* Introduction
  + ML is more like teaching employee than raising a child
  + In a sentence, ML provides a set of tools comp’s use to transform data into actionable knowledge
* Origins of ML
  + Recent advances in technology have created abundant and easily available data
  + ML is turning this data into intelligent action
  + Statistics, Data availability, and Computing prowess reinforce each other’s development
  + Data Mining
    - Concerned with obtaining novel insights from large data sets
    - Focuses: ML- Performing a known task; DM – search for hidden nuggets of info
    - ML – Teach robot to drive; DM – Figure out car safety trends
    - Note: DM requires ML, but not vice versa
* Uses / Abuses
  + Identify philanthropists, criminals, predict elections and natural disasters
  + Anecdotally, a US retailer started to send a woman pre-natal care product coupons, based on her previous purchases. The father contacted the company angry the company was encouraging teen pregnancy, only to call later to apologize since his daughter turned out actually pregnant.
  + Iffy issue, due to the invasion of privacy
* Ethical Considerations
  + Law lags technology significantly
  + ML algorithms can find awkward data – data it’s not really meant to (e.g. teenage mom)
* How do machines learn
  + Tom M Mitchell – Machine is said to learn if it can take experience and utilize it to improve performance on similar future experiences
  + 3 Main Components – Data, Abstraction, Generalization  
    
  + Aim is to manage a small set of key ideas than eidetically remember each case (limited usefulness)
* Abstraction – these set of key ideas, as a tool to interpret data
  + i.e. Magritte  
    This isn’t a pipe, it’s a *picture* of a pipe. This is the sort of abstraction we perform in ML

Data is summarized into a model, such as an equation, tree, if-else statements, or clusters

We ‘train’ these models using the data given

* Abstraction – cont.
  + This training is not bottom-up inductive reasoning – its training data onto our model
  + Newton fit equations to observational data, found relation
    - Re-investigating always present facts/data to gain new insight
* Generalization
  + Turning abstracted knowledge into a form that can be utilized for action.
  + Reducing possible number of abstracted theories into a manageable set
  + Use heuristics to eliminate unneeded theories – not perfect, but otherwise infeasibly slow
    - E.g. humans fear airplanes more than cars
    - Leads to bias
* Steps to Apply ML to Data
  + Collect data
  + Explore/prepare data
  + Train a model on the data
  + Evaluate Model
  + Improve Model (or change model altogether)
* Choosing an ML Algorithm
  + Predictive Models
    - Aim to predict a target feature (maybe in past, e.g. conception date)
    - Supervised Learning
    - Classification, for example (win lose. Earthquake next year, etc)
    - Also, numeric prediction (salary, forecasting)
  + Descriptive Model
    - Unsupervised learning
    - Often found in data mining
    - Pattern discovery 🡪 market basket analysis
    - Clustering 🡪 obtain data ‘demographics’
* ML in R
  + <http://cran.r-project.org/web/views/MachineLearning.html>
  + https://cran.r-project.org/web/packages/mlr/index.html