Jumping on the Analytics Bandwagon
How to Utilize the Data Within Your Organization’s Systems to Gain Valuable Insight

Overview
- Data vs Information
- Garbage In – Garbage Out
- Analytics to Intelligence

Background
Data vs Information

Data

Data is unprocessed facts and figures without any added interpretation or analysis.

- "The price of crude oil is $110 per barrel."

Company Data Examples
- Journal entries
- Customer lists
- Vendor lists
- Invoices
- Expenses
- Employee listings

Information

Information is data that has been interpreted so that it has meaning for the user.

- "The price of crude oil has risen from $100 to $115 per barrel." gives meaning to the data.

Company Information
- Financial statements
- A/R Aging
- A/P Aging
- Top Producing Wells
- Employee Retention

- Data is raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized.

- When data is processed, organized, structured or presented in a given context so as to make it useful, it is called information.
Computers, since they operate by logical processes, will unquestioningly process unintended, even nonsensical, input data ("garbage in") and produce undesired, often nonsensical, output ("garbage out").

- Data is not entered into the system at all.
GIGO

- Data is not entered correctly
- The purpose/usage of the field is not defined

Causation & Correlation

- **Correlation**: a relationship between two variables
- **Causation**: an act that occurs in such a way that something happens as a result

Correlation does not imply causation

Correlation Analysis

Analytics to Intelligence
Historical

- Descriptive Analytics: condense large amounts of data into smaller, more useful nuggets of information
- Diagnostic Analytics: determining if the correlation is causal

Forward Looking

- Predictive Analytics: probability of different outcomes
- Prescriptive Analytics: integrate tried-and-true predictive models into our repeatable processes to yield desired outcomes

Example - Input

- JADE Report
- Expense Type Mapping

Note: Names, dates, and amounts have been modified for presentation purposes.

Review the Data

- Review the data by skimming through it
- Performing various trial summarizations
- Identify anomalies or odd data

**PrePay Bids.** The amounts do negate out, however, the notation under the reversing entry is ‘update with actual invoice amount’, but the Transaction Date and Activity Date for both the entry and the reversing are the same dates.

<table>
<thead>
<tr>
<th>Description</th>
<th>PrePay</th>
<th>Real</th>
<th>Date</th>
<th>Activity Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>PrePay</td>
<td>Real</td>
<td>3/1/11</td>
<td>3/1/11</td>
</tr>
<tr>
<td>Item 2</td>
<td>PrePay</td>
<td>Real</td>
<td>3/2/11</td>
<td>3/2/11</td>
</tr>
<tr>
<td>Item 3</td>
<td>PrePay</td>
<td>Real</td>
<td>3/3/11</td>
<td>3/3/11</td>
</tr>
<tr>
<td>Item 5</td>
<td>PrePay</td>
<td>Real</td>
<td>3/5/11</td>
<td>3/5/11</td>
</tr>
</tbody>
</table>

These dates are examples and do not apply to the actual dataset.
Cost Per Property Analysis

1. Total Cost Per Property
   - Total cost per property helps make up accurate analysis of expenses.

<table>
<thead>
<tr>
<th>Category</th>
<th>Total Cost</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equipment</td>
<td>12,300,000</td>
<td>30%</td>
</tr>
<tr>
<td>Supplies</td>
<td>11,000,000</td>
<td>28%</td>
</tr>
<tr>
<td>Labor</td>
<td>10,000,000</td>
<td>25%</td>
</tr>
<tr>
<td>Total</td>
<td>33,300,000</td>
<td>100%</td>
</tr>
</tbody>
</table>

   Top 5 properties with the most expense:

<table>
<thead>
<tr>
<th>Property</th>
<th>Cost (Non-Overhead)</th>
<th>% Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Property A</td>
<td>12,300,000</td>
<td>37%</td>
</tr>
<tr>
<td>Property B</td>
<td>11,000,000</td>
<td>33%</td>
</tr>
<tr>
<td>Property C</td>
<td>8,000,000</td>
<td>24%</td>
</tr>
<tr>
<td>Property D</td>
<td>7,000,000</td>
<td>21%</td>
</tr>
<tr>
<td>Property E</td>
<td>5,000,000</td>
<td>15%</td>
</tr>
</tbody>
</table>

Cost Categories Analysis

3. Highest paid vendors - Overall risk:
   - 10% vendors total.
   - 60% of all payments went to the top 6 vendors.

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Cost Total</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vendor A</td>
<td>6,000,000</td>
<td>60%</td>
</tr>
<tr>
<td>Vendor B</td>
<td>1,000,000</td>
<td>10%</td>
</tr>
<tr>
<td>Vendor C</td>
<td>1,000,000</td>
<td>10%</td>
</tr>
<tr>
<td>Vendor D</td>
<td>1,000,000</td>
<td>10%</td>
</tr>
<tr>
<td>Vendor E</td>
<td>1,000,000</td>
<td>10%</td>
</tr>
<tr>
<td>Vendor F</td>
<td>1,000,000</td>
<td>10%</td>
</tr>
</tbody>
</table>

Different Views

Diagnostic Analysis

Duplicate Invoices
Tools

- Microsoft Excel
- Microsoft Access Database
- ACL and IDEA

Questions?

Reema Parappilly, CISA
Senior Manager, IT Advisory
reema.p@weaver.com
832.320.3254