Sports Analytics II

Produced by Dr. Mario

UNC STOR 390

Data Management

- Different Sources: What are some examples?
  - Quantitative and Qualitative
  - Objective and Subjective
  - Text and Images and Video
- Increasing Number of Sources
- Increasing Volume from Those Sources
- Data Comes Structured and Unstructured
  - Structured is Easier to Analyze
  - Unstructured is More Flexible

Data Management

- **Process of Data Management**
  - Standardization
    - Centralization
      - Integration
- **Standardization**
  - Purpose: To Make Combining Data Easy
  - Know All Sources of Data in the Organization
  - Create a Data Inventory: Variable, Description, Format
  - Different Departments, but Same Format

Data Management

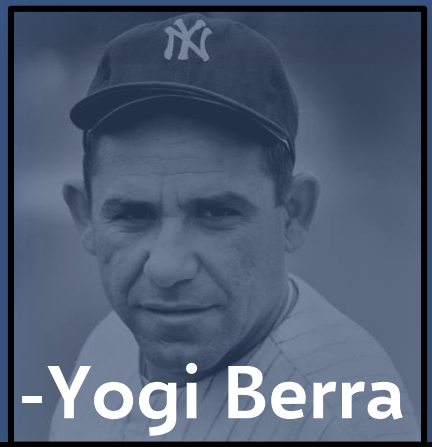
- Centralization
  - Purpose: To Make Acquiring Data Easy
  - Stored and Protected in the Same Location
  - Accessible by All Decision Makers
  - Continual Assessment of Data Quality Via Multiple Eyes

- Integration
  - Purpose: To Make Analyzing Data Easy
  - Merging/Linking Data According to Unique Identifiers
  - Examples: What Insights Could Be Learned?
    - Training Staff Data and Coaching Staff Data
    - Scouting Data and Play-by-Play Data
    - Marketing Data and Salary Data

Analytic Models

- Process of Predictive Analytics
  - Consider the Research Question or Problem
  - Identify or Create Dependent Variables of Interest
  - Incorporate All Information
  - Find Relationships (Linear/Nonlinear)
  - Evaluate Model, Report, and Repeat

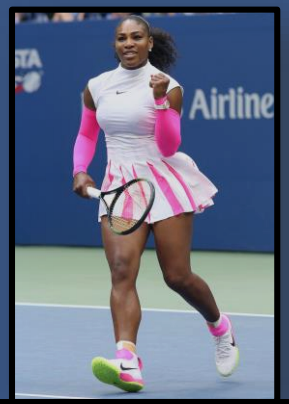
Prediction is difficult, especially about the future.



-Yogi Berra

Analytic Models

- Question Quality Influences Analysis Quality
  - Tangible Versus Intangible
    - How effective is Serena's serve when behind in sets?
    - How much stress does Serena put on her opponent?
  - Informational Versus Predictive
    - What information influences Serena's likelihood of winning a match?
    - How successful will Serena be next year?



Analytic Models

- Short Course on Sports Analytics
  - Mike Lewis, Goizueta Business School, Emory
  - Two Types of Projects
    - Long-Term
    - Short-Term
  - Analyses Should be Used to Evaluate Decision Biases
  - Big Data Problem
    - Number of Variables Increasing
    - Number of Players Remains the Same

Analytic Models

- Fundamentals on Sports Analytics
  - Mike Lewis, Goizueta Business School, Emory
  - Metrics AKA Quantifiable Measures
    - Understand the Past, Evaluate the Present, Predict the Future
    - Four Phase Process by Ben Alamar
      - Opportunity, Survey, Analysis, Communication
    - Metrics Must be Explainable and Testable
  - Statistical Model Essentials
    - Linear Regression
    - Generalized Linear Models

Analytic Models

Analytics don’t work at all.

It’s just some crap that people who were really smart made up to try to get in the game because they had no talent.

-Sir Charles Barkley

Analytic Models

- Five Questions for All Analyses – Ben Alamar
  - What was the thought process that led to the analysis?
  - What is the context of the result?
  - How much uncertainty is in the analysis?
  - How does the result inform the decision-making process?
  - How can we further reduce the uncertainty?

Analytic Models

Analytics will almost never outperform human judgment when it comes to individuals. What analytics are useful for is helping human decision makers self-correct.

-Mike Lewis

Information Systems

- Mechanisms for Data Delivery
- Organization and Presentation Matters
- “One Version of the Truth” – Ben Alamar
- Static: Automatically Generated Reports
- Interactive: Computer, Phone, Tablet, and Web Applications

Analytics In the Organization

- On-the-Field
  - Coaching
  - Player Development
  - Player Evaluation

- Off-the-Field
  - Sales
  - Marketing
  - Cap Management
  - Hiring
  - Public Relations and Social Media

Analytics Across the Organizations

- Sports Analytics Use Survey (2013)
- Sample of 27 People (NFL, MLB, NBA, EPL)
- How Many Different Sources?
  - 1-2 (6.7%)
  - 3-4 (33.3%)
  - 5-6 (13.3%)
  - >6 (46.71%)

Analytics Across the Organizations

- How Much Data is Centralized?
  - All (31.3%)
  - Most (37.4%)
  - Some (31.3%)

- How Much Data is Dependent on One Person?
  - Some (50.0%)
  - Most (43.7%)
  - All Data Centralized (6.3%)

Analytics Across the Organizations

- Is Data Checked for Errors?
  - Always (31.3%)
  - Usually (37.5%)
  - Sometimes (18.8%)
  - Occasionally (6.1%)
  - Rarely (6.3%)

- How Many Database Programmers are Employed?
  - 0 (37.5%)
  - 1-2 (50.0%)
  - 3-4 (0.0%)
  - >5 (12.5%)

Analytics Across the Organizations

- How Many Statistical Analysts are Employed?
  - 0 (20.0%)
  - 1-2 (66.6%)
  - 3-4 (0.0%)
  - >5 (13.3%)
- Roadblock: Difficulty Identifying Strong Applicant
- Difficulty in Both Hiring and Evaluating

Analytics Across the Organizations

- Clear Process for Hiring/Evaluating Analysts
  - Strongly Agree (13.3%/14.3%)
  - Somewhat Agree (13.3%/14.3%)
  - Neutral (26.7%/28.6%)
  - Somewhat Disagree (13.4%/21.4%)
  - Strongly Disagree (33.3%/21.4%)

- Analytic Resources in Line with Strategic Game Plan
  - Strongly Agree (26.7%)
  - Somewhat Agree (33.3%)
  - Neutral (33.3%)
  - Strongly Disagree (6.7%)

Final Inspiration

I am not cocky. I am 95% confident.

- Mahatma Mario