Comparison of Quantitative and Semi-Qualitative Methods in Job Satisfaction Analysis

UCSF University of California San Francisco

Matthew Jura, MS[1,2] ▪ Joanne Spetz, PhD[1]

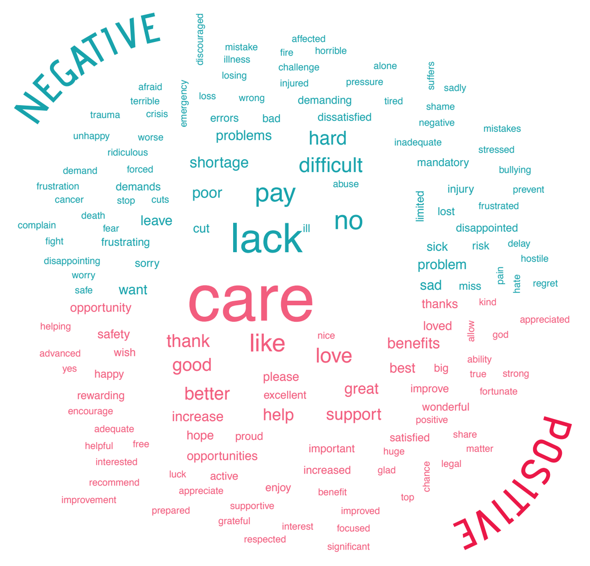
[1]Philip R. Lee Institute for Health Policy Studies, School of Medicine, UCSF; [2]Institute for Biomedical Informatics, National Yang-Ming University, Taiwan

RESEARCH OBJECTIVE

The main concern of healthcare leaders is achieving high work satisfaction and minimizing burnout among health service providers [1-2]. This research examines work satisfaction among registered nurses (RNs) in California using both quantitative and qualitative data. Sentiment analysis (SA) is used to evaluate affective state of textual information, and the results are compared with the satisfaction data obtained from a structured questionnaire.

STUDY DESIGN

- The data was obtained from the California Board of Registered Nursing biennial surveys (2006-2016)[3], which were sent to randomly-selected samples of RNs (response >50% per year).
- The instrument includes questions regarding employment, education, and demographics. It also includes 29 Likert-scale questions related to job satisfaction, including work organization, environment, professional development, and relationships with patients, other staff, and management.
- Each survey also contains an open-ended comment section; submitted comments are usually related to current employment experiences or reflect general attitudes towards the nursing profession.
- The job satisfaction questions were analyzed to identify groups of the respondents that exhibit similar opinions or satisfaction levels. We used hierarchical and k-means clustering algorithms (described in detail in the top right box).
- The open-ended comments were analyzed using sentiment analysis. This method helps to systematically obtain the emotional state of a subjective expression (described in the bottom right box).
- We evaluated the association of the sentiment score with the satisfaction groups (clusters) with linear regression model adjusted for demographic variables.



A cloud of the most frequent words observed in the comments, divided into positive and negative categories.

ANY QUESTIONS OR COMMENTS?

Feel free to leave us a comment or download the poster.

matthew.jura@ucsf.edu

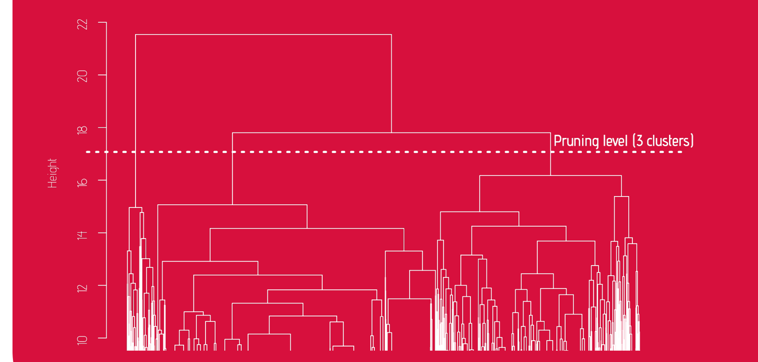
OPTIONAL READING

QUANTITATIVE ANALYSIS

Both hierarchical and k-means clustering are techniques which help to find similarities between observations across many features (variables). Both are "unsupervised" learning algorithms [4] – there is no reference value to compare the results and evaluate performance thereof. The number of the clusters is arbitrary. There is no one, correct answer to clustering.

HIERARCHICAL CLUSTERING

- Objective – find clusters based on the distance between the answers (given on 5 level Likert-scale).
- Algorithm – create a hierarchy:
  - Assign each item to its own cluster:
  - If we have N items we start with N clusters (bottom of the exhibit below).
  - Find the closest pair of clusters and merge them into a single cluster.
  - Calculate the distances (similarities) between the new cluster and each of the old clusters.
  - Repeat the previous steps until all items are clustered into a single cluster of size N (top of the tree).
- To find an arbitrary number of groups (clusters) we "prune" the tree at a set height.



K-MEANS

K-means is a centroid-based clustering technique. It is an iterative process of finding similarities between the observations based on the notion how the data points are close to the central value (in this case randomly chosen).

The algorithm minimizes total within-cluster variation. The within-cluster variation is defined as a sum of squared Euclidean distances between the items and the centroid of the cluster k.

$$D(C_k) = \sum_{x_i \in C_k} (x_i - \mu_k)^2$$

x_i is a data point belonging to the cluster C_k; μ_k is the mean value of the points belonging to the cluster.

- We specify a number of clusters (k=3).
- Set randomly the location of cluster centroids.
- Assign each observation to the closest centroid.
- For each of the k clusters recalculate the center value - mean - of all data points in the assigned cluster.
- Repeat the steps and minimize the distance between the observations and the centroid until the centroids do not change or the maximum number of iterations is reached.
- The total within-cluster variation is the total sum of squares:

$$\sum_{k=1}^{k} D(C_k) = \sum_{k=1}^{k} \sum_{x_i \in C_k} (x_i - \mu_k)^2$$

OPTIONAL READING

SEMI-QUALITATIVE ANALYSIS

SENTIMENT ANALYSIS [5]

- We use sentiment analysis (SA) algorithms for the semi-qualitative examination of the comments.
- Sentiment analysis helps to systematically obtain the emotional state of a subjective expression.
- Each comment is divided into words, which have assigned an emotion score (in this case a numerical value) based on the provided lexicon (AFINN) [6].
- Examples of positive or negative words are on the left-hand side of the box.
- At the end, the entire comment is given an overall score – the sum or the average of the word scores.
- SA fails to detect sarcasm or context of textual data.

ANALYSIS STEPS:

- Obtain the comments, split into words, remove punctuation and all unnecessary content, such as e-mail addresses, URIs, dates, numbers.
- Match words from comments with AFINN lexicon and obtain sentiment score.
- For AFINN, scores range from -5 to +5 (most negative emotion to the most positive emotion).
- Each word has assigned identification number matching it with a specific comment.
- For each comment we calculate summary statistics: the average of the score, the sum, or the % of negative words within the comment.
- Not all words exist in the AFINN lexicon. We can calculate the SA score based only on the "emotionally charged" content of the comment.
- We ignore the comments that are not more than 5 words long.
- In the analysis comparing SA with satisfaction clusters, we limit the data set to only those who answered questions about job-satisfaction and provided a comment.

All analyses were done in CRAN R [7].

KEY RESULTS

SATISFACTION GROUPS (CLUSTERS)

- We analyzed 24,543 cases – actively employed RNs who answered both profession satisfaction question and job satisfaction questionnaire.
- Using the clustering techniques we identified three major groups with the incremental level of satisfaction – low satisfaction (1), medium (2), high (3).

Table 1. Average Likert scores (1-5) for job satisfaction and profession satisfaction questions in the two clustering techniques.

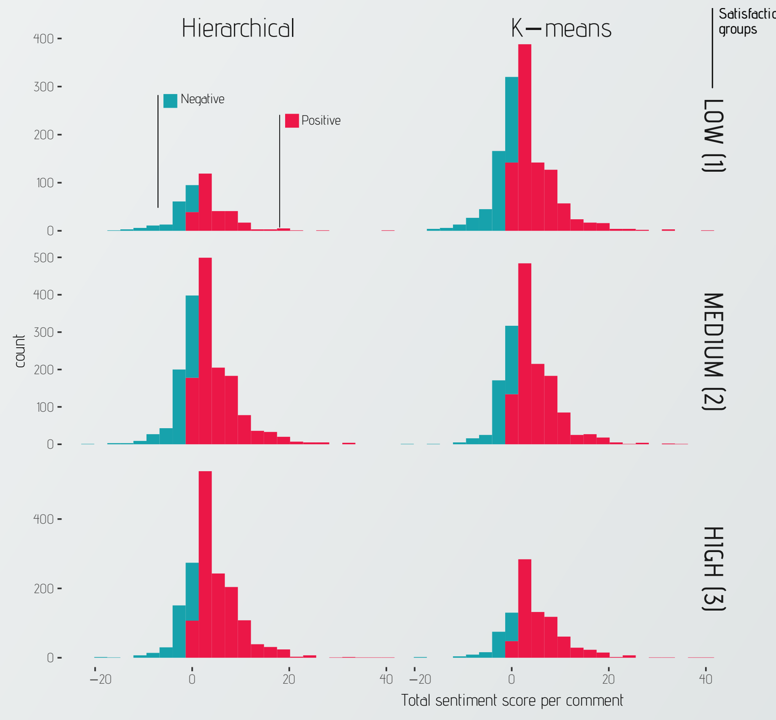
	Hierarchical		K-means	
Satisfaction Group	Job	Profession	Job	Profession
LOW (1)	2.7	3.1	3.3	3.5
MEDIUM (2)	3.8	3.9	4.2	4.1
HIGH (3)	4.5	4.3	4.8	4.5

The following figures show the distribution of the average Likert scores for individual satisfaction questions. They are arranged in ***five*** thematic concentrations and sorted by a difference between the scores in low satisfaction group (1) and high satisfaction group (3).

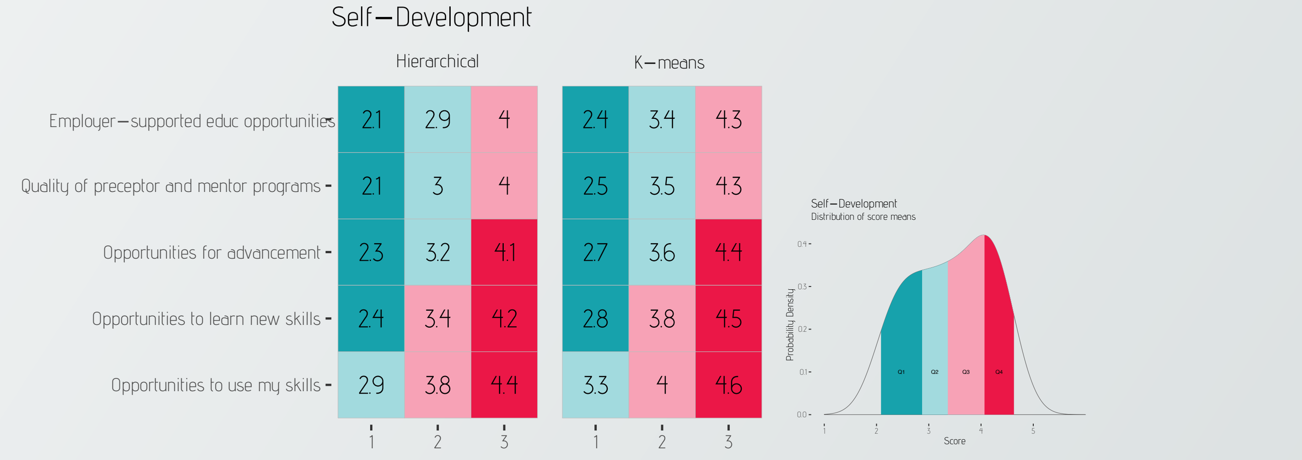
The most pronounced differences in opinions are related to Management – the low satisfaction group gives the lowest observed scores among all observed values; where the high satisfaction group on average expressed their high satisfaction falling into 4th quartile of score distribution. In addition, we observed the most diverging opinions in Work Environment in relation to adequacy of RN staff, Patient Care – involvement in patient care decisions, Self-Development - opportunities for advancement and learning new skills, and in Workplace Organization - recognition and meaningful work.

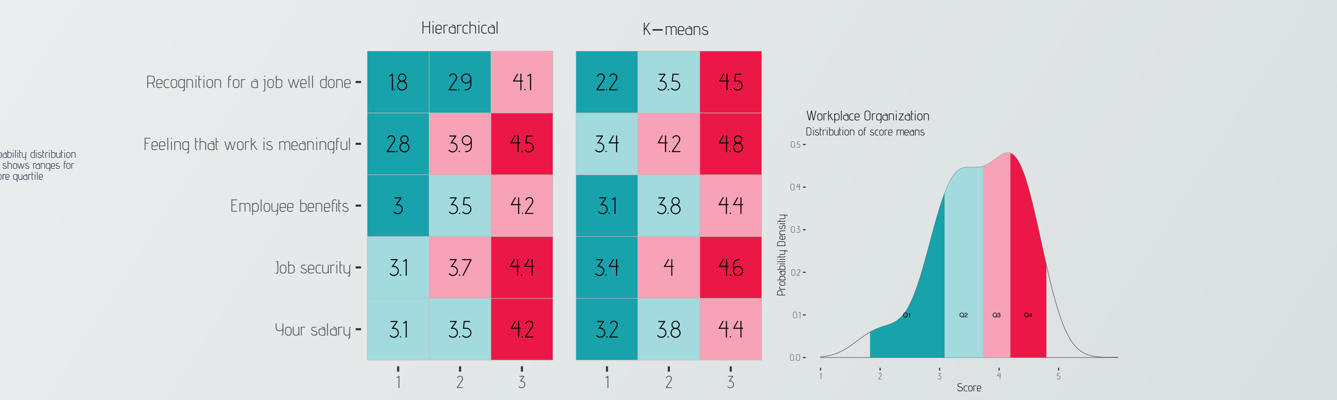
SENTIMENT ANALYSIS OF THE SURVEY COMMENTS

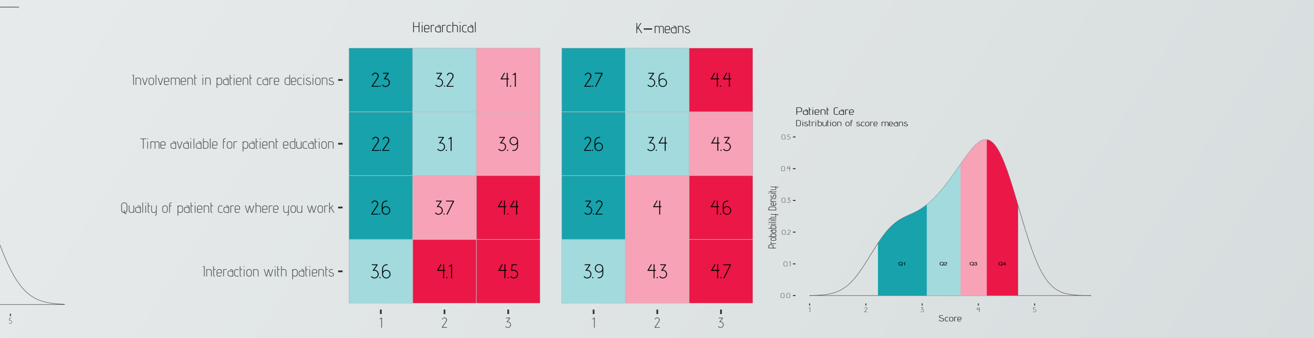
- Total: 5,650 comments analyzed.
- Average comment length is 60-70 words, longest reached more than 700 words.
- 4,117 found with overall positive, 1,533 with overall negative emotion;
- No substantial difference in average word count between positive and negative comments.
- Comments with lower emotion score occur more frequently in the lowest satisfaction groups.
- The overall positive comments are frequent in all satisfaction groups, however, from the distribution it is clear that scores are mainly driven by the presence of the negative comments.



Distribution of sentiment scores stratified by different clustering techniques (Hierarchical vs. K-means) and satisfaction clusters (1-LOW, 2- MEDIUM, 3 - HIGH).







SENTIMENT SCORES IN SATISFACTION GROUPS

Table 2. Adjusted linear regression models – measure the association between the satisfaction groups, scores and comments' sentiment score.

	Average sentiment score			
	(1)	(2)	(3)	(4)
AGE (YEARS)	0.001 (-0.003, 0.005)	-0.001 (-0.004, 0.003)	0.001 (-0.003, 0.005)	-0.0002 (-0.004, 0.004)
MALE/FEMALE	-0.16** (-0.31, -0.02)	-0.13* (-0.28, 0.01)	-0.13* (-0.27, 0.02)	-0.14* (-0.29, 0.001)
$10K INCOME	0.02*** (0.01, 0.03)	0.02*** (0.005, 0.03)	0.01 (-0.003, 0.02)	0.02*** (0.01, 0.03)
NON-WHITE/WHITE	0.19*** (0.10, 0.28)	0.21*** (0.12, 0.30)	0.18*** (0.09, 0.27)	0.22*** (0.13, 0.31)
MEDIUM SATISFACTION GROUP/LOW		0.36*** (0.26, 0.45)		
HIGH SATISFACTION GROUP/LOW		0.61*** (0.51, 0.72)		
PROFESSION *SATISFACTION* (1-5)			0.20*** (0.16, 0.24)	
JOB SATISFACTION (1-5)				0.19*** (0.15, 0.22)
CONSTANT	0.55*** (0.33, 0.77)	0.38*** (0.16, 0.60)	-0.13 (-0.38, 0.13)	-0.08 (-0.33, 0.18)
OBSERVATIONS	3,450	3,450	3,358	3,450

*$p<0.1$; **$p<0.05$; ***$p<0.01$

- Both increase of annual income by $10k and non-white race yield positive and statistically significant association with higher positive score of the questionnaire comments.
- Males, on average, left less positive comments in the questionnaire, after adjusting for all other coefficients.
- The coefficients for satisfaction groups as well as individual satisfaction questions yield similar positive associations (p<0.001). The higher satisfaction group (model 2) or job/profession satisfaction (models 3-4) the more positive sentiment score of the questionnaire comment.

CONCLUSIONS

- Nurses have divided opinions about their current work and nursing profession.
- Clustering techniques help to identify groups with similar expectations or opinions. This can enhance the decision-making process within organizations.
- Clustering analysis is not only limited to job satisfaction questions, the users may incorporate additional variables that may help to describe and separate groups (clusters).
- We observe association between the sentiment scores obtained from the comments and the satisfaction groups (clusters).
- Sentiment analysis allows for using qualitative data and efficiently extracting structured data points.
- SA helps to analyze the relationship between the textual feedback and other variables in the structured data in quantitative data analysis.
- SA may be used to screen large amounts of textual data as a preliminary analysis before selection for deeper and more labor consuming full qualitative evaluation.
- SA is useful in many other applications of opinion analysis (e.g., screening patient opinions).

REFERENCES:

1. Khalatbari, Javad, Shohreh Ghorbanshiroudi, and Mehdi Firouzbakhsh. "Correlation of Job Stress, Job Satisfaction, Job Motivation and Burnout and Feeling Stress." Procedia - Social and Behavioral Sciences 84 (2013): 860–63. https://doi.org/10.1016/j.sbspro.2013.06.662.
2. Stalpers, Dewi, Brigitte J M De Brouwer, Marian J Kaljouw, and Marieke J Schuurmans. "Associations between Characteristics of the Nurse Work Environment and Five Nurse-Sensitive Patient Outcomes in Hospitals: A Systematic Review of Literature." International Journal of Nursing Studies 52 (2015): 817–35. https://doi.org/10.1016/j.ijnurstu.2015.01.005.
3. Spetz, Joanne, Dennis Keane, Carolina Herrera, Lisel Blash, Lela Chu, Zachary Levin, Ulrike Muench, Matthew Jura, and Jacqueline Miller. "California Board of Registered Nursing: 2006, 2008, 2010, 2012, 2014, 2016 Surveys of Registered Nurses," California Board of Nursing
4. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning : With Applications in R. New York : Springer, [2013] ©2013, 2013. https://search.library.wisc.edu/catalog/9910207152902121.
5. Qazi, Atika, Ram Gopal Raj, Glenn Hardaker, and Craig Standing. "A Systematic Literature Review on Opinion Types and Sentiment Analysis Techniques." Internet Research 27, no. 3 (April 19, 2017): 608–30. https://doi.org/10.1108/IntR-04-2016-0086.
6. Nielsen, F. Å. "AFINN," March 2011. http://localhost/pubdb/p.php?6010.
7. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2018. http://www.R-project.org/.