On Generating High InfoQ with Bayesian Networks

Ron S. Kenett
Research Professor, University of Turin
Chairman and CEO, KPA Ltd.

www.kpa-group.com  ron@kpa-group.com

http://www.rss.org.uk/site/cms/contentEventViewEvent.asp?chapter=9&e=1543
Knowledge

Goals

Examples from:
- Customer surveys
- Risk management of telecom systems
- Monitoring of bioreactors
- Managing healthcare of diabetic patients

Primary Data
- Experimental
- Observational

Secondary Data
- Experimental
- Observational
Bayesian Networks

Learning

Estimating

Programmer’s nightmare:
1. “If the grass is wet, then it rained”
2. “If the sprinkler is on, the grass will get wet”
   Output: “If the sprinkler is on, then it rained”
Customer Surveys

Behaviors

Loyalty

Attitudes & Perceptions

Experiences & Interactions
Goal 1. Decide where to launch improvement initiatives
Goal 2. Highlight drivers of overall satisfaction
Goal 3. Detect positive or negative trends in customer satisfaction
Goal 4. Identify best practices by comparing products or marketing channels
Goal 5. Determine strengths and weaknesses
Goal 6. Set up improvement goals
Goal 7. Design a balanced scorecard with customer inputs
Goal 8. Communicate the results using graphics
Goal 9. Assess the reliability of the questionnaire
Goal 10. Improve the questionnaire for future use

\[
InfoQ(f, X, g) = U(f(X|g))
\]
The ABC 2010 Annual Customer Satisfaction Survey

Company: __________________________
Completed by: __________________________


Dear Customer,

For each of the following statements, please select a number indicating the extent of your agreement with the statement concerning your experience with ABC during 2010. Then, under “Importance Level”, select another number indicating the importance of the statement to you. If a certain statement is not relevant or not applicable, please select N/A.

---

### Equipment and System

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Strongly agree</th>
<th>Importance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

6. The equipment’s features and capabilities meet your needs.
7. Improvements and upgrades provide value.
8. Output quality meets or exceeds expectations.
9. Upkeep is acceptable.
10. For customers who purchased a system during 2010: ABC’s equipment proposal met your requirements.

11. Overall satisfaction level from the equipment:

<table>
<thead>
<tr>
<th>Very low</th>
<th>Very high</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

### Sales Support

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Strongly agree</th>
<th>Importance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

12. Verbal promises have been honored.
13. Sales personnel communicate frequently enough with you.
14. Sales personnel respond promptly to requests.
15. Sales personnel are knowledgeable about equipment.
16. Sales personnel are knowledgeable about market opportunities.

17. Overall satisfaction level from sales support:

<table>
<thead>
<tr>
<th>Very low</th>
<th>Very high</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

### Technical Support

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Strongly agree</th>
<th>Importance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

18. Technical support is available when needed.
19. The technical staff is knowledgeable.
20. The technical staff is well informed about the latest equipment updates/enhancements.
21. Parts are available when needed.
22. The remote support care center is valuable and meets your expectations.
23. Problems are resolved within the required time frame.
24. The technical staff is courteous and helpful.

25. Overall satisfaction level from technical support:

<table>
<thead>
<tr>
<th>Very low</th>
<th>Very high</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Bayesian Network Analysis of Customer Surveys

20% BOT12
39% BOT12
## Information Quality (InfoQ) of Integrated Analysis

<table>
<thead>
<tr>
<th>Models</th>
<th>f₁</th>
<th>f₂</th>
<th>f₃</th>
<th>f₄</th>
<th>....</th>
<th>N_f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals</td>
<td>g₁</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>4</td>
</tr>
<tr>
<td>g₂</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>g₃</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>g₄</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

**InfoQ(f,X,g) = u(f(X|g))**

| Ng | 1  | 2  | 3  | 3  |      |

<table>
<thead>
<tr>
<th>Goal</th>
<th>BN</th>
<th>CUB</th>
<th>Rasch</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 <strong>Decide</strong> where to launch improvement initiatives</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2 <strong>Highlight</strong> drivers of overall satisfaction</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>3 <strong>Detect</strong> positive or negative trends in customer satisfaction</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4 <strong>Identify</strong> best practices by comparing products or marketing channels</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>5 <strong>Determine</strong> strengths and weaknesses</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>6 <strong>Set up</strong> improvement goals</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>7 <strong>Design</strong> a balanced scorecard with customer inputs</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>8 <strong>Communicate</strong> the results using graphics</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>9 <strong>Assess</strong> the reliability of the questionnaire</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>10 <strong>Improve</strong> the questionnaire for future use</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Goal 1: Identify causes of risks that materialized

Goal 2: Design risk mitigation strategies

Goal 3: Provide a risk management dashboard
Bayesian Network of communication network data
Support \( (A \rightarrow B) \)

Confidence \( (A \rightarrow B) \) = \( \frac{\text{Support}(A \rightarrow B)}{\text{Support}(B)} \)

Lift \( (A \rightarrow B) \) = Confidence \( (A \rightarrow B) \) / Support \( (B) \)

HyperLift = more robust version of Lift

Data structure and data integration
Communication and construct operationalization

Probability of risks by types and severity

Analysis for customer PBX=92960 (Finance), based on 38 events.

Average Stats for same Business Line (Finance)

**Monitoring of bioreactor**


**Managing diabetic patients**
Control of bioreactor over time
Treatment of diabetic patient

Decisions
- Diet
- Dose
- Utility

Quality
Risk
Cost
## Decision Table

### Ca-based Therapy
- Treat hypercalcemia;
- Continue current therapy;
- Decrease vitamin D dose to achieve ideal Ca; Decrease Ca-based ...
- Decrease or discontinue vitamin D dose or Ca-based phosphate binders; ...

### Phosphate binder?
- Assess nutrition, discontinue phosphate binder if ...
- Being dietary counseling and restrict dietary phosphate; start or increase ...
- Being short-term AI-based phosphate binder use, then increase ...

### QB
- Increase QB;
- Keep QB;
- Decrease QB.

### Erythropoietin
- Keep the current dose;

### Iron Management
- Keep the treatment;

### Time
- ...

### Diet?
- ...

### Risk

#### Hospitalization Ratio
- **Pt. 2**
  - 1.0E-3 0 - 1.5
  - 0.00 1.5 - 3
  - 35.54 3 - 4.5
  - 58.64 4.5 - 6
  - 5.46 6 - 7.5
  - 0.31 7.5 - 9

- **Pt. 5**
  - 1.0E-3 0 - 1.5
  - 0.00 1.5 - 3
  - 28.86 3 - 4.5
  - 63.74 4.5 - 6
  - 9.94 6 - 7.5
  - 0.35 7.5 - 9

#### Mortality Ratio
- **Pt. 2**
  - 4.82 0 - 3
  - 16.19 3 - 6
  - 21.02 6 - 9
  - 5.43 9 - 12
  - 3.95 12 - 15
  - 2.30 16 - 21
  - 0.23 21 - 24

- **Pt. 5**
  - 9.71 0 - 3
  - 30.65 3 - 6
  - 19.54 6 - 9
  - 14.02 9 - 12
  - 11.59 12 - 15
  - 8.91 15 - 18
  - 4.10 18 - 21
  - 0.87 21 - 24

#### Mortality risk relative (Ca)
- **Pt. 2**
  - 1.40 0 - 1
  - 98.2 1 - 1.5
  - 0.30 1.5 - 2
  - 0.10 2 - 2.5
  - 0.00 2.5 - 3
  - 0.00 3 - 3.5
  - 0.00 3.5 - 8

- **Pt. 5**
  - 4.91 0 - 1
  - 99.6 1 - 1.5
  - 1.9 1.5 - 2
  - 0.1 2 - 2.5
  - 0.0 2.5 - 3
  - 0.0 3 - 3.5
  - 0.0 3.5 - 8

#### Mortality risk relative (B)
- **Pt. 2**
  - 18.00 0 - 1
  - 78.3 1 - 1.5
  - 0.59 1.5 - 2
  - 0.20 2 - 2.5

- **Pt. 5**
  - 18.50 0 - 1
  - 78.3 1 - 1.5
  - 0.59 1.5 - 2

#### Mortality risk relative (Diet)
- **Pt. 2**
  - 0.00 0 - 1
  - 42.14 1 - 1.5
  - 33.60 1.5 - 2
  - 9.24 2 - 2.5
  - 4.72 2.5 - 3
  - 3.61 3 - 3.5

- **Pt. 5**
  - 0.00 0 - 1
  - 42.14 1 - 1.5
  - 33.60 1.5 - 2
  - 9.24 2 - 2.5
  - 4.72 2.5 - 3
  - 3.61 3 - 3.5

#### Dose of erythropoietin
- **Pt. 2**
  - 12.63 0 - 1000
  - 44.93 1000 - 6000
  - 29.12 6000 - 11000
  - 9.25 11000 - 16000
  - 3.95 16000 - 21000
  - 0.09 21000 - 26000
  - 0.02 26000 - 31000
  - 0.01 31000 - 36000
  - 0.00 36000 - 41000

- **Pt. 5**
  - 2.97 0 - 1000
  - 10.57 1000 - 6000
  - 83.32 6000 - 11000
  - 2.18 11000 - 16000
  - 0.93 16000 - 21000
  - 0.02 21000 - 26000
  - 0.00 26000 - 31000
  - 0.00 31000 - 36000
  - 4.69E-4 36000 - 41000
1. Data resolution
2. Data structure
3. Data integration
4. Temporal relevance
5. Chronology of data and goal
6. Generalizability
7. Construct operationalization
8. Communication