Brewing Insights:

Utilizing Cluster Analysis for Strategic Beer Type Selection in the Craft Beer Industry

Motivation & Introduction

Background: Motivation

Analysis benefits a business owner in the beer industry

1. **Growing Market** in specialty craft beers

1. **Market Analysis** shows cluster analysis makes beer selection manageable

Background: Research Goals

Understand flavor profiles in the craft beer market

Use cluster analysis to identify distinct beer categories

<u>Goal completed:</u>

Business owner selects representative range of beers

Background: Dataset Overview

25 columns and 3197 rows

3197 Unique Beers

Data Cleaning:
Remove insignificant variables
+ Outliers

	Name	Style	Brewery	ABV	Astringency	Body	Alcohol	Bitter	Sweet	Sour	Salty	Fruits	Hoppy	Spices	Malty
0	Amber	Altbier	Alaskan Brewing Co.	5.3	13	32	9	47	74	33	0	33	57	8	111
1	Double Bag	Altbier	Long Trail Brewing Co.	7.2	12	57	18	33	55	16	0	24	35	12	84
2	Long Trail Ale	Altbier	Long Trail Brewing Co.	5.0	14	37	6	42	43	11	0	10	54	4	62
3	Doppelsticke	Altbier	Uerige Obergärige Hausbrauerei GmbH / Zum Uerige	8.5	13	55	31	47	101	18	1	49	40	16	119
4	Sleigh'r Dark Doüble Alt Ale	Altbier	Ninkasi Brewing Company	7.2	25	51	26	44	45	9	1	11	51	20	95

15 columns and 3167 rows

*Source: Kaggle's "Beer Profile and Ratings Data Set"

Noteworthy Visualization

```
Style
Lager - Adjunct                          45
Lager - European Pale                    43
Wheat Beer - Hefeweizen                  42
Stout - Irish Dry                        42
Lambic - Fruit                           42
                                         ..
Bière de Champagne / Bière Brut           8
Lager - India Pale Lager (IPL)            6
IPA - New England                         4
Brett Beer                                3
Sour - Gose                               3
Name: count, Length: 111, dtype: int64
```

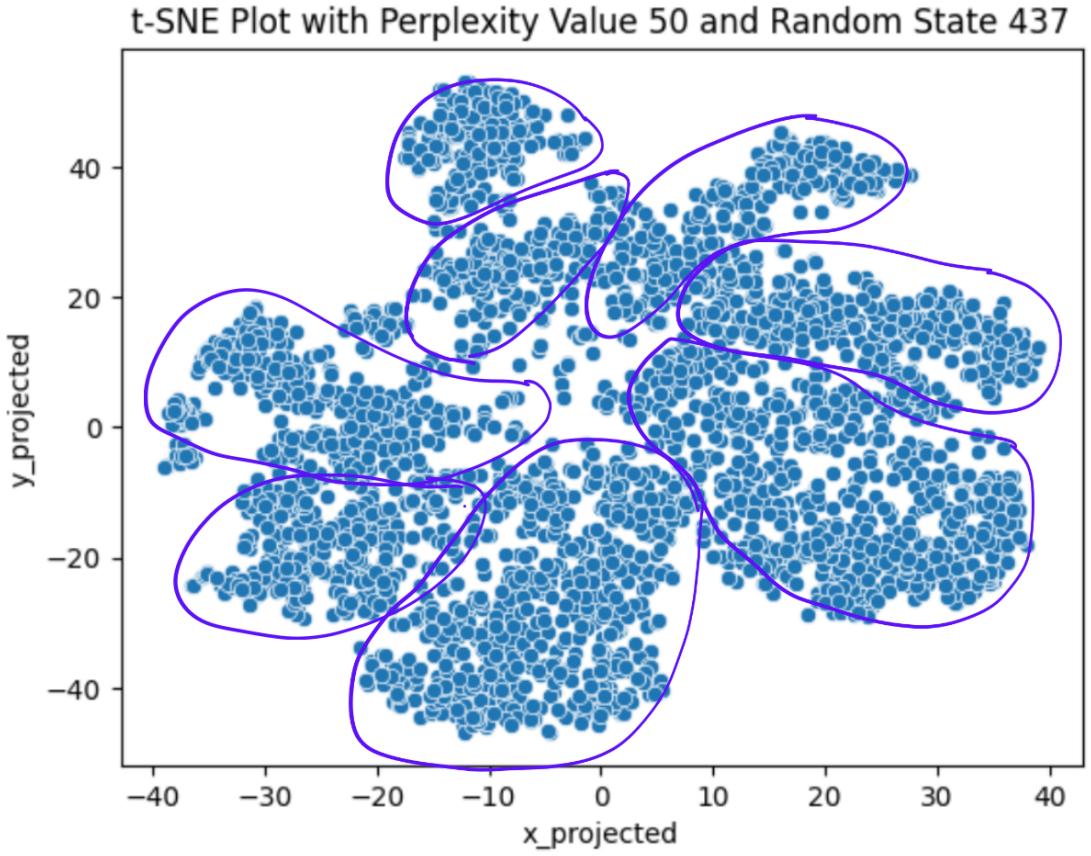
111 Distinct Beer Style

Must group into meaningful amount of distinct categories.

Clustering Structure

t-SNE & Hopkins Statistic

Clustering Structure



- Clusterable:

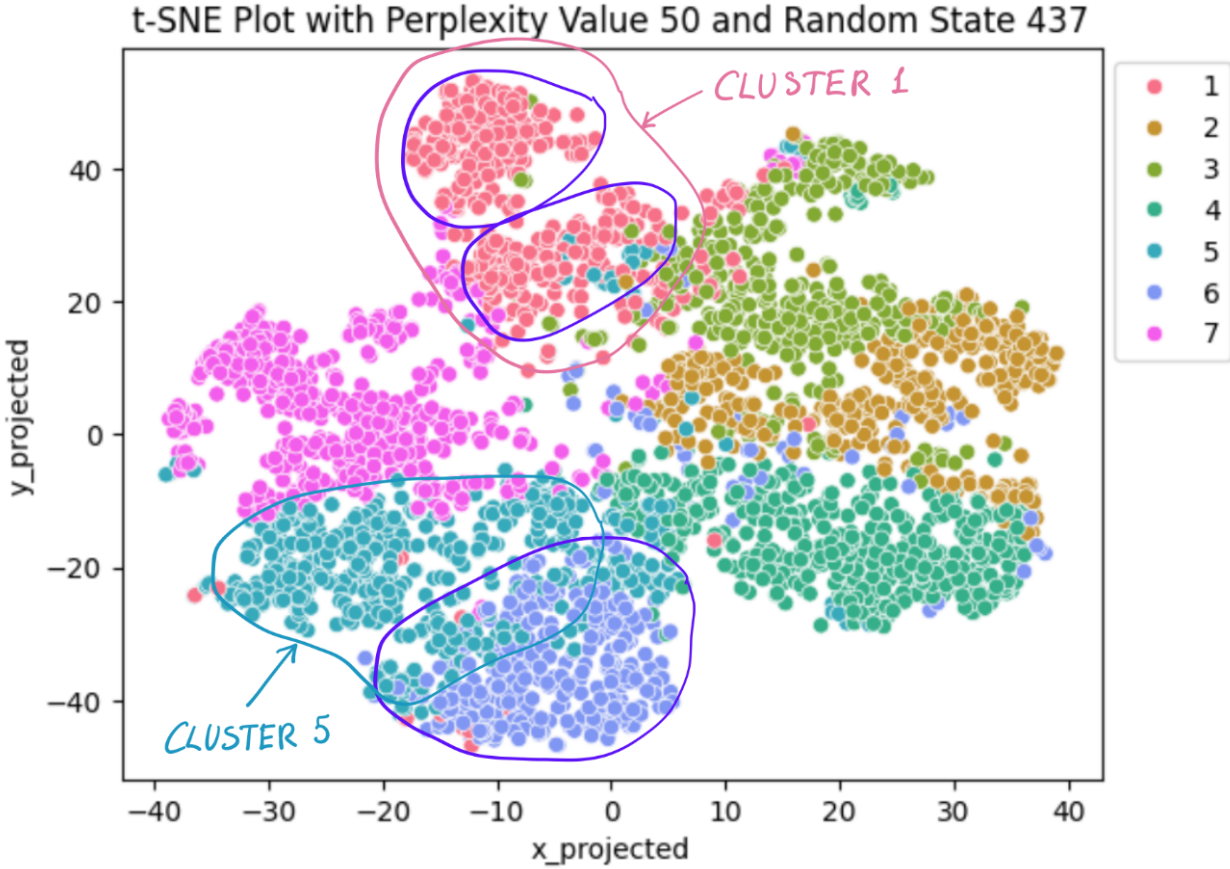
  Mean Hopkins Statistic = 0.1 ≈ 0

- Not well-separated
- Not well-balanced
- Shape: elliptical petals

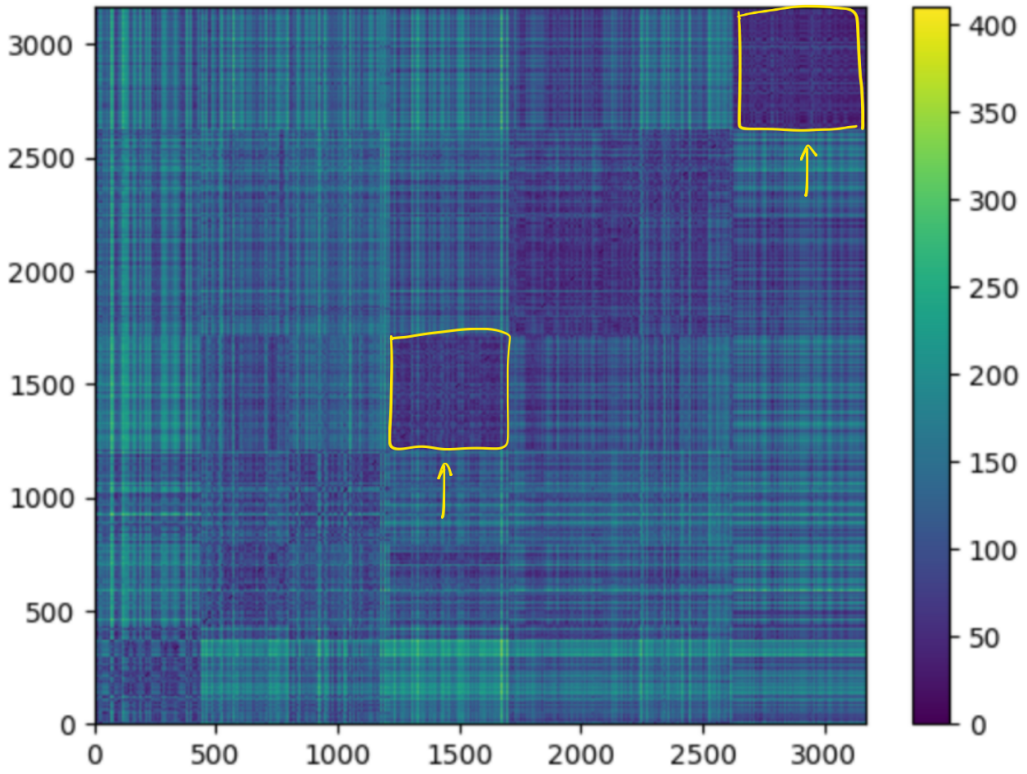
Clustering Algorithm #1

K-Medoids (k=7)

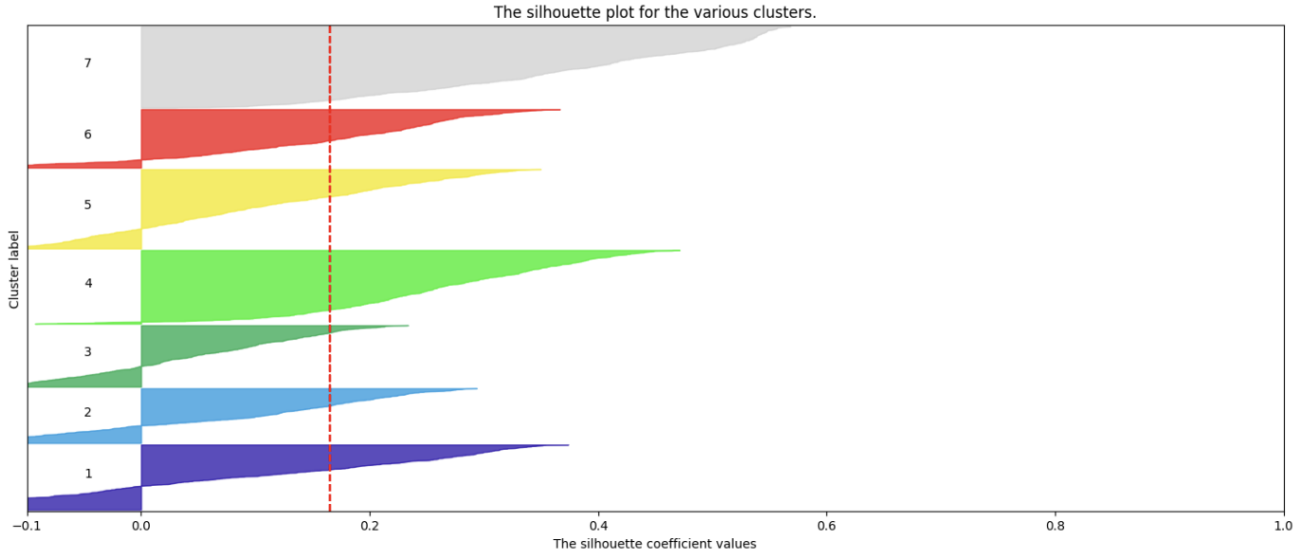
K-Medoid: Motivation and Results



K-Medoid: Additional Cluster Exploration



- Clusters are **close to each other**
- **2 clusters** with high cohesion/separation



- Clusters are **not well-separated**
- **Grey** and **Bright green –** the best cohesion/separation
- 5 clusters with **sharp** bar plots

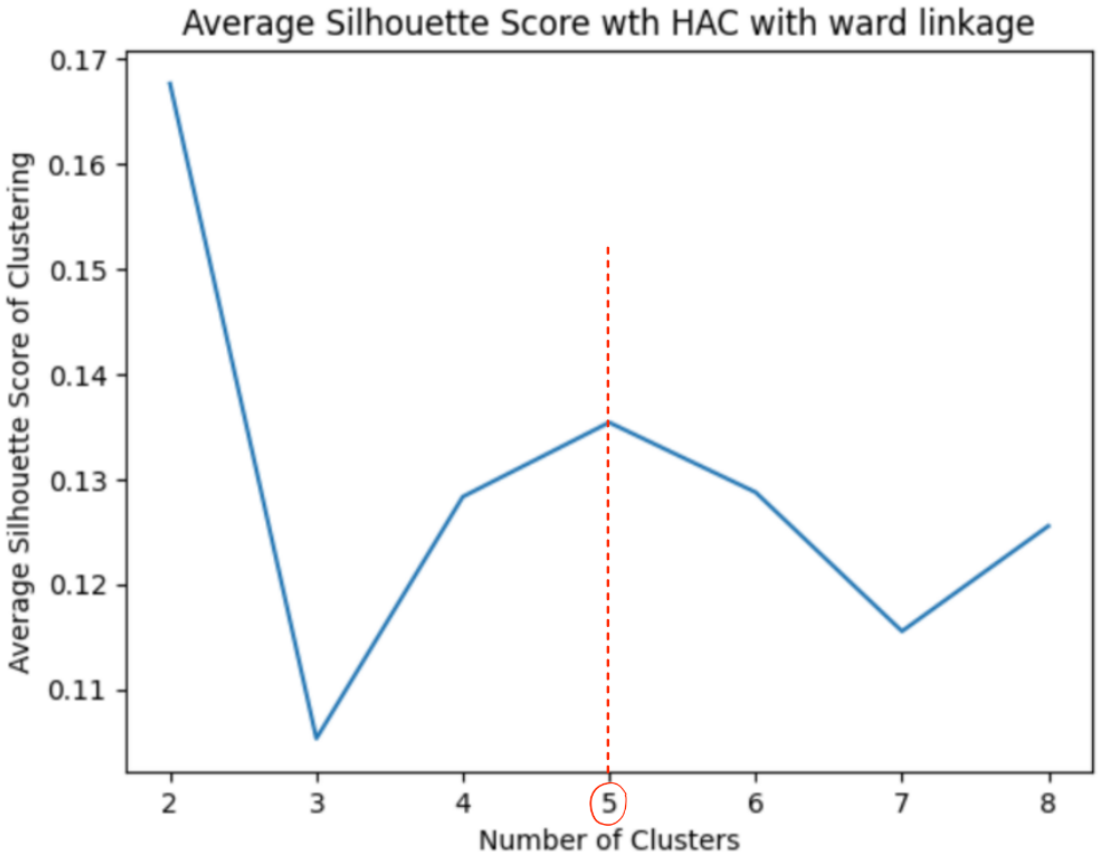
K-Medoids: Cluster Description

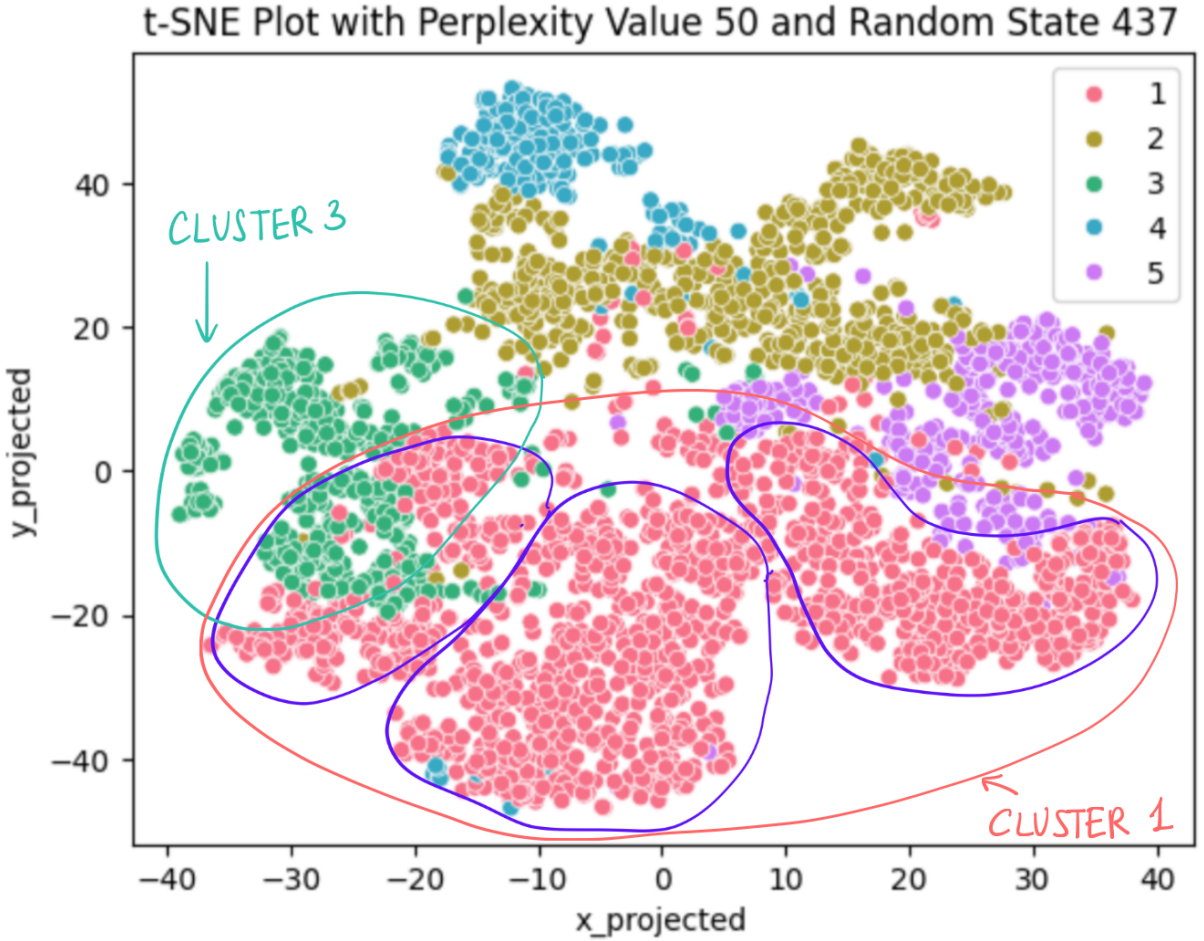
- CLUSTER 1: **Wheat Beer** (High Astringency/Sour/Fruity) e.g. Witbier, Lambic
- CLUSTER 2: **Strong Ale** (Highest ABV/Sweet/Malty) e.g. Barleywine, Bock
- CLUSTER 3: **Dark Beer** (High ABV/Sweet/Spicy & Fruity) e.g. Dunkel, Dubbel
- CLUSTER 4: **Roasted Ale** (Full Body/Bitter/Malty) e.g. Stout, Porter
- CLUSTER 5: **Pilsner** (High Astringency/Bitter & Salty/Hoppy)
- CLUSTER 6: **IPA** (High ABV/most Bitter/Hoppy & Fruity)
- CLUSTER 7: **Lager** (Light beer)

Clustering Algorithm #2

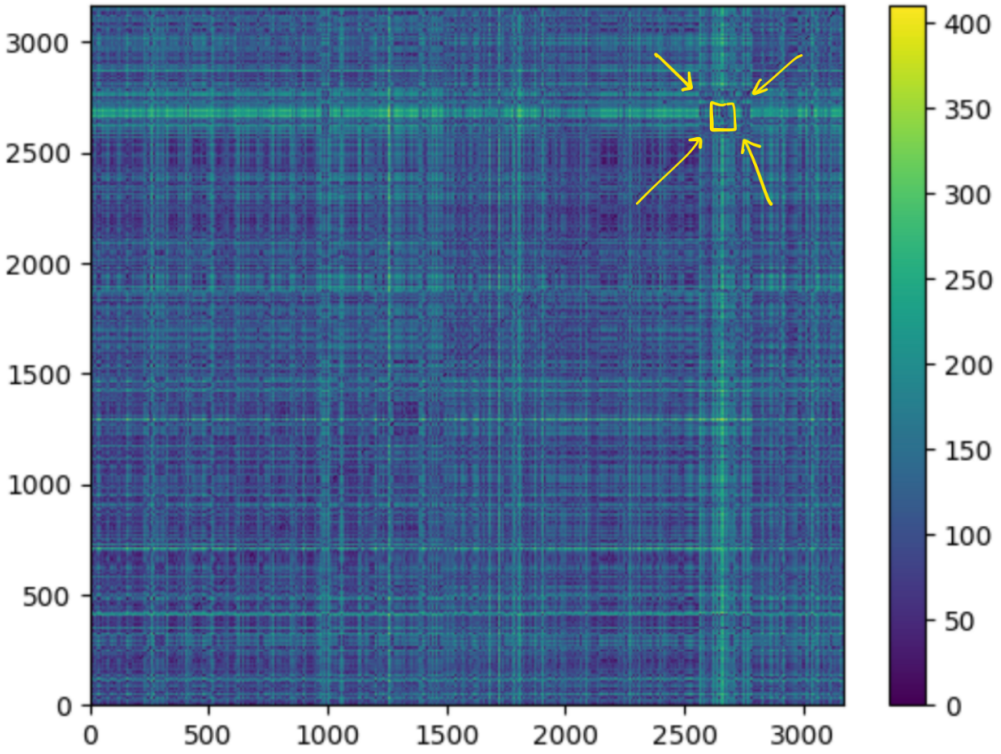
HAC Ward's linkage (k=5)

HAC with Ward's linkage: Motivation and Results

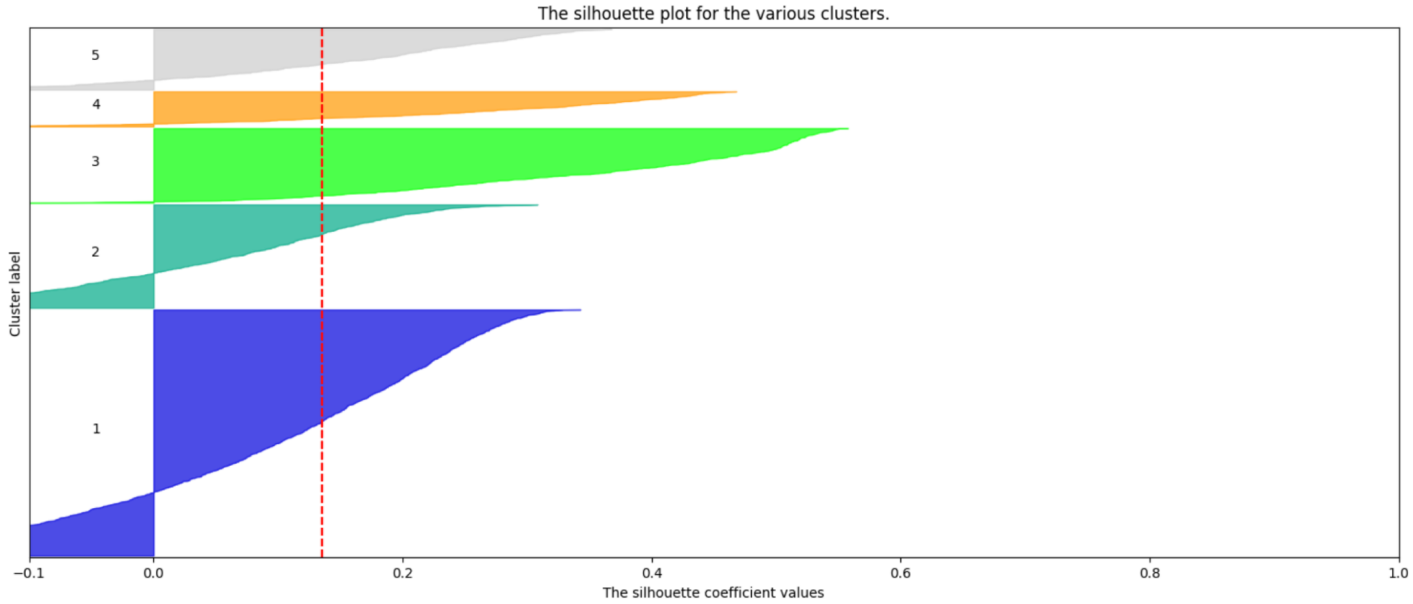




HAC with Ward's linkage: Additional Cluster Exploration



- Clusters are **close to each other**
- 1 cluster is **furthest away**



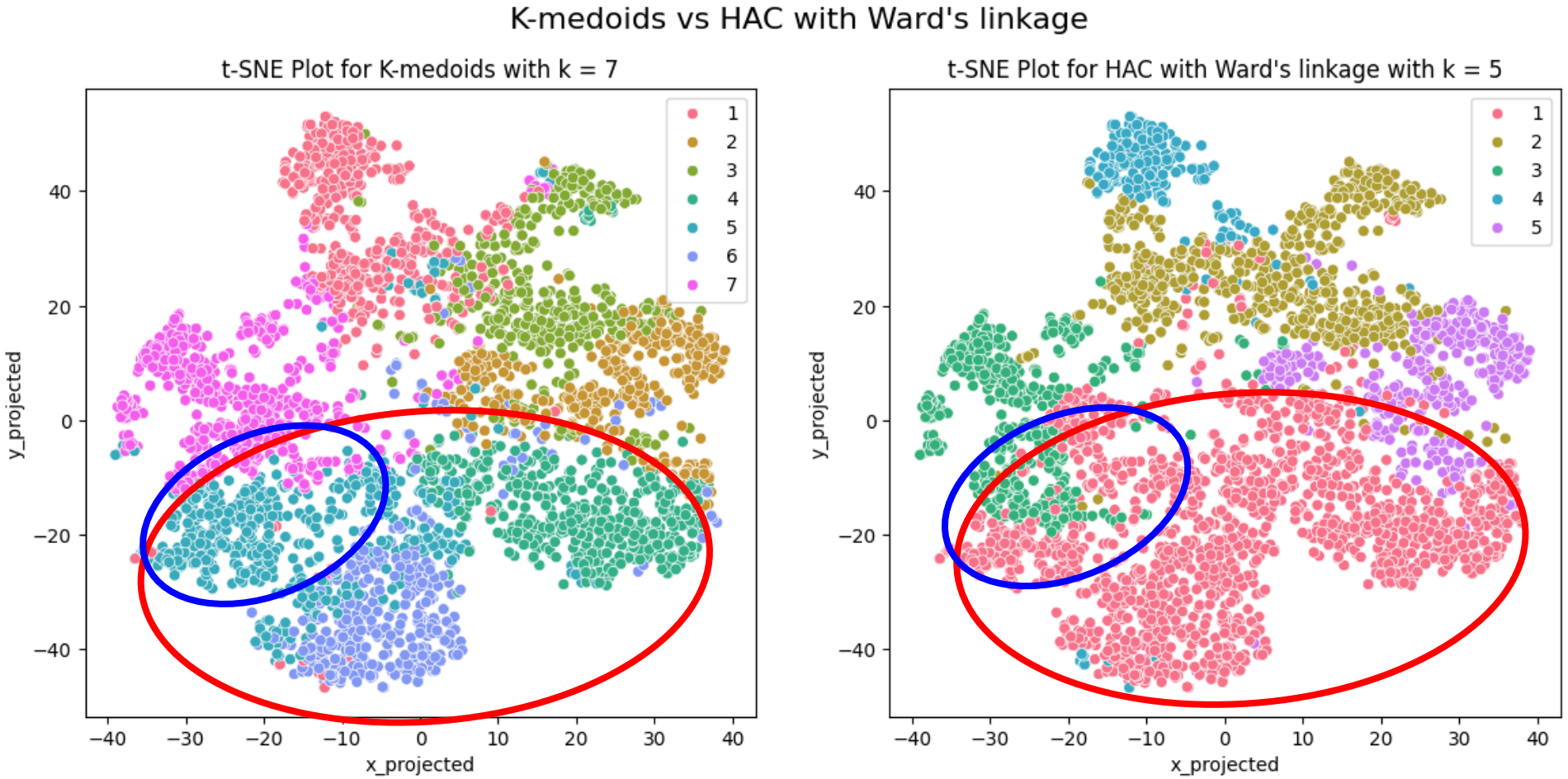
- Clusters are **not well-separated**
- **Orange** and **Bright green –** the best cohesion/separation
- **Blue** – the worst cohesion/separation

HAC with Ward's linkage: Cluster Description

- CLUSTER 1: **Ale** (Full Body/Bitter/Hoppy & Malty) e.g. Pale Ale, IPA, Stout
- CLUSTER 2: **Dark Beer** (High ABV/Sweet/Spicy & Fruity) e.g. Dunkel, Dubbel
- CLUSTER 3: **Lager** (Light beer)
- CLUSTER 4: **Wheat Beer** (High Astringency/Sour/Fruity) e.g. Witbier, Lambic
- CLUSTER 5: **Strong Ale** (Highest ABV and Full Body/Bitter & Sweet/Malty) e.g. Barleywine, Bock

Insights and Comparisons

Insights and Algorithm Comparisons



Insights and Algorithm Comparisons

Cluster #	K-medoids	HAC with Ward's linkage	Mouth Feel	Taste	Flavor
1	Wheat Beer	Wheat Beer	High Astringency	Sour	Fruity
2	Strong Ale	Strong Ale	Highest ABV	Sweet	Malty
3	Dark Beer	Dark Beer	High ABV	Sweet	Spicy & Fruity
4	Roasted Ale		Full Body	Bitter	Malty
5	IPA		High ABV	Most Bitter	Hoppy & Fruity
6	Pilsner		High Astringency	Bitter & Salty	Hoppy
7		Pale Ale	Full Body	Bitter	Hoppy & Malty
8	Lager	Lager	Light	Light	Light

Conclusions

Final Thoughts

Final Thoughts

3197 unique beers → 7 distinct clusters

<u>Recommendation</u>

Select 1 beer type from each category

Research Goal Completed!

Future Work

Shortcomings:

1. No perfect cluster separation
2. Overly simplified beer categories

Add Additional Variables – better match consumer trends

1. Sales
2. Demographics
3. Consumer Preferences over time