Fire Data Analytics Book Outline

As I have mentioned there are a few good but a bit dated books out there that can potentially be updated/adapted/recycled for a modern fire service audience.

* Berner, M. (2013) Statistics for Public Administration: Practical Uses for Better Decision Making. 2nd Ed. ICMA Press. (<http://legacy.icma.org/en/press/print/statistics_for_public_administration_practical_uses_for_better_decision_making>)
* USFA (2004) Fire Data Analysis Handbook. 2nd Ed. FA-266 (<https://www.usfa.fema.gov/downloads/pdf/publications/fa-266.pdf>)
* And IFSTA has expressed interest/willingness to potentially publish this sort of book, perhaps Jones and Bartlett might be interested as well.

So to frame the discussion of what might be included in a revised or new book I figured I would jot down an outline of some possible topics.

1. Types of Data
   1. Nominal
   2. Ordinal
   3. Scale / Ratio
2. Types of Measures
   1. Input Measures
   2. Output Measures
   3. Outcome Measures
3. Types of Assessment
   1. Formative Assessment
   2. Process Assessment
   3. Impact Assessment
   4. Outcome Assessment
4. Fire Service Data Types (overviews)
   1. Call Handling Data
   2. Response Time Data
   3. Operational Data
   4. Outcome Data
   5. EMS / Patient Care Report data
   6. Community Risk Assessment / Reduction Data
      1. Data to focus CRR activities (before)
      2. Data to document CRR activities (after)
   7. Non-Fire Service Data (but relevant to fire departments)
      1. Census Data
      2. Building & Occupancy Data
      3. Tax Assessor & Cadastral Data
      4. Public Health Data
      5. Marketing Data
5. Cleaning Data
   1. Error identification
   2. Data Curation
   3. Importance of data cleaning
6. Types of Analyses
   1. Descriptive Statistics
      1. “What happened”
   2. Inferential Statistics
      1. “Why did it happen”
   3. Exploratory Analyses
      1. “What am I missing”
   4. Predictive Analyses
      1. “what could happen”
   5. Rates
      1. Raw counts vs. rates
      2. Normalization
   6. Data Distribution
      1. Central Tendency
         1. Mean,
         2. Median
         3. Mode
      2. Distribution Curves
         1. Normal Curve
         2. Bi-Modal Curves
         3. Skewness
      3. Percentiles
         1. 90th Percentile vs. Averages
   7. Identifying Outliers
      1. Interquartile Range (IQR) – Upper Fence Method
      2. Cut-off Scores
         1. SME defined Cut-offs
      3. Should outliers be removed from a dataset?
         1. Excluding and ignoring
         2. Setting aside / Exception Reports
   8. Performance Analysis
      1. Response Times
         1. Turnout Times
         2. Travel Times
         3. Unit Response Times
         4. Total Response Times
         5. Effective Response Force / Effective Firefighting Force
         6. Reflex Time / Vertical Response Time
         7. Commit Time
      2. Operational Analysis
         1. High Service Utilization Analyses
            1. Mobile Integrated Health Care (“Extremely Satisfied Customers”)

How to identify high service level users that may not be the highest call volume address (a resident that calls 10 times vs. a nursing home or apartment building that calls 100 for different residents).

* + - * 1. Identifying repeat system problems (Repeat False Alarms)
      1. Deployment and Planning
         1. Determining how to determine the need to establish or move fixed facility or mobile resources.
    1. Common Fire Data Challenges
       1. Cancelled Enroute
       2. Differences between “Dispatched To” to “Situation Found”
       3. Firefighter Injuries
       4. Identifying “Extremely Satisfied Customers” (High Volume Service Users).
       5. XXX
  1. System Performance
     1. Unit Hour Utilization (UHU)
     2. Unit Reliability
        1. How is Reliability affected by Move-Ups and Alternative Deployment strategies
        2. Is “unit reliability” something a fire company can fix themselves?

1. Data Handling / Management????
   1. Data Quality
      1. GIGO
      2. Data Cleansing
   2. Connecting Data
      1. Linkage/Joins/Merging
         1. Primary Keys
         2. Foreign Keys
         3. Parent – Child Relationships
         4. Joins
            1. One to One
            2. One to Many
            3. Many to One
            4. Many to Many
            5. Inner Joins (left)
            6. Outer Joins (right)
         5. De-Duplication
      2. Data System Integrations
         1. Entity Relationships
         2. XXXX
2. Tools for Managing & Analyzing Data
   1. Data file formats
      1. CSV
      2. XLSX
      3. TXT
      4. XML
      5. JSON
         1. GeoJSON
   2. Spreadsheets
      1. Pivot Tables
   3. Statistical Packages
      1. SAS
      2. SPSS
      3. KNIME
      4. R
3. Geospatial Data Analysis – GIS Analysis
   1. Deployment Analysis
      1. Coverage polygons
      2. Response Polygons
      3. XXX
   2. Network Analysis
      1. Creating response territories
      2. Setting road features
         1. Joining and Disconnecting Roads
         2. Road Speed
            1. Acceleration/Deceleration
            2. Traffic Congestion
         3. Time Penalties
            1. Left Turns
            2. Traffic Control Devices
         4. Turn Radius
         5. Weight/Height/Size Limits
   3. Geocoding
      1. Street Centerlines
      2. Address Validation
   4. Spatial Statistics
      1. Hotspot Analysis
      2. Convex Hull
   5. Field Collection of Data
      1. Smart Device Applications
4. Data Visualizations
   1. Tables
   2. Graphs
   3. Charts
   4. Maps
   5. Infographics
   6. Business Intelligence Tools
   7. Pitfalls of Visualizations
      1. Less Is More
      2. Intelligibility
      3. Misleading representations
      4. Using and Abusing Color
      5. Interactive Visualizations
      6. Special Effects
5. Advanced Data Usage
   1. Performance Management
      1. Key Performance Indicators (KPIs)
         1. Leading Indicators
         2. Lagging Indicators
      2. Benchmarking
      3. Scorecards
      4. CompStat/CityStat programs
      5. Managing unintended consequences
   2. Dynamic Deployment
      1. System Status Management
      2. Peak Time Deployment
      3. XXXX
   3. Incident Prediction
      1. Fire Risk prediction
      2. Predictive Community Risk Reduction
         1. Targeted prevention efforts
            1. Identifying high risk populations

Data Sources

* + - * 1. Identifying high risk properties for responding firefighters

Building condition and Vacancy Data

SARA Title 3, Tier 2 Hazardous Materials Data

Occupancy Changes

Finding Missing Buildings

* + - * 1. Identifying where higher risk populations live
        2. Identifying where time spent with boots on the ground hits more at risk
        3. Identifying higher risk properties to first responders

Age

Construction

Special materials on site or usage

* 1. Open Data Portals
     1. Democratizing Data vs. Dangerous Data