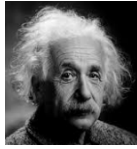


The Analytics Life Cycle



Food for Thought

- **Humans** are live predictive analytics entities
- **Politician's tricks** → single data point; single variable prediction
 - Does one case prove anything? Statistical power?
 - Is there really a gender pay gap?
 - ↑ \$1 minimum wage → Impact on unemployment?
- What doctors do? → diagnosis, treatment
- Why are you in the MS program? What are you predicting?
- Who in this class will be the first to become CEO?
- How much money will you make 5 years after graduation?
- Will you make more than \$200K 5 years after graduation?
- What's the profit impact of providing free WiFi?
- How to fight cybercrime, terrorism (Shining Path), fraud, etc.?
- Interesting cases: Target, Starbucks, Oakland A's

The Analytics Life Cycle

Step 1 – Problem Definition: **Formulate analytics question**

Question Type:

- (1) Quantitative (regression, regression trees, neural networks, etc.)
- (2) Classification (e.g., logistic regression, classification trees, etc.);



Step 2 – Data Work

Identify & gather data

Type: structured, unstructured and visual data

Pre-process data: cleanse, prepare, transform, format, etc.



Step 3 – Model Decisions: **Select method/approach**

Descriptive – familiarize with and analyze the data; unsupervised learning

Goals: identify patterns: descriptive statistics, correlation, cluster data, market basket association, data mining, etc.

Predictive – use existing data to predict outcomes; supervised learning:

Goals: Test hypotheses; Explain relationships; Predict outcomes
– i.e., accuracy

Prescriptive – decision models; optimization, etc.

Goals: inform best courses of actions



Step 4 – Analysis:

Analysis goals: model sophistication/parsimony; apply business domain knowledge; variance vs. bias; fitness vs. over-identification



Step 5 – Reporting: **Written, interactive, visual, etc.**

Learning Objectives

- The main **learning objective** in this course is for you to learn how to **select, justify** and **employ** the most appropriate model to answer a predictive analytics question
- We will cover all aspects of the analytics cycle, but with an emphasis on **predictive modeling**
- This is **NOT** as **Statistics** class, but it requires:
- A good understanding of **Statistics** and **Machine Learning**
- And other factors for developing sound predictive models



Developing Sound Predictive Models

Besides good knowledge of statistics and machine learning, developing a sound predictive model **also** depends on:

- What are your predictive modeling **goals**:
 - Inference
 - Interpretation and/or
 - Prediction
- Which **data sources** you use
- How do you prepare/**pre-process** the data
- Which **modeling method** you use
- Which features (i.e., **variables**) do you include in the model
- Having a good **business rationale** for you predictors selection



KOGOD SCHOOL
of
BUSINESS