**BI Overview**

3 types of analytics: Descriptive, Predictive, Perscriptive

Decision Making Process: Intel, Design, Choice (Normative/Satisfying/Elim-by-aspects)

BI Architecture: DW, Bus. Analytics (tools), Bus. Performance Mgmt. (analysis), UI (dashboard)

**Data Warehouse**

OLTP (online transaction processing systems) used to “run” business, while DW used to analyze/optimize business.

ETL (extract, transform, load) Messy -> Clean for DW

**Data Mining**

Iterative Process: nontrivial, novel, useful, understandable.

* Data:
  + Structured
    - Categorical
      * Nominal (no order, not measurable) [no median/%-ile]
      * Ordinal (rank order, but not interval – low/med/high) [median, but no addition]
    - Numerical
      * Interval (Difference meaningful, but ratios are not – no 0) [addition/mean, but no ratio]
      * Ratio (interval + clearly defined 0) [ratios possible]
  + Unstructured
    - Textual (NLP)
    - Multimedia (Computer Vision)
    - HTML/XML (JSON)

Types of Patterns: associations, predictions, clusters, sequential relationships

CRISP-DM: Bus. Understanding -> Data Understanding -> Data Prep -> Modeling -> Evaluation -> Deployment

**Linear Regression**

, where n – k – 1 = df

SST = SSR (explained) + SSE (not explained, minimized)

| | | |

**Logistic Regression**

Logit: | |