Human Capital Management:
Identify Retention Risk with Attrition Modeling

James E. Parry,
Client Technical Specialist
IBM SPSS Predictive Analytics

A leading provider of predictive analytic software, services and solutions

- **Software** – data collection, text and data mining, advanced statistical analysis and deployment technologies

- **Services** – implementation, training, consulting, and customization

- **Solutions** – combine software and services to deliver high-value line-of-business solutions; used for optimizing marketing campaigns, call center effectiveness, identification of fraudulent activity and more

- Over 40 years of experience and a broad customer base
  - 250,000 customers: 100 countries, 50 states, 100% of top universities
  - Widely used throughout US DoD and US Intelligence

- Non-proprietary approach
  - Non-intrusive integration (Services Orientated Architecture)
  - Database agnostic
  - Leverages existing operational software, IT investments, and custom analytic assets
Questions Data Mining Can Answer

**Commercial Sector**

- Reducing campaign costs and increasing customer conversions
- Decreasing customer churn
- Reducing fraud and improper payments
- Maximizing ROI on direct marketing campaigns
- Improving product offerings by understanding customer needs

**Public Sector**

- Reducing recruiting costs and increasing employee retention
- Decreasing institutional attrition
- Rare event /Anomaly detection
- Maximizing ROI on campaigns
- Improving public health and safety by understanding constituent needs
Business Analytics Components

Resource Planning and Analytics
- Planning
- Forecasting
- What if
- Data Capture

Predictive Analytics
- Data Collection (surveys)
- Statistics
- Modeling
- Deployment

Enterprise Business Intelligence
- Reporting
- Dashboarding
- Metrics
- Events
- Self Service
- Collaboration
- Statistics
- More…

Structured Data Sources

Unstructured Data Sources

Latency -> Static, On Time, Real Time
Enabling Predictive Analytics Across the Enterprise

**Capture**
Data Collection/Access delivers an accurate view of history, behaviors, attitudes, and opinions.

**Predict**
Predictive capabilities bring repeatability to ongoing decision making, and drive confidence in your results and decisions.

**Act**
Unique deployment technologies and methodologies maximize the impact of analytics in your operation.

---

**Data Collection/Access**

**Platform**

**Build Content**

**Deployment Technologies**

Text Mining  Data Mining  Statistics
Demonstration

- An HR manager has been responding to a spike in employee attrition. In an attempt to analyze the data, the manager has logged into the BI system and looked at factors that might be driving attrition. While some factors appear to influence attrition the more factors she adds to the charts, the more complex the picture becomes. What she needs to be able to do is:

  - Gain insight into main drivers of attrition.
  - Identify persons most likely to voluntarily separate or attrite based upon historical data.
  - Find a way to rank order the cases from most likely to attrite to least likely to attrite.
  - Generate a list of at-risk employees and distribute to employee’s managers.
BI Analyst: Cognos

- The analysis begins by logging into Cognos, creating and viewing employee data.
BI Analyst: Cognos

- Noticing a stark rise in attrition the BI decides to dig into the data for answers:
BI Analyst: Cognos

- Searching for key drivers of attrition, the analyst looks at job position by commute time by year.

<table>
<thead>
<tr>
<th>Position</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative Assistant</td>
<td>0.442</td>
<td>3.557</td>
<td>1.293</td>
<td>5.806</td>
<td>2.985</td>
<td></td>
</tr>
<tr>
<td>Director/Supervisor</td>
<td>24.779</td>
<td>76.285</td>
<td>4.741</td>
<td>14.194</td>
<td>29.851</td>
<td></td>
</tr>
<tr>
<td>Executive Management</td>
<td>8.407</td>
<td>3.953</td>
<td>0.784</td>
<td>3.879</td>
<td>14.839</td>
<td>9.701</td>
</tr>
<tr>
<td>Non-managerial/Professional</td>
<td>62.832</td>
<td>16.206</td>
<td>99.216</td>
<td>89.655</td>
<td>64.516</td>
<td>54.478</td>
</tr>
<tr>
<td>Office Services/Clerical</td>
<td>3.540</td>
<td></td>
<td></td>
<td>0.431</td>
<td>0.645</td>
<td>2.985</td>
</tr>
</tbody>
</table>
BI Analyst: Cognos

- The analysis becomes more complex but no clear pattern emerges
Cognos and SPSS Integration

- Integration at the Data Model level
Cognos and SPSS Integration

- Integration at the Data Model level
Cognos and SPSS Integration

- Integration at the Data Model level
- Thin client integration
  - Socialize analytics
  - Distribute the power of analytics
Cognos and SPSS Integration

- Integration at the Data Model level
- Thin client integration
  - Socialize analytics
  - Distribute the power of analytics
Cognos and SPSS Integration

- Integration at the Data Model level
- Thin client integration
  - Socialize analytics
  - Distribute the power of analytics
- Creates a self-service ROI engine using Predictive Analytics
Cognos and SPSS Integration

- Integration at the Data Model level
- Thin client integration
  - Socialize analytics
  - Distribute the power of analytics
- Creates a self-service ROI engine using Predictive Analytics

Historical data is used to build a predictive model of employee attrition
Cognos and SPSS Integration

- Use different models to solve different problem sets
- Prediction, Segmentation, Association and Rules-based Models

When predicting attrition an auto-classifier will be run
Cognos and SPSS Integration

- Specify the outcome or target variable
- Use business expertise to select custom fields for prediction

Data are now available for a predictive model
Cognos and SPSS Integration

- Specify the outcome or target variable
- Use business expertise to select custom fields for prediction

Inclusion and exclusion rules are available to specify which employees to analyze
Cognos and SPSS Integration

- Evaluate predictor importance
- Use standard evaluation charts such as the distribution chart to gauge accuracy

Work-life satisfaction, drive time and location are top drivers here
Cognos and SPSS Integration

- Validate the model on test data

Choose which fields to return in the analysis
Cognos and SPSS Integration

- Validate the model on test data

Data are returned with predicted attrition and propensity scores.
Cognos and SPSS Integration

- Score directly back to the Cognos BI server
- Eliminates reformatting data, reduces data handling errors
- Enhances data security

Just enter BI Server and DB credentials
Cognos and SPSS Integration

- Score directly back to the Cognos BI server
- Eliminates reformatting data, reduces data handling errors
- Enhances data security

A new table will be created called HCM_Scored
Cognos and SPSS Integration

- Score directly back to the Cognos BI server
- Eliminates reformatting data, reduces data handling errors
- Enhances data security

This table will be rendered directly in Cognos
Cognos and SPSS Integration

- Score directly back to the Cognos BI server
- Eliminates reformatting data, reduces data handling errors
- Enhances data security

<table>
<thead>
<tr>
<th>Propensity Group</th>
<th>Propensity</th>
<th>Rank</th>
<th>ID</th>
<th>Position</th>
<th>Estimated Income</th>
<th>Length of Hire</th>
<th>Benefit</th>
<th>Training</th>
<th>Degree</th>
<th>Facility</th>
<th>DriveTime</th>
<th>Enjoyable</th>
<th>Challenging</th>
</tr>
</thead>
<tbody>
<tr>
<td>98% or more</td>
<td>0.99351547</td>
<td>1</td>
<td>1,581</td>
<td></td>
<td>89,459.9</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.9929214</td>
<td>2</td>
<td>2,845</td>
<td></td>
<td>89,459.9</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.9929214</td>
<td>2</td>
<td>3,089</td>
<td></td>
<td>89,459.9</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.99258984</td>
<td>4</td>
<td>2,825</td>
<td></td>
<td>89,459.9</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.99197368</td>
<td>5</td>
<td>509</td>
<td></td>
<td>90,478.6</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.99033685</td>
<td>6</td>
<td>3,555</td>
<td></td>
<td>90,478.6</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.99024295</td>
<td>7</td>
<td>623</td>
<td></td>
<td>74,460.7</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.98983969</td>
<td>8</td>
<td>431</td>
<td></td>
<td>35,000</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98% or more</td>
<td>0.98967893</td>
<td>9</td>
<td>1,978</td>
<td></td>
<td>35,000</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From Data . . .

Human Capital Management - Business Analytics

• **Gain insight into main drivers of attrition.**
• **Identify persons most likely to voluntarily separate or attrite based upon historical data.**
• **Find a way to rank order the cases from most likely to attrite to least likely to attrite.**
• **Generate a list of at-risk employees and distribute to employee’s managers.**
To...Data Driven Decisions

- Why don’t my employees find the workplace enjoyable?
- Should we consider telecommuting?
- Why are some facilities experiencing higher attrition rates than others?

・Gain insight into main drivers of attrition.
・Identify persons most likely to voluntarily separate or attrite based upon historical data.
・Find a way to rank order the cases from most likely to attrite to least likely to attrite.
・Generate a list of at-risk employees and distribute to employee’s managers.
Data Driven Decisions

- Gain insight into main drivers of attrition.
- Identify persons most likely to voluntarily separate or attrite based upon historical data.
- Find a way to rank order the cases from most likely to attrite to least likely to attrite.
- Generate a list of at-risk employees and distribute to employee’s managers.
Data Driven Decisions

- Gain insight into main drivers of attrition.
- Identify persons most likely to voluntarily separate or attrite based upon historical data.
- Find a way to rank order the cases from most likely to attrite to least likely to attrite.
- Generate a list of at-risk employees and distribute to employee’s managers.
Customer Story

U.S. Military Academy – West Point

Personnel and Human Resources

Challenge – Produce the highest quality career officers

- Needed a way to improve recruiting of career officers
- Needed a way to reduce cadet attrition
- Over 20 years of historical data in different formats
- Current analysis operations performed by math department are time consuming and reactive

Solution

- SPSS Modeler for determining characteristics that lead to higher attrition and for modeling “propensity to attrite”
- SPSS Text Analytics for analysis of student survey data to use as additional inputs into modeling algorithms

Benefits

- Improved retention of current cadets
- Improved identification of recruits more likely to exceed minimum length of service contract
Customer Story
*U.S. Army Recruiting Command*

**Personnel and Human Resources**

Challenge – *Improve recruiting effectiveness and efficiency*

- Improve recruiting through better targeted marketing
- Prioritization of leads – over 600,000 leads per year
- Reduce drop-outs from Delayed Entry Program

Solution

- SPSS Modeler for modeling to classify individuals most likely to enlist, finish boot camp, complete the Delayed Entry Program
- SPSS Text Analytics for analysis of unstructured data from recruit interviews and surveys

Benefits

- Before using SPSS Modeler, on average, an Army recruiter had to contact 100 leads to get 4 contracts. After using Modeler, a recruiter had to contact only 20 leads to get 2-3 contracts
- All leads were prioritized in order to provide the maximum gains from the recruitment model.
Customer Story

**U.S. Army Reserve - OCAR**

**Personnel and Human Resources**

**Challenge – Reduce attrition and determine reasons for reserve attrition**

- Reserve soldiers have careers and responsibilities outside of the U.S. Army, making **high attrition rates** an ongoing challenge
- Need to determine the characteristics that lead to attrition and the types and **levels of incentives** that can aid in retaining a soldier

**Solution**

- SPSS Modeler used for classify soldiers at risk of attrition, including the analysis of military occupational skills (MOS) in classifying attrition
- SPSS Modeler to create models for incentive planning.

**Benefits**

- **Predicted attrition** using demographic data for army reservists.
- Created a predictive model to analyze **why reservists leave** and used this model for scoring the likely for attrition of candidates on a weekly basis.
- Modeled the soldier **incentive types and levels** that would minimize cost and attrition
Thank You!