Analysis : Beer Review

Gabriel Simões
University of Lisbon - Faculty of Sciences
Lisbon, Portugal
fc59662@alunos.fc.ul.pt

Jorge Caldeira
University of Lisbon - Faculty of Sciences
Lisbon, Portugal
fc58214@alunos.fc.ul.pt

José Maria Diogo
University of Lisbon - Faculty of Sciences
Lisbon, Portugal
fc55937@alunos.fc.ul.pt

Abstract

Online consumer platforms and vast public datasets have created an unprecedented opportunity to analyze complex socio-economic phenomena and intricate market dynamics. Among these data sources, public beer review repositories provide a rich, micro-level view into consumer preference, product characteristics, and the structure of the breweries market. This study investigates the relationships within a large dataset of beer reviews from BeerAdvocate, focusing specifically on the Boston Beer Company (Samuel Adams) brewery. By applying techniques of social network analysis to this structured review data, we constructed two primary graph models: one based on reviewer similarity across multiple rating attributes (overall, taste, aroma, appearance, palate) as well as the style of the beer, and another based on different beers with shared reviewers. The data was filtered to the 10 most-reviewed beer styles to ensure analytical focus and depth. Our objective is to uncover patterns of consumer clustering, product similarity, and reviewer influence as revealed by these network structures. The study is guided by the hypothesis that constructing similarity graphs based on identical multi-attribute ratings within a specific style allows for the identification of cohesive product characteristics, whereas the reviewer-based graphs are designed to isolate and characterize differentiated consumer behavioral segments. A comparative analysis of the style-specific networks demonstrated substantial differences in network cohesion, with structures ranging from relatively interconnected graphs (e.g., Bock, showing 1,286 edges) to examples of extreme sparsity and fragmentation (e.g., Milk / Sweet Stout, featuring 0 edges). This variation is an example of the dramatically unequal distribution of shared reviewer history across the selected beer styles. These results show a theoretical methodology for dissecting multi-dimensional consumption data, offering critical insights into product congruence, reviewer heterogeneity, and the structural role of strong versus weak ties within product-focused online communities, highlighting the versatility of these user groups.

Keywords

Data Analysis, Beer Reviews, Network Analysis, Consumer Preference, BeerAdvocate

1 Introduction

Online social networks and digital consumption platforms, such as dedicated review sites and e-commerce platforms, have fundamentally transformed how products are consumed, evaluated, and discussed in modern society. These platforms generate immense volumes of user-created content, including multi-dimensional product reviews that capture opinions on various subtle aspects, such as: aroma, appearance, taste, and overall quality. [1]

However, extracting meaningful relationships and structural patterns from raw review data, particularly when analyzing the interconnectedness of consumer segments and product attributes, presents a significant challenge. Reviews are complex, containing explicit numerical scores as well as plain-text feedback, which often includes conflicting sentiments across different product dimensions. Furthermore, accurately modeling user perception requires acknowledging that consumer tastes evolve over time as they gain experience or expertise, a phenomenon known to affect ratings in areas like movies, wines, and beers. This complexity needs an approach beyond simple statistical aggregation, requiring a methodology capable of modeling the dynamic relationships between users, their preferences, and the products they review. [2]

This study investigates a large, publicly available repository of beer reviews, with a specialized focus on the products from the Boston Beer Company (Samuel Adams) brewery. This targeted dataset, characterized by multi-aspect ratings and a large community of reviewers, provides an ideal environment to apply sophisticated network analysis to model both product similarity and reviewer behavior. [1]

The core purpose of this investigation is to apply Social Network Analysis and graph-theoretic techniques to uncover the structural foundations of consumer behavior and product/attribute relationships within this focused ecosystem. While recent research has explored converting review text into knowledge graphs for improved prediction of review ratings, our approach leverages graph modeling not for prediction, but for structural analysis. [3] We constructed two primary network types: Similarity Graph, where nodes (reviews) are connected by identical multi-attribute ratings within the same beer style, and Reviewer-Based Graph, where nodes are connected if they were left by the same user. By subjecting the generated network structures to in-depth structural analysis, this research aims to isolate key patterns in reviewer segmentation, product congruence, and the prevalence of strong versus weak ties within this domain-specific community.

2 Literature Review

The BeerAdvocate beer review dataset has become a surprisingly widely used source for investigating behavioral phenomena and advanced Natural Language Processing (NLP) applications (Jacobsen, 2015; Braun & Timpe; McAuley, Leskovec, & Jurafsky, 2012; McAuley & Leskovec, 2013; Zhang et al., 2020). Its richness derives from the large volume of reviews but also from the detailed structure of each entry, which includes multi-aspect ratings and human-written text describing each entry.

In the field of Behavioral Economics, studies explored the so-called "mimicry effect," investigating how user ratings were influenced by the expert reviews published by the site's founders, known as

The Bros (Jacobsen, 2015). The evaluations were analyzed within controlled temporal windows (for example, six months before and after the expert review), measuring the Rating Difference and the discrepancy between the expert's rating and the average of the users' previous evaluations (Jacobsen, 2015).

At the same time, the dataset became a central resource in Natural Language Processing research, particularly in tasks involving rating prediction based on review text (Braun & Timpe). Studies analyzing the 1,586,602 reviews (Braun & Timpe; Jacobsen, 2015) uncovered clear linguistic relationships between textual content and the assigned ratings. These analyses motivated the formulation of the problem both as a classification task and a regression task (Braun & Timpe), using models such as Multinomial Naive Bayes (Braun & Timpe).

Beyond traditional approaches, the dataset also inspired more specialized studies, including research on learning attitudes and attributes in multi-aspect analysis systems — exploring how specific passages of text relate to sub-ratings such as aroma or flavor (McAuley et al., 2012). In more recent work, the dataset was used to evaluate multi-aspect sentiment analysis frameworks supported by neural networks, such as the Sentiment-Aspect Attribution Module (SAAM), which integrates CNNs and RNNs to improve sentiment and rating prediction accuracy (Zhang et al., 2020).

Thus, the literature demonstrates that the BeerAdvocate dataset is simultaneously a behavioral (Jacobsen, 2015), linguistic, and computational resource (Braun & Timpe; McAuley et al., 2012; McAuley & Leskovec, 2013; Zhang et al., 2020) of high relevance, enabling the exploration of social dynamics, cognitive effects, and advanced methodologies in machine learning and textual data processing.

3 Methodology

3.1 Dataset Description

The data set was initially chosen from the Stanford Network Analysis Platform (SNAP) , specifically seeking a large review corpus that included multi-aspect ratings, similar to those analyzed in existing network research. Although the original SNAP data was unavailable, a matching dataset was sourced from Kaggle, which functions as a large community for data scientists, offering datasets and tools for machine learning research. The selected Beer Reviews from BeerAdvocate dataset consists of around 1.5 million beer reviews spanning over a decade up to November 2011. Each review includes ratings in terms of five distinct aspects: appearance, aroma, palate, taste, and overall impression, besides the profile name of the user. The entire collection is an undirected network composed of 1,586,259 (edges) and 99,438 total entities (nodes, both users and beers) The full data resides in a single file called beer_review.csv.

3.2 Network Construction

The initial dataset was highly extensive, containing reviews for over sixty-six thousand unique beers and thirty-three thousand unique users. To conduct a deep analysis focused on a constrained community, the data was filtered following three key steps. First, during the initial data cleaning and focusing phase, the raw data was cleaned by dropping entries where the Alcohol by Volume (beer_abv) was not recorded and replacing missing reviewer profiles (review_profilename) with the placeholder "anonymous." Second, the analysis was narrowed in a brewery selection step to the brewery with the highest volume of reviews, which was identified as the Boston Beer Company (Samuel Adams). This subset comprises 38,812 reviews, providing a robust community of 7,251 unique reviewers and 97 unique beers across 50 distinct beer styles. Finally, the working dataset underwent a topical constraint filter, restricting the data to the top 10 beer styles by review count, thereby focusing computational resources on the most active segments of the data. This sequential process resulted in a final core dataset of 21,012 reviews across 10 styles, including Bock, American Pale Wheat Ale, and American IPA.

For the subsequent analysis, each individual review served as a node in the graph, carrying all associated metadata (e.g., beer_style, review_overall, review_profilename). Two distinct network models were constructed using the NetworkX library:

- Reviewer-Similarity Graph (S2): Edges connect any two reviews if they share the same beer style and have identical numerical ratings across all five dimensional aspects: review_overall, review_taste, review_aroma, review_appearance, and review_palate. This graph models latent product congruence based on strict user agreement.
- Reviewer-Based Graph (S3): Edges connect any two reviews if they were submitted by the same reviewer (review_profilename). This structure models shared reviewer engagement, revealing user activity patterns and potentially delineating strong ties between the reviewer and the products they consume.

Furthermore, separate subgraph analyses were performed for each of the top 10 beer styles, using the same Reviewer-Based Graph logic to investigate cohesion within specific consumer segments (e.g., Bock_reviewers.gexf).

3.3 Analytical Methods

To gain an in-depth understanding of the constructed network topologies and to enable meaningful comparison, the graphs will be subjected to a deep analysis using established SNA techniques. These techniques are essential for evaluating the structure, efficiency, and robustness of the communities modeled by the graphs. The primary analytical methods employed span fundamental graph-level metrics, node-level centrality measures, and community detection algorithms. The analysis begins with the macroscopic assessment of the constructed networks through fundamental metrics. These include the number of Nodes and Edges, which are foundational indicators of network size and complexity. The Average Degree is computed to reflect the mean number of connections per node, providing insight into overall network activity, while Density measures the ratio of actual connections to the maximum possible connections, quantifying the general cohesion of the network structure. Moving to individual node influence, centrality metrics are employed to quantify the structural role of individual nodes (reviews) within the network. Degree Centrality represents the number of direct connections a node possesses, thus indicating its immediate connectivity within the community. Closeness Centrality calculates how quickly a node can reach all other nodes in the network, efficiently identifying points that facilitate rapid information access. Betweenness Centrality identifies nodes that frequently act as critical bridges or intermediaries lying on the shortest paths between other nodes, thereby highlighting gatekeepers of flow within the network. For a more nuanced measure of influence, Eigenvector Centrality is calculated, which attributes higher scores to nodes connected to other highly-scoring, influential nodes within the network, identifying nodes with connections to important parts of the structure. Finally, for identifying latent substructures and social grouping, community detection is performed. The Louvain Algorithm, widely used for its efficiency, iteratively optimizes Modularity: a metric that quantifies the quality of a partition by measuring how much more densely connected nodes within a community are compared to random expectation. Furthermore, the analysis may utilize the Leiden Algorithm, an enhancement of the Louvain method known for producing high-quality partitions and ensuring all identified communities are well-connected.

4 Results and Discussion

To make the visualization of these graphs more intuitive, we used Gephi's functions to organize the nodes according to the type of beer. Below are the graphs corresponding to reviews connected by common reviewers and by common ratings:

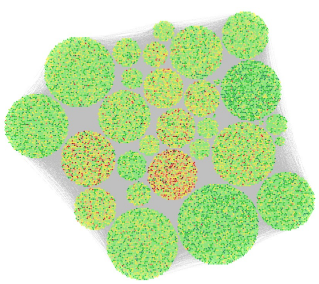

Figure 1: Reviews connected by multiple review parameters in common

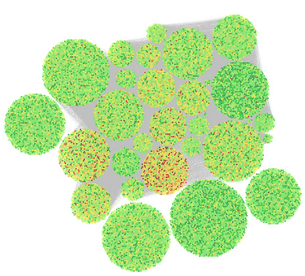

Figure 2: Reviews connected by reviewers in common

4.1 Descriptive Analysis

We observed in the graph of user reviews that the number of connected components is 5,923, with a highly skewed size distribution. Most users have only 1–3 reviews, while some reach up to 26 reviews, forming larger cliques. This asymmetry indicates that most users contribute with only a few reviews, but there is a minority of heavy reviewers who evaluate many beers, forming larger components. Here we see that the largest review clusters (with 26 reviews) are authored by the users Duhast500, mikesgroove, TheManiacalOne, and Brenden. The graph that relates rating similarities shows a highly fragmented network structure, composed of 7,641 connected components. This fragmentation results from the edge-formation criterion in which two reviews are connected only when they share identical values across all rating parameters. The distribution of component sizes displays a highly asymmetric pattern, where most components contain only a small number of nodes, representing rare or unique rating profiles. In contrast, some components reach sizes above 100 nodes, grouping reviews with extremely frequent rating patterns. An analysis was performed on the average overall scores of the reviews in the 500 largest components and in 500 isolated components, and we observe that the average remains similar, with the isolated components being those that concentrate ratings of 1 to 2 (Graph 1). The largest connected component corresponds to the most common rating in our subset: the American IPA style, with an overall review of 3.5, taste 4, and aroma, appearance, and palate of 3.5.

4.2 Comparative Analysis

Average clustering coefficient: The average clustering coefficient is calculated by taking into account how "closed" the neighborhoods of the graph's nodes are, and how close they are to forming cliques. In both graphs this value is exactly 1, and all nodes have a clustering coefficient equal to 1, which implies that whenever two reviews share a common neighbor, they are also directly connected to each other. Structurally, each connected component is a clique, all possible triangles exist because all clusters represent complete graphs. The exact equality of the parameters forces complete connectivity among all reviews that share the same profile or user.

Degree distribution: The similarity graph shows an average degree around 12, but with a long tail, including nodes with degree above 130, which belong to the largest components, while the degree distribution of the user graph shows an average degree of around 7, but with a long tail reaching values around 25. This means that most reviews are in small components, while some belong to users with many reviews and therefore have high degree. In these cliques, the degree of each node is simply the size of the component minus one, so high degree nodes represent extremely frequent rating profiles, while nodes with degree 0 or very low correspond to almost unique profiles.

Eigenvector centrality: Eigenvector centrality measures the structural "influence" of a node, considering not only how many connections it has but also how important its neighbors are. Both graphs show most nodes with values very close to zero and a few nodes with high values close to 1, which reflects the fact that structural importance is concentrated in the largest components. Within each clique, all nodes are equivalent, but larger components contribute more to the principal eigenvectors.

Classic modularity and Leiden modularity: Modularity measures how good a partition of the graph into communities is, comparing the fraction of edges inside communities with what would be expected in a random graph with the same degree. In both graphs the number of communities matches the number of components, with very high modularity values. This indicates that each connected component is recognized as a dense community separate from the others, maximizing modularity due to the strong concentration of edges. The distribution of community sizes follows the distribution of the components.

Distance analysis: In network science, a distance analysis studies how far nodes are from each other along the graph's edges, using the length of shortest paths. From these pairwise distances, several global measures are derived, such as eccentricity of each node, the graph's diameter, radius, and average path length. Since each component represents a complete graph, the average path length is always equal to 1. Any pair of connected reviews is linked by a single edge.

Minimum spanning tree: A minimum spanning tree is a way of "connecting" all vertices of a graph with the minimum possible number of edges and the lowest total cost. The user and similarity graphs contain 15,089 and 13,371 edges respectively, both values much lower than the total number of edges in the connected graphs, which shows that a large part of the connections is redundant in terms of global connectivity.

Density: Density measures how "full" of edges a graph is in relation to the maximum theoretically possible for the same number of nodes. Both graphs are extremely sparse, even though each user component is internally dense. The graph referring to rating similarities has a density of 0.000592 and the user graph has 0.000336. This means that, compared to the number of possible pairs of reviews, the similarity graph achieves about twice the fraction of connections. Intuitively, with 0.000592 there are slightly larger cliques, which indicates greater global cohesion and more observed relationships per possible pair of reviews.At first glance, it may seem counterintuitive that the graph with fewer components ends up having lower density. In both cases under study, the graphs are unions of isolated cliques. However, in the less dense graph, a substantial portion of the vertices are concentrated in a few large components. This means the number of possible vertex pairs grows much faster than the actual number of edges formed. As a result, the graph with fewer components—but with more vertices packed into a handful of large blocks—accumulates a very high number of potential pairs compared to the total edges, which drives the overall density down.

5 Conclusion

Both graphs produce components that are complete cliques. This property implies significant redundancy of edges and removes internal hierarchies. The clustering coefficient equal to 1 confirms the absence of open triads. The degree distribution shows distinct averages with long tails that reach values above 130, representing extremely frequent rating profiles. Eigenvector centrality is concentrated in the largest components, where high values reflect the disproportionate contribution of larger cliques to the principal eigenvectors. The high modularity coincides with the number of connected components, indicating that each component constitutes a dense and isolated community. The average path length equal to 1 results from the structure of complete cliques, where any pair of connected nodes is directly linked. The global densities are very low, reflecting extreme fragmentation. This approach has significant limitations. Treating components as complete cliques removes structural nuances, confining variability to the number and size of the components. However, analysing datasets as large as this one through network theory presents considerable challenges. Some of these include understanding and correctly converting tabular data into graphs that capture the patterns needed to answer the research questions, the lack of computational capacity to make truly interesting inferences, and the redundancies that may arise when computational constraints force the use of increasingly small subsets.

References

[1] Mcauley, Julian, et al. Learning Attitudes and Attributes from Multi-Aspect Reviews. 2012.

[2] Mcauley, Julian, and Jure Leskovec. From Amateurs to Connoisseurs: Modeling the Evolution of User Expertise through Online Reviews. 2013.

[3] de Vink, A.J.W. “ReviewGraph: A Knowledge Graph Embedding Based Framework for Review Rating Prediction with Sentiment Features.” Arxiv.org, 2024,arxiv.org/html/2508.13953v1. Accessed 8 Dec. 2025.

[4] https://www.beeradvocate.com/ . Accessed on 8 Dec. 2025

[5]Braun, Benjamin, and Robert Timpe. Text Based Rating Predictions from Beer and Wine Reviews.

[6] Jacobsen, Grant D. “Consumers, Experts, and Online Product Evaluations: Evidence from the Brewing Industry.” Journal of Public Economics, vol. 126, June 2015, pp. 114–123, https://doi.org/10.1016/j.jpubeco.2015.0 Accessed 8 Dec 2025.

[7] Zhang, Yifan, et al. “Multi-Aspect Sentiment Analysis with Latent Sentiment-Aspect Attribution.” ArXiv.org, 2020, arxiv.org/abs/2012.08407. Accessed 8 Dec. 2025.

[8] https://github.com/JorgeCaldeira/ADC-Project-FCUL/

A Additional Figures

	brewery_name	total_entries	unique_beer_count
634	Boston Beer Company (Samuel Adams)	38812	97
1793	Dogfish Head Brewery	33800	100
4457	Stone Brewing Co.	33022	119
4273	Sierra Nevada Brewing Co.	28637	121
375	Bell's Brewery, Inc.	24975	96
4058	Rogue Ales	23450	90
2056	Founders Brewing Company	19955	130
4896	Victory Brewing Company	19407	107
2893	Lagunitas Brewing Company	16771	71
210	Avery Brewing Company	16044	93

Figure 3: List of Breweries with the most reviews (and different beers

	beer_style	review_count
16	Bock	3371
8	American Pale Wheat Ale	2805
49	Witbier	2566
46	Vienna Lager	2418
26	Fruit / Vegetable Beer	2259
35	Märzen / Oktoberfest	1893
6	American IPA	1631
42	Schwarzbier	1494
17	Czech Pilsener	1336
33	Milk / Sweet Stout	1239

Figure 4: Top 10 most reviewed beer styles from Boston Beer Company

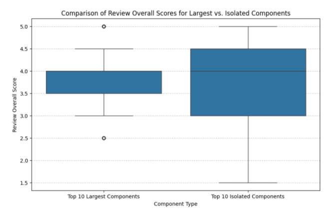

Figure 5: Graph 1 - Comparison of Review Overall Scores for Largest vs. Isolated Components