Data Analytics, Artificial Intelligence & Deep Learning: Potential Impact On Workforce Management
HR’s Image

• Finance, Marketing & Operations are using scientific methods & technology more
  – In analyzing data
  – In formulating recommendations
  – In supporting recommendations
• HR is often viewed as having lagged behind
• Opinions often used rather than data
• Recommendations often cost resources but do not promise tangible results (i.e., if we spend on training we think people will be better)
Evidence-Based Management

• Origin in evidence-based medicine
• AOM & SIOP focus
• Premise: workforce management needs a major transformation
  – Decisions based on credible evidence
  – Decisions made using scientific method
  – Decisions that support the business
What “evidence” is used in decisions?

- Intuition
- It worked at my last job
- Pop literature
- Guru counsel
- “Best” practice
- We have always done this
- Research
- Reward $ystems, Inc.
Two Approaches To Finding & Applying Evidence

- **Deductive** *(theory > findings)*
  - Identify testable theory *(more pay > less turnover)*
  - Formulate hypothesis *(if – then causation)*
  - Test hypothesis
  - Assess, refine/confirm/reject theory

- **Inductive** *(evidence > findings)*
  - Accumulate evidence
  - Analyze patterns
  - Develop theory
Deductive Is Hard Work

• Formulating feasible hypotheses
  – “Above market pay gets/keeps top talent”

• Defining all the underlying assumptions
  – “People value pay”
  – “They believe their pay is above market”
  – Other…

• Constructing a test of the assumptions
  – *Running tests on impact of pay is risky*
And Inductive Is “In The News”

• Literature is full of “big data/analytics”
• All the stories are about successes (surprise?)
• Plus, trolling massive data sources is easy
• And, the chaotic environment makes theory formulation difficult
• And, above all, analytics is the “new best thing”
But Danger Lurks

• *If sufficiently tortured data will confess to anything*

• We are subject to cognitive bias
  – We more readily accept information that agrees with what we believe/want to be so
  – We see things in clouds when there are only clouds/random patterns
  – We develop conclusions from inadequate samples
  – We often assume correlation = causation
Analyzing Data Requires Human Judgment/Intuition

• Challenge is to turn
  – Data into information
  – Information into knowledge
  – Knowledge into action

• Data can be analyzed using quantitative methods to yield patterns/relationships

• But a lot of subjective intervention is unavoidable
Serious Mistake

• Using measures just because they are quantitative
• Assuming something that can be counted is important, whether it is or not
• Shying away from qualitative measures requiring subjective judgments
• Q: Aren’t many of the most critical decisions made based on judgments?
Data Analytics Can Drive Out “Soft Stuff”

• Book/movie “Moneyball” was example of the value of using metrics to make better decisions (induction)

• But movie “Trouble With The Curve” illustrated that experience and knowledge (deduction) can be a valuable partner with data and analytics
The “Soft Stuff” Counts

- Ryan Leaf was picked ahead of Peyton Manning in the NFL draft (*Ryan who?*)
- Leaf failed not on athletic ability, that can be measured objectively
- Leaf failed on personality, that can only be measured subjectively
But Relying Only On Soft Stuff May Not Be Wise

• The most widely used selection tool is the one-to-one unstructured interview
• The least valid predictor is the one-to-one unstructured interview
• Using some hard data may prompt use of more valid process
Data Analytics

• Determining what you need
  – Historical data
  – Trend data
  – Predictive data
  – Prescriptive data

• Knowing what type of relationship you need
  – Correlation
  – Causation
Evaluating Data

• From relevant contexts?
• Source of data valid?
• Age of data impact usefulness?
• Accuracy of data questionable?
• Quantitative or qualitative data?
• Based on all causal factors?
• Evaluator have the qualifications and the neutrality to analyze impartially?
Sears: An Experiment With Balance

• Organization wanted to be a compelling to: Work, Shop and Invest
• Explored impact of employee satisfaction
• Used design that established correlation but also indicated causation
Findings

• 5 point improvement in employee attitudes (measured by survey) -> 1.3 point improvement in customer satisfaction -> 0.5% increase in revenue growth

• The factors leading to improvement in employee attitudes were:
  – Quality of supervision
  – Quality of job/workplace design
  – Quality of employee training
  – Quality of teamwork
  – Quality of work context (pay & benefits; promotion opportunity; fair treatment)
But What If There Is No Data?

• Data can be found if an organization wants evidence about
  – What has happened
  – What is happening

• Innovation/invention may involve conditions that have never been

• So if the future will not be like the past/present is there useful data?
The Tools
Analyzing Data

• Determine what type of data you need
  – Historical data
  – Trend data
  – Predictive data
  – Prescriptive data

• Know what type of relationship you need
  – Correlation
  – Causation
Example Of Predictive Model: Assessing Workforce Viability Against Future Requirements

Today’s Workforce:
- Attract: External
- Develop: Internal
- Legal/Reg. Environment
- Performance: Internal
- Role Design: Internal

Required Workforce:
- Retirement
- Turnover
- Competitive Arena
- Technology
- Promotion/Transfer

Identify sources of supply and losses, project net gains or losses and then determine whether human capital will be adequate.
## Predictive & Prescriptive Model: Flow Analysis: Control Room Operators

<table>
<thead>
<tr>
<th></th>
<th>Entry Level</th>
<th>Journey Level</th>
<th>Senior/Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current Staff</strong></td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td><strong>Current Demand</strong></td>
<td>2</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td><strong>Current Gaps</strong></td>
<td>-2</td>
<td>6</td>
<td>-4</td>
</tr>
<tr>
<td><strong>Demand: 1 year out</strong></td>
<td>3 (+1)</td>
<td>12 (-)</td>
<td>6 (-)</td>
</tr>
<tr>
<td><strong>Losses projected: next year</strong></td>
<td>1</td>
<td>-1</td>
<td>4</td>
</tr>
<tr>
<td><strong>Gaps: 1 year out</strong></td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td><strong>Demand: 3 years out</strong></td>
<td>3 (-)</td>
<td>14 (+2)</td>
<td>8 (+2)</td>
</tr>
<tr>
<td><strong>Losses projected: next 3 years</strong></td>
<td>4</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td><strong>Gaps: 3 years out</strong></td>
<td>4</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

Progression through levels must be projected as well to determine staffing needs.

Reward $ystems, Inc.
Analyzing Large Data Sets

• When data points are numerous they need to be clustered into like “chunks”
• Often using an average (or median) disguises the true patterns
• Using histograms (frequency distributions) can enlighten the analyst
  “Getting into the weeds” can be key to understanding what the data says
Employee Attitude Survey

• Common error: report “averages” rather than frequency distributions: what does the average tell you in data below?

<table>
<thead>
<tr>
<th>A</th>
<th>V</th>
<th>E</th>
<th>R</th>
<th>A</th>
<th>G</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“Helping Organizations Succeed Through People” 24
Single Factor Linear Regression: Analysis Tool

- The relationship between two variables
- Common example:
  - Plot internal job value & external rates
  - Analyze outliers & reconcile
  - Determine degree of correlation
  - Develop formula to explain relationship
  - Use formula as policy line ($y = a + bx$)
Scatter gram: Useful Formula For Explaining Relationship?

\[ Y = aX + b \]
Scattergram: What Data Is Used To Calculate Formula?
The Line/Formula Fits The Data: How Do We Avoid This Trap?
Scatter Chart – Current Paid Wages and Survey Data

Monetary Value

Wage Survey Trend Line

Relative Job Value

Current Company Trend Line

- Average company rate for job
- Average market rate for job

Reward $ystems, Inc.

“Helping Organizations Succeed Through People” 29
Non-Linear Regression

• If the model that best explains the data is not a linear function other statistical techniques must be used
  – Ex., Executive pay analysis often results in a curve (exponential function) when relating pay levels to scope data such as size

• Many curve-fitting packages are available and should be considered

• Danger: Forcing data into linear model produces a false relationship
Multiple Factor Regression

\[ Y (\$) = a + bx_1 + cx_2 + dx_3 \]

\( x_1 \) = factor 1 (education)

\( x_2 \) = factor 2 (e.g., years of experience)

\( x_3 \) = factor 3 (e.g., job grade)

\( r^2 \) = coefficient of determination (degree to which factors explain variance - % of actual)
Multiple Factor Regression Has Many HR Applications

- Statistically testing factors impacting selection of new hires
- Identifying potential departures
- Explaining pay rates... what factors are influencing pay (e.g., performance ratings, longevity in job, grade)
- Widely used in testing statistically for discrimination
An Example Of An Application

• Analysis done using multiple regression, finding what factors impact pay rates. The following are found to do have statistically significant impact:
  – Education level (positive correlation)
  – Years of experience (positive correlation)
  – Gender (males paid more)
  – Ethnicity (whites paid more)
  – Job grade (positive correlation)

• What would your reaction be? What additional information would you want before you called your in-house/out-house counsel?
Reverse Regression

• Harris Trust case on gender-based pay discrimination
  – Organization found pay for women averaged about the same as for men: therefore, conclusion was that there was no problem
  – Statistician showed that women were more qualified as a group, therefore should have been paid more than men on average
  – Lesson: look at everything; from every angle... and then evaluate patterns
Correlation & Causation

• If there is a high correlation between A and B it does not establish causation
  – A and B may both be caused by C
  – The correlation may be accidental
  – Causation unclear… taller people tend to weigh more, but does one cause the other (or do they co-vary)?
  – B may cause A or A cause B
What Causes What?

• Taco Bell found:
  – Stores in top quartile in customer satisfaction had better financial performance
  – Stores in top quartile in employee turnover (it was low) had double the sales and 55% higher profits than stores in lowest quartile

• What can be concluded from this?
• What was the direction of causation?
Causation: The Direction Counts

• Does satisfaction impact performance?

• Does performance impact satisfaction?

• What do you think research shows?
Factors Impacting Likelihood Of High Performance: Correlation

• Both intelligence and conscientiousness have been shown to be correlated to performance
• Which do you think has the best correlation?
• Which can be impacted by actions the organization can take?
• So what do you do with the evidence?
Trend Analysis

• Be aware of trends in data over time
  – Cyclicality/seasonality
  – Upward or downward pattern
  – Extreme values: timing measurements of improvement (regression to mean)

• Particularly important in HRM for:
  – Scenario planning; forecasting future
  – Understanding how significant changes are
  – Incentive plan target setting
Example: Establishing New Cash Incentive Plan

• Need to set “baseline” to determine the size of the incentive fund, based on the performance level
• Need to determine if implementing the plan made a positive difference
Last 20 Periods:
Where To Set Baseline?
Last 20 Periods: Where To Set Baseline?

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   | X | X | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   | X |   |   | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   | X |   | X |   | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   | X |   |   |   | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   | X | X |   | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   | X |   | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   | X |   | X | X | X |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

Reward $ystems, Inc.

“Helping Organizations Succeed Through People” 42
Last 20 Periods:
Where To Set Baseline?
Evaluating Interventions

• We attempt to intervene
  – To fix problems
  – To improve performance
  – To change what has been to what we want

• Evaluating an interventions requires
  – Understanding what has been
  – Determining if the desired change has occurred
Did “X” Make A Difference?
Did “X” Make A Difference?

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

“Helping Organizations Succeed Through People”
Did “X” Make A Difference?
Did “X” Make A Difference?

Reward $ystems, Inc.

“Helping Organizations Succeed Through People”
But Are We Sure It Was “X”?

• Did X cause all of the difference?
• What else might have contributed?
• Would X make the same difference the next time? Under different conditions?
Validity Of Measures

• Measures used must avoid
  – Contamination: including factors that are not related to what you intend to measure (e.g., rewarding sales person on $ volume, which includes price increases over which she had no control, in addition to units sold)
  – Deficiency: not including factors that are a part of what is being measured (e.g., appraising performance on quantity without considering quality)
Example Of Poor Measure For Determining Incentives

• Sales reps for manufacturer paid on total sales volume ($)
• Included were parts orders to repair machines (up to 50% of volume for some of the reps)
• Company needed new machine sales
• Sales rep not involved in parts orders
• So paying for parts volume is ________?
The Two Types Of Validity

• *Internal validity*: data analysis establishes relationship within the context of the study

• *External validity (generalizability)*: establishes that the relationship holds across different contexts
Elusiveness Of External Validity

• Example: lab research in “Drive”: Impact of external rewards on intrinsic rewards finding only holds for throwing tennis balls at targets (even if internally valid)

• Question: is that generalizable to employee doing a job they would not do for entertainment… and would they do it for thirty years?

• Problem is *relevance to the context*
Challenge: Appropriate Scaling

• 5 point performance appraisal scale may not produce useful data
  – Some jobs are “pass – fail”

• Do differences make a difference?
  – Need to know the return on improved performance
There is more value in improving Sweepers than Mickey Mouse.

Source: “Beyond HR” by J. Boudreau & P. Ramstad
Finding Relevant Data (Determining Market Rates)

- Board mandates CEO be paid “competitively”
- HR meets with Directors and CEO to define relevant competitive market
- HR finds existing data (surveys) or gathers own through proxy data or custom survey
- Finding: Pay is somewhat below average of data sample
But Finding Based On Data Outside The Relevant Range

Comparing To “Competitors”  Y = you; T = them

<table>
<thead>
<tr>
<th>C</th>
<th>E</th>
<th>O</th>
<th>P</th>
<th>A</th>
<th>Y</th>
<th>S</th>
<th>A</th>
<th>L</th>
<th>E</th>
<th>V</th>
<th>O</th>
<th>L</th>
<th>U</th>
<th>M</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

**********Relevant Range**********

Reward $ystems, Inc.

“Helping Organizations Succeed Through People” 57
Is Compensation Competitive?

• Survey shows what was paid, but do you know:
  – What award was based on (criteria)?
  – How difficult performance standards were?
  – How your performance compares to competitors?
  – How actual payout compared to target and to maximum award (how plan is designed)?
  – How much variance has there been over time?
  – What your plan would have paid if your performance was the same as theirs?
Evaluating Turnover

• How do you measure turnover?
• What is the acceptable range of turnover?
• Is it the same for all occupations/roles?
• What type of turnover?
• What is the impact of the turnover?
• What can be done about it?
Evaluate Turnover

Is It Too High? What Are The Implications?

Total Turnover 29%

Internal Functional Dysfunctional
12%
8%
4%

External
17%

Voluntary
14%

Functional
4%

Dysfunctional
10%

Involuntary 3%

Avoidable (Dysfunctional) 8%

Unavoidable 2%
Is This An Objective Analysis?
Not Totally

• Is internal turnover good or a problem?
  – Someone has to make an assessment
• Is turnover to the outside “voluntary”?
  – Managers can make someone want to leave
• Is turnover to the outside “functional”?  
  – Was there a celebration when person left?
• Was turnover avoidable?
  – Could actions be taken earlier to avoid it?
Using Technology To Understand/Predict

• Knowing when critical talent is prone to leave is very valuable
  – Anyone using modeling to get at this?
  – How about person to person methods?

• Evaluating patterns of lost talent
  – Are there time-related patterns?
  – Are there occupational patterns?
  – Are there geographical patterns?
How Important Will Analytics/AI/ Big Data Be?

• AI systems have defeated Chess & Go masters
  – But there are rules and a finite set of moves
• Machine learning has been able to use both correlation and causation to project results
  • But what happens when you have a unique set of circumstances that have not existed before?
• We need to humbly accept these systems can do things humans cannot... or do things better
• But we need to have the wisdom to know when the human element is needed to create...
Bottom Line

• Data is often plentiful
• Data can support almost anything if it is manipulated or selection is poor
• Value of AI/Machine learning varies depending on the application
• Subjective opinions can be valid and provide unique insights
• We are not going to be replaced by machines!
And A Lot Depends On The Quality Of Your Analysts

Don’t be afraid to step out of the box and to use wizards!

Reward $ystems, Inc.
Must Haves In Your Library

• “A Short Introduction To Strategic HRM” – Cascio & Boudreau
• “Becoming The Evidence-Based Manager” – Latham
• “Beyond HR” – Boudreau & Ramstad
• “Applied Psychology In HRM” – Cascio & Aguinis