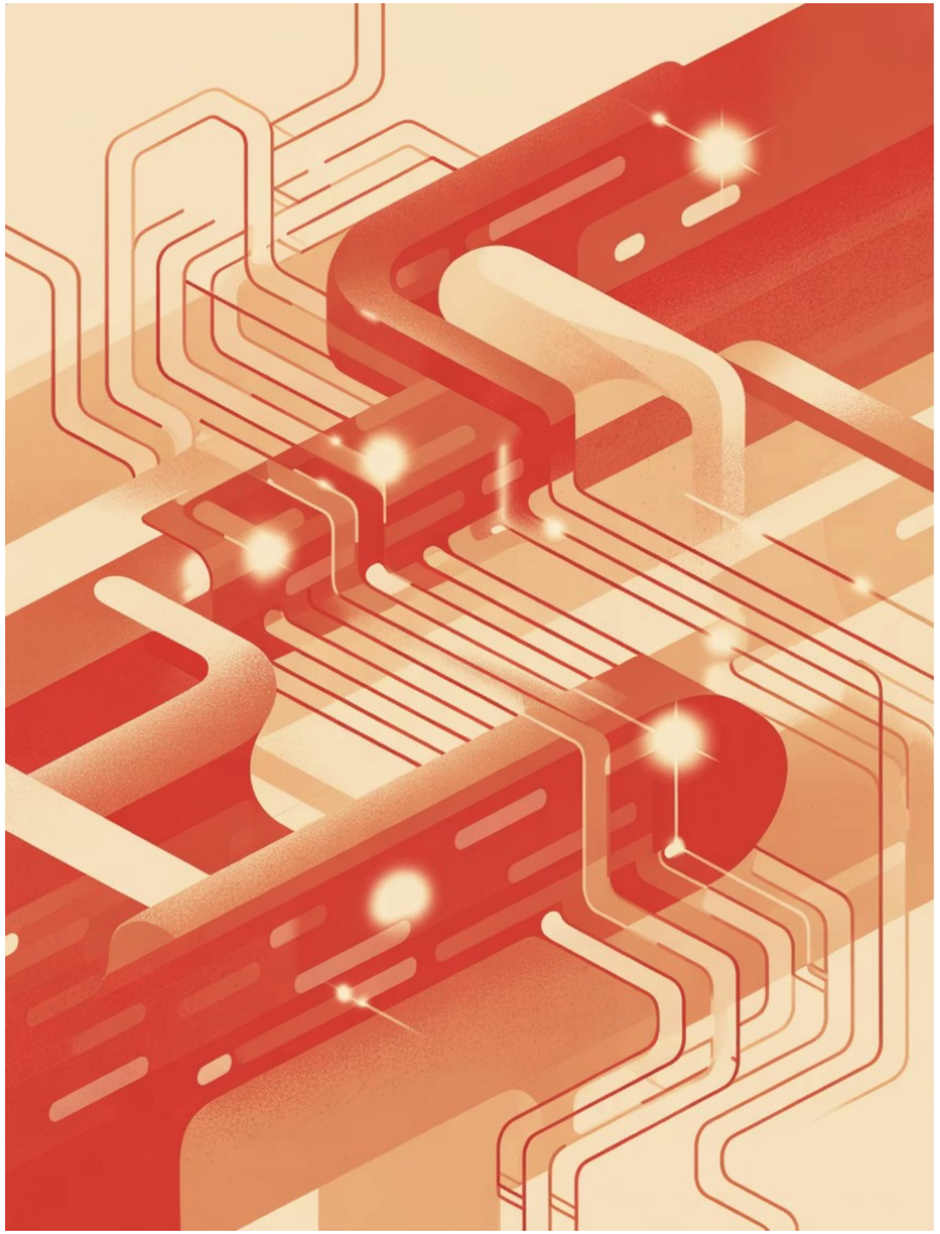


# Mining Frequent Patterns

**Presented by:** Arijit Bera

**College:** Dr.Sudhir Chandra Sur Institute of Technology & Sports Complex

**Roll No:** 25500123046

**Sec:** A

**Stream:** CSE

**Year:** 3rd

**Sem:** 6th

**Sub:** Data Warehousing & Data Mining(PECIT602B)

# What is Frequent Pattern Mining?

Frequent pattern mining is a core concept in data mining that focuses on identifying common itemsets or subsequences that occur frequently within a dataset. This powerful technique was first introduced by Agrawal et al. in 1993, specifically for discovering compelling associations in transaction data.

## Discovering Associations

It helps uncover hidden relationships between items in large transaction databases.

## Real-world Example

A classic example: customers often buy {Diapers, Beer} together, revealing interesting purchasing habits.

# Why is Frequent Pattern Mining Important?

Frequent pattern mining is crucial because it acts as a foundational pillar for various advanced analytical techniques, revealing hidden relationships that drive strategic decisions across diverse industries.

- Reveals invaluable hidden relationships in vast datasets, such as in comprehensive market basket analyses.
- Forms the bedrock for advanced applications like association rule mining, classification, clustering, and sophisticated recommendation systems.
- Boasts wide-ranging applications spanning retail analytics, web usage mining, bioinformatics research, and many other data-intensive fields.

# Key Concepts in Frequent Pattern Mining

Understanding these core concepts is essential for effective frequent pattern mining, forming the language of data associations.

## Itemset

A collection of one or more items, such as {Milk, Bread}, typically bought together in a transaction.

## Support

The frequency of an itemset within the dataset, expressed as the fraction of transactions containing it.

## Confidence

The conditional probability that item Y is purchased when item X is purchased, represented as X → Y.

## Frequent Itemset

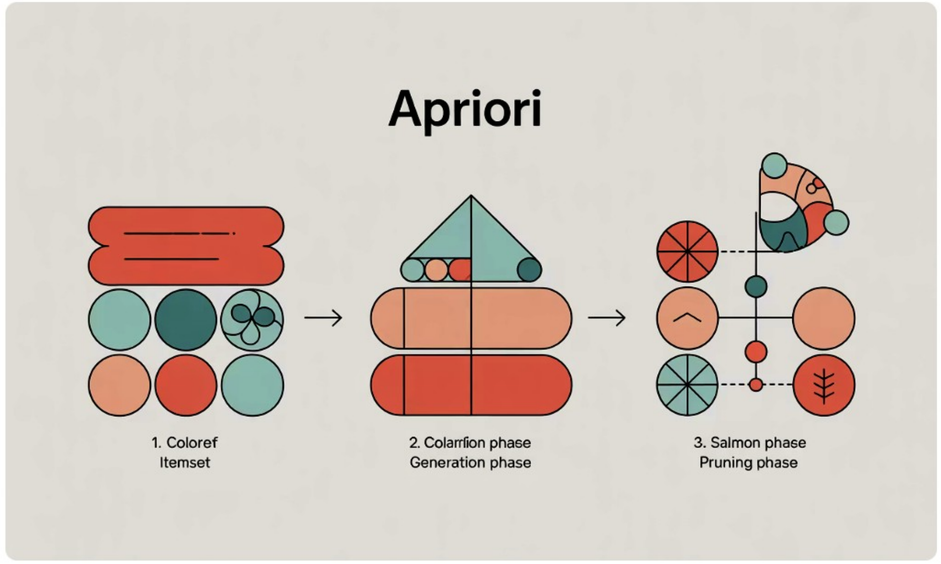
An itemset whose support value exceeds a predefined minimum support threshold, indicating its significance.

# Popular Frequent Pattern Mining Algorithms

Two prominent algorithms, Apriori and FP-Growth, offer distinct approaches to discovering frequent patterns, each with its own advantages.

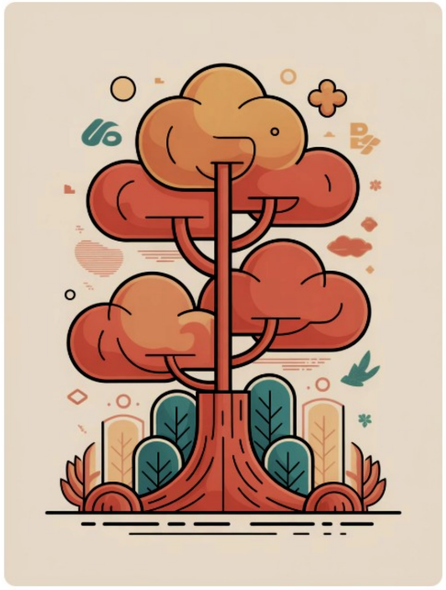
## Apriori Algorithm

- Generates candidate itemsets by joining frequent (k-1) itemsets.
- Employs the downward closure property to prune infrequent candidates efficiently.
- Often requires multiple scans of the database, which can become resource-intensive for very large datasets.




## FP-Growth Algorithm

- Constructs a compact Frequent Pattern Tree (FP-tree) from the database.
- Mines frequent patterns directly from the FP-tree without the need for candidate generation.
- Demonstrates superior efficiency for larger and denser datasets compared to Apriori.



# Summary: Unlocking Data Insights

Frequent pattern mining is an indispensable tool for extracting valuable knowledge from extensive datasets, offering profound insights into underlying relationships and behaviours.

**Identify Frequent Items**

**Choose Efficient Algorithms**

**Drive Informed Decisions**

### Uncover Patterns

Discover common item combinations and sequences.

### Algorithm Choice

Leverage Apriori or FP-Growth based on data characteristics.

### Drive Analytics

Essential for advanced data mining and informed decisions.