**Market Basket Analysis**

In this project , we are going to develop ,**Market Basket Analysis using Association Rule Mining** .

In Common Man terms we try to develop a Program that will predict the frequent items purchased by the customers in a given supermarket.

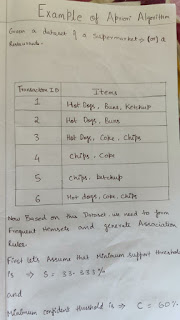
We apply a particular algorithm known as Apriori Algorithm which comes under **Association Rule Mining**on a given Dataset , to predict the Frequent itemsets based on its Confidence , Support etc.

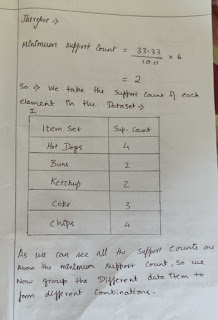
First Let us Look at What Apriori Algorithm is:

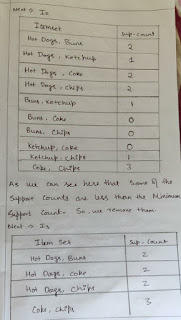
**APRIORI ALGORITHM**

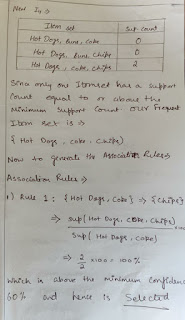
**Apriori algorithm** is used for finding frequent itemsets in a dataset for boolean association rule. Name of the algorithm is Apriori because it uses prior knowledge of frequent itemset properties. We apply an iterative approach or level-wise search where k-frequent itemsets are used to find k+1 itemsets.

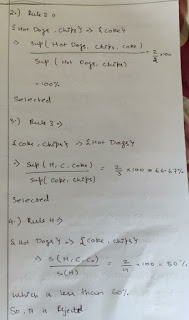
Lets take a look at an example to understand Apriori Algorithm Better:

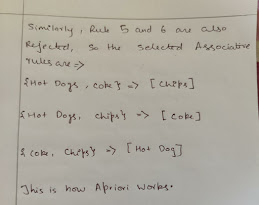
[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

So this is how we use apriori Algorithm to develope Associative rules for our Market Basket Analysis.

So now we will See the Coding implementation of it:

We  will be Implementing this project in PYTHON.

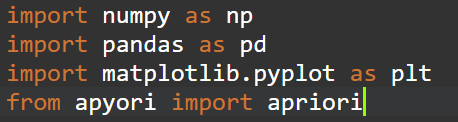
**IMPLEMENTATION**

**Dataset Description**

* Different products given 7500 transactions over the course of a week at a retail store
* So our database is that of a Retail Store OR supermarket.
* Based on this database we will calculate our Association Rules..
* We have library(**apyori**) to calculate the association rule using Apriori.

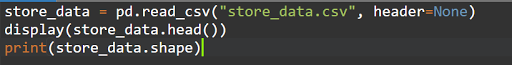
**First Step is to :**

**Import the Library:**

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

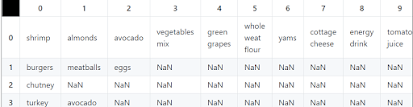
**Next:**

**Read data and Display:**

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

So this reads our dataset and Displays it.

For Example:

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

**Next we :**

**Preprocessing on Data**

* Here we need a data in form of list for Apriori Algorithm.

[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)

1. **Now to apply the Algorithm:**

* We have provide min\_support, min\_confidence, min\_lift, and min length of sample-set for find rule.
* **Measure 1: Support.**

This says how popular an itemset is, as measured by the proportion of transactions in which an itemset appears.

* **Measure 2: Confidence.**

This says how likely item Y is purchased when item X is purchased, expressed as {X -> Y}. This is measured by the proportion of transactions with item X, in which item Y also appears.

* **Measure 3: Lift.**

This says how likely item Y is purchased when item X is purchased, while controlling for how popular item Y is.

* We have seen this in the above example.
* **Now we derive the Frequent item sets & Association Rules:**
* **We use the Following syntac and code for generating the Frequent Item Sets:**

**[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)**

**//code:**

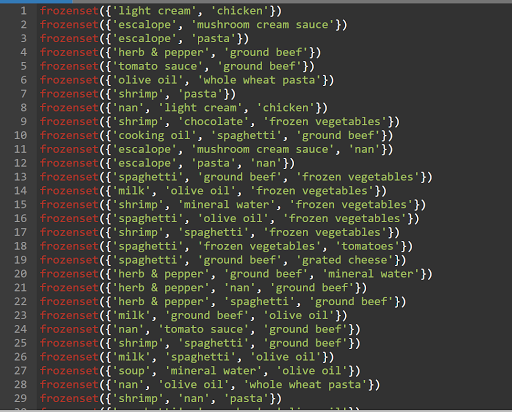
association\_rules = apriori(records, min\_support=0.0045, min\_confidence=0.2, min\_lift=3, min\_length=2)

association\_results = list(association\_rules)

* **To Print the Frequent Item sets we:**

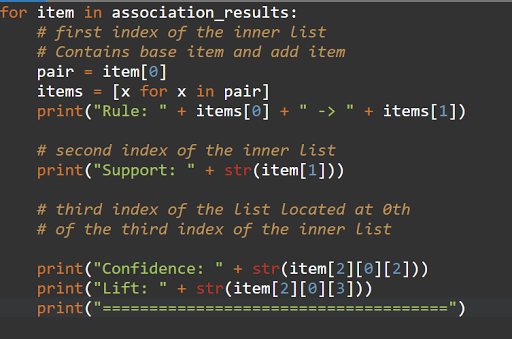
**[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)**

* **The Frequent Itemsets are:**

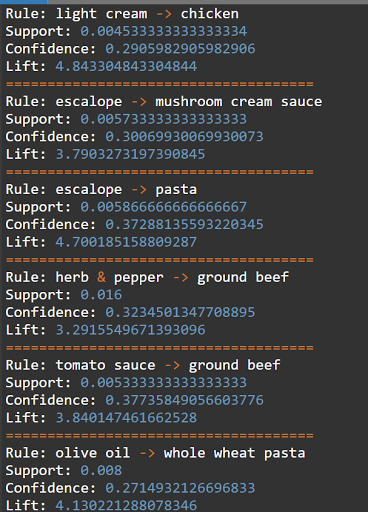
**[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)**

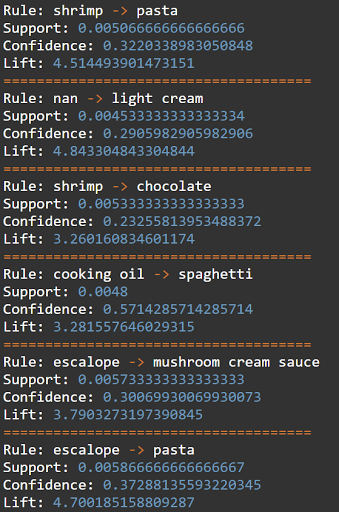
Now based on these frequent item sets , We will generate Associative rules using Support , Confidence and Lift:

* **We find the Associative rules:**

**[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)**

* **So the Associative rules we generate are:**

**[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)**

**[](https://www.blogger.com/blog/post/edit/9075165755386437816/8369239820851701279)**

So we have successfully completed our project and have derived our desired output of Associative rules and frequent itemsets. using Apriori Algorithm.

Hence we have successfully completed Market Basket Analysis.