Population Health Modeling: Day 11

Yale SCHOOL OF PUBLIC HEALTH
Data Science and Data Equity

Check-In

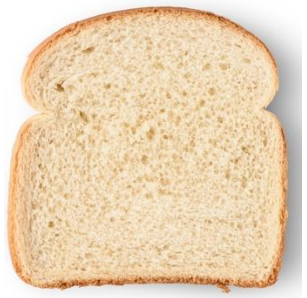
Brief Review

- Basic principles for public health research and modeling
- Outbreak investigations
- Everything you need to know about *Streptococcus pneumoniae*
- Introduce time series analysis
- Adapting your data, time series analysis
- Introduce hierarchical modeling
- Adapting your data, hierarchical modeling
- Presenting your data and results, round 1

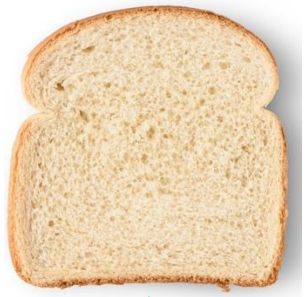
Brief Overview

- Introduce Market Basket Analysis
  - Walk-through example
- DIY Market Basket Analysis

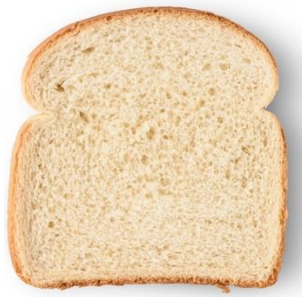
Association rules mining: Basics



Association rules mining: Basics

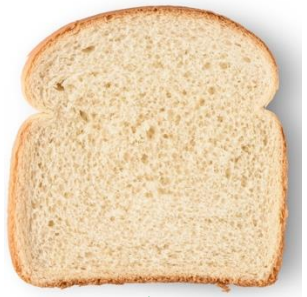


Association rules mining: Basics

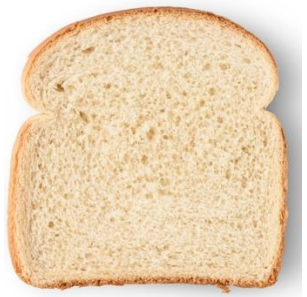


This is a rule:
Bread + Peanut Butter

Association rules mining: Basics



Association rules mining: Basics

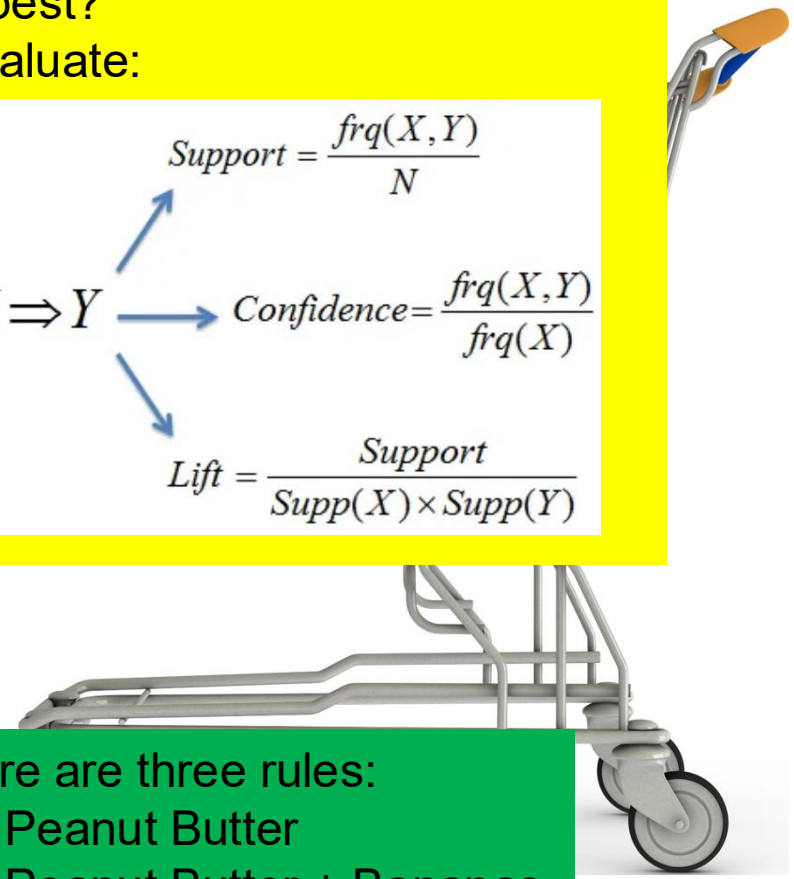


Which rule is best?
3 criteria to evaluate:
Support
Confidence
Lift

$$Support = \frac{frq(X,Y)}{N}$$

$$X \Rightarrow Y \rightarrow Confidence = \frac{frq(X,Y)}{frq(X)}$$

$$Lift = \frac{Support}{Supp(X) \times Supp(Y)}$$



Now there are three rules:
Bread + Peanut Butter
Bread + Peanut Butter + Bananas
Bread + Peanut Butter + Jelly

Deciding which rules are best

For a rule X => Y, where X is bread and Y is peanut butter:

Support: how often an item X appears in all transactions

Support{X} = X/N

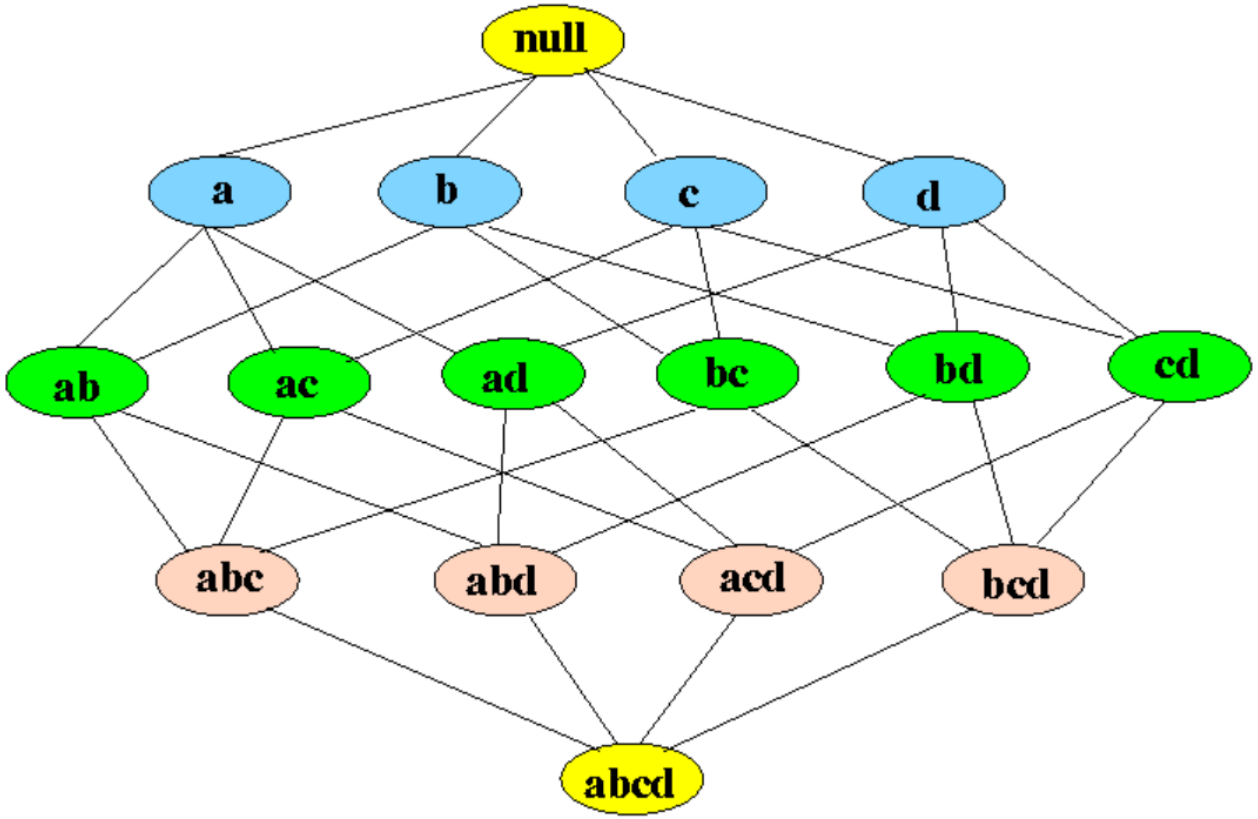
Confidence: how often Y is purchased given that X is purchased in a transaction

Confidence {X=>Y} = Support {X,Y}/ Support{X}

Lift: how often Y is purchased given that X is purchased, controlling for popularity of Y

Lift {X=>Y} = Support {X,Y}/(Support{X} * Support{Y})

Many Rules are possible!



https://www.datacamp.com/community/tutorials/market-basket-analysis-r

Market Basket Analysis for Health/Biology

frontiers in Pharmacology

Front Pharmacol. 2021; 12: 641530.
Published online 2021 Apr 27. doi: 10.3389/fphar.2021.641530
PMCID: PMC811089
PMID: 3398667

Uncovering Modern Clinical Applications of Fuzi and Fuzi-Based Formulas: A Nationwide Descriptive Study With Market Basket Analysis

Chi-Jung Tai, [1,2] Mohamed El-Shazly, [3,4] Yi-Hong Tsai, [1] Dezső Csupor, [5] Judit Hohmann, [5] Yang-Chang Wu, [6,7] Tzyy-Guey Tseng, [8] Fang-Rong Chang, [1,9,10,11,*] and Hui-Chun Wang [1,9,10,11,*]

ELSEVIER

Journal of Biomedical Informatics
Volume 43, Issue 6, December 2010, Pages 891-901

An automated technique for identifying associations between medications, laboratory results and problems

Adam Wright [a,b], Elizabeth S. Chen [c,d], Francine L. Maloney [e]

International Journal of

Extracting Diagnosis Patterns in Elect Records using Association Rule

Stephen M. Kang'ethe
School of Computing and Informatics, University of Nairobi
P. O. Box 30197 – 00100 Nairobi, Kenya

Pete
School of Computin
P. O. Box 3019

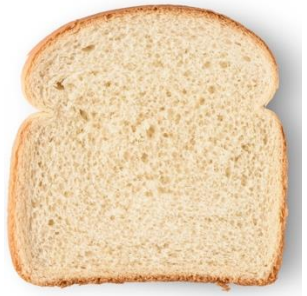
Correlates of Nonrandom Patterns of Serotype Switching in Pneumococcus

FREE

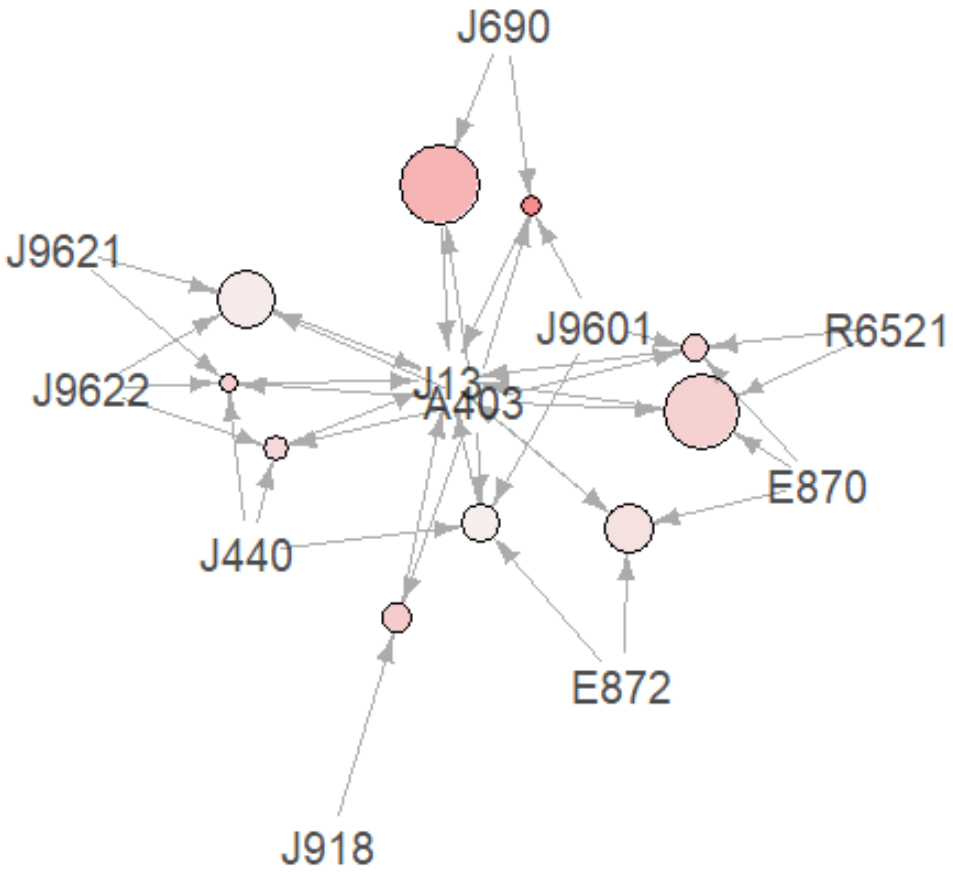
Shreyas S Joshi, Mohammad A Al-Mamun, Daniel M Weinberger

The Journal of Infectious Diseases, Volume 221, Issue 10, 15 May 2020, Pages 1669–1676,
https://doi.org/10.1093/infdis/jiz687
Published: 25 December 2019 **Article history**

Find New Associations!



E.g. Identifying clinical factors associated with recurrence of pneumonia



ICD-10 code	Description
A403	Sepsis due to *Streptococcus pneumoniae*
J9601	Acute respiratory failure with hypoxia
R6521	Severe sepsis with septic shock
E870	Hyperosmolality and hypernatremia
E872	Acidosis
J918	Pleural effusion in conditions classified elsewhere
J440	Chronic obstructive pulmonary disease with (acute) lower respiratory infection
J9622	Acute and chronic respiratory failure with hypercapnia
J9621	Acute and chronic respiratory failure with hypoxia
J690	Pneumonitis due to inhalation of food and vomit

Market Basket Analysis Example

- Perforating Acute Otitis Media in children

- Pneumococcal AOM vs other bacterial or viral causes
  - VT Pneumococcal AOM vs other bacterial or viral causes

- Pneumococcal AOM with NP carriage
  - Of pneumococcus, of other bacteria