Let's discuss a general question first.

Imagine you have created a **diary milk chocolate product**  and 1000 chocolates already sold. Now you want to know what's the feedback of the diary milk?

So here you will ask to the customers about the product taste, color, packaging and the cost satisfaction with respect to quality. But asking 1000 people and collecting the information is so time consuming. So other method can be a social media where generally people share their opinion ex Twitter. This is the reason behind choosing Twitter for further analysis.

Now coming to the paper the issue is similar to the above example.

Issue of the paper is "Identification of supply chain management issues in food industry."

Data used: 3weeks of Twitter data

**Industry of focus**:. Food(beef)

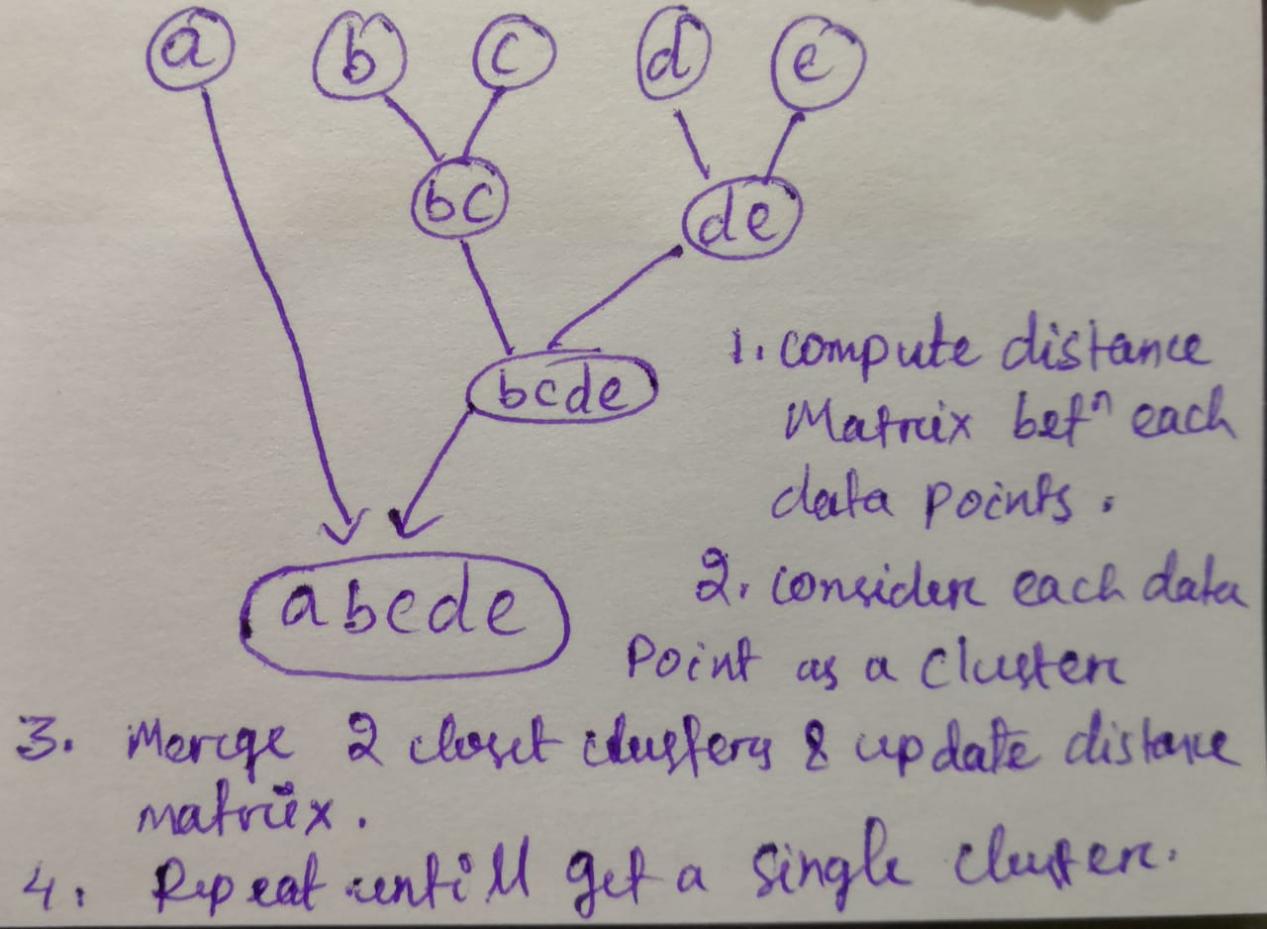
Approach:. Svm, hierarchical clustering with boot starp resampling algorithms

**Expected result**-- cluster of words which will identify , recommend n predict to supply chain management for future decision making

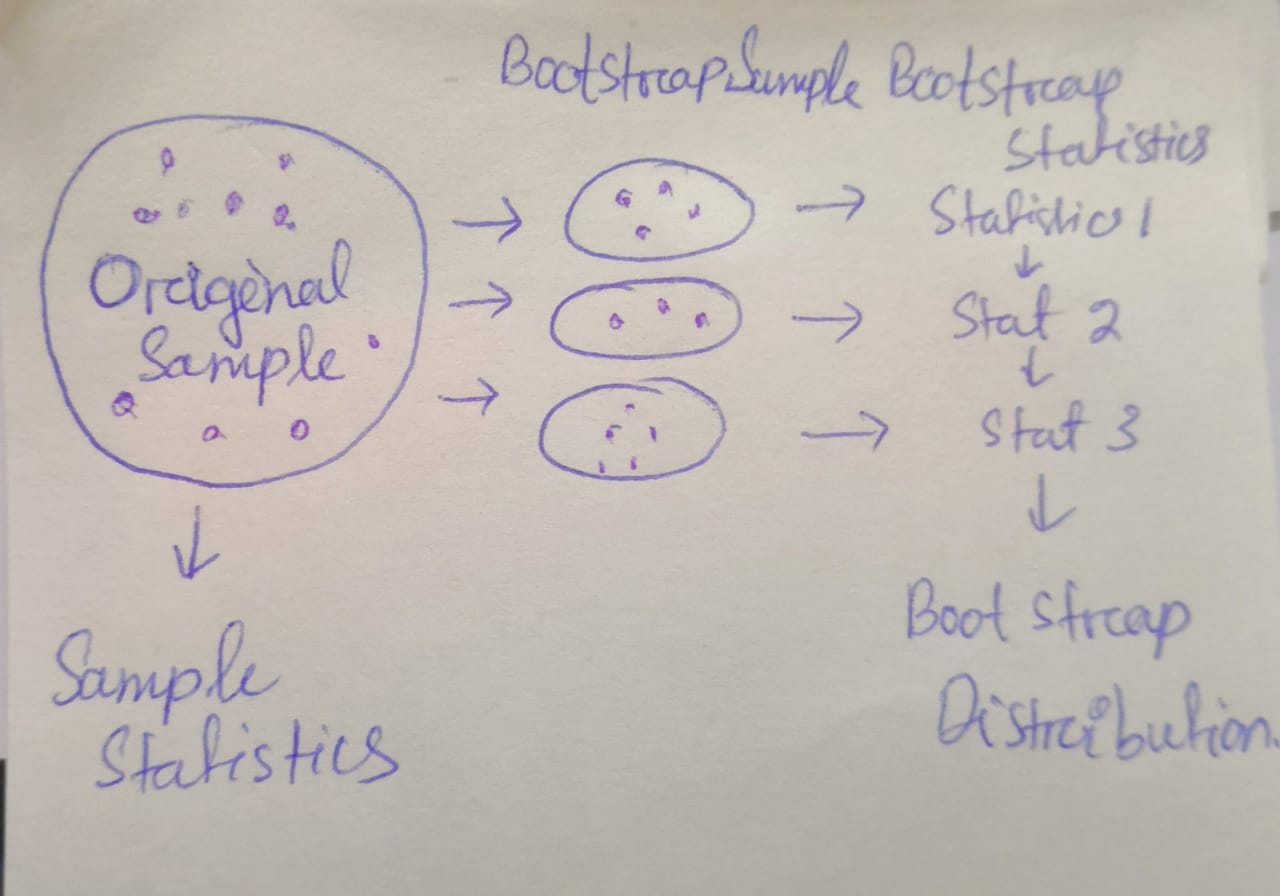
For identifying a consumers view social media is the cheap, diverse and full of large audiences. Even real time data can also be captured over this Twitter platform. Previously NVivo software was used to collect data but due to cost n accuracy issues that method of collection doesn't work.

**Methods**---

From Twitter posts by using streaming APis they collected the keyword list and got the data set in json, txt format. After data collection they used sentiment analysis to find +ve and-ve response of the consumers through tweets. Finally used svm algo to segregate the tweets for further process. They used p valued resampling to create a hierarchical cluster of tweets.



This is the simple way to explain hierarchical clustering.



This is the bootstrapping method.

Total tweets was 1338638 using beef n steak word. Only English language is considered.

**Results-**-

Most frequently used -ve words like, beef were smell, deal, color, taste, Bone with 0.95 significance.

+Ve tweets like fresh, dish, celebrate, share...

Analysis from.country wise like USA, Australia, usa has been noted.

**Issue identification from county wise analysis-**----

Bad flavor n unpleasant smell due to lack of oxidation,

Discolor of beefs

Hard texture and presence of foreign body

**Conclusion:**

Main reason for dissatisfaction of consumers are

1. Color

2. Food safety

3. Smell

4. Flavor

Few are satisfied due to proper promotion, deals n particular beef products like drinks.

**Limitations in models:**

During hierarchical cluster they found few tweets r not related to beef supply chain so they couldn't explain properly.

Future work: they want to implement for other food supply chain like lamb, pork.

**What I observed to change** --

Here they lacked to create a recommendation for food supply chain,

In algo hierarchical cluster will definitely add other key words due to lack in use of NER n tokenization.

For normalizing and improvising the text extraction we can first preprocess the data for both emotion analysis and feature extractions.

After finding the correct words for consumer behavior we can use svm for classification.

Another point even if they calculated country wise +ve and -ve response, they could have ensemble it but they didn't even generalize it which we can modify and the solution to work.