Analysis of Juror Questionnaires and the Assignment of Juror Ratings using Machine Learning

David M. Caditz, Ph.D.

January 24, 2023

1 Introduction

Courtroom jury selection describes a process by which a jury is selected and impaneled for a trial. Most venues in the United States allow for participation of the litigating parties by voir dire, moves to strike for cause and the exercise of peremptory challenges. The outcome of a jury trial may be influenced, and even determined, by the composition of the jury and litigants typically invest significant time and effort in the jury selection process.

Jury selection typically begins with the venire members submitting responses to a court-provided juror questionnaire, and, in some cases, a supplemental questionnaire containing questions provided by the parties. Studies show that juror questionnaires are viewed positively by judges, attorneys and jurors [1]. As such, jury selection statutes and procedures promote the use of questionnaires. The American Bar Association, in its report 'Principles for Juries and Jury Trials', recommends [2]:

> Before voir dire begins, the court and parties, through the use of appropriate questionnaires, should be provided with data pertinent to the eligibility of jurors and to matters ordinarily raised in voir dire.... All completed questionnaires should be provided to the parties in sufficient time before the start of voir dire to enable the parties to adequately review them before the start of that examination.

Numerous states have adopted such wording by statute or procedure [9, 10]. Despite these prodecures and recommendations, trial lawyers often consider evaluation of questionnaires, among other aspects of jury selection, to be constrained by court imposed timelines, resulting in a rushed and sometimes hectic analysis [3, 4, 5, 6].

Jury questionnaires comprise a few dozen up to, in some cases, several hundred questions [7]. At the same time, the venire may comprise up to several hundred members. The wealth of information available from juror questionnaires, if effectively utilized, can inform and guide courtroom-based voir dire,

dismissals for cause and peremptory challenges. However, the need to review and evaluate potentially thousands of question responses within a court-imposed timeline may pose significant challenges for attorneys and jury consultants.

Lawyers typically approach questionnaire evaluation in an ad hoc manner, going through the stack of completed questionnaires one-by-one, and assigning ratings to prospective jurors based on intuition or "gut feeling". This approach is a poor use of a lawyer's time, and it risks the introduction of bias and inconsistent rating assignments. This paper describes a robust and time-efficient analysis method known as *Cluster Analysis*, and its application to the evaluation of jury questionnaires. Cluster analysis is a common technique with wide application in fields of pattern recognition, anomaly detection, machine learning, and questionnaire evaluation. The method inverts the typical one-by-one rating process by first grouping venire members according to their similarities. Ratings are then assigned to members by group, either automatically or, if desired, manually, using the representative characteristics of each group. The grouping process can be completed by computer algorithm in a matter of seconds, resulting in a more consistent and time-efficient evaluation method.

Section 2 describes the data preparation needed before applying cluster analysis. This includes culling of irrelevant data, scoring the relevant data, and handling missing or ambiguous data. The goal is to produce a *Scoring Summary Sheet* that quantifies the relevant characteristics of the venire members. Section 3 provides an overview of the grouping method. The method applies a well-known data clustering algorithm to the data contained in the Scoring Summay Sheet. Ratings are then applied to the clusters based on their representative characteristic values. A simple example using two scores is provided.

The goal of this work is to provide sufficient information to attorneys and consultants, should they wish to engage with a data scientist or computer programmer to implement the proposed cluster analysis method.

2 Data Preparation

The application of cluster analysis to jury questionnaire data requires that the questionnaire responses be properly scored, irrelevant data are excluded, and missing data is properly handled. The following sections describe the process of data preparation.

2.1 Response Scoring

Each question on the the juror questionnaire can be categorized according to response type. Appendix A summarizes some common reponse types with example questions and responses. Ordinal and Interval responses are easily described using an ordered numeric scale, e.g., from 1 to 10. Nominal and Dichotomous responses, on the other hand, have no inherent order. The respondent is simply asked to select from among a fixed set of categories. Open ended responses allow the respondent full control over the content of their responses, possibly

with character, word or space limits. Examples of each question type are given in A.

Open ended responses can provide more detailed or nuanced insights into the views of the respondent, however, they also require more effort on the part of the evaluator to read and understand. One method for efficiently analyzing open ended questions is outlined in Appendix B.

In the case of jury selection, litigants are not impartial survey evaluators because they have preferences over the responses. Litigants therefore can rate responses on a numeric scale according to preference. This effectively turns every response into an ordinal response. For the analysis below, I assume that the each relevant response has been assigned a numeric 'preference score', with higher scores indicating preferable responses.

2.2 Data Culling

Some information collected by the jury questionnaire may be redundant and/or irrelevant to a litigant's juror preferences. For example, a litigant my judge that a juror's age, occupation, or whether they live in an apartment or house has no bearing on how they may decide a given case. The litigant is therefore indifferent to the response. In general, if a characteristic cannot be numerically scored according to preference, it is likely irrelevant, and should be removed from the analysis.

2.3 Missing Data

Gaps or missing data may prevent assigning scores to some venire members. This may occur if a venire member did not respond to a relevant question, or if their response is ambiguous. Missing scores may be substituted with the average score for that question over the remaining venire. For the subsequent analysis, it is assumed that all such gaps are filled using average values and each venire member is represented by a complete set of relevant scores.

2.4 Scoring Summary Sheet

Once the scoring is completed, it is convenient to create a *Scoring Summary Sheet*. A Scoring Summary Sheet summarizes the relevant numeric responses for the venire. Each column represents the scores for a given question and each row represents the responses of one venire member. A Scoring Summary Sheet can be constructed using a spreadsheet program, for example, Microsoft Excel. If the court provides questionnaire data in electronic format or in a scannable paper format, it should be simple to import the data. If only standard paper copies are provided, it is a straightforward task to transcribe the data into electronic format. In either case, responses should be coded according to their preference on a numeric scale, typically from 1 to 10 with higher scores indicating more preferable responses. Table 1 provides an example Scoring Summary Sheet

showing scores for the responses C_1 through C_N for each of the N relevant questions.

Juror ID	C_1 Employment	C_2 Education	C_3 Age	... ...	C_N
1	5	1	4	...	6
2	2	5	2	...	4
3	1	2	1	...	2
⋮	⋮	⋮	⋮	...	⋮
99	8	10	9	...	6
100	10	8	10	...	7

Table 1: Example Scoring Summary Sheet

3 Clustering Analysis and Juror Ratings

It may be tempting at this point to dive in, examine the Scoring Summary Sheet directly, row by row, and assign juror ratings based on some judgment of the scores. In fact, this is likely the most common approach used by attorneys and consultants. However, doing so would constitute a significant waste of time and effort, and would likely result in inconsistent ratings and the introduction of evaluator bias. Here I describe what I believe to be a superior method of assigning ratings based on well-known techniques in survey analysis and data science.

3.1 Venire Clustering

As mentioned above, a venire member's set of scores is represented by a row on the Scoring Summary Sheet, each row comprising N such scores. We wish to group venire members by the similarity of their scores. This can be done using a computer algorithm known as a clustering algorithm, for example the K-means algorithm [8]. In the following, I will use a simple case of $N = 2$ scores by considering only the first two questions, C_1 and C_2, in Table 1. However, the clustering procedure is valid for any number, N, of scores. Figure 1 shows a plot of the set of scores for venire of 100 members, based on the first two columns of the Scoring Summary Sheet in Table 1. Figure 1 is called a *Scoring Plot* for the example venire.

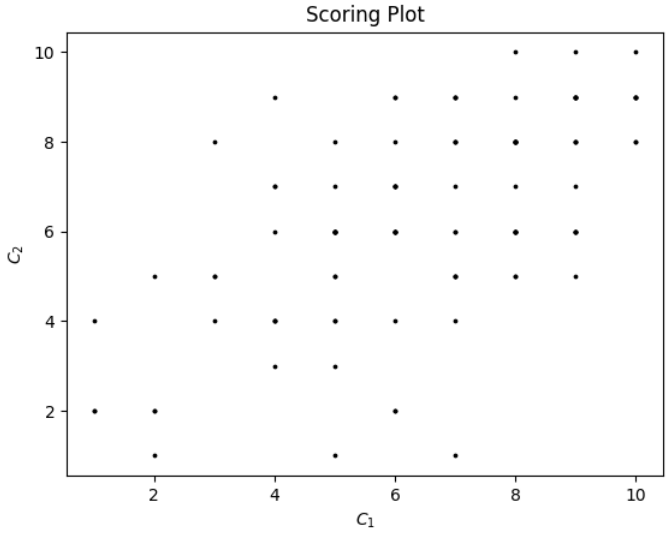

Figure 1: Scoring plot of two dimensional (two questions) juror score data for 100 venire members.

Applying the K-means algorithm to the data in Figure 1 results in a set of

groups or clusters shown in Figure 2. The algorithm partitions the data into 10 distinct clusters. In this figure, the clusters are color coded in order to show which points belong to which cluster. Clusters are described by their centroids, shown as crossed circles in Figure 2. The centroid represents the average of the scores for its cluster. Table 2 provides a tabular version of the K-means results. The first two columns in Table 2 show the centroid positions of the 10 clusters generated from the data shown in Figure 1.

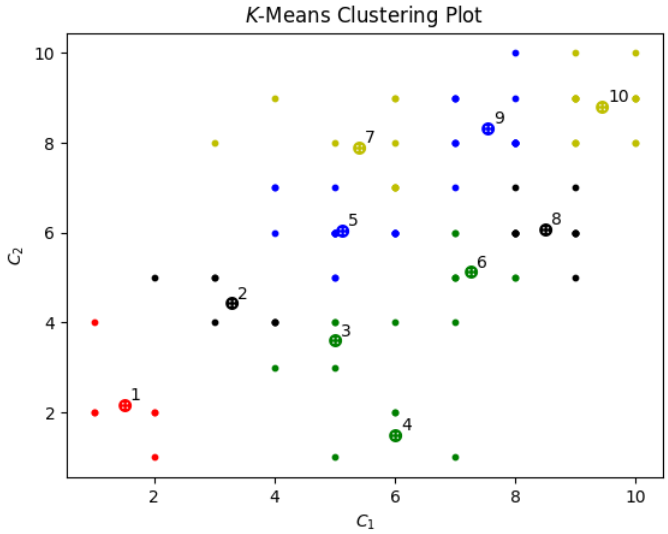

Figure 2: Plot of juror score data for 100 venire members (same as Figure 1). Cluster centroids (large crossed circles) are determined using the K-means algorithm with $K = 10$. Centroids are labeled with the juror rating on a scale of 1 to 10, based on their distance from the plot's origin. Items are color coded to show cluster assignments.

3.2 Juror Rating

Each cluster in Figure 2 describes a group of jurors with similar characteristics. For example, the venire members in group 1 all have scores of 2 or below for questions C_1 and C_2 while the members in group 10 all have scores above 8. The centroid of each cluster represents the average or 'representative' juror of the corresponding cluster. The third column in Table 2 shows the distance of each centroid from the plot's origin. Because higher scores C_1 and C_2 are more preferable, centroids farther from the origin represent more preferable jurors. The fourth column in Table 2 shows the centroid rating on a scale of 1 to 10, with centroids farther from the origin receiving higher ratings. These ratings can be calculated automatically, using a simple computer code, or if preferred, manually by the trial lawyer, based on the values of the underlying scores.

$\mathbf{C_1}$	$\mathbf{C_2}$	**Distance**	**Rating**
1.5	2.17	2.64	1
3.29	4.43	5.51	2
5.	3.6	6.16	3
6.	1.5	6.18	4
5.11	6.06	7.93	5
7.25	5.13	8.88	6
5.4	7.9	9.56	7
8.5	6.08	10.45	8
7.53	8.33	11.23	9
9.44	8.81	12.91	10

Table 2: K-means centroid positions, distance of centroid from plot origin, and cluster ratings

Finally, all jurors in a given cluster are assigned a rating equal to that of the cluster centroid. The entire procedure, from clustering to juror rating, can be completed by computer algorithm in a matter of seconds.

4 Conclusion

It has been said that jury selection is a social science, not a hard science, and consequently, intuition, gut feeling, and emotion play a primary role. However, this is not a reason to discard rigorous analysis when and where it is called for. Sorting, searching, and grouping data are better done by computer algorithm than by intuition and gut feeling. Jury questionnaire evaluation begins with organization and quantification of the data, a strength of computer algorithms, and ends with human interpretation and insight.

With proper pre-trial preparation, automated questionnaire evaluation is a quick and straightforward process. The outcome of the questionnaire evaluation is a set of pre-voir dire juror ratings. These ratings will inform the voir dire and allow litigators to focus questions toward specific jurors in line with their jury selection strategies. Consistent questionnaire evaluation and a clear presentation of the questionnaire data provide a solid foundation for voir dire, where the attorney's skills, experience and intuition can play its part.

A Appendix

Table A1: Common Questionnaire Response Types with Examples

Type	**Example Question**	**Example Response**
Ordinal	Large corporations commonly terminate employees without good cause	Strongly agree, agree, neutral, disagree, strongly disagree
Nominal	How would you describe your occupation	Service, trade, professional ,.....
Interval	What is your age?	30-35, 35 - 40, 40-45, ...
Dichotomous	Have you served on a jury before?	Yes/No
Open ended	Do you have strong feelings about awarding money to the winning party? If so, why?	String response (see Appendix B)

B Appendix

B.1 Handling Open Ended Questions

Open ended questions provide no fixed responses and the submitted responses may require careful reading to understand and score. One method to quickly and efficiently handle open-ended questions, assuming they are in electronic form, uses a text scanning tool (e.g., https://seoscout.com/tools/text-analyzer) to detect keywords within the responses. Keyword contained in each response are then used to rate its favorability.

For example, consider the following open ended question:

> What are your feelings about awarding money to the winning party in a personal injury lawsuit?

Table A2 shows some possible juror responses together with keywords extracted from each response using computer-based text scanning code. Experienced litigators will be able to anticipate the relevant keywords for common open ended questions. These relevant keywords are arranged in columns as shown in Table A3. Each subsequent row specifies with a check mark whether a relevant keyword was found in the corresponding juror's response. Keyword frequencies are then used to assess the favorability of each juror. The entire process can be performed automatically using specialized computer code, however, ambiguous cases may be read and assessed manually.

Juror ID	Response	Extracted Keywords
1	People don't deserve money for doing nothing. People hope they will get lucky and get money by filing frivolous lawsuits. People should work for their money, not file lawsuits.	don't deserve nothing work lucky frivolous
2	If someone is wrongfully injured, they deserve money for pain and suffering. A large award may deter the person from doing it again. Money can help the injured person. Large awards can act as a deterrent. Paying money will punish the person who caused the injury.	wrongfully injured deserve pain suffering deter punish caused
⋮	⋮	⋮

Table A2: Open ended responses and keywords

	⟵ Better for Defense			Better for Plaintiff ⟶		
Juror ID	**frivolous**	**work**	**deserve**	**injured**	**wrongful**	**deter**
1	✓	✓	✓			
2			✓	✓	✓	✓
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table A3: Juror's use of relevant keywords

References

[1] "Seventh circuit american jury project - final report," 2008. [Online]. Available: https://www.uscourts.gov/sites/default/files/seventh_circuit_american_jury_project_final_report_0.pdf

[2] P. L. Refo, "Principles for juries and jury trials." [Online]. Available: https://www.uscourts.gov/sites/default/files/aba_principles_for_juries_and_jury_trials_2005.pdf

[3] N. F. Engstrom, "The trouble with trial time limits," *The Georgetown Law Journal*. [Online]. Available: https://law.stanford.edu/wp-content/uploads/2018/05/The-Trouble-with-Trial-Time-Limits.pdf

[4] A. Stuart, "Don't let the court ruin your jury selection," *The Advocate*. [Online]. Available: https://www.advocatemagazine.com/article/2018-january/don-t-let-the-court-ruin-your-jury-selection

[5] W. Veen and C. Friedman, "Voir dire under time restraints," *Plaintiff*. [Online]. Available: https://www.plaintiffmagazine.com/recent-issues/item/voir-dire-under-time-restraints

[6] L. Blue and R. Hirschhorn, "A guide to analyzing jury questionnaires in civil cases." [Online]. Available: https://www.cebjury.com/documents/AnalysisGuide.pdf

[7] A. Ford, "Simpson jury question list probes range of attitudes," *Los Angeles Times*. [Online]. Available: https://www.latimes.com/archives/la-xpm-1994-10-01-mn-45233-story.html

[8] k-means clustering, "k-means clustering — Wikipedia, the free encyclopedia," 2023, [Online; accessed 2-January-2023]. [Online]. Available: https://en.wikipedia.org/wiki/K-means_clustering

[9] "California code of civil procedure," 2021. [Online]. Available: https://codes.findlaw.com/ca/code-of-civil-procedure/ccp-sect-222-5/

[10] A. Pfau, "Implementing civil voir dire law and rules." [Online]. Available: https://ww2.nycourts.gov/sites/default/files/document/files/2018-06/ImplementingVoirDire2009.pdf