

Transactions

```

coffee, milk, sugar, ...
...

```

map: create patterns

```

(coffee) -> 1
(milk) -> 1
(coffee, milk) -> 1
...

```

map: create patterns

```

...
...
(coffee, milk, eggs) -> 1
...

```

map: create patterns

```

(coffee, milk) -> 1
...
(coffee, milk) -> 1

```

Pattern format:
 (key) -> value
 key: tuple of products
 value: 1 (# times seen in the transaction)

reduce: combining (sum) values of each key

```

(coffee) -> 20
...
(coffee, milk) -> 17
...

```

```

...
(milk) -> 50
...
(coffee, milk, eggs) -> 8

```

```

(sugar) -> 12
...
(coffee, sugar) -> 5

```

Frequent Pattern format:
 (key) -> value
 key: tuple of products
 value: # times seen in the entire dataset (combined)

map: subpatterns

```

(coffee) -> (None, 20)
...
(coffee, milk) -> (None, 17)
(milk) -> (coffee, 17)
(coffee) -> (milk, 17)
...

```

map: subpatterns

```

...
(milk) -> (None, 50)
...
(coffee, milk, eggs) -> (None, 8)
(milk, egg) -> (coffee, 8)
(coffee, egg) -> (milk, 8)
(coffee, milk) -> (egg, 8)

```

map: subpatterns

```

(sugar) -> (None, 12)
...
(coffee, sugar) -> (None, 8)
(sugar) -> (coffee, 8)
(coffee) -> (sugar, 8)

```

Subpattern format:
 key: tuple of products
 value: tuple(prod, # times seen)

reduce: grouping the subpatterns of each key

```

(coffee) -> [(None, 20), (milk, 17), (sugar, 8)]
...
(coffee, milk) -> [(egg, 8), (None, 17)]
...

```

```

(milk) -> [(coffee, 17), (None, 50)]
...
(sugar) -> [(None, 12), (coffee, 8)]
(coffee, milk, eggs) -> [(None, 8)]

```

Rule format:
 key: tuple of products
 value: list of tuples

map: association rules**map: association rules**

```

(coffee) -> [(milk, 0.85), (sugar, 0.4)]
...
(coffee, milk) -> [(egg, 0.47)]
...
(milk) -> [(coffee, 0.34)]
...
(sugar) -> [(coffee, 0.66)]
(coffee, milk, eggs) -> []

```

Association Rules:
 key: tuple of products
 value: list of tuples with confidence value