

Contextual Legal Definitions - KModes Clustering

Madeline Kaufman

Project 2

Introduction

In today's complex legal landscape, a significant challenge lies in the widespread lack of understanding of laws and regulations. This issue disproportionately affects individuals, especially those from historically undereducated, resource-limited, and marginalized communities. For them, the absence of basic legal knowledge often leads to greater vulnerability and injustice, as they lack the means to access legal assistance.

The legal system is riddled with artificial and longstanding barriers that hinder the democratization of legal knowledge. The system's barriers prevent widespread understanding of laws, creating a dependency on paid legal services. This undermines informed citizenship and perpetuates inequality. Empowering everyone with basic legal understanding is crucial for a fairer, more informed society.

Herein lies a key challenge that I hope to start addressing with this project. While existing systems effectively provide answers to user queries and accurate citations, they often lack the contextual depth provided by legal terminology definitions. These definitions are critical to enhancing user understanding and bridging gaps in legal knowledge.

The issue is that these definitions are dispersed across various nodes in the database, making it difficult to incorporate them effectively into responses. As well as not all definitions are relevant to every query, and including too many could overwhelm users and detract from the clarity of the information provided. To address this challenge, it is necessary to identify and organize these definitions, determining their scope and relevance to specific contexts.

To begin, I plan to use KModes clustering to group similar definitions together, classify nodes as definitions or non-definitions, and assess the potential scope of each definition. This approach will help streamline their integration into the system, ensuring users receive clearer, more contextually relevant information.

Importing the relevant libraries

```
In [2]: from kmodes.kmodes import KModes
```

```
In [59]: import pandas as pd
import re
import numpy as np
import matplotlib.pyplot as plt
from sklearn.decomposition import PCA
import seaborn as sns
sns.set()
from IPython.display import display, HTML
from IPython.display import display_html

from itertools import chain, cycle
from sklearn.metrics import silhouette_score
from sklearn.preprocessing import OneHotEncoder
from sklearn.metrics import davies_bouldin_score
from sklearn.manifold import TSNE
from sklearn.decomposition import PCA
```

Loading in Data

```
In [4]: data = pd.read_csv('ca_hsc_mads.csv', on_bad_lines='skip')
```

The data is quite extensive and while I continue to evolve this algorithm and make it better to solving our problem I will only begin by running it on a small data table within our database. This data table is Californias Health and Safety Code.

For the sake of this project and to simplify understanding of the dataset the only relevant variables we will be looking at and necessary for the clustering is the `node_text` and encoded category pattern columns. The nodes that contain different text could potentially contain one sentence to multiple sentences of different legal codes within Californias Healthy & Safety Iesgislation. The columns of encoded node patters contain values 0 and 1. 0 being no this node does not match the text searched text pattern. Or 1 meaning yes this node does match the searched text pattern.

```
In [76]: data.head()
```

Out[76]:

		id	node_type	top_level_title
0	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...		sub_content_node	HSC {CA/statu
1	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...		sub_content_node	HSC {CA/statu
2	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...		sub_content_node	HSC {CA/statu
3	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...		sub_content_node	HSC {CA/statu
4	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...		sub_content_node	HSC {CA/statu

Searching for Patterns & Encoding them

To start I automate the process of identifying specific types of legal definitions and terms within a large collection of text, this is so when I run the KModes model it will hopefully help by clustering them into specific definition clusters.

In [5]:

```
df = pd.DataFrame(data)

pattern = {}

# Defining node patterns
patterns = {
    'chapter': "For purposes of this chapter, the following definitions apply",
    'chapter2': r'For purposes of this chapter, "[^"]*" means',
    'section': "For purposes of this section",
    'article': "The definitions in this article",
    'article2': r'For purposes of this article,\s*the following definitions app
    'possible_definition': r'"[^"]*" \s+ means',
    'possible_definition2': r'"[^"]*" \s+ includes',
    'possible_definition3': r'"[^"]*" \s+',
    'possible_definition4': r'The word \s+ "[^"]*" ',
    'possible_definition5': r'"[^"]*" \s+ has the same meaning',
    'code_definitions': 'Definitions',
```

```
}  
  
# Function to search for a pattern in node  
def contains_pattern(text, pattern):  
    if isinstance(text, str):  
        return 1 if re.search(pattern, text, re.IGNORECASE) else 0  
    else:  
        return 0  
  
for pattern_name, pattern in patterns.items():  
    df[pattern_name] = df['node_text'].apply(lambda x: contains_pattern(x, patt  
print(df)
```

	id	node_type \
0 1	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
2	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
3	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
4	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
...	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
45693
45694	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
45695	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
45696	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
45697	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	sub_content_node
	CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11/...	addendum_node

	top_level_title	sibling_nodes
0	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
1	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
2	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
3	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
4	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
...
45693	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
45694	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
45695	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
45696	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
45697	HSC	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...

	internal_references \
0	{}
1	{}
2	{}
3	{CA/statutes/code=HSC/DIVISION=20/CHAPTER=6.11...
4	{}
...	...
45693	{}
45694	{}
45695	{}
45696	{}
45697	{}

	node_text \
0	(b) A notice to comply shall be the only means... \n (c) (1) A person who receives a notice to c...
1	(d) If a notice to comply is issued, a single ...
2	(e) If a person who receives a notice to compl...
3	(f) This section may not be construed as doing...
4	...
...	(i) All administrative penalties collected fro...
45693	(j) The UPA shall consult with the district at...
45694	\n (k) (1) A unified program agency may suspen...
45695	(l) This section does not do any of the follow...
45696	(Amended by Stats. 2007, Ch. 626, Sec. 22. Eff...
45697	

	citation	chapter	chapter2	section	article \
0	Cal. HSC § 25404.1.2(b)	0	0	0	0
1	Cal. HSC § 25404.1.2/(c)	0	0	0	0
2	Cal. HSC § 25404.1.2(d)	0	0	0	0
3	Cal. HSC § 25404.1.2(e)	0	0	0	0
4	Cal. HSC § 25404.1.2(f)	0	0	0	0
...
45693	Cal. HSC § 25404.1.1(i)	0	0	0	0

45694	Cal. HSC § 25404.1.1(j)	0	0	0	0
45695	Cal. HSC § 25404.1.1/(k)	0	0	0	0
45696	Cal. HSC § 25404.1.1(l)	0	0	0	0
45697	Cal. HSC § 25404.1.1Addendum	0	0	0	0

	article2	possible_definition	possible_definition2 \
0	0	0	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
...
45693	0	0	0
45694	0	0	0
45695	0	0	0
45696	0	0	0
45697	0	0	0

	possible_definition3	possible_definition4	possible_definition5 \
0	0	0	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
...
45693	0	0	0
45694	0	0	0
45695	0	0	0
45696	0	0	0
45697	0	0	0

	code_definitions
0	0
1	0
2	0
3	0
4	0
...	...
45693	0
45694	0
45695	0
45696	0
45697	0

[45698 rows x 18 columns]

In [6]:

```
def extract_quoted_text(s, patterns):
    if not isinstance(s, str):
        return 0

    for key, pattern in patterns.items():
        if re.search(pattern, s, re.IGNORECASE):
            return 1 # Return 1 for 'possible definition'

    return 0 # Return 0 for 'probably not a definition'

# Defining patterns
```

```

patterns = {
    'chapter': "For purposes of this chapter, the following definitions apply",
    'chapter2': r'For purposes of this chapter, "[^"]*" means',
    'section': "For purposes of this section",
    'article': "The definitions in this article",
    'article2': r'For purposes of this article,\s*the following definitions app
    'possible_definition': r'"[^"]*" \s+ means',
    'possible_definition2': r'"[^"]*" \s+ includes',
    'possible_definition3': r'"[^"]*" \s+',
    'possible_definition4': r'The word \s+ "[^"]*" ',
    'possible_definition5': r'"[^"]*" \s+ has the same meaning',
    'code_definitions': 'Definitions',
}

# Applying the function to the DataFrame
df = df.copy()
df['definition_label'] = df['node_text'].apply(lambda x: extract_quoted_text(x,

```

Checking Patterns

Analyzing the count of the extracted patterns provides a valuable overview of their distribution across the nodes. This quantitative analysis reveals the prevalence of certain legal terms or structures within the dataset, offering insights into the focus areas or scopes. The reason this is important is so that when I analyze the performance of the clusters I can reference back to this to see how well the model accurately clustered the definitions into something that might be useful for us.

```

In [7]: columns_to_exclude = ['id', 'node_type', 'top_level_title', 'sibling_nodes', 'i
encoded_columns = [col for col in df.columns if col not in columns_to_exclude]

# for column in encoded_columns:
#     count = df[column].sum()
#     print(f"Count of 1s in {column}: {str(count)}")

```

```

Count of 1s in chapter: 16
Count of 1s in chapter2: 19
Count of 1s in section: 540
Count of 1s in article: 6
Count of 1s in article2: 12
Count of 1s in possible_definition: 3099
Count of 1s in possible_definition2: 254
Count of 1s in possible_definition3: 4417
Count of 1s in possible_definition4: 19
Count of 1s in possible_definition5: 234
Count of 1s in code_definitions: 342
Count of 1s in definition_label: 4742

```

```

In [8]: df_encoded = df[encoded_columns]

```

KModes Model

The k-modes model is a clustering algorithm, similar to k-means, but specifically designed for categorical data. Unlike k-means, which uses means for clustering and is suitable for numerical data, k-modes uses modes. In this context, a mode represents the most frequent category in a set. The algorithm clusters data by minimizing the dissimilarities between data points in a cluster and the mode of that cluster. This is achieved through an iterative process where each data point is assigned to the cluster with the nearest mode, and then the modes are updated based on the current cluster memberships. The k-modes model is particularly useful in scenarios where data is not numerical, and traditional methods like k-means are not applicable. It's effective for segmenting data into distinct groups based on categorical attributes, offering valuable insights in fields like market segmentation, pattern recognition, and social science research.

In [9]:

```
n_clusters = 12

# Initialize the K-Modes model
km = KModes(n_clusters=n_clusters, init='Huang', n_init=4, verbose=1)
# Fit the model
clusters = km.fit_predict(df_encoded)
# Add the cluster labels to your original DataFrame
df['cluster'] = clusters
```

```
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 0, cost: 364.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 2, iteration: 1/100, moves: 921, cost: 512.0
Run 2, iteration: 2/100, moves: 12, cost: 512.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 3, iteration: 1/100, moves: 174, cost: 359.0
Run 3, iteration: 2/100, moves: 0, cost: 359.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 4, iteration: 1/100, moves: 878, cost: 1527.0
Best run was number 3
```

Preview of all the clusters

In [51]:

```
for i in range(n_clusters):
    print(f"Cluster {i}:")
    print(df[df['cluster'] == i]['node_text'])
```


Cluster 0:

- 01 (b) A notice to comply shall be the only means... \n (c)
- 2 (l) A person who receives a notice to c...
- 3 (d) If a notice to comply is issued, a single ...
- 4 (e) If a person who receives a notice to compl...
- (f) This section may not be construed as doing...
- 45693 ...
- 45694 (i) All administrative penalties collected fro...
- 45695 (j) The UPA shall consult with the district at...
- 45696 \n (k) (l) A unified program agency may suspen...
- 45697 (l) This section does not do any of the follow...
- (Amended by Stats. 2007, Ch. 626, Sec. 22. Eff...

Name: node_text, Length: 40956, dtype: object

Cluster 1:

- 239 (b) Notwithstanding the definition of "covered...
- 359 (e) Subdivision (c) does not apply to any of t...
- 367 \n (a) (l) Except as provided in paragraph (2)...
- 467 (a) A household hazardous waste collection fac...
- 468 (b) For the purposes of this section, "sharps ...
- 45273 ...
- 45317 "Agent," as used in this chapter, means a pers...
- 45478 (b) For the purposes of this chapter, "traveli...
- 45503 A person may not purport to be a CACO in this ...
- 45523 A person may not hold himself or herself out t...
- \n (d) (l) A manufacturer or an owner who viol...

Name: node_text, Length: 882, dtype: object

Cluster 2:

- 92 \n (d) (l) For the purpose of administering an...
- 962 Subject to an appropriation by the Legislature...
- 1832 (b) A violation or noncompliance by a federal ...
- 1898 (b) Upon issuing an order of quarantine pursua...
- 2061 (b) The following materials are not hazardous ...
- 44197 ...
- 44423 (c) For purposes of this section, a foster fam...
- 44497 (b) The State Department of Social Services ma...
- 44575 (f) The conviction of a licensee or the person...
- 45303 (c) The department may determine that no porti...
- For purposes of this section, the following te...

Name: node_text, Length: 128, dtype: object

Cluster 3:

- 6 \n (a) (l) All aspects of the unified program ...
- 32 (a) For purposes of this chapter, the followin...
- 43 (a) Notwithstanding Section 25117.2, "manageme...
- 44 (b) "Perchlorate" means all perchlorate-contai...
- 45 (c) "Perchlorate material" means perchlorate a...
- 45677 ...
- 45678 (r) "Release" means any spilling, leaking, pum...
- 45679 (s) "Responsible party" means a person describ...
- 45680 (t) "Site designation committee" means the com...
- 45682 (u) "State board" means the State Water Resour...
- \n (a) (l) All aspects of the unified program ...

Name: node_text, Length: 2800, dtype: object

Cluster 4:

- 95 (a) For the purposes of this section, the foll...
- 143 (c) "Major appliance" has the same meaning as ...
- 144 (d) "Materials that require special handling" ...
- 249 (a) For purposes of this section, "electronic ...
- 320 (j) "Owner or operator" has the same meaning g...
- 45225 ...
- (c) "Hazardous substance removal" has the same...

- 45461 (d) "Animal control officer" has the same mean... (d)
- 45486 "Code enforcement officer" has the same me... (i)
- 45579 "Hazardous material" has the same meaning ...
- 45670 (i) "Hazardous material" has the same meaning ...

Name: node_text, Length: 180, dtype: object

Cluster 5:

- 437 (c) "Paint" includes both oil-based paint and ...
- 1088 "License" includes, but is not limited to any,...
- 1104 "Natural resources" includes, but is not limit...
- 1109 (b) "Hazardous waste" includes, but is not lim...
- 1132 "Restricted hazardous waste" includes both of ...
- ...
- 44508 (c) All donors of sperm shall be screened and ...
- 44527 (a) Whenever there is a reasonable possibility...
- 44557 "Distribution" includes sale and exchange.
- 44579 As used in this chapter, "person" includes: la...
- 45525 (f) For purposes of this section, "provides fo...

Name: node_text, Length: 175, dtype: object

Cluster 6:

- 399 On and after July 1, 2017, a manufacturer shal...
- 793 \n (e) (1) The advisory committee shall recomm...
- 805 \n (e) (1) The advisory committee shall recomm...
- 1643 (d) For purposes of this section, "recycle" an...
- 2975 (h) For purposes of this section, the followin...
- ...
- 42454 (a) On or before January 31 of each year, the ...
- 43228 (b) Subsequent to being hired pursuant to subd...
- 43552 \n (c) (1) "Community living support services,...
- 44794 (k) For purposes of this section, "serious ill...
- 45251 (d) This section shall not be construed to app...

Name: node_text, Length: 64, dtype: object

Cluster 7:

- 121 (a) For purposes of this section, the followin...
- 238 (a) For purposes of this section, the followin...
- 413 (a) For purposes of this section, "spent dry c...
- 473 (b) For purposes of this section, "surplus hou...
- 504 (a) For purposes of this section, the followin...
- ...
- 44780 (b) For purposes of this section, the followin...
- 45092 (a) For purposes of this section, the followin...
- 45096 (b) For purposes of this section, the followin...
- 45127 (c) The subject or subject's conservator or gu...
- 45516 (d) For purposes of this section, the followin...

Name: node_text, Length: 299, dtype: object

Cluster 8:

- 42 For purposes of this article, the following de...
- 310 For purposes of this article, the following de...
- 331 For purposes of this article, the following de...
- 434 For purposes of this article, the following de...
- 719 (c) For purposes of subdivision (b), the follo...
- ...
- 44537 Unless the context requires otherwise, the fol...
- 44752 For the purposes of this chapter, the followin...
- 44857 As used in this chapter, the following definit...
- 45038 As used in this chapter, the following definit...
- 45222 For purposes of this chapter, the following de...

Name: node_text, Length: 174, dtype: object

Cluster 9:

- 5963 As used in this article the word "taxes" shall...
- 20966 The boat shall have the word "lifeboat" plainl...

22058 (c) Biohazardous waste, except as provided in ... (d)
22069 Label sharps containers with the words "sh... (e) Each
22262 procedure area shall have a sharps wa...
23151 Any food is misbranded if it is offered for sa...
23203 (a) By identifying each organically produced i...
24130 (a) An internet food service intermediary that...
24132 \n (c) (1) A microenterprise home kitchen oper...
33726 Lettering on this label shall contrast with th...
33846 Lettering on the label shall be black on a red...
34220 (b) In the case of a guardianship or conservat...
34261 (d) All manufactured homes, mobilehomes, comme...
34268 (b) A manufactured home, mobilehome, commercia...
35395 (a) No person or public agency shall advertise...
35518 (c) Nothing in this article shall be construed...
35519 (d) Any person or public agency using the word...

Name: node_text, dtype: object

Cluster 10:

3102 For purposes of this chapter, the following de...
3190 For purposes of this chapter, the following de...
7445 For purposes of this chapter, the following de...
7667 For purposes of this chapter, the following de...
10321 For purposes of this chapter, the following de...
20618 For purposes of this chapter, the following de...
21252 For purposes of this chapter, the following de...
22328 For purposes of this chapter, the following de...
22416 For purposes of this chapter, the following de...
23604 For purposes of this chapter, the following de...
25527 For purposes of this chapter, the following de...
28574 For purposes of this chapter, the following de...
45349 For purposes of this chapter, the following de...

Name: node_text, dtype: object

Cluster 11:

939 For purposes of this article, the following de...
1315 For purposes of this article, the following de...
1566 For purposes of this article, the following de...
1993 For purposes of this article, the following de...
2185 For purposes of this article, the following de...
2257 For purposes of this article, the following de...
7702 For purposes of this article, the following de...
10022 For purposes of this article, the following de...
14290 For purposes of this article, the following de...
26696 For purposes of this article, the following de...

Name: node_text, dtype: object

Looking at Distribution of nodes in clusters

In [11]:

```
cluster_sizes = df['cluster'].value_counts()  
cluster_sizes
```

Out[11]: cluster

```
0    40956
3     2800
1     882
7     299
4     180
5     175
8     174
2     128
6      64
9      17
10     13
11     10
```

Name: count, dtype: int64

Looking at centroids that determed clusters

```
In [12]: kmodes = KModes(n_clusters)
kmodes.fit(df_encoded)

# Get the centroids
centroids = kmodes.cluster_centroids_
# Examine each centroid
for i, centroid in enumerate(centroids):

    print(f"Centroid {i}: {centroid}")
```

```
Centroid 0: [0 0 0 0 0 0 0 0 0 0 0]
Centroid 1: [0 0 1 0 0 1 0 1 0 0 0 1]
Centroid 2: [0 0 0 0 0 1 0 1 0 0 0 1]
Centroid 3: [0 0 0 0 0 0 0 1 0 0 0 1]
Centroid 4: [0 0 0 0 0 0 0 0 0 0 0 1 1]
Centroid 5: [0 0 0 0 0 1 0 1 0 0 1 1]
Centroid 6: [0 0 0 0 0 0 1 1 0 1 1 1]
Centroid 7: [0 0 1 0 0 0 0 1 0 0 1 1]
Centroid 8: [0 0 0 0 0 0 1 1 0 0 0 1]
Centroid 9: [0 0 0 0 0 0 0 1 1 0 0 1]
Centroid 10: [0 0 1 0 0 0 0 0 0 0 0 1]
Centroid 11: [0 0 0 0 1 0 0 0 0 0 0 1 1]
```

```
In [74]: df2 = df.copy()
```

Silhouette Score & DB

```
In [14]: score = silhouette_score(df_encoded, clusters, metric='hamming')
score
```

Out[14]: 0.9888633236660732

```
In [15]: db_score = davies_bouldin_score(df_encoded, clusters)
print("The Davies-Bouldin score is:", db_score)
```

The Davies-Bouldin score is: 0.3783047609587911

Silhouette Score: 0.9888633236660732

Davies Bouldin Score: 0.3783047609587911

Both of these scores together indicate a highly effective clustering outcome. The high Silhouette Score points to clear, well-separated clusters, and the low Davies-Bouldin Score suggests that these clusters are both compact and distinct from each other. This combination of scores is typically indicative of a successful clustering model where the chosen clusters are both meaningful and well-defined.

TSNE Plot

Cluster Separation: Most clusters appear well-separated, especially clusters 0, 1, 9, and 11, which are farther from the center. This indicates good distinction between different groups in the data.

Cluster Density: Clusters like 3 and 5 show a tight grouping of points, suggesting high similarity within these clusters. Conversely, clusters like 2 and 8 seem more spread out, indicating more variation within those clusters.

Outliers: There are some potential outliers, particularly within clusters 2, 8, and 11, where points are distant from their cluster centers. These could represent nodes that don't fit well with others in their assigned cluster or could be special cases.

Overlap: Some clusters, such as 6, 7, and 8, have regions that are close to each other, indicating some potential overlap in the node's characteristics.

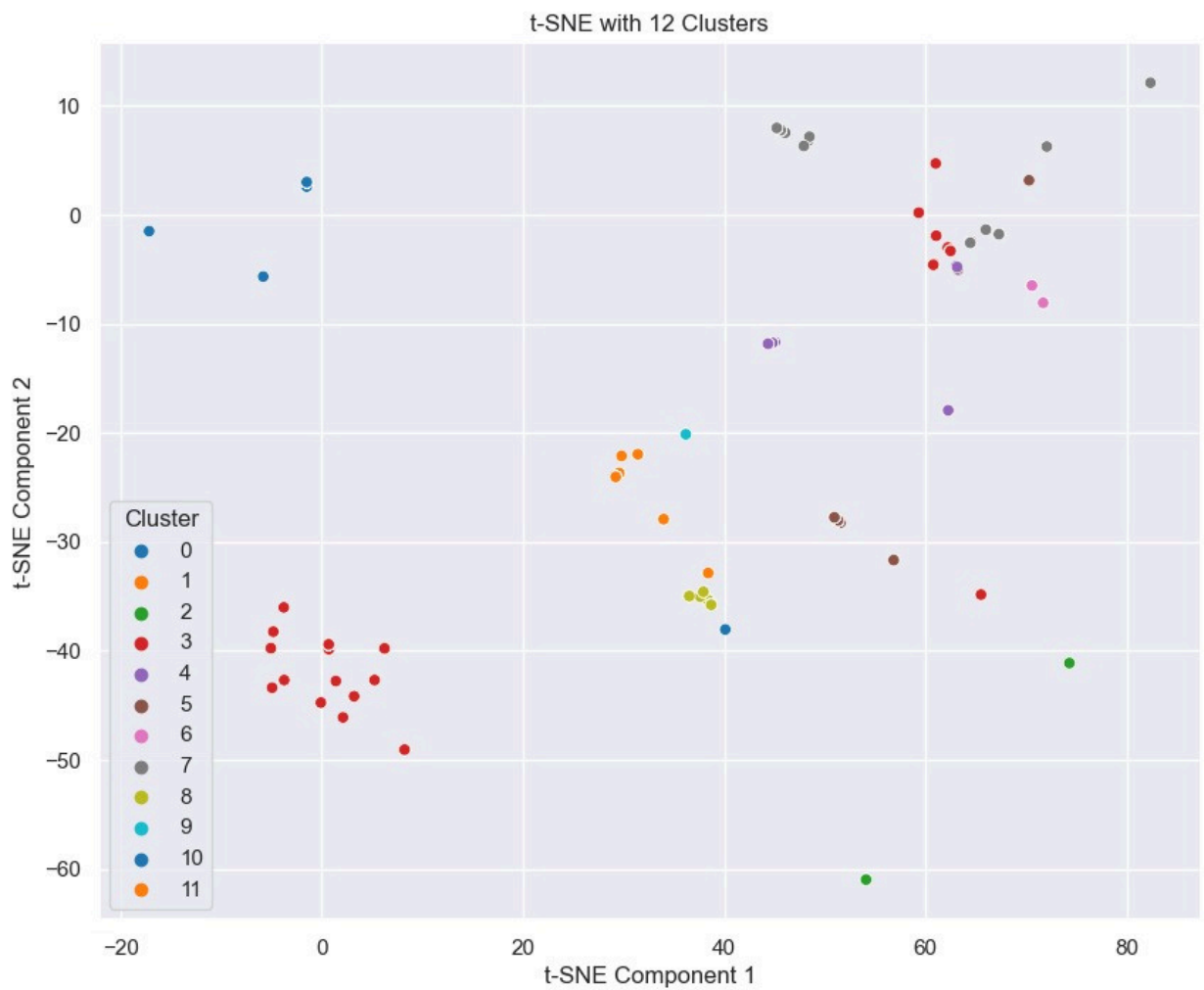
Given the high silhouette score, these clusters most likely represent meaningful groupings, despite a few areas of overlap and some spread in certain clusters.

In [16]:

```
tsne = TSNE()
df_tsne = tsne.fit_transform(df_encoded)      # df_encoded should be your pre-proce

# Create a scatterplot
plt.figure(figsize=(10, 8))                  # You can adjust the figure size
sns.scatterplot(x=df_tsne[:, 0], y=df_tsne[:, 1], hue=df['cluster'], palette='t
plt.title('t-SNE with 12 Clusters')
plt.xlabel('t-SNE Component 1')
plt.ylabel('t-SNE Component 2')
plt.legend(title='Cluster')
```

Out[16]:

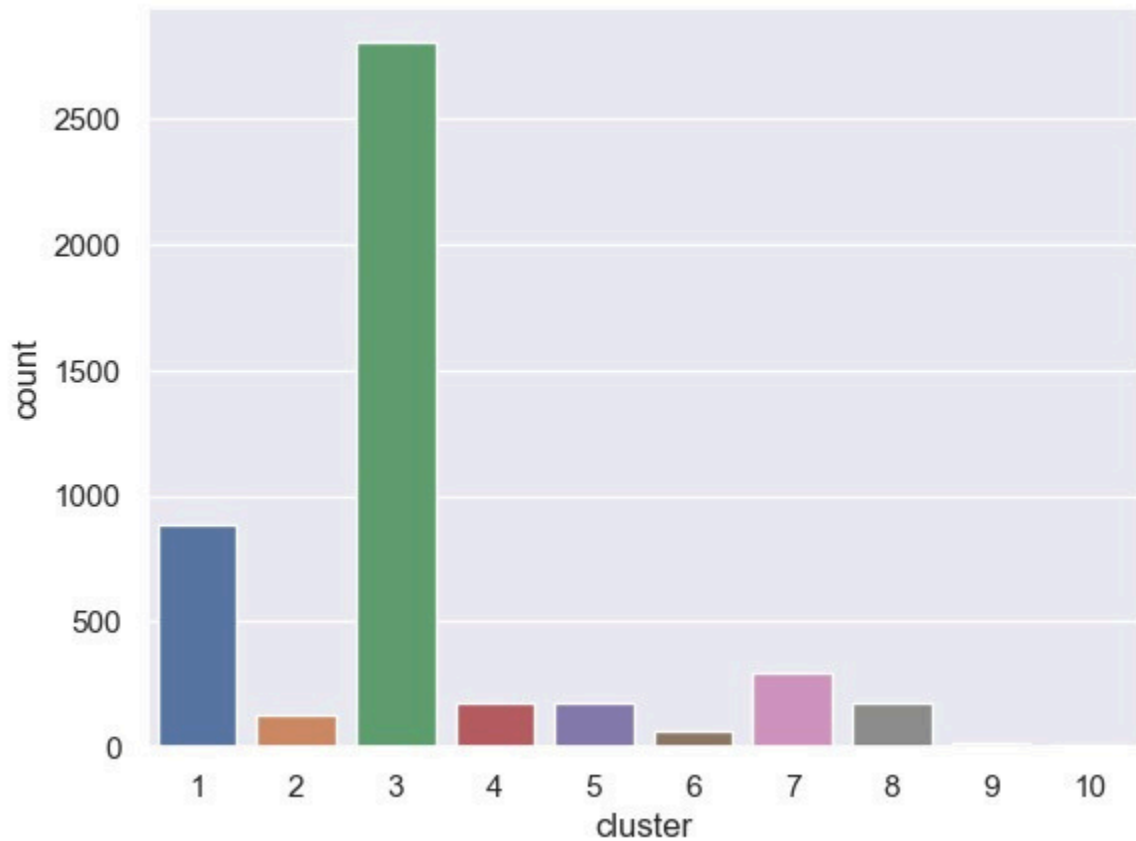


Cluster Node Count Spread Visual

In [20]:

```
filtered_df = df[(df['cluster'] != 11) & (df['cluster'] != 0)]  
  
sns.countplot(x='cluster', data=filtered_df)
```

Out[20]:



In [22]:

```
pattern_columns = [  
    'chapter', 'chapter2', 'section', 'article', 'article2',  
    'possible_definition', 'possible_definition2', 'possible_definition3',  
    'possible_definition4', 'possible_definition5', 'code_definitions'  
]  
  
# Dictionary to store cross-tabulation DataFrames  
cross_tabs = {}  
# Generate cross-tabulation for each pattern column  
for i, pattern_column in enumerate(pattern_columns, start=1):  
    cross_tabs[f'cross{i}'] = pd.crosstab(df['cluster'], df[pattern_column])  
  
for name, cross_tab in cross_tabs.items():  
    print(f"{name}:\n{cross_tab}\n")
```

cross1:
chapter 0 1
cluster 40956 0
0 882 0
1 128 0
2 2798 2
3 180 0
4 175 0
5 64 0
6 298 1
7 174 0
8 17 0
9 0 13
10 10 0
11

cross2:
chapter2
cluster 40956 0 1
0 882 0
1 128 0
2 2782 18
3 180 0
4 175 0
5 64 0
6 298 1
7 174 0
8 17 0
9 13 0
10 10 0
11

cross3:
section
cluster 40956 0 1
0 882 0
1 0 128
2 2800 0
3 155 25
4 151 24
5 0 64
6 0 299
7 174 0
8 17 0
9 13 0
10 10 0
11

cross4:
article
cluster 40956 0 1
0 882 0
1 128 0
2 2799 1
3 180 0
4 175 0
5 64 0
6 299 0
7 169 5
8

9 10 11	17	0
	13	0
cross5:	10	0
article2		
cluster		
0		
1	40956 ⁰	1
2	882	0
3	128	0
4	2798	2
5	180	0
6	175	0
7	64	0
8	299	0
9	174	0
10	17	0
11	13	0
	0	10

cross6:		
possible_definition		
cluster		
0		
1	40956 ⁰	1
2	882	0
3	128	0
4	0	2800
5	180	0
6	175	0
7	64	0
8	0	299
9	174	0
10	17	0
11	13	0
	10	0

cross7:		
possible_definition2		
cluster		
0		
1	40956 ⁰	1
2	882	0
3	128	0
4	2737	63
5	178	2
6	0	175
7	64	0
8	285	14
9	174	0
10	17	0
11	13	0
	10	0

cross8:		
possible_definition3		
cluster		
0		
1	40956 ⁰	1
2	0	882
3	128	0
4	0	2800
	0	180

```

5 6 7          0 175
8              0 64
9              0 299
10             174 0
11             0 17
              13 0
cross9:
              10 0
possible_definition4
cluster
0
1             40956 0 1
2             882 0
3             128 0
4             2798 2
5             180 0
6             175 0
7             64 0
8             299 0
9             174 0
10            0 17
11            13 0
              10 0

cross10:
possible_definition5
cluster
0
1             40956 0 1
2             882 0
3             128 0
4             2773 27
5             0 180
6             175 0
7             64 0
8             272 27
9             174 0
10            17 0
11            13 0
              10 0

cross11:
code_definitions
cluster
0
1             40956 0 1
2             878 4
3             127 1
4             2746 54
5             178 2
6             173 2
7             59 5
8             222 77
9             0 174
10            17 0
11            0 13
              0 10

```

In [23]:

```

pattern_columns = [
    'chapter', 'chapter2', 'section', 'article', 'article2',
    'possible_definition', 'possible_definition2', 'possible_definition3',

```

```

'possible_definition4', 'possible_definition5', 'code_definitions'
]

cross_tabs = {}
for i, pattern_column in enumerate(pattern_columns, start=1):
    cross_tabs[f'cross_{i}'] = pd.crosstab(df['cluster'], df[pattern_column])

cluster_dataframes = {f'cluster_{i}': pd.DataFrame() for i in range(12)}

for i in range(12):

    for name, cross_tab in cross_tabs.items():

        cluster_row = cross_tab.loc[i].to_frame().T
        cluster_row.index = [name]
        cluster_dataframes[f'cluster_{i}'] = pd.concat([cluster_dataframes[f'cl

```

```

In [24]: cluster0 = cluster_dataframes['cluster_0'].drop(cluster_dataframes['cluster_0']
cluster1 = cluster_dataframes['cluster_1'].drop(cluster_dataframes['cluster_1']
cluster2 = cluster_dataframes['cluster_2'].drop(cluster_dataframes['cluster_2']
cluster3 = cluster_dataframes['cluster_3'].drop(cluster_dataframes['cluster_3']
cluster4 = cluster_dataframes['cluster_4'].drop(cluster_dataframes['cluster_4']
cluster5 = cluster_dataframes['cluster_5'].drop(cluster_dataframes['cluster_5']
cluster6 = cluster_dataframes['cluster_6'].drop(cluster_dataframes['cluster_6']
cluster7 = cluster_dataframes['cluster_7'].drop(cluster_dataframes['cluster_7']
cluster8 = cluster_dataframes['cluster_8'].drop(cluster_dataframes['cluster_8']
cluster9 = cluster_dataframes['cluster_9'].drop(cluster_dataframes['cluster_9']
cluster10 = cluster_dataframes['cluster_10'].drop(cluster_dataframes['cluster_1
cluster11 = cluster_dataframes['cluster_11'].drop(cluster_dataframes['cluster_1

```

```

In [27]: def background(row):

    highlight = 'background-color: mediumseagreen;'
    highlight2 = 'background-color: palegreen;'
    default = ""

    # must return one string per cell in this row
    if row[1] > row[0]:
        return [highlight2, default]
    elif row[1] < row[0]:
        return [highlight, default]
    else:
        return [default, default]

```

```

In [28]: styled_cluster0 = cluster0.style.apply(background, subset=[1, 0], axis=1)
styled_cluster1 = cluster1.style.apply(background, subset=[1, 0], axis=1)
styled_cluster2 = cluster2.style.apply(background, subset=[1, 0], axis=1)
styled_cluster3 = cluster3.style.apply(background, subset=[1, 0], axis=1)
styled_cluster4 = cluster4.style.apply(background, subset=[1, 0], axis=1)
styled_cluster5 = cluster5.style.apply(background, subset=[1, 0], axis=1)
styled_cluster6 = cluster6.style.apply(background, subset=[1, 0], axis=1)
styled_cluster7 = cluster7.style.apply(background, subset=[1, 0], axis=1)
styled_cluster8 = cluster8.style.apply(background, subset=[1, 0], axis=1)
styled_cluster9 = cluster9.style.apply(background, subset=[1, 0], axis=1)

```

```
styled_cluster10 = cluster10.style.apply(background, subset=[1, 0], axis=1)
styled_cluster11 = cluster11.style.apply(background, subset=[1, 0], axis=1)
```

Cross Tabulation

Cross1 - "For purposes of this chapter, the following definitions apply"

Cross2 - "For purposes of this chapter, ' ' means"

Cross3 - "For purposes of this section"

Cross4 - "The definitions in this article"

Cross5 - "For purposes of this article, the following definitions apply:"

Cross6 - "' ' means"

Cross7 - "' ' includes"

Cross8 - "' '"

Cross9 - "The word ' '"

Cross10 - "' ' has the same meaning"

Cross11 - "Definitions"

As observed before each cluster contains these total records

```
| cluster | count | |-----|-----| | 0 | 40956 | | 3 | 2800 | | 1 | 882 | | 7 | 299 | | 4 | 180 | | 5 |
175 | | 8 | 174 | | 2 | 128 | | 6 | 64 | | 9 | 17 | | 10 | 13 | | 11 | 10 |
```

In [53]:

```
def display_side_by_side(*args,titles=cycle([''])):
    html_str=""
    for df,title in zip(args, chain(titles,cycle(['<br>'])) ):
        html_str+='{<th style="text-align:center"><td style="vertical-align:top"
html_str+=f'<h2 style="text-align: center;">{title}</h2>'
html_str+=df.to_html().replace('table','table style="display:inline"')
html_str+='{</td></th>}'
    display_html(html_str,raw=True)

display_side_by_side(styled_cluster0, styled_cluster1, styled_cluster2, styled_
display_side_by_side(styled_cluster6, styled_cluster7, styled_cluster8, styled_
```

Cluster 0

	0	1
--	---	---

Cluster 1

	0	1
cross8	0	882
cross11	878	4

Cluster 2

	0	1
cross3	0	128
cross11	127	1

Cluster 3

	0	1
cross1	2798	2
cross2	2782	18
cross4	2799	1
cross5	2798	2
cross6	0	2800
cross7	2737	63
cross8	0	2800
cross9	2798	2
cross10	2773	27
cross11	2746	54

Cluster 4

	0	1
cross3	155	25
cross7	178	2
cross8	0	180
cross10	0	180
cross11	178	2

Cluster 5

	0	1
cross3	151	24
cross7	0	175
cross8	0	175
cross11	173	2

Cluster 6

	0	1
cross3	0	64
cross8	0	64
cross11	59	5

Cluster 7

	0	1
cross1	298	1
cross2	298	1
cross3	0	299
cross6	0	299
cross7	285	14
cross8	0	299
cross10	272	27
cross11	222	77

Cluster 8

	0	1
cross4	169	5
cross11	0	174

Cluster 9

	0	1
cross8	0	17
cross9	0	17

Cluster 10

	0	1
cross1	0	13
cross11	0	13

Cluster 11

	0	1
cross5	0	10
cross11	0	10

CrossTab Examination on only crosses marked 1 or 'yes'

Cluster 0 - This cluster of nodes are not definitions

Cluster 1 - These nodes could potentially host text that is meaningful in searching for definitions however, need to further investigate because there is potential that words within quotes are not necessarily definition nodes however 4 nodes aligned with text pattern Definitions. Will need to look at manually for more information.

Cluster 2 - These nodes are definitions sections of scope sections and potentially entire Code Definition

Cluster 3 - high likelihood this cluster is definition scope article based on cross tabulation

Cluster 4 - These nodes are definitions, unsure of scope, possibly section based on overlap of patterns w/ section pattern

Cluster 5 - These are potentially definitions or context definitions, with high probability of being scope Section.

Cluster 6 - These are potentially definitions or context definitions, with high likelihood of being scope Section. Unlike cluster 5 however, cluster 6 has fewer overlapping patterns.

Cluster 7 - These are most likely a mix of definitions with a majority being of scope Section and a few outliers of scope chapter.

Cluster 8 - High likelihood this cluster is would be the determining nodes of scope article.

Cluster 9 - High likelihood the nodes in this cluster is definitions of scope article

Cluster 10 - This cluster hosts the determining nodes of scope article of scope 'chapters'. This is highly valuable because could lead to the nodes that are specific definitions underneath that scope. Looking at the cluster 10 summary this matches.

Cluster 11 - This cluster is definitions of scope 'articles'. This is highly valuable because could lead to the nodes that are specific definitions underneath those nodes. Looking at the cluster 11 summary this matches.

Manually looking through each cluster to see patterns of clusters/definitions

In [60]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df[df['cluster'] == 0]['node_text'])
```


0 (b) A notice to comply shall be the only means by which a UPA may cite a minor violation, unless the person cited fails to correct the violation or fails to submit the certification of correction within the time period prescribed in the notice, in which case the UPA may take any enforcement action, including imposing a penalty, as authorized by this chapter.

1

(c) (1) A person who receives a notice to comply detailing a minor violation shall have not more than 30 days from the date of the notice to comply in which to correct any violation cited in the notice to comply. Within five working days of correcting the violation, the person cited or an authorized representative shall sign the notice to comply, certifying that any violation has been corrected, and return the notice to the UPA. (2) A false certification that a violation has been corrected is punishable as a misdemeanor. (3) The effective date of the certification that any violation has been corrected shall be the date that it is postmarked.

2

(d) If a notice to comply is issued, a single notice to comply shall be issued for all minor violations noted during the inspection, and the notice to comply shall list all of the minor violations and the manner in which each of the minor violations may be brought into compliance.

3

(e) If a person who receives a notice to comply pursuant to subdivision (a) disagrees with one or more of the alleged violations listed on the notice to comply, the person shall provide the UPA a written notice of disagreement along with the returned signed notice to comply. If the person disagrees with all of the alleged violations, the written notice of disagreement shall be returned in lieu of the signed certification of correction within 30 days of the date of issuance of the notice to comply. If the issuing agency takes administrative enforcement action on the basis of the disputed violation, that action may be appealed in the same manner as any other alleged violation under Section 25404.

1.1.

4

(f) This section may not be construed as doing any of the following: (1) Preventing the reinspection of a facility to ensure compliance with this chapter or to ensure that minor violations cited in a notice to comply have been corrected and that the facility is in compliance with those laws and regulations within the jurisdiction of the UPA. (2) Preventing the UPA from requiring a person to submit necessary documentation needed to support the person's claim of compliance pursuant to subdivision (c). (3) Restricting the power of a city attorney, district attorney, county counsel, or the Attorney General to bring, in the name of the people of California, any criminal proceeding otherwise authorized by law. (4) Preventing the UPA from cooperating with, or participating in, a proceeding specified in paragraph (3).

...

45693

(i) All administrative penalties collected from actions brought by a UPA pursuant to this section shall be paid to the UPA that imposed the penalty, and shall be deposited into a special account that shall be expended to fund the activities of the UPA in enforcing this chapter.

45694

(j) The UPA shall consult with the district attorney, county counsel, or city attorney on the development of policies to be followed in exercising the authority delegated pursuant to this section as it relates to the authority of the UPA to issue orders.

45695 (k) (1) A unified program agency may suspend or revoke any unified program facility permit, or an element of a unified program facility permit, for not paying the permit fee or a fine or penalty associated with the permit in accordance with the procedures specified in this subdivision. (2) If a per

mittee does not comply with a written notice from the unified program agency to the permittee to make the payments specified in paragraph (1) by the required date provided in the notice, the unified program agency may suspend or revoke the permit or permit element. If the permit or permit element is suspended or revoked, the permittee shall immediately discontinue operating that facility or function of the facility to which the permit element applies until the permit is reinstated or reissued.\n (3) A permittee may request a hearing to appeal the suspension or revocation of a permit or element of a permit pursuant to this subdivision by requesting a hearing using the procedures provided in subdivision (d).

45696

(l) This section does not do any of the following:\n (1) Otherwise affect the authority of a UPA to take any other action authorized by any other provision of law, except the UPA shall not require a person to pay a penalty pursuant to this section and pursuant to a local ordinance for the same violation.\n (2) Restrict the power of a city attorney, district attorney, county counsel, or the Attorney General to bring, in the name of the people of California, any criminal proceeding otherwise authorized by law.\n (3) Prevent the UPA from cooperating with, or participating in, a proceeding specified in paragraph (2).

45697

(Amended by Stats. 2007, Ch. 626, Sec. 22. Effective January 1, 2008.)

Name: node_text, Length: 40956, dtype: object

In [61]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 1]['node_text'])
```

239 (b) Notwithstanding the definition of "covered electronic device" in paragraph (2) of subdivision (a), the obligations of the department established in subdivisions (c) to (f), inclusive, apply only to covered electronic devices specified in subparagraph (A) of paragraph (1) of subdivision (g) of Section 42463 of the Public Resources Code.

359

(e) Subdivision (c) does not apply to any of the following:\n (1) A person whose ordinary course of business does not include the sale of lead-acid batteries.\n (2) A person that does not sell lead-acid batteries directly to consumers, such as over-the-counter, but instead removes nonfunctional or damaged batteries and installs new lead-acid batteries as a part of an automotive repair dealer service.\n (3) A business that removes lead-acid batteries and installs new lead-acid batteries as a part of roadside services. "Roadside services," for purposes of this paragraph, means the services performed upon a motor vehicle for the purpose of transporting the vehicle or to permit it to be operated under its own power, by or on behalf of a motor club holding a certificate of authority pursuant to Chapter 2 (commencing with Section 12160) of Part 5 of Division 2 of the Insurance Code.

367 \n (a) (1) Except as provided in paragraph (2), the lead-acid battery fees imposed pursuant to Sections 25215.25 and 25215.35 shall be collected by the California Department of Tax and Fee Administration in accordance with the Fee Collection Procedures Law (Part 30 (commencing with Section 55001) of Division 2 of the Revenue and Taxation Code). For the purposes of this section, the reference to "feepayer" shall include a dealer and manufacturer.\n (2) Notwithstanding the petition for redetermination and claim for refund provisions of the Fee Collection Procedures Law (Article 3 (commencing with Section 55081) of Chapter 3 of, and Article 1 (commencing with Section 55221) of Chapter 5 of, Part 30 of Division 2 of the Revenue and Taxation Code), the California Department of Tax and Fee Administration shall not do either of the following:\n (A) Accept or consider any petition for redetermination of fees determined under this article if the petition is founded upon the grounds that a battery is or is not a lead-acid battery, as defined in Section 25215.1. The California Department of Tax and Fee Administration shall forward to the department any petition for redetermination that is based on those grounds.\n (B) Accept or consider a claim for refund of fees paid pursuant to this article, if the claim for refund is founded upon the grounds that a battery is or is not a lead-acid battery, as defined in Section 25215.1. The California Department of Tax and Fee Administration shall forward to the department any claim for refund that is based on these grounds.

467

(a) A household hazardous waste collection facility that has a permit issued under Section 25218.8 may operate as a "home-generated sharps consolidation point," as defined in subdivision (b) of Section 117904, if the facility is approved by the enforcement agency as a point of consolidation pursuant to Section 117904 and the facility complies with the provisions of that section.

468

(b) For the purposes of this section, "sharps waste" has the meaning defined in Section 40190.5 of the Public Resources Code.

...

45273

"Agent," as used in this chapter, means a person acting in accordance with Title 9 (commencing with Section 2295) of Part 4 of Division 3 of the Civil Code for purposes of managing, operating, leasing, or performing a similar function with respect to a building subject to this chapter.

45317

(b) For the purposes of this chapter, "traveling circus or carnival" does not include any fair regulated under Chapter 4 (commencing with Section 19400) of Division 8 of the Business and Professions Code, or any rodeo, horse, or school

l event. 45478 A person may not purport to be a CACO in this state or use the title “certified animal control officer” in this state unless the person holds a valid certificate of registration pursuant to this chapter.

45503

A person may not hold himself or herself out to be a Certified Code Enforcement Officer in this state or use the title “Certified Code Enforcement Officer” in this state unless the person holds a certificate of registration pursuant to this chapter.

45523

\n (d) (1) A manufacturer or an owner who violates this section shall be guilty of an infraction punishable by a fine of one hundred dollars (\$100) for the first violation of this section, five hundred dollars (\$500) for the second violation of this section, and one thousand dollars (\$1,000) for the third violation, or any subsequent violation, of this section.\n (2) For purposes of this subdivision, a violation means the failure to mark an individual trash receptacle or storage container as required pursuant to subdivision (a), (b), or (c).\n (3) For purposes of this subdivision, “owner” shall not mean a local governmental entity.

Name: node_text, Length: 882, dtype: object

In [62]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 2]['node_text'])
```

92 \n (d) (1) For the purpose of administering and enforcing this article, an authorized representative of the department, upon obtaining consent or after obtaining an inspection warrant pursuant to Title 13 (commencing with Section 1822.50) of Part 3 of the Code of Civil Procedure, may, upon presenting appropriate credentials and at a reasonable time, do any of the following:\n (A) Enter a factory, warehouse, or establishment where jewelry is manufactured, packed, held, or sold; enter a vehicle that is being used to transport, hold, or sell jewelry; or enter a place where jewelry is being held or sold.\n (B) Inspect a factory, warehouse, establishment, vehicle, or place described in subparagraph (A), and all pertinent equipment, raw material, finished and unfinished materials, containers, and labeling in the factory, warehouse, establishment, vehicle, or place. In the case of a factory, warehouse, or establishment where jewelry is manufactured, packed, held, or sold, this inspection shall include any record, file, paper, process, control, and facility that has a bearing on whether the jewelry is being manufactured, packed, held, transported, sold, or offered for sale or for promotional purposes in violation of this article.\n \n (2) (A) An authorized representative of the department may secure a sample of jewelry when taking an action authorized pursuant to this subdivision. If the representative obtains a sample prior to leaving the premises, he or she shall leave a receipt describing the sample obtained.\n (B) The department shall return, upon request, a sample that is not destroyed during testing when the department no longer has any purpose for retaining the sample.\n (C) A sample that is secured in compliance with this section and found to be in compliance with this article that is destroyed during testing shall be subject to a claim for reimbursement.\n (3) An authorized representative of the department shall have access to all records of a carrier in commerce relating to the movement in commerce of jewelry, or the holding of that jewelry during or after the movement, and the quantity, shipper, and consignee of the jewelry. A carrier shall not be subject to the other provisions of this article by reason of its receipt, carriage, holding, or delivery of jewelry in the usual course of business as a carrier.\n (4) An authorized representative of the department shall be deemed to have received implied consent to enter a retail establishment, for purposes of this section, if the authorized representative enters the location of that retail establishment where the public is generally granted access.

962

Subject to an appropriation by the Legislature for purposes of this section, the department shall include in each Priority Product Work Plan, commencing with the 2024–26 Priority Product Work Plan, in addition to any other information that the department is required to include pursuant to Section 69503.4 of Title 22 of the California Code of Regulations, or any successor regulation, a brief description of all of the following information:

1832

(b) A violation or noncompliance by a federal hazardous waste facility, pursuant to Section 6961 of Title 42 of the United States Code, shall, for purposes of this section, be limited to a violation or noncompliance caused by an action or inaction within the boundaries identified in Part B of the federal hazardous waste permit application, pursuant to Section 270.14 of Title 40 of the Code of Federal Regulations, for that facility.

1898

(b) Upon issuing an order of quarantine pursuant to subdivision (a), the authorized agent shall notify the person who owns the hazardous waste, or the owner or lessee of the vehicle in which the wastes are transported, of all of the following:\n (1) The hazardous waste has been subject to a quarantine order because the hazardous waste is, or is suspected of being, stored, transported, disposed of, or handled in violation of this chapter.\n (2) No person shall remove, transfer, or dispose of the hazardous waste until permission for removal, transfer, or disposal is given by an authorized agent of the department or by a court.\n (3) The person so notified may request, and shall be granted, an immediate hearing before a person designated by the director to review the validity

y of the authorized agent's order. For purposes of this section, an immediate hearing shall be held within 24 hours after a hearing is requested by the person subject to the order.

2061

(b) The following materials are not hazardous wastes for purposes of this section:\n (1) Hazardous materials that are recycled, and used onsite, and are not transferred offsite.\n (2) Aqueous waste treated in a treatment unit operating, or that subsequently operates, pursuant to a permit-by-rule, or pursuant to Section 25200.3 or 25201.5. However, hazardous waste generated by a treatment unit treating waste pursuant to a permit-by-rule, by a unit that subsequently obtains a permit-by-rule, or other authorization pursuant to Section 25200.3 or 25201.5 is hazardous waste for purposes of this section.

...

44197

(c) For purposes of this section, a foster family agency is defined in paragraph (4) of subdivision (a) of Section 1502.

~~44473~~ The State Department of Social Services may license, subject to the following conditions, an Adult Residential Facility for Persons with Special Health Care Needs to provide 24-hour services to up to five adults with developmental disabilities who have special health care and intensive support needs, as defined in subdivisions (f) and (g) of Section 4684.50 of the Welfare and Institutions Code.\n (1) The State Department of Developmental Services shall be responsible for granting the certificate of program approval for an Adult Residential Facility for Persons with Special Health Care Needs (ARFPSHN). The State Department of Social Services shall not issue a license unless the applicant has obtained a certification of program approval from the State Department of Developmental Services.\n (2) The State Department of Social Services shall ensure that the ARFPSHN meets the administration requirements under Article 2 (commencing with Section 1520) including, but not limited to, requirements relating to fingerprinting and criminal records under Section 1522 and administrator certification requirements of an adult residential facility pursuant Section 1562.3, including, but not limited to, the following:\n (A) Successfully complete a department-approved administrator certification training program requiring a minimum of 35 hours of instruction conducive to learning in which participants are able to simultaneously interact with each other as well as with the instructor, and that provides training on a uniform core of knowledge under Section 1562.3.\n (B) Unless an extension is granted to the applicant by the department, an applicant for an administrator's certificate shall, within 60 days of the applicant's completion of instruction, pass the examination provided for in Section 1562.3.\n (C) Submit an application for administrator certification to the department to include:\n (i) An administrator certification application.\n (ii) A certificate of completion of the administrator certification training program required pursuant to this section.\n (iii) The fee for processing an administrator certification application, including the issuance of the administrator certificate, as specified in 1562.3.\n (iv) Documentation that the applicant has passed the examination.\n (3) The State Department of Social Services shall administer employee actions under Article 5.5 (commencing with Section 1558).\n (4) The regional center shall monitor and enforce compliance of the program and health and safety requirements, including monitoring and evaluating the quality of care and intensive support services. The State Department of Developmental Services shall ensure that the regional center performs these functions.\n (5) The State Department of Developmental Services may decertify any ARFPSHN that does not comply with program requirements. When the State Department of Developmental Services determines that urgent action is necessary to protect clients of the ARFPSHN from physical or mental abuse, abandonment, or any other substantial threat to their health and safety, the State Department of Developmental Services may request the regional center or centers to remove the clients from the ARFPSHN or direct the regional center or centers to obtain alternative services for the consumers within 24 hours.\n (6) The Sta

te Department of Social Services may initiate proceedings for temporary suspension of the license pursuant to Section 1550.5.\n (7) The State Department of Developmental Services, upon its decertification, shall inform the State Department of Social Services of the licensee's decertification, with its recommendation concerning revocation of the license, for which the State Department of Social Services may initiate proceedings pursuant to Section 1550.\n (8) The State Department of Developmental Services and the regional centers shall provide the State Department of Social Services all available documentation and evidentiary support necessary for any enforcement proceedings to suspend the license pursuant to Section 1550.5, to revoke or deny a license pursuant to Section 1551, or to exclude an individual pursuant to Section 1558.\n (9) The State Department of Social Services Community Care Licensing Division shall enter into a memorandum of understanding with the State Department of Developmental Services to outline a formal protocol to address shared responsibilities, including monitoring responsibilities, complaint investigations, administrative actions, and closures.\n (10) The licensee shall provide documentation that, in addition to the administrator requirements set forth under paragraph (4) of subdivision (a) of Section 4684.63 of the Welfare and Institutions Code, the administrator, prior to employment, has completed a minimum of 35 hours of initial training in the general laws, regulations and policies and procedural standards applicable to facilities licensed by the State Department of Social Services under Article 2 (commencing with Section 1520).\n (11) An administrator certificate issued under this section shall expire every two years, on the anniversary date of the initial issuance of the certificate.\n (12) An administrator certificate issued under this section shall be renewed every two years and renewal shall be conditional upon the certificate holder submitting documentation of completion of 40 hours of continuing education related to the uniform core of knowledge specified in Section 1562.3. No more than one-half of the required 40 hours of continuing education necessary to renew the certificate may be satisfied through self-paced courses. All other continuing education hours shall be completed in an instructional setting conducive to learning in which participants must be able to simultaneously interact with each other as well as with the instructor. For purposes of this section, an individual who is an adult residential facility administrator and who is required to complete the continuing education hours required by the regulations of the State Department of Developmental Services, and approved by the regional center, shall be permitted to have up to 24 hours of the required continuing education course hours credited toward the 40-hour continuing education requirement of this section. Community college course hours approved by the regional centers shall be accepted by the department for certification.

44497

(f) The conviction of a licensee or the person in charge of the tissue bank of any crime which is substantially related to the qualifications or duties of the licensee or the person in charge of the tissue bank or which is substantially related to the functions of the tissue bank. For purposes of this section, a conviction means a plea or verdict of guilty or a conviction following a plea of nolo contendere. An action to revoke or suspend the license may be taken when the time for appeal has elapsed or the judgment of conviction has been affirmed on appeal or when an order granting probation is made suspending the imposition of sentence, notwithstanding a subsequent order pursuant to Section 1203.4 of the Penal Code permitting withdrawal of a plea of guilty and entry of a plea of not guilty, or setting aside the verdict of guilty, or dismissing the accusation, information or indictment. The director shall take into account all competent evidence of rehabilitation furnished by the licensee or person in charge of the tissue bank.

44575

(c) The department may determine that no portion of any later editions or amendments shall become effective for purposes of this section. The department shall determine that no portion of a later edition or amendment shall become effective for purposes of this section whenever the department has not received a

copy of the later edition or amendment by the date it is published by the American Association of Blood Banks.

45303

For purposes of this section, the following terms have the following meanings:

Name: node_text, Length: 128, dtype: object

In [63]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 3]['node_text'])
```


6

(a) (1) All aspects of the unified program related to the adoption and interpretation of statewide standards and requirements shall be the responsibility of the state agency which is charged with that responsibility under existing law. For underground storage tanks, that agency shall be the State Water Resources Control Board. The California regional water quality control boards shall have responsibility for the issuance of variances pursuant to subdivision (b) of Section 25299.4. The Department of Toxic Substances Control shall have the sole responsibility for the issuances of variances from the requirements of Chapter 6.5 (commencing with Section 25100) and the regulations adopted pursuant thereto, for the determination of whether or not a waste is hazardous or nonhazardous, for the determination of whether or not a person is eligible to be deemed to be operating pursuant to a permit-by-rule, conditional authorization, or conditional exemption pursuant to Chapter 6.5 (commencing with Section 25100) or the regulations adopted by the department, and for the suspension and revocation of permits-by-rule, conditional authorizations, and conditional exemptions.

(2) Except as provided in paragraphs (1) and (3), those aspects of the unified program related to the application of statewide standards to particular facilities, including the issuance of unified program facility permits, the review of reports and plans, environmental assessment, compliance and correction, and the enforcement of those standards and requirements against particular facilities, shall be the responsibility of the unified program agencies.

(3) (A) Except in those jurisdictions for which the UPA has been determined by the department, in accordance with regulations adopted pursuant to subparagraph (C), to be qualified to implement the environmental assessment and removal and remediation corrective action aspects of the unified program, the department shall have sole responsibility and authority under the unified program for all of the following:

- (i) Implementing and enforcing the requirements of paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25200.10 and 25200.14, and the regulations adopted by the department to implement those sections. As a pilot program in up to 10 counties, pending the adoption and implementation of regulations pursuant to subparagraph (C), the department may delegate to the CUPA, through a delegation agreement, responsibility and authority for implementing and enforcing the requirements of Section 25200.14.
- (ii) The issuance of orders under Section 25187 requiring removal or remedial action.
- (iii) The issuance of orders under Section 25187.1.

(B) Notwithstanding subparagraph (A), a UPA may issue an order under Section 25187 specifying a schedule for compliance or correction and imposing an administrative penalty for any violation of the requirements of Chapter 6.5 (commencing with Section 25100) listed in paragraph (1) of subdivision (c) of Section 25404, or the requirements of any permit, rule, regulation, standard or requirement issued or adopted pursuant to the requirements of Chapter 6.5 (commencing with Section 25100) listed in paragraph (1) of subdivision (c) of Section 25404, if one of the following applies:

- (i) The order does not require removal or remedial action.
- (ii) The only removal or remedial actions required by the order are those actions determined to be necessary to address an imminent and substantial endangerment based upon a finding by the UPA pursuant to subdivision (f) of Section 25187.

(C) The department shall adopt emergency regulations specifying the criteria and procedures for implementing paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25187, 25187.1, 25200.10, and 25200.14, including criteria and procedures for determining whether or not a unified program agency is qualified to implement the environmental assessment and removal and remediation corrective action portions of the unified program under paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25187, 25187.1, 25200.10, and 25200.14. The criteria for determining whether a unified program agency is qualified shall, at a minimum, include consideration of the following factors:

- (i) Adequacy of the technical expertise possessed by the unified program agency.
- (ii) Adequacy of staff resources.
- (iii) Adequacy of budget resources and funding mechanisms.
- (iv) Training requirements.
- (v) Past performance in implementing and enforcing requirements related to environmental assessments, and removal and remediation corrective actions.
- (vi) Recordkeeping and accounting system

s.\n (D) The regulations adopted by the department pursuant to subparagraph (C) shall include provisions to ensure coordinated and consistent application of paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25187, 25187.1, 25200.10, and 25200.14, when both the department and the unified program agency are, or will be, implementing and enforcing the requirements of one or more of these sections at the same facility.\n (E) For purposes of subparagraph (D), "facility" means the entire site that is under the control of the owner or operator.\n (F) If the department is designated as a unified program agency, the department is deemed qualified to implement all of the following:\n (i) The environmental assessment, removal and remedial action, and corrective action aspects of the unified program.\n (ii) Paragraph (3) of subdivision (c) of Section 25300.3, Sections 25200.10, 25200.14, 25187, and 25287.1, and the regulations adopted by the department to implement those provisions.

32

(a) For purposes of this chapter, the following terms shall have the following meanings:\n \n (1) (A) "Certified Unified Program Agency" or "CUPA" means the agency certified by the secretary to implement the unified program specified in this chapter within a jurisdiction.\n (B) "Participating Agency" or "PA" means a state or local agency that has a written agreement with the CUPA pursuant to subdivision (d) of Section 25404.3, and is approved by the secretary, to implement or enforce one or more of the unified program elements specified in subdivision (c), in accordance with Sections 25404.1 and 25404.2.\n (C) "Unified Program Agency" or "UPA" means the CUPA, or its participating agencies to the extent each PA has been designated by the CUPA, pursuant to a written agreement, to implement or enforce a particular unified program element specified in subdivision (c). The UPAs have the responsibility and authority to implement and enforce the requirements listed in subdivision (c), and the regulations adopted to implement the requirements listed in subdivision (c), to the extent provided by Chapter 6.5 (commencing with Section 25100), Chapter 6.67 (commencing with Section 25270), Chapter 6.7 (commencing with Section 25280), Chapter 6.95 (commencing with Section 25500), and Sections 25404.1 to 25404.2, inclusive. After a CUPA has been certified by the secretary, the unified program agencies and the state agencies carrying out responsibilities under this chapter shall be the only agencies authorized to enforce the requirements listed in subdivision (c) within the jurisdiction of the CUPA.\n (2) "Department" means the Department of Toxic Substances Control.\n (3) "Minor violation" means the failure of a person to comply with a requirement or condition of an applicable law, regulation, permit, information request, order, variance, or other requirement, whether procedural or substantive, of the unified program that the UPA is authorized to implement or enforce pursuant to this chapter, and that does not otherwise include any of the following:\n (A) A violation that results in injury to persons or property, or that presents a significant threat to human health or the environment.\n (B) A knowing, willful, or intentional violation.\n (C) A violation that is a chronic violation, or that is committed by a recalcitrant violator. In determining whether a violation is chronic or a violator is recalcitrant, the UPA shall consider whether there is evidence indicating that the violator has engaged in a pattern of neglect or disregard with respect to applicable regulatory requirements.\n (D) A violation that results in an emergency response from a public safety agency.\n (E) A violation that enables the violator to benefit economically from the noncompliance, either by reduced costs or competitive advantage.\n (F) A class I violation, as provided in Section 25110.8.5.\n (G) A violation that hinders the ability of the UPA to determine compliance with any other applicable local, state, or federal rule, regulation, information request, order, variance, permit, or other requirement.\n (4) "Secretary" means the Secretary for Environmental Protection.\n (5) "Unified program facility" means all contiguous land and structures, other appurtenances, and improvements on the land that are subject to the requirements listed in subdivision (c).\n (6) "Unified program facility permit" means a permit issued pursuant to this chapter. For purposes of this chapter, a unified program facility permit encompasses the permitting requirements of Section 25284, and permit

or authorization requirements under a local ordinance or regulation relating to the generation or handling of hazardous waste or hazardous materials, but does not encompass the permitting requirements of a local ordinance that incorporates provisions of the California Fire Code or the California Building Code.

43

(a) Notwithstanding Section 25117.2, "management" means disposal, storage, packaging, processing, pumping, recovery, recycling, transportation, transfer, treatment, use, and reuse.

44

(b) "Perchlorate" means all perchlorate-containing compounds.

45

(c) "Perchlorate material" means perchlorate and all perchlorate-containing substances, including, but not limited to, waste perchlorate and perchlorate-containing waste.

...

45677

(r) "Release" means any spilling, leaking, pumping, pouring, emitting, emptying, discharging, injecting, escaping, leaching, dumping, or disposing into the environment on blighted property.

45678

(s) "Responsible party" means a person described in subdivision (a) of Section 25323.5 of this code or subdivision (a) of Section 13304 of the Water Code.

45679

(t) "Site designation committee" means the committee established pursuant to Section 25261.

45680

(u) "State board" means the State Water Resources Control Board.

45682

(a) (1) All aspects of the unified program related to the adoption and interpretation of statewide standards and requirements shall be the responsibility of the state agency which is charged with that responsibility under existing law. For underground storage tanks, that agency shall be the State Water Resources Control Board. The California regional water quality control boards shall have responsibility for the issuance of variances pursuant to subdivision (b) of Section 25299.4. The Department of Toxic Substances Control shall have the sole responsibility for the issuances of variances from the requirements of Chapter 6.5 (commencing with Section 25100) and the regulations adopted pursuant thereto, for the determination of whether or not a waste is hazardous or nonhazardous, for the determination of whether or not a person is eligible to be deemed to be operating pursuant to a permit-by-rule, conditional authorization, or conditional exemption pursuant to Chapter 6.5 (commencing with Section 25100) or the regulations adopted by the department, and for the suspension and revocation of permits-by-rule, conditional authorizations, and conditional exemptions.

(2) Except as provided in paragraphs (1) and (3), those aspects of the unified program related to the application of statewide standards to particular facilities, including the issuance of unified program facility permits, the review of reports and plans, environmental assessment, compliance and correction, and the enforcement of those standards and requirements against particular facilities, shall be the responsibility of the unified program agencies.

(3) (A) Except in those jurisdictions for which the UPA has been determined by the department, in accordance with regulations adopted pursuant to subparagraph (C), to be qualified to implement the environmental assessment and removal and remediation corrective action aspects of the unified program, the department shall have sole responsibility and authority under the unified program for all of the following:

(i) Implementing and enforcing the requirements of paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25200.10 and 25200.14, and the regulations adopted by the department to implement those sections. As a pilot program in up to 10 counties, pending the adoption and implementation of regulations pursuant to subparagraph (C), the department may delegate to the CUPA, through a delegation agreement, responsibility and author

ity for implementing and enforcing the requirements of Section 25200.14.\n (i i) The issuance of orders under Section 25187 requiring removal or remedial action.\n (iii) The issuance of orders under Section 25187.1.\n (B) Notwithstanding subparagraph (A), a UPA may issue an order under Section 25187 specifying a schedule for compliance or correction and imposing an administrative penalty for any violation of the requirements of Chapter 6.5 (commencing with Section 25100) listed in paragraph (1) of subdivision (c) of Section 25404, or the requirements of any permit, rule, regulation, standard or requirement issued or adopted pursuant to the requirements of Chapter 6.5 (commencing with Section 25100) listed in paragraph (1) of subdivision (c) of Section 25404, if one of the following applies:\n (i) The order does not require removal or remedial action.\n (ii) The only removal or remedial actions required by the order are those actions determined to be necessary to address an imminent and substantial endangerment based upon a finding by the UPA pursuant to subdivision (f) of Section 25187.\n (C) The department shall adopt emergency regulations specifying the criteria and procedures for implementing paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25200.10 and 25200.14, including criteria and procedures for determining whether or not a unified program agency is qualified to implement the environmental assessment and removal and remediation corrective action portions of the unified program under paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25187, 25187.1, 25200.10, and 25200.14. The criteria for determining whether a unified program agency is qualified shall, at a minimum, include consideration of the following factors:\n (i) Adequacy of the technical expertise possessed by the unified program agency.\n (ii) Adequacy of staff resources.\n (iii) Adequacy of budget resources and funding mechanisms.\n (iv) Training requirements.\n (v) Past performance in implementing and enforcing requirements related to environmental assessments, and removal and remediation corrective actions.\n (vi) Recordkeeping and accounting systems.\n (D) The regulations adopted by the department pursuant to subparagraph (C) shall include provisions to ensure coordinated and consistent application of paragraph (3) of subdivision (c) of Section 25200.3 and Sections 25187, 25187.1, 25200.10, and 25200.14, when both the department and the unified program agency are, or will be, implementing and enforcing the requirements of one or more of these sections at the same facility.\n (E) For purposes of subparagraph (D), "facility" means the entire site that is under the control of the owner or operator.\n (F) If the department is designated as a unified program agency, the department is deemed qualified to implement all of the following:\n (i) The environmental assessment, removal and remedial action, and corrective action aspects of the unified program.\n (ii) Paragraph (3) of subdivision (c) of Section 25200.3, Sections 25200.10, 25200.14, 25187, and 25287.1, and the regulations adopted by the department to implement those provisions.

Name: node_text, Length: 2800, dtype: object

In [64]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 4]['node_text'])
```

95 (a) For the purposes of this section, the following terms mean the following:\n (1) "Local publicly owned electric utility" has the same meaning as defined in Section 224.3 of the Public Utilities Code.\n (2) "Retail seller" has the same meaning as set forth in subdivision (j) of Section 399.12 of the Public Utilities Code.\n (3) "Transportation electrification" has the same meaning as set forth in Section 237.5 of the Public Utilities Code.

143

(c) "Major appliance" has the same meaning as defined in Section 42166 of the Public Resources Code.

144

(d) "Materials that require special handling" has the same meaning as defined in Section 42167 of the Public Resources Code.

249

(a) For purposes of this section, "electronic device" has the same meaning as a "covered electronic device," as defined in subparagraph (A) of paragraph (1) of subdivision (g) of Section 42463 of the Public Resources Code. "Covered electronic device" does not include a covered battery-embedded product, as defined in subparagraph (B) of paragraph (1) of subdivision (g) of Section 42463 of the Public Resources Code.

320

(j) "Owner or operator" has the same meaning given in Section 9601(20) of Title 42 of the United States Code and any person that previously met that definition or is the legal successor to a person that meets the definition or previously met the definition.

...

45225

(c) "Hazardous substance removal" has the same meaning as used in Section 7058.7 of the Business and Professions Code.

45461

(d) "Animal control officer" has the same meaning as defined in paragraph (8) of subdivision (d) of Section 241 of the Penal Code.

45486

(d) "Code enforcement officer" has the same meaning as defined in Section 829.5 of the Penal Code.

45579

(i) "Hazardous material" has the same meaning as defined in subdivision (d) of Section 25260.

45670

(i) "Hazardous material" has the same meaning as defined in subdivision (d) of Section 25260.

Name: node_text, Length: 180, dtype: object

In [65]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 5]['node_text'])
```

437 (c) "Paint" includes both oil-based paint and recyclable latex paint that is collected in accordance with this article.

1088

"License" includes, but is not limited to any, permit, registration, or certification issued by any local, state, or federal agency for the generation, transportation, treatment, storage, recycling, disposal, or handling of hazardous waste.

1104

"Natural resources" includes, but is not limited to, disposal site capacity and substances which are hazardous waste, or which are in hazardous waste, the reuse of which is technologically and economically feasible.

1109

(b) "Hazardous waste" includes, but is not limited to, RCRA hazardous waste.

1132

"Restricted hazardous waste" includes both of the following:

...

44508

(c) All donors of sperm shall be screened and found nonreactive as required under subdivision (a), except in the following instances:

- (1) A recipient of sperm, from a sperm donor known to the recipient, may waive a second or other repeat testing of that donor if the recipient is informed of the requirements for testing donors under this section and signs a written waiver.
- (2) A recipient of sperm may consent to therapeutic insemination of sperm or use of sperm in other assisted reproductive technologies even if the sperm donor is found reactive for hepatitis B, hepatitis C, syphilis, HIV, or HTLV if the sperm donor is the spouse of, partner of, or designated donor for that recipient. The physician providing insemination or assisted reproductive technology services shall advise the donor and recipient of the potential medical risks associated with receiving sperm from a reactive donor. The donor and the recipient shall sign a document affirming that each person comprehends the potential medical risks of using sperm from a reactive donor for the proposed procedure and that each consents to it. Copies of the document shall be placed in the medical records of the donor and the recipient.
- (3) (A) Sperm whose donor has tested reactive for syphilis may be used for the purposes of insemination or assisted reproductive technology only after the donor has been treated for syphilis. Sperm whose donor has tested reactive for hepatitis B may be used for the purposes of insemination or assisted reproductive technology only after the recipient has been vaccinated against hepatitis B.
- (B) (i) Sperm whose donor has tested reactive for HIV or HTLV may be used for the purposes of insemination or assisted reproductive technology for a recipient testing negative for HIV or HTLV only after the donor's sperm has been effectively processed to minimize the likelihood of transmission through the sperm for that specific donation and if informed and mutual consent has occurred.
- (ii) The department shall adopt regulations regulating facilities that perform sperm processing, pursuant to this subparagraph, that prescribe standards for the handling and storage of sperm samples of carriers of HIV, HTLV, or any other virus as deemed appropriate by the department. The department may propose to adopt, as initial regulations, the most relevant and up-to-date recommendations published by the American Society for Reproductive Medicine. Notice of the department's proposed adoption of the regulations shall be posted on the department's Internet Web site for at least 45 days. Public comment shall be accepted by the department for at least 30 days after the conclusion of the 45-day posting period. If a member of the public requests a public hearing during the 30-day comment period, the hearing shall be held prior to the adoption of the regulations. If no member of the public requests a public hearing, the regulations shall be deemed adopted at the conclusion of the 30-day comment period. Comments received shall be considered prior to the adoption of the final initial regulations. The department may modify any recommendations published by the American Society for Reproductive Medicine. Adoption of initial regulations by the department pursuant

ant to this subdivision shall not be subject to the rulemaking requirements of Chapter 3.5 (commencing with Section 11340) of Part 1 of Division 3 of Title 2 of the Government Code and written responses to public comments shall not be required. Updates to the regulations shall be adopted pursuant to the same process. Until the department adopts these regulations, facilities that perform sperm processing pursuant to this section shall follow facility and sperm processing recommendations for the reduction of viral transmission developed by the American Society for Reproductive Medicine. This section does not prevent the department from monitoring and inspecting facilities that process sperm to ensure adherence to the regulations, or, until regulations are adopted, to the recommendations set forth by the American Society for Reproductive Medicine.

(iii) Before insemination or other assisted reproductive technology services are performed, the physician providing the services shall inform the recipient of sperm from a spouse, partner, or designated donor who has tested reactive for HIV or HTLV of all of the following:

- (I) That sperm processing may not eliminate all of the risks of HIV or HTLV transmission.
- (II) That the sperm may be tested to determine whether or not it is reactive for HIV or HTLV.
- (III) That the recipient shall provide documentation to the physician providing insemination or assisted reproductive technology services prior to treatment that she has established an ongoing relationship with another physician to provide for her medical care during and after completion of fertility services.
- (IV) The most relevant and up-to-date recommendations published by the American Society for Reproductive Medicine regarding followup testing for HIV and HTLV after use of sperm from an HIV or HTLV reactive donor and have the recommendations regarding followup testing be documented in the recipient's medical record.
- (v) The physician providing insemination or assisted reproductive technology services shall also verify, and document in the recipient's medical record, that the donor of sperm who tests reactive for HIV or HTLV is under the care of a physician managing the HIV or HTLV.
- (vi) The physician providing insemination or assisted reproductive technology services shall recommend to the physician who will be providing ongoing care to the recipient recommended followup testing for HIV and HTLV according to the most relevant and up-to-date guidelines published by the American Society for Reproductive Medicine, which shall be documented in the recipient's medical record.
- (vii) If the recipient becomes HIV or HTLV positive, the physician assuming ongoing care of the recipient shall treat or provide information regarding referral to a physician who can provide ongoing treatment of the HIV or HTLV.

(4) A recipient of sperm donated by a sexually intimate partner of the recipient for reproductive use may waive a second or repeat testing of that donor if the recipient is informed of the donor testing requirements of this section and signs a written waiver. For purposes of this paragraph, "sexually intimate partner of the recipient" includes a known or designated donor to whose sperm the recipient has previously been exposed in a nonmedical setting in an attempt to conceive.

44527

(a) Whenever there is a reasonable possibility, as determined by a physician and surgeon or doctor of podiatric medicine, that a blood transfusion may be necessary as a result of a medical or surgical procedure, the physician and surgeon or doctor of podiatric medicine, by means of a standardized written summary as most recently developed or revised by the State Department of Public Health pursuant to subdivision (e), shall inform, either directly or through a nurse practitioner, certified nurse midwife, or a physician assistant, who is licensed in the state and authorized to order a blood transfusion, the patient of the positive and negative aspects of receiving autologous blood and directed and nondirected homologous blood from volunteers. For purposes of this section, the term "autologous blood" includes, but is not limited to, predonation, intraoperative autologous transfusion, plasmapheresis, and hemodilution.

44557

"Distribution" includes sale and exchange.

44579

As used in this chapter, "person" includes: laboratory, firm, association, cor

poration, copartnership, and educational institution. 45525 (f) For purposes of this section, “provides for compensation” includes contains that are rented or provided as part of a service for which the owner or manufacturer receives compensation.

Name: node_text, Length: 175, dtype: object

In [66]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 6]['node_text'])
```


399 On and after July 1, 2017, a manufacturer shall place a recycling symbol consistent with the requirements of Section 103(b)(1) of the Federal Mercury Containing and Rechargeable Battery Management Act, Pub. L. No. 104-142 (1996) (42 U.S.C. 14301(b)(1)) and either "Pb" or the words "lead," "return," and "recycle" on all replacement lead-acid batteries sold in California. For purposes of this section, an entity that engages another party to manufacture batteries on its behalf shall be deemed the manufacturer.

~~797~~ (e) (1) The advisory committee shall recommend to the secretary that the extension be approved if the advisory committee determines that there are no brake friction materials that are safe and available for individual or multiple vehicle models, classes, platforms, or other vehicle-based categories identified in the application. (2) The advisory committee shall recommend to the secretary that the extension not be approved if the advisory committee determines that alternative brake friction materials are safe and available for individual or multiple vehicle models, classes, platforms, or other vehicle-based categories identified in the application. (3) For purposes of this section, "safe and available" shall mean all of the following: (A) The brake system for which the alternative brake friction material is manufactured meets applicable federal safety standards, or if no federal standard exists, a widely accepted safety standard. (B) Acceptable alternative brake friction materials are commercially available for the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (C) Adequate industry testing and production capacity exists to supply the alternative brake friction materials for use on the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (D) The alternative brake friction material is technically feasible for use on the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (E) The alternative brake friction materials meet customer performance expectations, including noise, wear, vibration, and durability for the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (F) The alternative acceptable brake friction material is economically feasible with respect to the industry and the cost to the consumer for the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (4) The advisory committee shall provide relevant data to the department and the board concerning the potential impacts of the extension on California watersheds for purposes of the report required pursuant to Section 25250.65.

805

(e) (1) The advisory committee shall recommend to the secretary that the extension be approved if the advisory committee determines that there are no brake friction materials that are safe and available for individual or multiple vehicle models, classes, platforms, or other vehicle-based categories identified in the application. (2) The advisory committee shall recommend to the secretary that the extension not be approved if the advisory committee determines that alternative brake friction materials are safe and available for individual or multiple vehicle models, classes, platforms, or other vehicle-based categories identified in the application. (3) For purposes of this section, "safe and available" shall mean all of the following: (A) The brake system for which the alternative brake friction material is manufactured meets applicable federal safety standards, or if no federal standard exists, a widely accepted safety standard. (B) Acceptable alternative brake friction materials are commercially available for the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (C) Adequate industry testing and production capacity exists to supply the alternative brake friction materials for use on the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (D) The alternative brake friction material is technically feasible for use on the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application. (E) The alternative brake friction mat

erials meet customer performance expectations, including noise, wear, vibration, and durability for the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application.\n (F) The alternative acceptable brake friction material is economically feasible with respect to the industry and the cost to the consumer for the individual or multiple vehicles, classes, platforms, or vehicle-based categories identified in the application.\n (4) The advisory committee shall provide relevant data to the department and the board concerning the potential impacts of the extension on California watersheds for purposes of the report required pursuant to Section 25250.65.

1643

(d) For purposes of this section, "recycle" and "recycling" shall have the same meaning as set forth in subdivision (a) of Section 25121.1.

2975

(h) For purposes of this section, the following definitions apply:\n (1) "Local officer" has the meaning provided for in Section 101480.\n (2) "Unified program agency" has the meaning provided for in Section 25404.\n (3) "Water replenishment district" has the meaning provided for in Section 60012 of the Water Code.

...

42454

(a) On or before January 31 of each year, the licensee of a licensed residential care facility for the elderly shall prepare a document disclosing its average monthly rate increases, inclusive of rates for living units and service fees, for each of the previous 3 years. For purposes of this section, "service fees" do not include fees for optional services or services provided by a third party. The licensee shall disclose the average amount of the increase, as well as the average percentage of increase. Newly licensed facilities without three years of resident rate increase history shall disclose the average increase for the years during which the facility has been serving residents. This section does not apply to newly licensed facilities with no current residents.

43228

(b) Subsequent to being hired pursuant to subdivision (a), a teacher shall make satisfactory progress towards meeting the educational requirement for a fully qualified teacher, as specified in departmental regulations. For purposes of this section, "satisfactory progress" shall mean completion, with passing grades, of a minimum of two units each semester or the equivalent number of units each quarter until the educational requirement is satisfied. Six of the required semester or equivalent number of quarter units of early childhood education from an accredited university or college shall be completed during the next two consecutive regular semesters or equivalent quarters.

43552

\n (c) (1) "Community living support services," for purposes of this section, are voluntary and chosen by persons with disabilities in accordance with their preferences and goals for independent living. "Community living support services" may include, but are not limited to, any of the following:\n (A) Supports that are designed to develop and improve independent living and problem solving skills.\n (B) Education and training in meal planning and shopping, budgeting and managing finances, medication self-management, transportation, vocational and educational development, and the appropriate use of community resources and leisure activities.\n (C) Assistance with arrangements to meet the individual's basic needs such as financial benefits, food, clothing, household goods, and housing, and locating and scheduling for appropriate medical, dental, and vision benefits and care.\n (D) When needed, assistance with independent activities of daily living or personal care.\n (2) "Supportive housing," for purposes of this section, is rental housing that has all of the following characteristics:\n (A) It is affordable to people with disabilities.\n (B) It is independent housing in which each tenant meets all of the following conditions:\n (i) Holds a lease or rental agreement in their own name and is responsible for pay

ing their own rent.\n (ii) Has their own room or apartment and is individually responsible for arranging any shared tenancy.\n (C) It is permanent, wherein each tenant may stay as long as they pay their share of rent and comply with the terms of their lease.\n (D) It is tenancy housing under which supportive housing providers are required to comply with applicable state and federal laws governing the landlord-tenant relationship.\n (E) Participation in services or any particular type of service is not required as a condition of tenancy.

44794

(k) For purposes of this section, "serious illness" shall mean a condition that may result in death, regardless of the estimated length of the patient's remaining period of life.

45251

(d) This section shall not be construed to apply to owners of a building or part of a building within a residential common interest development or association, if the owners comply with the provisions of subdivision (d) of Section 25915.2. For purposes of this section, "association" and "common interest development" are defined in Sections 4080 and 4100 of the Civil Code.

Name: node_text, Length: 64, dtype: object

In [67]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 7]['node_text'])
```

121 (a) For purposes of this section, the following definitions apply:\n (1) "Mobi lity option" has the same meaning as defined in Section 44124.\n (2) "Zero-emission or near-zero-emission vehicle incentive program" means a program to provide incentives to an individual for the purchase of a light-duty zero-emission or near-zero-emission vehicle.

238

(a) For purposes of this section, the following definitions shall apply:\n (1) "Electronic device" means a video display device, as defined in Section 42463 of the Public Resources Code, with a screen size of greater than four inches.\n (2) "Covered electronic device" has the same meaning as a covered electronic device, as defined in paragraph (1) of subdivision (g) of Section 42463 of the Public Resources Code.\n (3) "Manufacturer" and "retailer" have the same meaning as set forth in Section 42463 of the Public Resources Code.

413

(a) For purposes of this section, "spent dry cell battery containing zinc electrodes" means an alkaline or zinc-carbon battery, that meets all of the following conditions:\n (1) It is an enclosed device or sealed container consisting of one or more voltaic or galvanic cells, electrically connected to produce electric energy, of any shape, including, but not limited to, button, coin, cylindrical, or rectangular, and designed for commercial, industrial, medical, institutional, or household use.\n (2) It contains an electrode comprised of zinc or zinc oxide or a combination thereof, and a liquid starved or gelled electrolyte.\n (3) It does not contain any constituent, other than zinc or zinc oxide, that would cause it to be classified as a hazardous waste pursuant to this chapter.\n (4) It is discarded by the user.

473

(b) For purposes of this section, "surplus household consumer product" means a household consumer product that cannot or will not be sold to a consumer through that product's primary market.

504

(a) For purposes of this section, the following definitions shall apply:\n (1) "Halocarbon chemicals" means chemical compounds which contain carbon, and one or more halogens, and which may include hydrogen, including, but not limited to, trichloroethane, tetrachloroethylene, methylene chloride, halogenated benzenes, and carbon tetrachloride.\n (2) "Aromatic hydrocarbon chemicals" means chemical compounds containing carbon and hydrogen and at least one six-carbon ring containing double bonds, including, but not limited to, benzene, toluene, and naphthalene.\n (3) "Sewage disposal system" means a septic tank, cesspool, sewage seepage pit, leachline, or other structure into which sewage is drained for purposes of disposal and which is not connected to a municipal treatment works.

...

44780

(b) For purposes of this section, the following definitions apply:\n (1) "Patient's representative" means either a person designated by the patient as their representative or a person acting on the patient's behalf under the authority of the Long-Term Care Patient Representative Program pursuant to Chapter 3.6 (commencing with Section 9260) of Division 8.5 of the Welfare and Institutions Code.\n (2) "Verbal" means spoken or signed language.

45092

(a) For purposes of this section, the following definitions apply:\n (1) "ARRA" means Title III of Division B of the federal American Recovery and Reinvestment Act of 2009 or any amendment to that federal law extending federal premium assistance to qualified beneficiaries, as defined in Section 1366.21 of this code or Section 10128.51 of the Insurance Code.\n (2) "Employer" means an employer as defined in Section 1366.21 of this code or an employer as defined in Section 10128.51 of the Insurance Code.

45096

(b) For purposes of this section, the following definitions apply:\n (1) "Drug masking product" means synthetic urine or any other substance designed to be added to human urine or human hair for the purpose of defrauding an alcohol or drug screening test.\n (2) "Synthetic urine" means any substance that is designed to simulate the composition, chemical properties, physical appearance, or physical properties of human urine.

~~(c)~~ The subject or subject's conservator or guardian, or other representative, as specified in Section 24175, is informed both verbally and within the written consent form, in nontechnical terms and in a language in which the subject or the subject's conservator or guardian, or other representative, as specified in Section 24175, is fluent, of the following facts of the proposed medical experiment, which might influence the decision to undergo the experiment, including, but not limited to:\n (1) An explanation of the procedures to be followed in the medical experiment and any drug or device to be utilized, including the purposes of the procedures, drugs, or devices. If a placebo is to be administered or dispensed to a portion of the subjects involved in a medical experiment, all subjects of the experiment shall be informed of that fact; however, they need not be informed as to whether they will actually be administered or dispensed a placebo.\n (2) A description of any attendant discomfort and risks to the subject reasonably to be expected.\n (3) An explanation of any benefits to the subject reasonably to be expected, if applicable.\n (4) A disclosure of any appropriate alternative procedures, drugs, or devices that might be advantageous to the subject, and their relative risks and benefits.\n (5) An estimate of the expected recovery time of the subject after the experiment.\n (6) An offer to answer any inquiries concerning the experiment or the procedures involved.\n (7) An instruction to the subject that he or she is free to withdraw his or her prior consent to the medical experiment and discontinue participation in the medical experiment at any time, without prejudice to the subject.\n (8) The name, institutional affiliation, if any, and address of the person or persons actually performing and primarily responsible for the conduct of the experiment.\n (9) The name of the sponsor or funding source, if any, or manufacturer if the experiment involves a drug or device, and the organization, if any, under whose general aegis the experiment is being conducted.\n (10) The name, address, and phone number of an impartial third party, not associated with the experiment, to whom the subject may address complaints about the experiment.\n (11) The material financial stake or interest, if any, that the investigator or research institution has in the outcome of the medical experiment. For purposes of this section, "material" means ten thousand dollars (\$10,000) or more in securities or other assets valued at the date of disclosure, or in relevant cumulative salary or other income, regardless of when it is earned or expected to be earned.

45516

(d) For purposes of this section, the following definitions shall apply:\n (1) "Agent or representative" may include, but is not limited to, a licensed contractor that is performing work on, or has completed work on, the commercial property or business that is the subject of the inspection.\n (2) "Local government" means a city, including a charter city, county, or city and county.\n (3) "Local official" means a code enforcement officer, as defined in Section 829.5 of the Penal Code, of a city, county, or a city and county, a county health officer described in Article 1 (commencing with Section 101025) of Chapter 2 of Part 3 of Division 101, a city health officer described in Article 4 (commencing with Section 101450) of Chapter 4 of Part 3 of Division 101, or the designated agent of those health officers.\n (4) "Person with a potential financial interest in the outcome of the inspection" is any of the following:\n (A) A person who makes an offer to the owner to physically remediate for compensation potential violations of a state statute or regulation or local ordinance found in the course of an inspection by a local official.\n (B) A person who offers to provide or provides compensation to a local official in exchange for recommending the remediation services of a specific person, providing the contact information for those services to the owner or agent of the owner of the premise

s or business inspected, or providing the name of the owner of the property or business or agent or representative of the owner to the person who offers the remediation services.

Name: node_text, Length: 299, dtype: object

In [68]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 8]['node_text'])
```

42

For purposes of this article, the following definitions shall apply:

310

For purposes of this article, the following definitions shall apply:

331

For purposes of this article, the following definitions shall apply:

434

For purposes of this article, the following definitions shall apply:

719 (c) For purposes of subdivision (b), the following definitions apply:

y:\n (2) The baseline year is either of the following, whichever is applicable:

e:\n (A) For the initial report, the baseline year is the calendar year selected by the generator for which substantial hazardous waste generation, or onsite or offsite management, data is available prior to 1991.\n (B) For all subsequent reports, the baseline year is the current reporting year of the immediately preceding report.\n (1) The current reporting year is the calendar year immediately preceding the year in which the report is to be prepared.

...

44537 Unless the context requires otherwise, the following definitions shall apply to this chapter: 44752 For the purposes of this chapter, the following definitions apply:

44857

As used in this chapter, the following definitions shall apply:

45038

As used in this chapter, the following definitions shall apply:

45222

For purposes of this chapter, the following definitions shall apply:

Name: node_text, Length: 174, dtype: object

In [69]:

```
with pd.option_context('display.max_colwidth', None):  
    display(df2[df2['cluster'] == 9]['node_text'])
```

5963 As used in this article the word "taxes" shall include, but without limitation, all levies on an ad valorem basis upon land or real property. As used in this article, "taxes" shall not include any amounts of money deposited in a Sales and Use Tax Compensation Fund pursuant to Section 97.68 of the Revenue and Taxation Code or a Vehicle License Fee Property Tax Compensation Fund pursuant to Section 97.70 of the Revenue and Taxation Code.

20966

The boat shall have the word "lifeboat" plainly printed or painted upon it. It shall be used for no purpose other than for the saving of life or for other cases of emergency.

22058

(c) Biohazardous waste, except as provided in subdivision (b), shall be bagged in accordance with subdivision (b) of Section 118275 and placed for storage, handling, or transport in a rigid container that may be disposable, reusable, or recyclable. Containers shall be leak resistant, have tight-fitting covers, and be kept clean and in good repair. Containers may be recycled with the approval of the enforcement agency. Containers may be of any color and shall be labeled with the words "Biohazardous Waste" or with the international biohazard symbol and the word "BIOHAZARD" on the lid and sides so as to be visible from any lateral direction. Containers shall comply with United States Department of Transportation requirements when prepared for transport offsite from the facility.

22069

(d) Label sharps containers with the words "sharps waste" or with the international biohazard symbol and the word "BIOHAZARD."

22262

(e) Each procedure area shall have a sharps waste container that meets the following requirements:
(1) The sharps waste container shall be portable, if portability is necessary to ensure that the sharps waste container is within arm's reach of the practitioner.
(2) The sharps waste container shall be labeled with the words "sharps waste" or with the international biohazard symbol and the word "BIOHAZARD."
(3) All sharps waste produced during the process of tattooing, body piercing, or the application of permanent cosmetics shall be disposed by either of the following methods:
(A) Removal and disposal by a licensed waste hauler. Materials shall be disposed of at a licensed treatment facility or removed and transported through a mail-back system authorized by the State Department of Public Health.
(B) As solid waste, after being disinfected by a method approved by the department pursuant to paragraph (3) of subdivision (a) of Section 118215.
(4) Documentation of proper disposal of sharps waste shall be maintained for three years and shall be available for inspection at the request of the enforcement officer.

23151

Any food is misbranded if it is offered for sale under the name of another food, or if it is an imitation of another food for which a definition and standard of identity has been established by regulation and its label does not bear, in type of uniform size and prominence the word "imitation," and immediately following, the name of the food imitated.

23203

(a) By identifying each organically produced ingredient in the ingredient statement with the word "organic" or with an asterisk or other reference mark that is defined below the ingredient statement to indicate the ingredient is organically produced.

~~24130~~ 24130 internet food service intermediary that lists or promotes a microenterprise home kitchen operation on its internet website or mobile application shall meet all of the following requirements:
(1) Be registered with the department. A registration, once issued, is nontransferable. A registration shall be valid only for the person and type of business specified by that registration, and unless suspended or revoked for cause by the department.
(2) Prior to the listing or publication of a microenterprise home kitchen operation

n's offer of food for sale, clearly and conspicuously post on its internet web site or mobile application the requirements for the permitting of a microenterprise home kitchen specified in this chapter, which shall be written at the high school level and be provided by the department.

(3) Clearly and conspicuously post on its internet website or mobile application the fees associated with using its platform and fees associated with third-party delivery service pursuant to paragraph (2) of subdivision (b) of Section 114367.5 in a manner that allows both the consumer and the microenterprise home kitchen operation to see and understand the amount being charged for the services provided by the internet food service intermediary. The internet food service intermediary shall notify the microenterprise home kitchen operation of any changes to these fees exceeding a 2-percent increase in writing and no later than one month before the changes take effect.

(4) Clearly and conspicuously post on its internet website or mobile application whether or not it has liability insurance that would cover any incidence arising from the sale or consumption of food listed or promoted on its internet website or mobile application.

(5) Provide a dedicated field on its platform for a microenterprise home kitchen operation to post the permit number and the name of the enforcement agency that issued the permit.

(6) Clearly and conspicuously post on its internet website or mobile application how a consumer can contact the internet food service intermediary through its internet website or mobile application if the consumer has a food safety or hygiene complaint and a link to the department's internet website that contains information for how to file a complaint with the enforcement agency.

(7) Submit the name and permit number of a microenterprise home kitchen operation to the enforcement agency that issued the permit to the microenterprise home kitchen operation if the internet food service intermediary receives, through its internet website or mobile application, three or more unrelated individual food safety or hygiene complaints in a calendar year from consumers that have made a purchase through its internet website or mobile application. The internet food service intermediary shall submit this information to the enforcement agency within two weeks of the third complaint received.

(8) If it is notified by the enforcement agency of significant food safety related complaints from a verified consumer that has made a purchase through its internet website or mobile application, submit to the enforcement agency the name and permit number of the microenterprise home kitchen operation where the food was purchased, and a list of consumers who purchased food on the same day from that microenterprise home kitchen operation through its internet website or mobile application.

(9) Prior to the listing or publication of a microenterprise home kitchen operation's offer of food for sale, obtain consent from the microenterprise home kitchen operation to make the disclosures to government entities required pursuant to this section.

(10) Shall not permit the use of the word "catering" or any variation of that word in a listing or publication of a microenterprise home kitchen operation's offer of food for sale.

(11) Shall not use, or knowingly facilitate the use of, a third-party delivery service for food produced by the microenterprise home kitchen operation, except as authorized pursuant to paragraph (2) of subdivision (b) of Section 114367.5.

24132

(c) (1) A microenterprise home kitchen operation that advertises to the public, including, but not limited to, advertising by website, internet, social media platform, newspaper, newsletter, or other public announcement, shall include all of the following within the advertisement:

(A) Name of the enforcement agency that issued the permit.

(B) Permit number.

(C) Statement that the food prepared is "Made in a Home Kitchen" in a clear and conspicuous font and location within a written advertisement and an audible and comprehensible manner in a verbal advertisement.

(2) A microenterprise home kitchen operation shall not use the word "catering" or any variation of that word in an advertisement relating to the microenterprise home kitchen operation's offer of food for sale.

33726

Lettering on this label shall contrast with the label's background and shall b

e not less than one-quarter inch in height, except for the word "WARNING" which shall be not less than one-half inch in height. 33846

Lettering on the label shall be black on a red background and not less than 1/4 inch in height except for the word "WARNING" which shall be not less than 1/2 inch in height.

34220

(b) In the case of a guardianship or conservatorship, the manufactured home, mobilehome, commercial coach, truck camper, or floating home shall be registered in the name of the person or persons designated as the conservators or guardians, as evidenced by documentation of that status deemed adequate by the department. The name shall be followed by the word "guardian" or "conservator," whichever is appropriate. Transfer of ownership or encumbrance of a manufactured home, mobilehome, commercial coach, truck camper, or floating home so registered shall require the signatures of all designated conservators or guardians.

34261

(d) All manufactured homes, mobilehomes, commercial coaches, truck campers, and floating homes registered, on or before January 1, 1985, in the names of two or more persons as tenants in common, as provided in subdivision (b), shall be considered to be the same as if the names of the tenants in common were separated by the word "or," as provided in subdivision (b).

34268

(b) A manufactured home, mobilehome, commercial coach, truck camper, or floating home may be registered in the names of two or more persons as tenants in common. If the names of the tenants in common are separated by the word "and," each tenant in common may transfer his or her individual interest in the manufactured home, mobilehome, commercial coach, truck camper, or floating home without the signature of the other tenant or tenants in common. However, the signature of each tenant in common shall be required to transfer full interest in the title to a new registered owner. If the names of the tenants in common are separated by the word "or," any one of the tenants in common may transfer full interest in the title to the manufactured home, mobilehome, commercial coach, truck camper, or floating home to a new registered owner without the signature of the other tenant or tenants in common. The signature of each tenant in common is required in all cases to encumber the title to the manufactured home, mobilehome, commercial coach, truck camper, or floating home.

35395

(a) No person or public agency shall advertise itself as, or hold itself out as, providing emergency medical services, by using in its name or advertising the word "emergency," or any derivation thereof, or any words which suggest that it is staffed and equipped to provide emergency medical services, unless the person or public agency satisfies one of the following requirements:\n (1) Is a general acute care hospital providing approved standby, basic, or comprehensive emergency medical services regulated by this chapter.\n (2) Meets all of the following minimum standards:\n (A) Emergency services are available in the facility seven days a week, 24 hours a day.\n (B) Has equipment, medication, and personnel experienced in the provision of services needed to treat life-, limb-, or function-threatening conditions.\n (C) Diagnostic radiology and clinical laboratory services are provided by persons on duty or on call and available when needed.\n (D) At least one physician who is trained and experienced in the provision of emergency medical care who is on duty or on call so as to be immediately available to the facility.\n (E) Medical records document the name of each patient who seeks care, as well as the disposition of each patient upon discharge.\n (F) A roster of speciality physicians who are available for referral, consultation, and speciality services is maintained and available.\n (G) Policies and procedures define the scope and conduct of treatment provided, including procedures for the management of specific types of emergencies.\n (H) The quality and appropriateness of emergency services are evaluated at least annually as part of a quality assurance program.\n (I) Provides information to the public that describes the capabilities of the facility, including the s

cope of services provided, the manner in which the facility complies with the requirements of this section pertaining to the availability and qualifications of personnel or services, and the manner in which the facility cooperates with the patient's primary care physician in followup care.\n (J) Clearly identifies the responsible professional or professionals and the legal owner or owners of the facility in its promotion, advertising, and solicitations.\n (K) Transfer agreements are in effect at all times with one or more general acute care hospitals which provide basic or comprehensive emergency medical services where in patients requiring more definitive care will be expeditiously transferred and receive prompt hospital care. Reasonable care shall be exercised to determine whether an emergency requiring more definitive care exists and the person seeking emergency care shall be assisted in obtaining these services, including transportation services, in every way reasonable under the circumstances.

35518

(c) Nothing in this article shall be construed to:\n (1) Prohibit a physician in private practice, an outpatient department of a general acute care hospital whether located on or off the premises of the hospital, or other entity authorized to offer medical services from advertising itself as, or otherwise holding itself out as, providing urgent, immediate, or prompt medical services, or from using in its name or advertising the words "urgent," "prompt," "immediate," any derivative thereof, or other words which suggest that it is staffed and equipped to provide urgent, prompt, or immediate medical services.\n (2) Prohibit prehospital emergency medical care personnel certified pursuant to, or any state or local agencies established pursuant to, this division, or any emergency vehicle operating within the emergency medical services system from using the word "emergency" in the title, classification, or designation of the personnel, agency, or vehicle.

35519

(d) Any person or public agency using the word "emergency" or any derivation thereof hereof in its name or advertising on January 1, 1987, but which would be prohibited from using the word or derivation thereof by this article, shall have until January 1, 1988, to comply with this article.

Name: node_text, dtype: object

In [70]:

```
with pd.option_context('display.max_colwidth', None):
    display(df2[df2['cluster'] == 10]['node_text'])
```

3102 For purposes of this chapter, the following definitions apply:

3190 For purposes of this chapter, the following definitions apply:

7445 For purposes of this chapter, the following definitions apply:

7667 For purposes of this chapter, the following definitions apply:

10321 For purposes of this chapter, the following definitions apply:

20618 For purposes of this chapter, the following definitions apply:

21252 For purposes of this chapter, the following definitions apply:

22328 For purposes of this chapter, the following definitions apply:

22416 For purposes of this chapter, the following definitions apply:

23604 For purposes of this chapter, the following definitions apply:

25527 For purposes of this chapter, the following definitions apply:

28574 For purposes of this chapter, the following definitions apply:

45349 For purposes of this chapter, the following definitions apply:

Name: node_text, dtype: object

In [71]:

```
with pd.option_context('display.max_colwidth', None):
    display(df2[df2['cluster'] == 11]['node_text'])
```

939 For purposes of this article, the following definitions apply: For
1315 purposes of this article, the following definitions apply: For purposes
1566 of this article, the following definitions apply:
1993 For purposes of this article, the following definitions apply:
2185 For purposes of this article, the following definitions apply:
2257 For purposes of this article, the following definitions apply:
7702 For purposes of this article, the following definitions apply:
10022 For purposes of this article, the following definitions apply:
14290 For purposes of this article, the following definitions apply:
26696 For purposes of this article, the following definitions apply:

Name: node_text, dtype: object

In [73]:

```
specific_words = ['article', 'section', 'chapter']

for cluster in df['cluster'].unique():
    print(f"Cluster {cluster} specific words count:")
    cluster_data = df[df['cluster'] == cluster]
    all_text = ' '.join(cluster_data['node_text'].astype(str)).lower()
    word_counts = {word: all_text.count(word) for word in specific_words}
    print(word_counts)
```

```
Cluster 0 specific words count:
{'article': 3176, 'section': 23396, 'chapter': 6972}
Cluster 3 specific words count:
{'article': 221, 'section': 2115, 'chapter': 510}
Cluster 8 specific words count:
{'article': 83, 'section': 86, 'chapter': 113}
Cluster 2 specific words count:
{'article': 30, 'section': 355, 'chapter': 31}
Cluster 4 specific words count:
{'article': 8, 'section': 257, 'chapter': 11}
Cluster 7 specific words count:
{'article': 22, 'section': 799, 'chapter': 84}
Cluster 1 specific words count:
{'article': 115, 'section': 1066, 'chapter': 294}
Cluster 6 specific words count:
{'article': 3, 'section': 163, 'chapter': 16}
Cluster 5 specific words count:
{'article': 14, 'section': 186, 'chapter': 31}
Cluster 11 specific words count:
{'article': 10, 'section': 0, 'chapter': 0}
Cluster 10 specific words count:
{'article': 0, 'section': 0, 'chapter': 13}
Cluster 9 specific words count:
{'article': 5, 'section': 8, 'chapter': 2}
```

Issues with Cluster Model

It is not always intuitive as to which specific definition scope the nodes are. Some of the clusters are very nice and output wonderful clusters that will be very helpful in next steps when we go to extract the definition nodes to use in our program.

However, after analysis of cluster model and cross tabulation although the model itself performed wonderfully, the analysis concluded that maybe more thought must go into identifying patterns. For example, although all words within quotes reference some definition, it does not always mean that it is the actual definition but rather could be the

word within a law or section that is referencing somewhere in the section, article, chapter that the word is defined somewhere.

This gets tricky when we go to extract the definitions hence my attempt to start solving it.

Especially after finding out that not all words between quotes mean definite definition, it could be just a reference to the word that is contained in a definition elsewhere. This effects the method of creating the clusters because any cluster that contains nodes that were selected with pattern " " [possible definition3] are not necessarily defintions that will be useful for what we need them to be (ie. they are not actual definitions of the word).

Next Steps

The next steps which I have already began but decided to exlcude from this project because it is not complete and still deciding on the logic to include would be to subcluster the clusters above. Initially will need to look over text patterns originally defined and make some adjustments. Then from there subcluster each cluster to futher see if I can narrow down exactly which nodes belong to which definition scope.

Thank You