Visualization for Modeling and Presenting Results from Survey Data

William S. Cleveland, Lorraine Denby, Chuanhai Liu
Bell Labs, Lucent Technologies
600 Mountain Avenue, Murray Hill, New Jersey 07974 USA
wsc@lucent.com, ld@lucent.com, liu@lucent.com

1. Introduction

Today's businesses are often driven by surveys. Competitive customer satisfaction surveys provide feedback from customers of all the companies playing in a particular industry on the primary process areas of the business. Employee satisfaction surveys enable management to obtain viewpoints of the company employees. Consumer surveys obtain opinions of the marketplace. Important business decisions are based on the results gleaned from these surveys.

The biggest single problem in the analysis and presentation of such data is the failure to come to terms with the complexity of the data. Results are typically univariate looks on a question by question basis. Occasionally, results are bivariate. The full multivariate complexity is seldom attacked. Part of the reason is a lack of tools to do so. In this paper we will describe some of the common problems and discuss solutions and tools. Visualization plays an important role.

2. Common Analysis and Presentation Pitfalls

1) Univariate or bivariate analysis and presentation of results: Often the survey responses are summarized as the mean response or the % in agreement for each question. This practice ignores the structure between the questions. Tables and bargraphs do not help one find patterns between the questions or structure related to the characteristics of the respondents.

2) No indication of variability: Reports for upper management, along with those generated for the analysts, most often contain no indication of the variability of the estimates. Sometimes the sample size might be listed as the only hint of precision of the estimate.

3) Questions need to be benchmarked on the scale: It is easiest to explain this point with an example. Suppose question responses are on a scale of 1 to 10 (dissatisfied to satisfied). A response of 7 on satisfaction with the price of a product might indicate about as much satisfaction as one can expect but a response of 7 on satisfaction with product quality might be unacceptable.

4) Missing data: Skip patterns, fatigue, sensitive questions, poorly phrased questions all cause incomplete responses.

5) Rater effects: Individuals use scales differently. Raters position themselves on different points of the rating scale and use up different amounts of the scale, creating both random location and scale effects.

6) Verbatims: Verbatims are often simply listed in the reports word-for-word as they were recorded. Little to no analysis or summarization is attempted. The verbatim is rarely related to the satisfaction score.

3. Solutions

Bayesian hierarchical models were used for handling the missing data and for fitting the random location and scale effects (Cleveland, Denby and Liu 1998). The model described there is

$$r_{uj} = \sum_{k=1}^{a} \alpha_{(k)} v_{(k)uj} + \sum_{k=1}^{b} \beta_{(k)u} w_{(k)uj} + \gamma_u \epsilon_{uj}. \qquad (1)$$

The $\alpha = (\alpha_{(1)}, \ldots, \alpha_{(b)})$ are fixed-effect parameters. The vectors of random-effects location parameters $\beta_u = (\beta_{(1)u}, \ldots, \beta_{(b)u})$ are i.i.d. variables with multivariate distribution F_β. The random-effects scale parameters γ_u are i.i.d. variables with univariate distribution F_γ. The ϵ_{uj} are independent normal error terms with mean zero. Further the three collections of random variables — β_u, γ_u, and ϵ_{uj} — are mutually independent.

There are several restrictions that we add to make (1) well defined, or identifiable. The first, already stated, is that $E(\epsilon_{uj}) = 0$. Second, we suppose that the random-effects explanatory variables are contained in the space spanned by the fixed-effect explanatory variables and take $E(\beta_{(k)u}) = 0$.

The main technology for determining the location-effect distributions and the scale-effect distributions is the unit regressions described in (Cleveland, Denby and Liu 1998). First the dependent variable is corrected for the fixed effects and any variances dependent on the fixed effects. The residuals from the fixed effects model are used to obtain estimates of the random location effects, $\hat{\beta}_u$. Correcting for the random location effects we can obtain estimates of random scale effects, standardized residuals and studentized residuals. The distributions of these values are deconvolved to obtain an understanding of the random location and scale distributions.

Trellis displays (Becker, Cleveland, Shyu 1996) have proven useful in addressing both modeling identification and checking issues and the presentation problems listed in Section 2. Trellis is a framework for the visualization of multivariate data. The trellis display lays out the data in rows, columns and pages of panels. Each panel shows a subset of the data graphed with a one or two dimensional display technique, such as barplot, scatterplot, line plot, etc. Each panel shows the relationship of the plotted variables conditional on the values of the variables noted in the stip label at the top of each panel. The order of the panels is important and is chosen depending on the context of the display.

Trellis displays are essential in the model building phase to identify the location-effect distributions and the scale-effect distributions and in checking the model. By its very nature, survey data is not univariate or bivariate, but multivariate. So, displaying the results falls naturally into this framework. In the next section we will concentrate primarily on the presentation aspect of survey results.

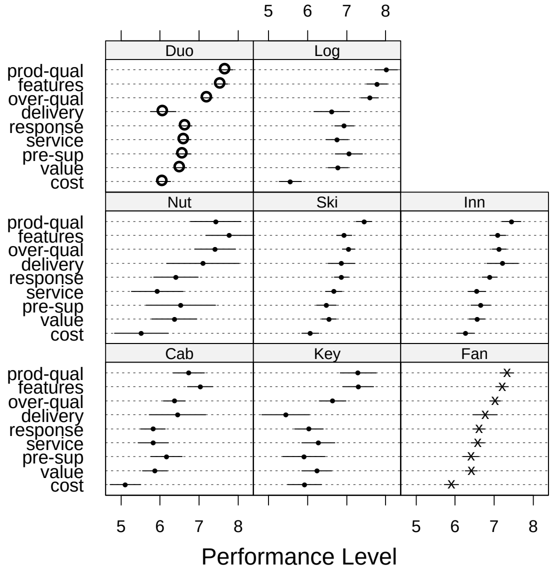
4. Some Examples

In this section we describe trellis displays that proved useful in the presentation of customer satisfaction and employee satisfaction survey results. These displays show not only the point estimates, but also the variability. Benchmarking of the questions and companies or employee group is clearly evident. The results are organized to make comparisons straightforward. The last display shows binned verbatims conditional on the satisfaction score and market, thus, aiding in the interpretation.

For the customer satisfaction survey the respondents rate one or more suppliers on 9 areas of the business process. The responses are on a 10 point scale (1=dissatisfied and 10=satisfied). There is a skip pattern where only purchasing managers rated `delivery` and only designers rated `pre-sup`. Missing data also occurred at random. Bayesian hierarchical models were fit to this data (Clark, Cleveland, Denby and Liu 1998) and the trellis displays show the posterior means and confidence interval based on the posterior variance.

Figure 1 shows the competitive profiling display (Cleveland, Denby and Liu 1999). Each panel shows the estimated mean values across the questions for one company. The panels are ordered by the company mean across the questions. `Cab` has the lowest average and is placed in the bottom left corner, then `Key`, `Fan`, etc. to `Log` in the top right corner with the highest average. The questions are ordered bottom to top in the panels according to their average over the companies. We are particularly interested in `Fan` and `Duo` so their estimates are shown as an "X" and an open circle (color is useful here also) for easy identification.

The consistent southwest to northeast pattern in each panel means that there is an industry wide ordering of the process areas. For most companies the largest and second largest are `prod-qual`



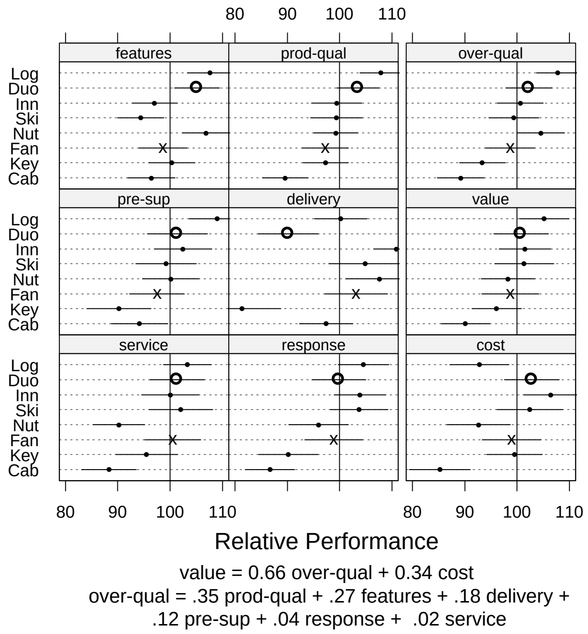


Figure 1. Question Levels by Company. ***Figure 2. Company Ratios by Question.***

and `features`, followed by `over-qual`. The smallest is usually `cost` and the second smallest or close to that position is `value`. For this industry attributes of physical product (`features` and `prod-qual`) are rated high. In the middle are the interactions of the customer with the company personnel (`delivery`, `response`, `service`, and `pre-sup`). A mixture of the attributes in these two groups is represented by `over-qual` which falls close to `prod-qual` so it seems to be linked closest to this attribute in the minds of the customers. At the bottom are `cost` and `value` which are associated with the pricing of the product and services and the associated value to the customer. The attribute that breaks the ordering most often is `delivery`, but it has the smallest sample size since it was one of the attributes that was skipped by many of the respondents.

Obviously if a company has a rating of 6.5 for `cost` it would be at the top of the industry but a rating of 6.5 for `prod-qual` is at the bottom of the industry. This points to the need for benchmarking. Figure 2 shows the competitive profile display for a company's performance relative to its competitors. Here the ordering of the panel shows the inherent relationship of the questions. The right column contains `value` which is modeled as a linear combination of `over-qual` and `cost`, it is placed in the middle of the row with `over-qual` above and `cost` below. The other six attributes are ordered from bottom left corner to top right in the left and middle columns in order of their impact on `over-qual` since they are known to be drivers of `over-qual`. The weights of each attribute in the `over-qual` equation and of `over-qual` and `cost` in the `value` equation are noted at the bottom of the plot. The confidence intervals help in determining which companies are statistically significantly above or below par for a question of interest. Thus, these confidence intervals can help management avoid making decision based on spurious variation.

The next example is from an employee satisfaction survey. In Figure 3 a panel shows the percent answering favorably about one department versus the other departments within a vice presidential area for a particular question. The vertical lines show two benchmarks: the corporate value and a value for other similar companies. The questions are ordered by the outside norm. This display clearly shows that the department of interest is significantly worse than the norm for `competitive position` and nearly so for `customer relationships`. It is above the norm for all other questions. In addition, this department is significantly better than the other departments in this vice presidential area on many of these questions (those where the two confidence intervals in a panel do not overlap).

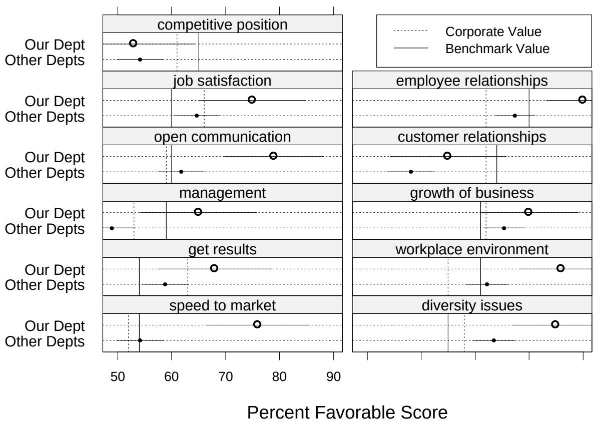


Figure 3. Results by Question.

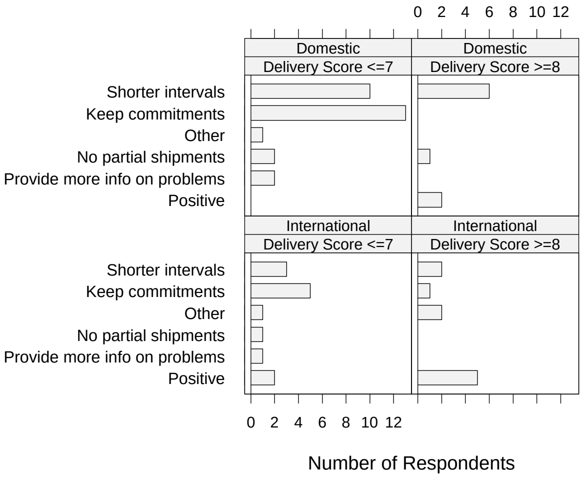


Figure 4. Binned Verbatims by Market and Delivery Score.

The last example shows an analysis of verbatims from the customer satisfaction survey. Sometimes the verbatims are binned, but rarely are they related to the satisfaction score of the related process area. Figure 4 shows Pareto charts of the verbatims conditional on the `delivery` satisfaction score and the market. Comparing the domestic charts for those unsatisfied (score ≤ 7) versus those satisfied (score ≥ 8) it is clear that keeping commitments is very important in whether the customer is satisfied or not. A similar relationship is seen in the international charts. Shortening the delivery interval is requested but does not relate to satisfaction.

5. Summary

The displays and analysis techniques discussed here correct many of the pitfalls that are currently seen in the reporting and analysis of customer and employee satisfaction results. An important feature is the clear display of variability.

REFERENCES

Many of these papers are available on web site: http://cm.bell-labs.com/stat/doc/

Clark, Linda A., William S. Cleveland, Lorraine Denby, and Chuanhai Liu (1998), "Modeling Customer Survey Data," in *Case Studies in Bayesian Statistics IV*, C. Gatsonis et al., eds., pp. 3-57. New York: Springer.

Becker, Richard A., William S. Cleveland, and Ming J. Shyu (1996), "The Design and Control of Trellis Display," *Journal of Computational and Statistical Graphics*, 5, 123–155.

Cleveland, William S., Lorraine Denby, and Chuanhai Liu (1998), "Random Scale Models."

Cleveland, William S., Lorraine Denby, and Chuanhai Liu (1998), "Competitive Profiling Displays: Multivariate Graphs for Customer Satisfaction Data," *Marketing Research*, Spring 1999 issue.

RÉSUMÉ

Le principal et unique probleme en analyse et représentation de données complexes est la difficulté de pouvoir conclure. Souvent, les compte-rendus sont de simples catalogues de tableaux et d'histogrammes. Aucune variabilité n'est représentée. Les données manquantes sont ignorées. Les effets duent a une modification des limites de l'échelle des valeurs et de son interprétation par l'individu sont masquées. La structure complete des variations est rarement prise en compte. Dans cet article, on abordera quelques problemes standarts et nous proposerons des solutions et des outils. La visualisation tiendra un rôle important.