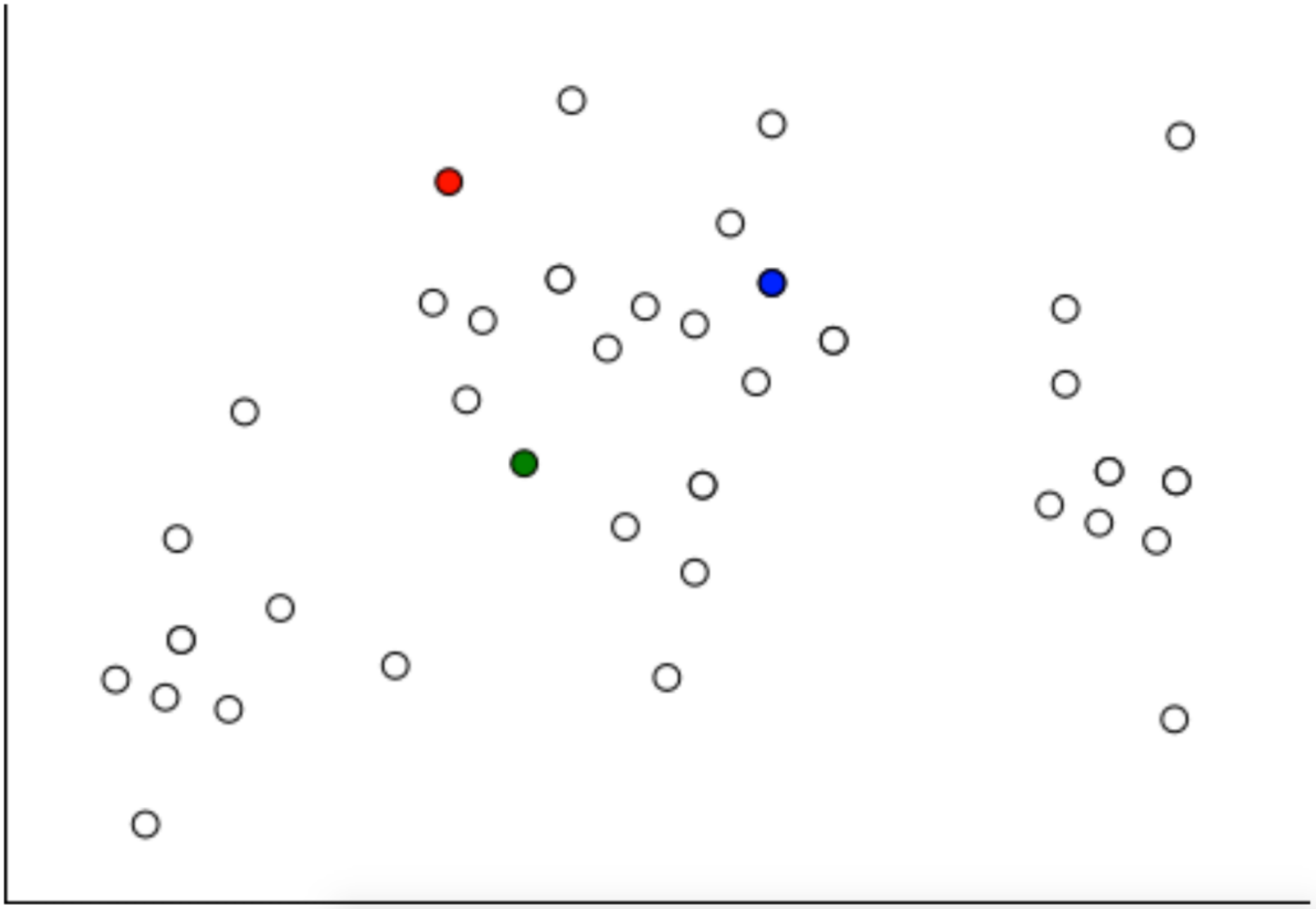
DECISION TREES

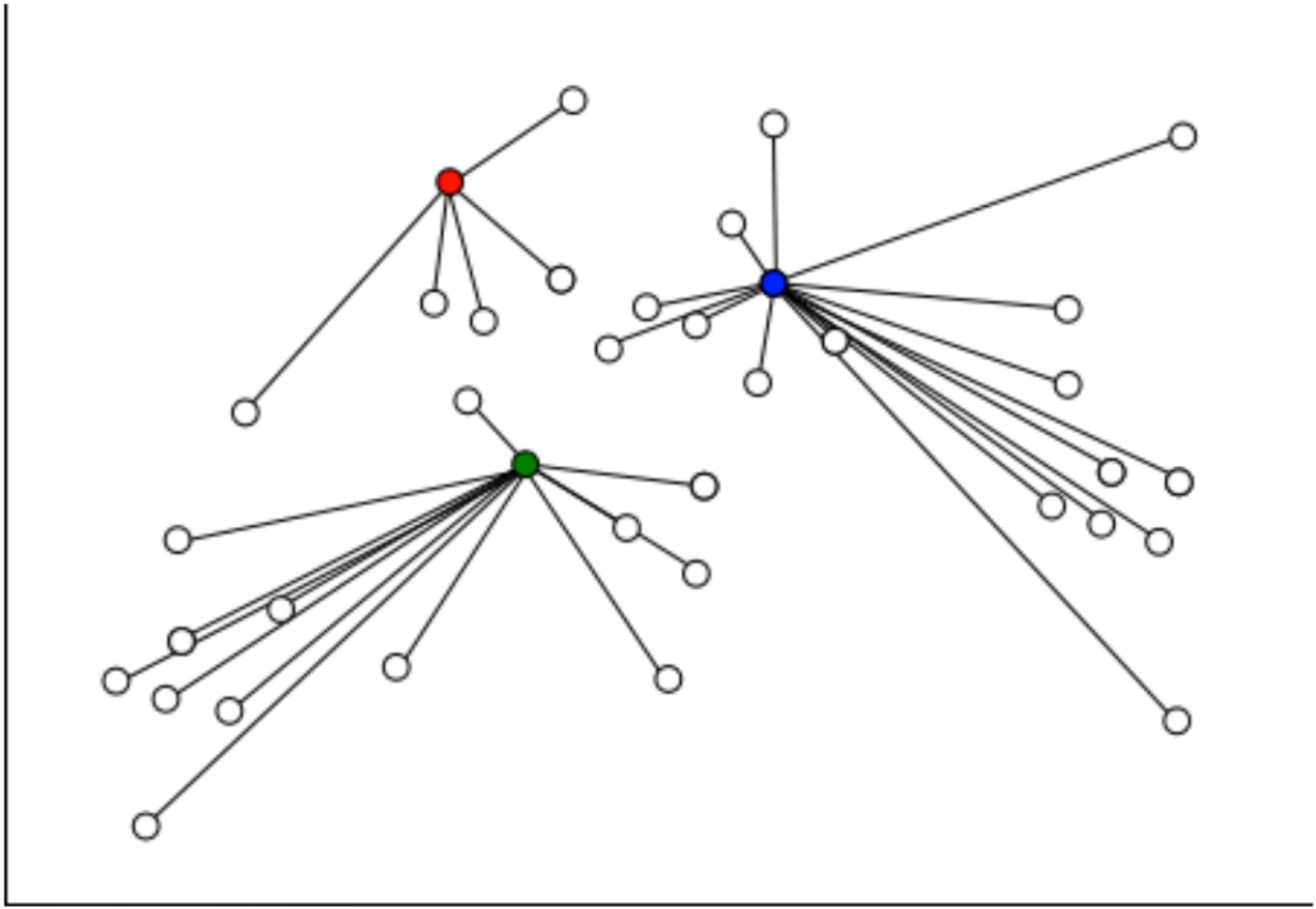
- "human" model structure
- doubles as an intro to information theory
- … to the notebook!

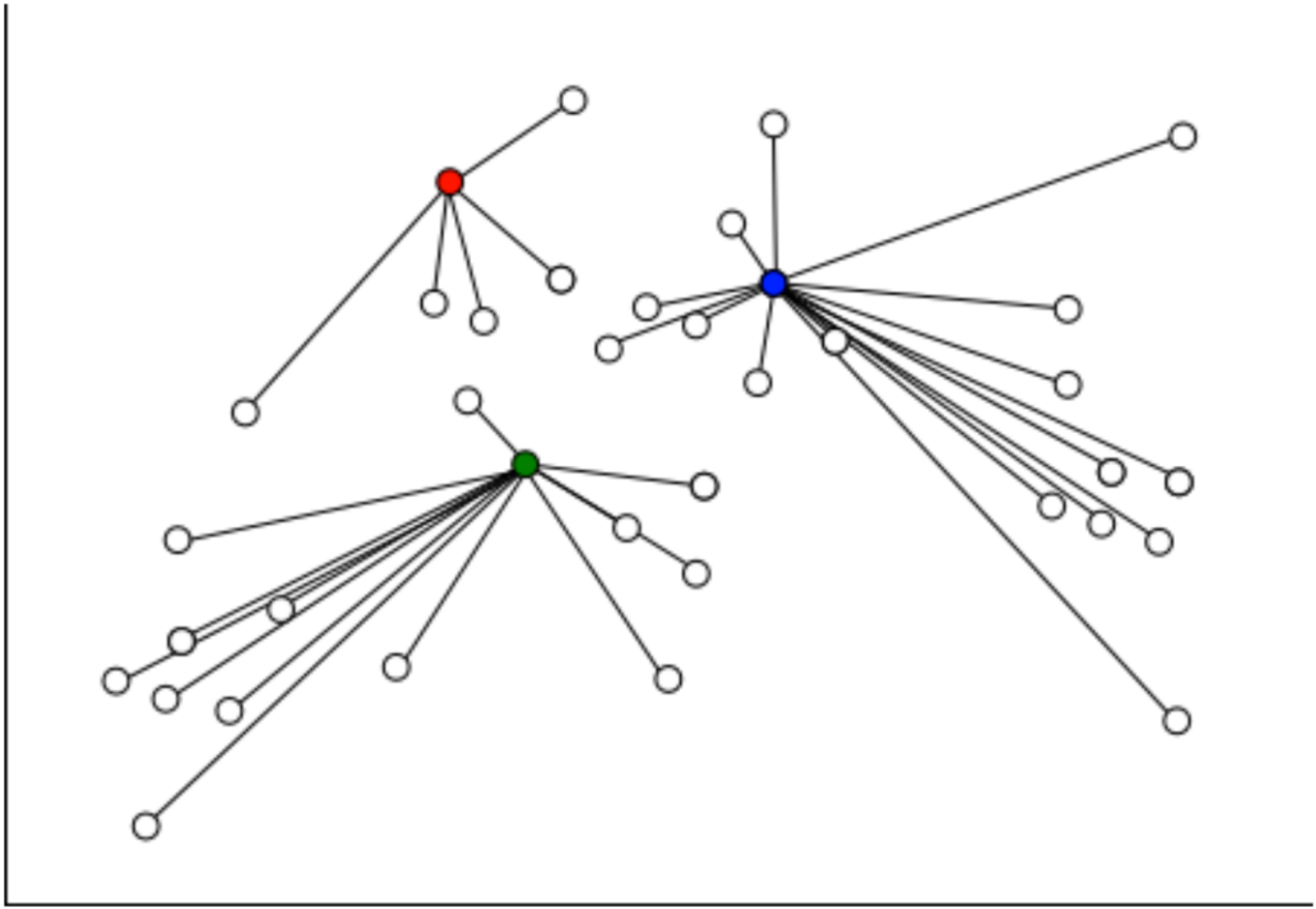
K-MEANS

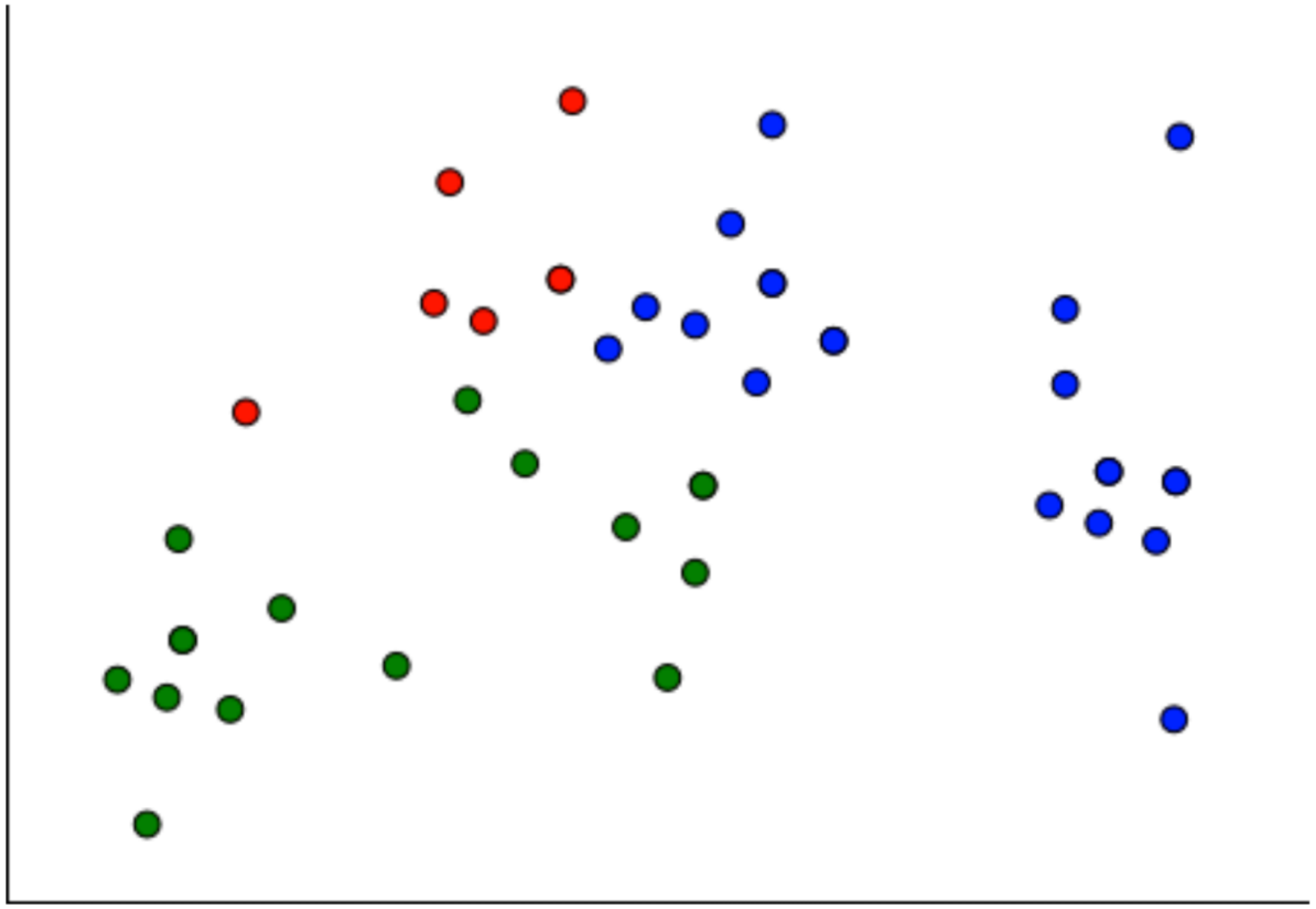
- "unsupervised learning"
- following slides taken from Ryan Smith: https://github.com/bixbyr/kmeans

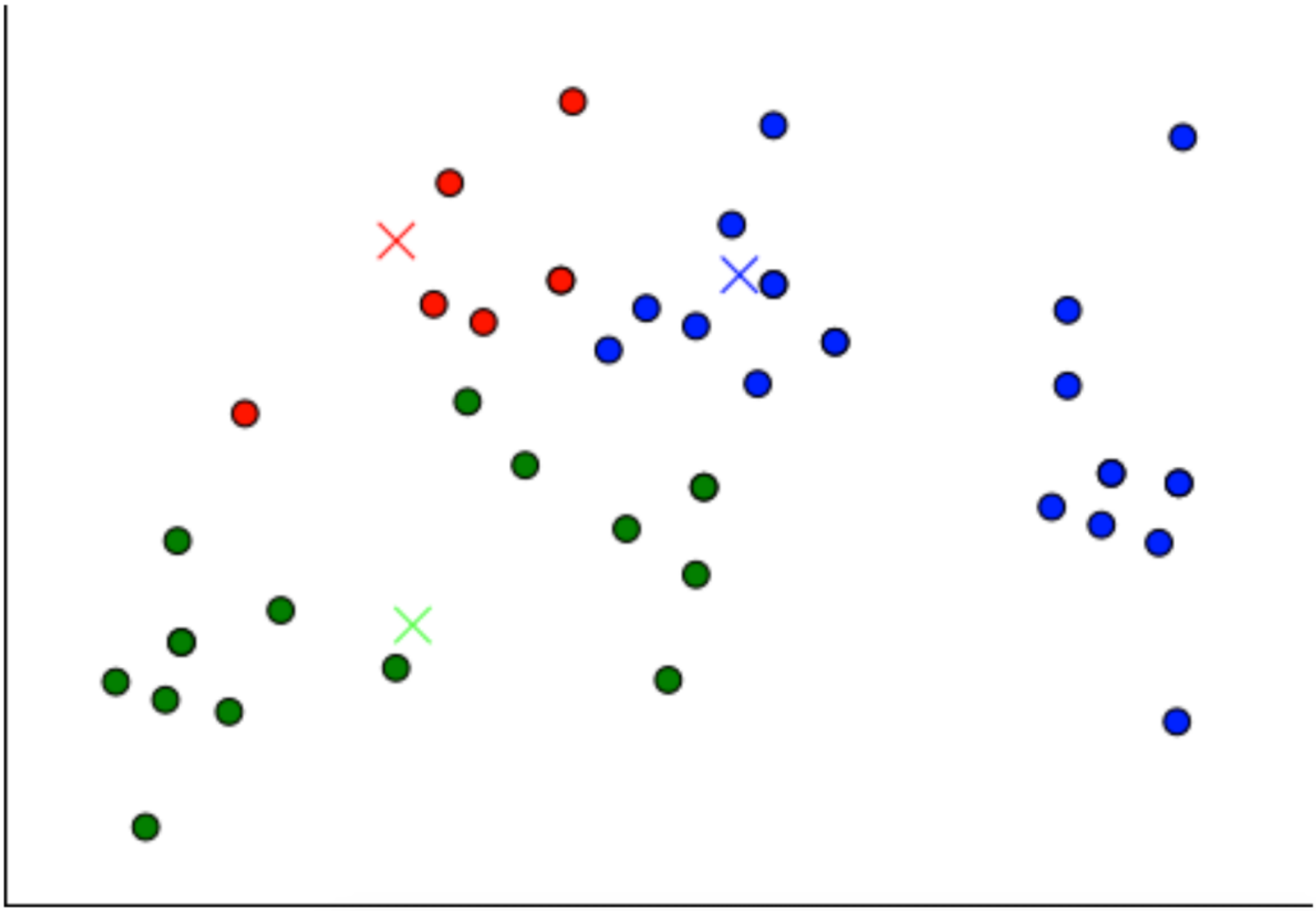


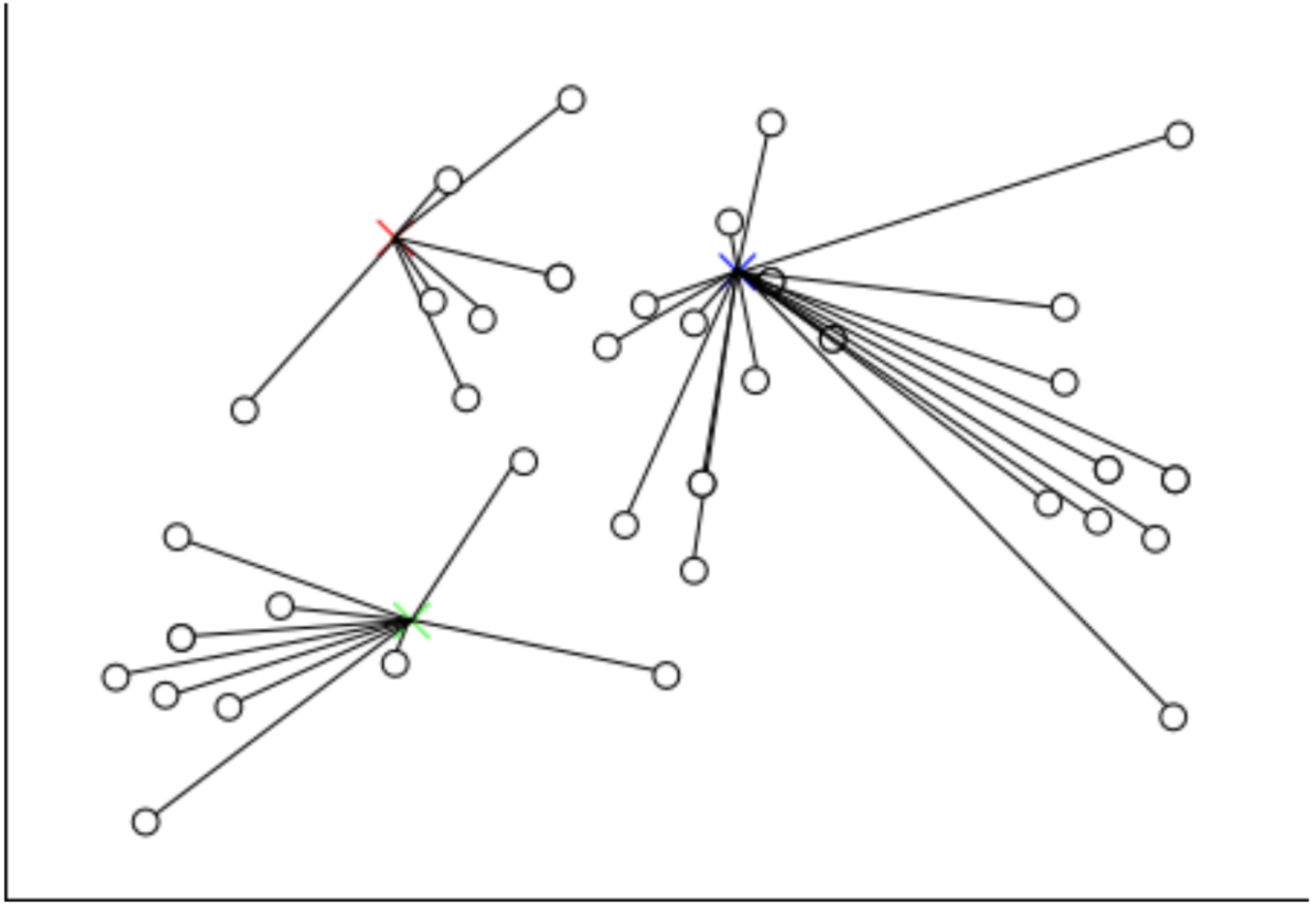












AND AGAIN

- back to the notebook!

FURTHER READING

- http://xkcd.com/1374/
- http://xkcd.com/882/