

DATA MINING

Assignment 1

Group:

Naren Surampudi: 2016AAPS0206H
V Aditya Srikanth 2016A7PS0091H
Prateek Dasgupta 2016A1PS0602H

Dataset:

groceries.csv

Preprocessing:

Apriori: None on CSV, CSV read using Pandas.

FP Tree: Remove non-frequent itemsets, rearrange in descending order of support values, truncate None values

Frequent Itemsets:

Support: 200- **121**

Support: 300- **61**

Support: 400- **40**

Association Rules:

Support: 200, Confidence: 30%- **36**

Support: 300, Confidence: 20%-

Support: 400, Confidence: 10%- **16**

Maximal Frequent Itemsets with Support:

Support: 200-

('bottled beer', 'whole milk'): 201,
('bottled water', 'other vegetables'): 244,
('bottled water', 'rolls/buns'): 238,
('bottled water', 'soda'): 285,
('bottled water', 'whole milk'): 338,
('bottled water', 'yogurt'): 226,
('brown bread', 'whole milk'): 248,
('citrus fruit', 'other vegetables'): 284,
('citrus fruit', 'whole milk'): 300,
('citrus fruit', 'yogurt'): 213,
('domestic eggs', 'other vegetables'): 219,
('domestic eggs', 'whole milk'): 295,
('frankfurter', 'whole milk'): 202,
('frozen vegetables', 'whole milk'): 201,
('fruit/vegetable juice', 'whole milk'): 262,

('margarine', 'whole milk'): 238,
('newspapers', 'whole milk'): 269,
('other vegetables', 'fruit/vegetable juice'): 207,
('other vegetables', 'pastry'): 222,
('other vegetables', 'rolls/buns'): 419,
('other vegetables', 'sausage'): 265,
('other vegetables', 'shopping bags'): 228,
('other vegetables', 'whole milk', 'yogurt'): 219,
('pastry', 'rolls/buns'): 206,
('pip fruit', 'other vegetables'): 257,
('pip fruit', 'tropical fruit'): 201,
('pip fruit', 'whole milk'): 296,
('pork', 'other vegetables'): 213,
('pork', 'whole milk'): 218,
('rolls/buns', 'yogurt'): 338,
('root vegetables', 'other vegetables', 'whole milk'): 228,
('root vegetables', 'rolls/buns'): 239,
('root vegetables', 'yogurt'): 254,
('sausage', 'rolls/buns'): 301,
('shopping bags', 'whole milk'): 241,
('soda', 'other vegetables'): 322,
('soda', 'pastry'): 207,
('soda', 'rolls/buns'): 377,
('soda', 'sausage'): 239,
('soda', 'shopping bags'): 242,
('soda', 'whole milk'): 394,
('soda', 'yogurt'): 269,
('tropical fruit', 'other vegetables'): 353,
('tropical fruit', 'rolls/buns'): 242,
('tropical fruit', 'root vegetables'): 207,
('tropical fruit', 'soda'): 205,
('tropical fruit', 'whole milk'): 416,
('tropical fruit', 'yogurt'): 288,
('whipped/sour cream', 'other vegetables'): 284,
('whipped/sour cream', 'whole milk'): 317,
('whipped/sour cream', 'yogurt'): 204,
('whole milk', 'beef'): 209,
('whole milk', 'butter'): 271,
('whole milk', 'curd'): 257,
('whole milk', 'pastry'): 327,
('whole milk', 'rolls/buns'): 557,
('whole milk', 'sausage'): 294

Support: 300-

('bottled water', 'whole milk'): 338,
('citrus fruit', 'whole milk'): 300,
('other vegetables', 'soda'): 322,
('other vegetables', 'whole milk'): 736,
('pastry', 'whole milk'): 327,
('rolls/buns', 'other vegetables'): 419,
('rolls/buns', 'sausage'): 301,

('rolls/buns', 'soda'): 377,
('rolls/buns', 'whole milk'): 557,
('root vegetables', 'other vegetables'): 466,
('root vegetables', 'whole milk'): 481,
('tropical fruit', 'other vegetables'): 353,
('tropical fruit', 'whole milk'): 416,
('whole milk', 'soda'): 394,
('whole milk', 'whipped/sour cream'): 317,
('yogurt', 'other vegetables'): 427,
('yogurt', 'rolls/buns'): 338,
('yogurt', 'whole milk'): 551

Support: 400-

('other vegetables', 'rolls/buns'): 419,
('other vegetables', 'root vegetables'): 466,
('other vegetables', 'whole milk'): 736,
('other vegetables', 'yogurt'): 427,
('rolls/buns', 'whole milk'): 557,
('root vegetables', 'whole milk'): 481,
('whole milk', 'tropical fruit'): 416,
('whole milk', 'yogurt'): 551

Closed Frequent Itemsets with Support:

Support: 200-

('bottled water', 'other vegetables'): 244,
('bottled water', 'rolls/buns'): 238,
('bottled water', 'soda'): 285,
('bottled water', 'whole milk'): 338,
('brown bread', 'whole milk'): 248,
('citrus fruit', 'other vegetables'): 284,
('citrus fruit', 'whole milk'): 300,
('domestic eggs', 'whole milk'): 295,
('fruit/vegetable juice', 'whole milk'): 262,
('margarine', 'whole milk'): 238,
('newspapers', 'whole milk'): 269,
('other vegetables', 'rolls/buns'): 419,
('other vegetables', 'sausage'): 265,
('other vegetables', 'shopping bags'): 228,
('other vegetables', 'whole milk'): 736,
('other vegetables', 'yogurt'): 427,
('pip fruit', 'other vegetables'): 257,
('pip fruit', 'whole milk'): 296,
('rolls/buns', 'yogurt'): 338,
('root vegetables', 'other vegetables'): 466,
('root vegetables', 'rolls/buns'): 239,
('root vegetables', 'whole milk'): 481,
('root vegetables', 'yogurt'): 254,
('sausage', 'rolls/buns'): 301,
('shopping bags', 'whole milk'): 241,
('soda', 'other vegetables'): 322,

('soda', 'rolls/buns'): 377,
('soda', 'sausage'): 239,
('soda', 'shopping bags'): 242,
('soda', 'whole milk'): 394,
('soda', 'yogurt'): 269,
('tropical fruit', 'other vegetables'): 353,
('tropical fruit', 'rolls/buns'): 242,
('tropical fruit', 'whole milk'): 416,
('tropical fruit', 'yogurt'): 288,
('whipped/sour cream', 'other vegetables'): 284,
('whipped/sour cream', 'whole milk'): 317,
('whole milk', 'butter'): 271,
('whole milk', 'curd'): 257,
('whole milk', 'pastry'): 327,
('whole milk', 'rolls/buns'): 557,
('whole milk', 'sausage'): 294,
('whole milk', 'yogurt'): 551

Support: 300-

None

Support: 400-

None

Redundant Association Rules

Support: 200, Confidence: 30%

('other vegetables', 'root vegetables'), 'whole milk'): 0.4892703862660944,
(('whole milk', 'other vegetables'), 'root vegetables'): 0.30978260869565216,
(('whole milk', 'root vegetables'), 'other vegetables'): 0.47401247401247404,
(('yogurt', 'other vegetables'), 'whole milk'): 0.5128805620608899,
(('yogurt', 'whole milk'), 'other vegetables'): 0.397459165154265