

# Association Rule Mining

## Basic Theory and Practice

Takuma Kimura

# Contents at a Glance

Section 1. What is Association Rule Mining?

Section 2. Basic Metrics

Section 3. Apriori Algorithm

Section 4. Market Basket Analysis with Python

# **Section 1**

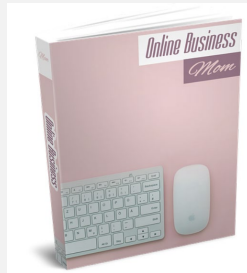
## **What is Association Rule Mining?**

# **1. What is Association Rule Mining?**

# A Case of E-Commerce Site



## Recommended for you

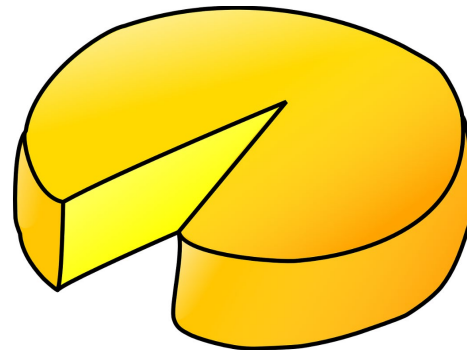
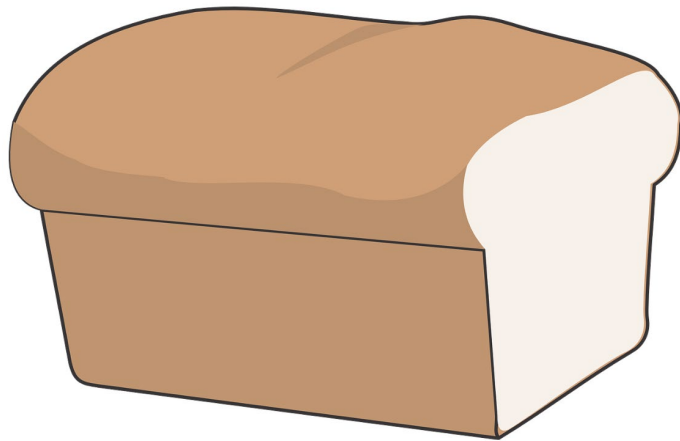


## Your purchase history



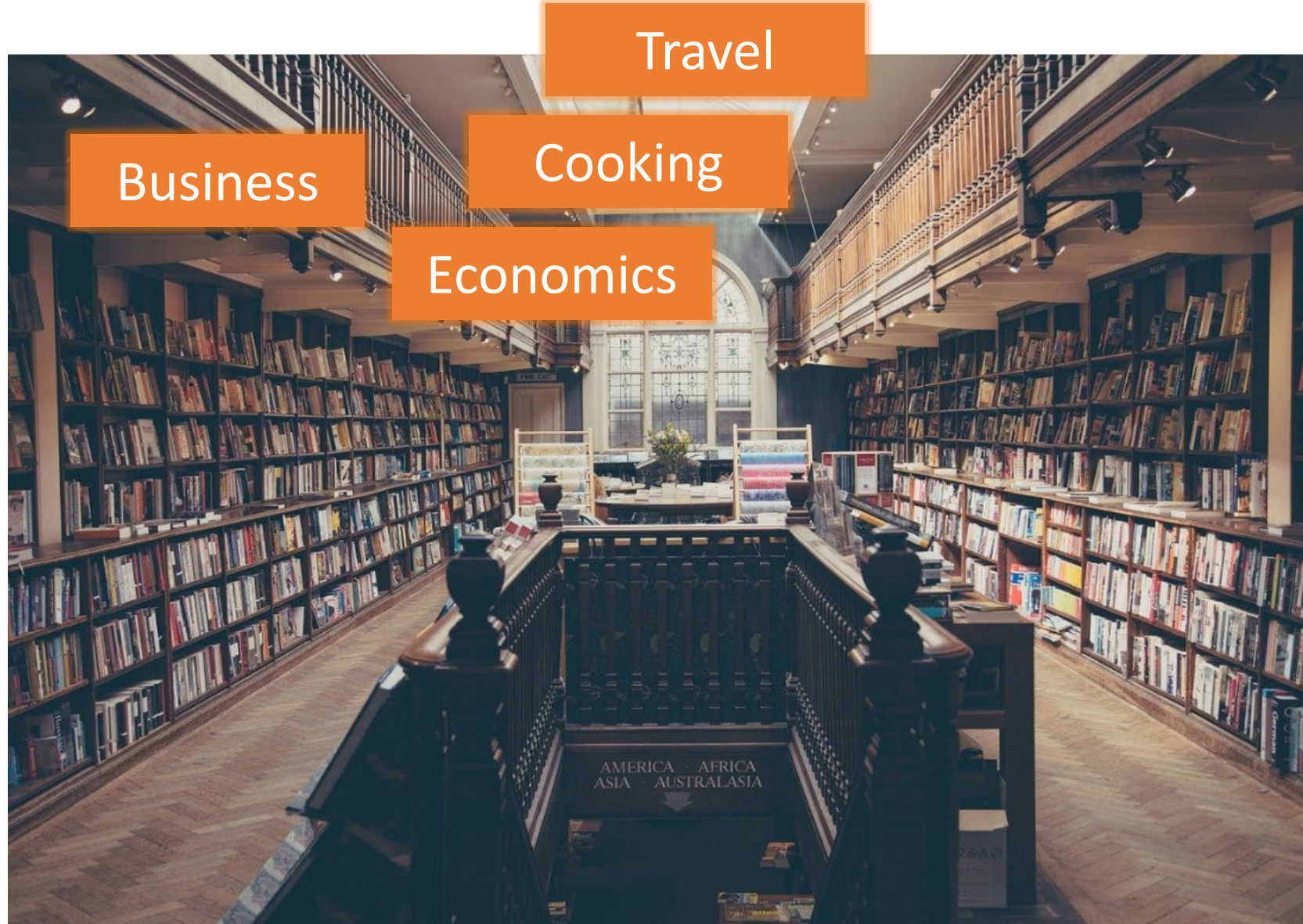
# What is Association Rule Mining?

- Association Rule Mining is a technique to discover association between items.
- It can reveal what products are being **purchased together**.





# Use Case: Bookstore





# Beer and Diaper?



&



?

\*Wilke, J. R. Retailing: Supercomputers manage holiday stock. The Wall Street Journal. 1992-12-23

# Usefulness of Association Rule Mining

We can do cross-selling effectively!

## Sources of Pictures

- p.3** Andrea Piacquadio from Pixel (Left)  
ThisIsEngineering from Pixel (Right)
- p.4** Javier-Rodriguez from Pixabay
- p.5** Clker-Free-Vector-Images from Pixabay
- p.6** Foundry Co from Pixabay
- p.7** Alexander Lesnitsky from Pixabay (Beer)  
PublicDomainPictures from Pixabay (Diaper)

## **2. Comparison with Similar Methods**

# Difference between Clustering and Association Rule Mining

- **Clustering:**

Find clusters within a **sample** based on **similarities**.

- **Association Rule Mining:**

Find association between **variables** based on **co-occurrence**.

# Example: Clustering

Variable

	Bread	Butter	Cheese	Milk
Customer 1		✓		✓
Customer 2	✓		✓	
Customer 3	✓		✓	✓
Customer 4	✓	✓		✓
Customer 5			✓	

Sample

# Example: Association Rule Mining

	Bread	Butter	Cheese	Milk
Transaction 1		✓		✓
Transaction 2	✓		✓	
Transaction 3	✓		✓	✓
Transaction 4	✓	✓		✓
Transaction 5			✓	



# Clustering? Association Rule Mining?

Association rule mining

	Bread	Butter	Cheese	Milk
1		✓		✓
2	✓		✓	
3	✓		✓	✓
4	✓	✓		✓
5			✓	

Clustering

# Collaborative Filtering? Association Rule Mining?

- Association rule mining does not aim to reveal each customer's preference.
- Association rule mining does not analyze historical data.

# **3. Basic Terms**

# Item and Itemset

Transaction ID	List of purchased items
T1	cheese, ham
T2	bread, milk
T3	bread, milk, ham
T4	bread, cheese, ham
T5	milk
T6	bread, milk

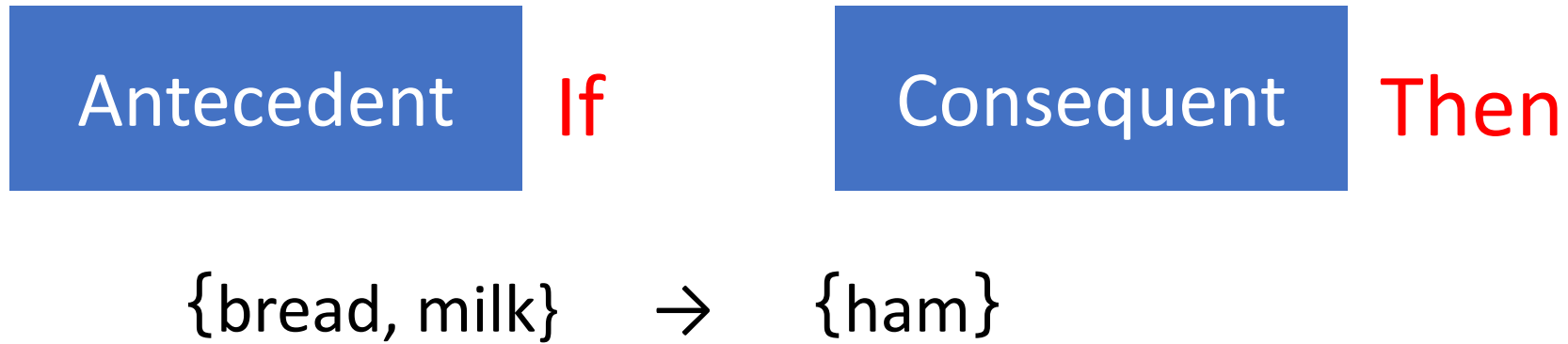
Transaction T1

Item:

- cheese
- ham
- bread
- milk

Itemset: {bread, milk}

# Antecedent and Consequent



Association rule mining reveals **co-occurrence**.

**NOT** causality!

# Basic Metrics of Association

- Support
- Confidence
- Lift

# **4. Basic Math: Set Theory**



# What is Set?

**Set:** A collection of **elements**

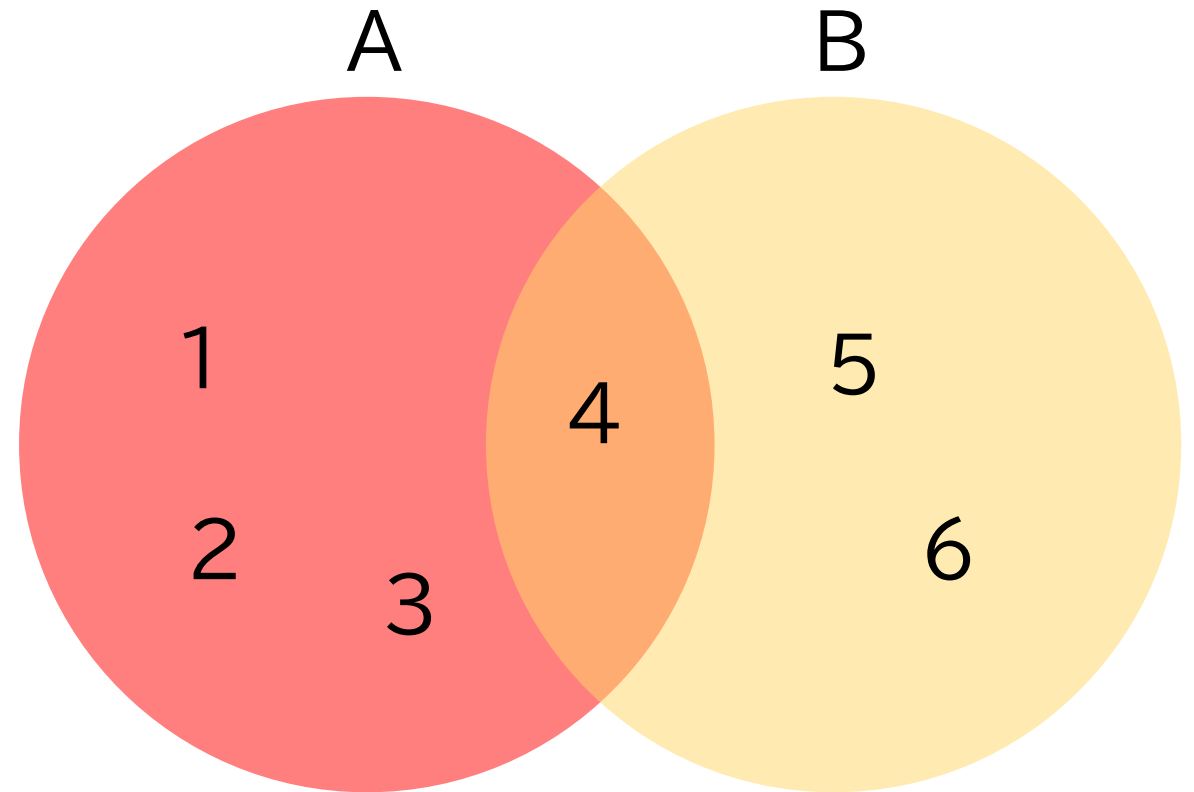
Elements

A {1, 2, 3, 4}

B {4, 5, 6}

$1 \in A$

↙ Belongs to



# Intersection and Union

Intersection:  $A \cap B$

$$A \cap B \{4\}$$

Union:  $A \cup B$

$$A \cup B \{1, 2, 3, 4, 5, 6\}$$

