Capstone Project Report

Name: Pooja Ramdas Kadam

Course: Al & ML (Batch - 4)

Problem Statement

Apriori is a statistical algorithm for implementing associate rule mining, that primarily relies on three components: Life, Support and Confidence. Using this algorithm try to find the rules that describe the relation between each of the products that were brought by the customers

Prerequisites

Along with Python below packages needed to be installed

Apyori

Pandas

Dataset Used

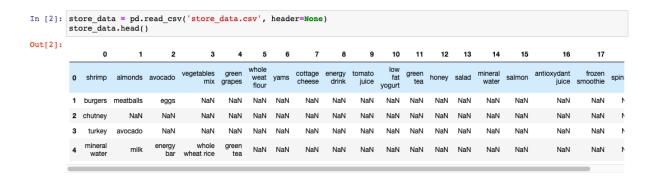
https://drive.google.com/file/d/1y5DYn0dGoSbC22xowBq2d4po6h1JxcTQ/view

Implementation

Import required libraries and load data

In [1]: import numpy as np
import pandas as pd
from apyori import apriori

Load data



Apply Apriori

```
In [18]: results = list(apriori(store_data_list, min_support=0.005))
                            print(len(results))
In [19]: results
                            [RelationRecord(items=frozenset({'almonds'}), support=0.020397280362618318, ordered_statistics=[OrderedStatistic(items_base=frozenset(), items_add=frozenset({'almonds'}), confidence=0.020397280362618318, lift=1.0)]),
                              RelationRecord(items=frozenset({'antioxydant juice'}), support=0.008932142381015865, ordered_statistics=[OrderedSt
                             atistic(items_base=frozenset(), items_add=frozenset({'antioxydant juice'}), confidence=0.008932142381015865, lift=
                             1.0)]),
                              RelationRecord(items=frozenset({'avocado'}), support=0.03332888948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]), support=0.03332888948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]), support=0.03332888948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]), support=0.03332888948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]), support=0.03332888948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]]), support=0.0333288948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]]), support=0.0333288948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]]), support=0.0333288948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]]), support=0.0333288948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]]), support=0.0333288948140248, ordered_statistics=[OrderedStatistic(items=frozenset(avocado')]]), support=0.0333288948, ordered_statistic(items=frozenset(avocado')]), support=0.033328894, 
                            ms_base=frozenset(), items_add=frozenset({'avocado'}), confidence=0.03332889948140248, lift=1.0)]),
RelationRecord(items=frozenset({'bacon'}), support=0.008665511265164644, ordered_statistics=[OrderedStatistic(items=frozenset({'bacon'}), support=0.008665511265164644, ordered_statistics=[OrderedStatistic(items=frozenset({'bacon'}), support=0.008665511265164644, ordered_statistics=[OrderedStatistic(items=frozenset({'bacon'}), support=0.008665511265164644, ordered_statistics=[OrderedStatistic(items=frozenset({'bacon'}), support=0.008665511265164644, ordered_statistics=[OrderedStatistic(items=frozenset({'bacon'}), support=0.008665511265164644, ordered_statistics=[OrderedStatistics]
                             s_base=frozenset(), items_add=frozenset({'bacon'}), confidence=0.008665511265164644, lift=1.0)]),
                            S_base=102enset(), tems_add=102enset({ bacob(), support=0.010798560191974404, ordered_statistics=[OrderedStatistic(items_base=frozenset(), items_add=frozenset({ barbecue sauce})), confidence=0.010798560191974404, lift=1.0)]),
                             RelationRecord(items=frozenset({'black tea'}), support=0.014264764698040262, ordered_statistics=[OrderedStatistic (items_base=frozenset(), items_add=frozenset({'black tea'}), confidence=0.014264764698040262, lift=1.0)]), RelationRecord(items=frozenset({'blueberries'}), support=0.009198773496867084, ordered_statistics=[OrderedStatisti
                             c(items_base=frozenset(), items_add=frozenset({'blueberries'}), confidence=0.009198773496867084, lift=1.0)])
                                RelationRecord(items=frozenset({'body spray'}), support=0.011465137981602452, ordered statistics=[OrderedStatistic
                             (items_base=frozenset(), items_add=frozenset({'body spray'}), confidence=0.011465137981602452, lift=1.0)]),
                            RelationRecord(items=frozenset(('brownies')), support=0.03372883615517931, ordered_statistics=[OrderedStatistic(items_base=frozenset(), items_add=frozenset(('brownies')), confidence=0.03372883615517931, lift=1.0)]),
```