

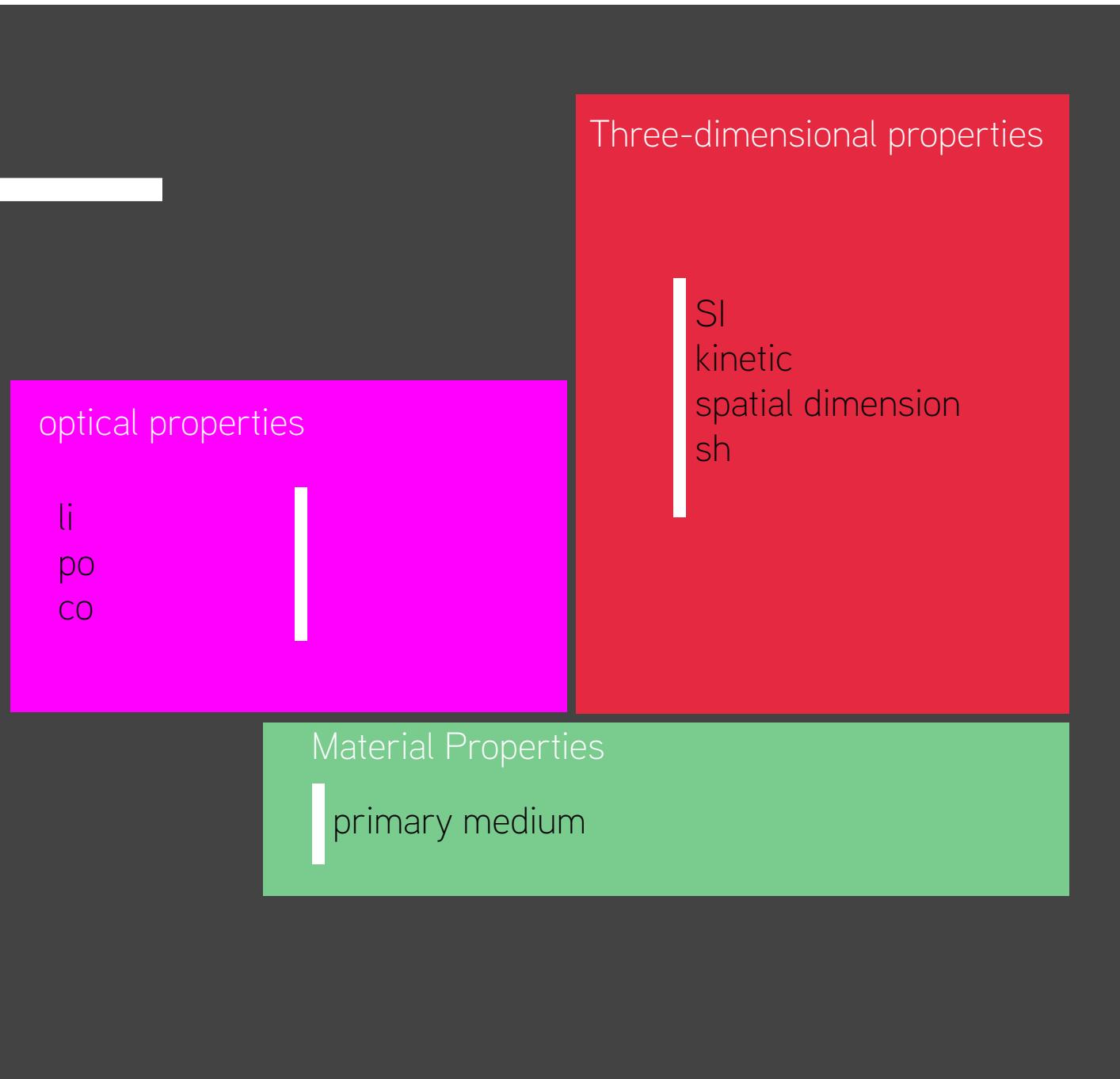
# Unsupervised Curator

Paik, Soonk

## Feature Selection

8 features were selected for clustering.

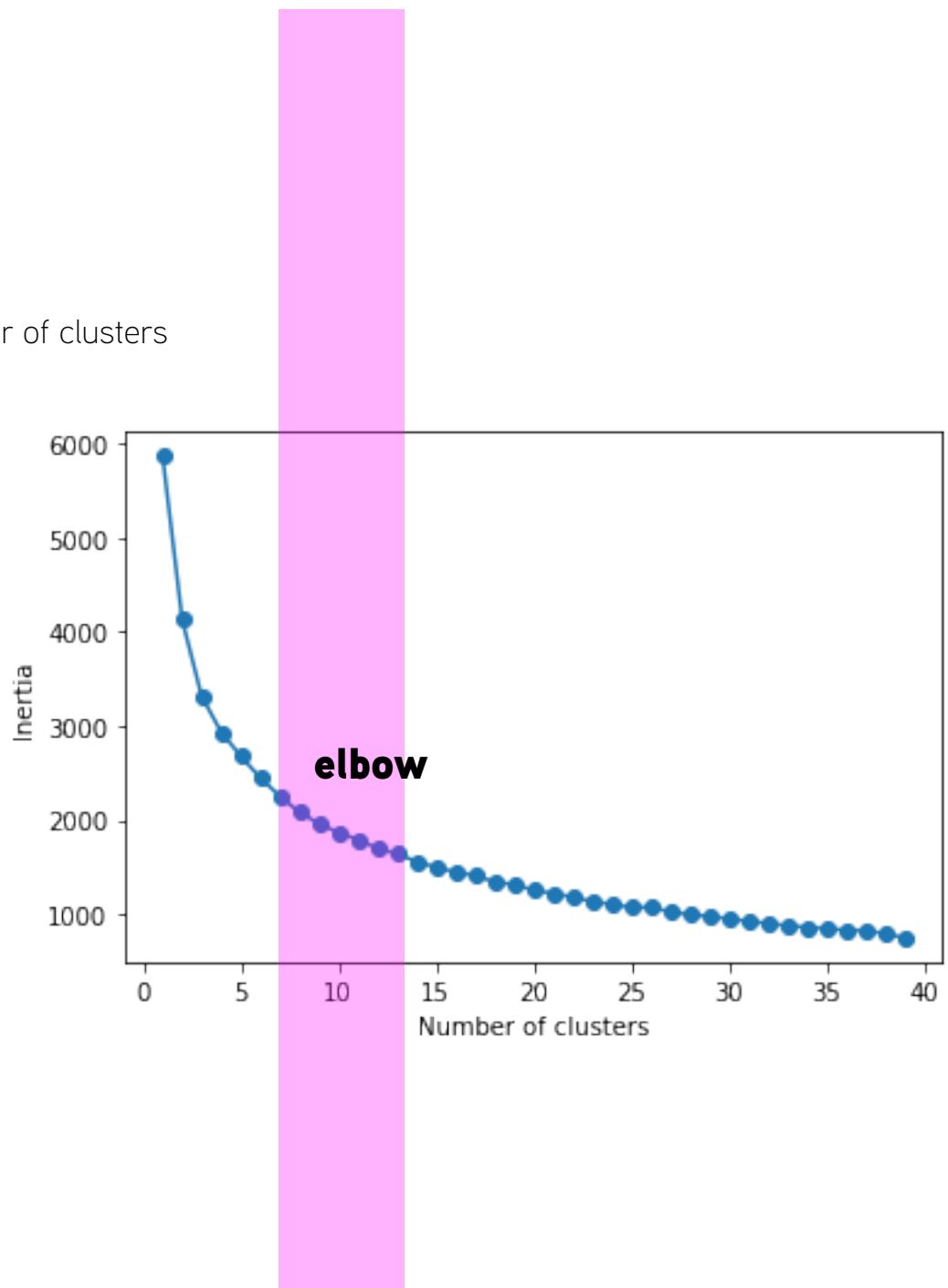
Especially for primary medium,  
unique numbers are assigned to each  
primary medium.  
(Since the name of the medium is not  
integer that transforming the string to  
integer was necessary)



## Kmeans

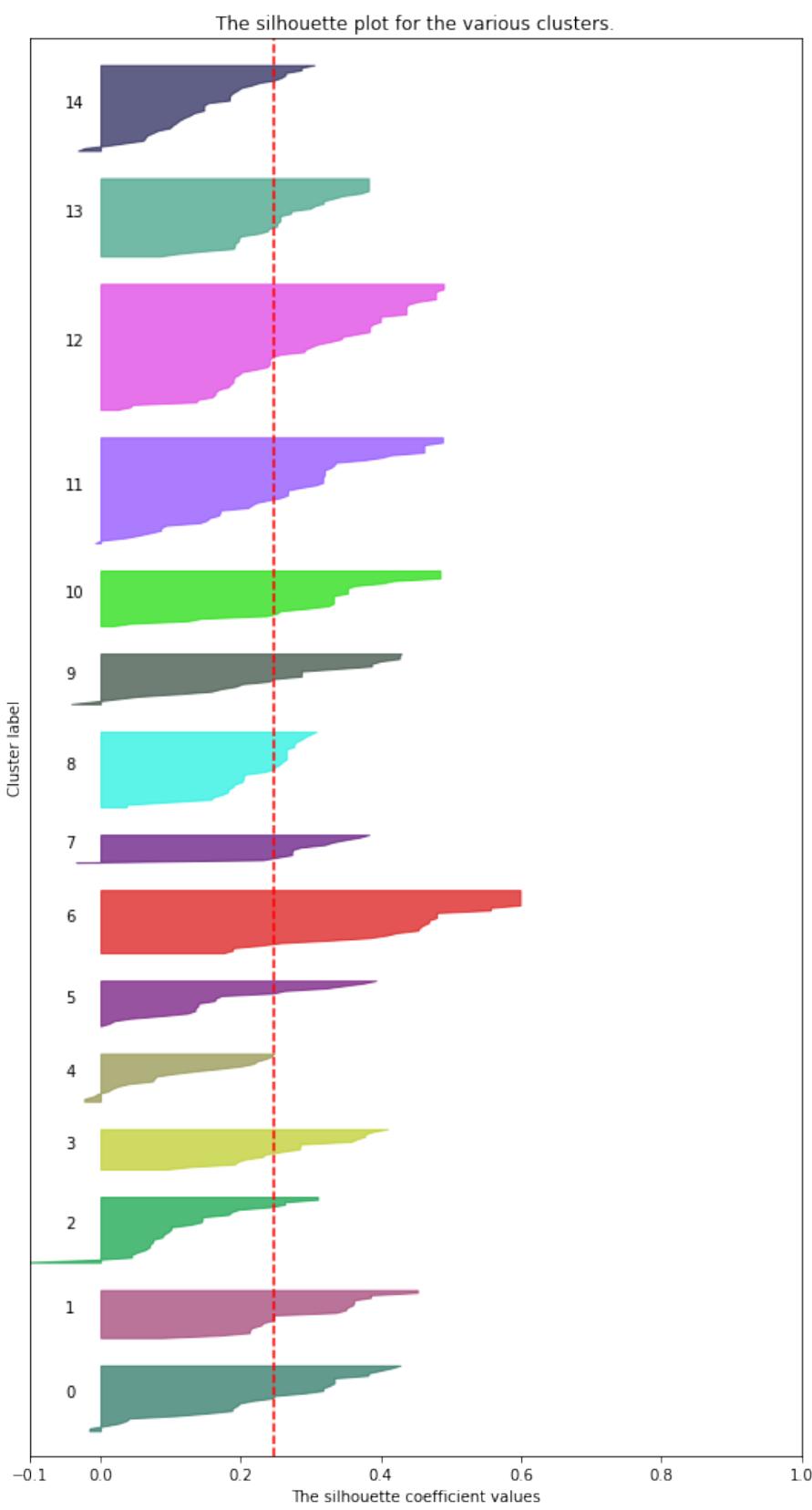
After plotting the number of clustering and the Inertia value, Elbow area ranges around 10-15.

```
# first attempt at fitting K means to view change in Inertia  
  
# container to store inertia scores over iterations  
distortions = []  
  
# fit KMeans iteratively to begin to assess the appropriate number of clusters  
for i in range(1, 40):  
    km = KMeans(n_clusters=i)  
    km.fit(X)  
    distortions.append(km.inertia_)  
  
# visualize change in inertia  
plt.plot(range(1, 40), distortions, marker='o')  
plt.xlabel('Number of clusters')  
plt.ylabel('Inertia')  
plt.show()
```

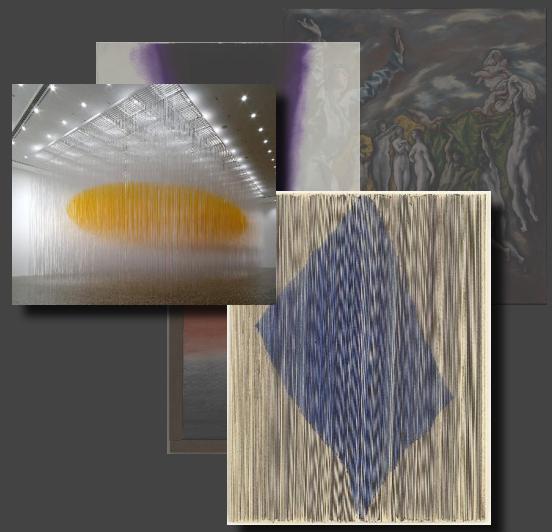


## Silhouette Plotting

Silhouette Plotting shows the detailed distribution of the clustering



Gallery of Falling Structure



Gallery of Monochrome Figures



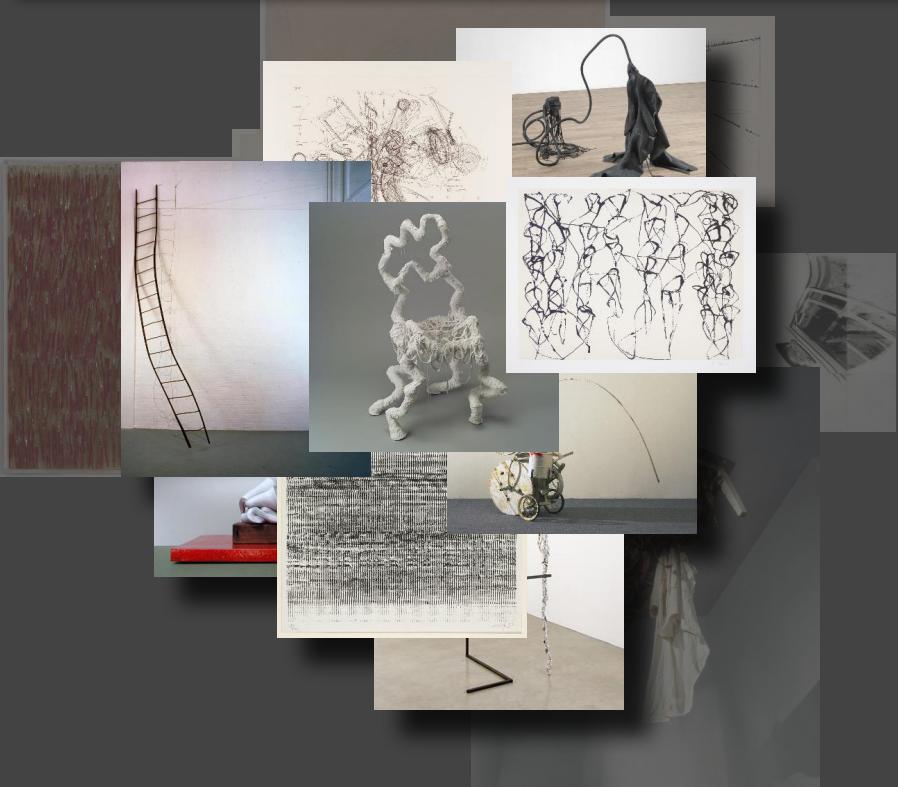
Gallery of Vivid Surreal/Abstract



Gallery of Distorted Perspective



Gallery of the Tangled



Gallery of Emptiness



Difference btw Human's perspective  
& ML's perspective

The graph shows how closely and precisely clustering are made.  
And the previous slide shows the evaluation by human's perspective (my perspective).  
Supervised labeling and ML's classification has dramatic discrepancies.

For the Hidden patterns that human cannot recognize were captured by ML.

