

# Market Basket Analysis

BY: NATHAN CHRISTOPHER MENON

# AGENDA

- Objective
- Shopping Data
- Insights
- Association Rule Mining
- Thank You

# OBJECTIVE

- Analyze the dataset and derive an action plan for Mr & Mrs Khanna to increase their sales revenue
- Use Association Rule Mining to derive insights

| Product             | ProductID | Season    | Price |
|---------------------|-----------|-----------|-------|
| Summer Cap          | P01       | Summer    | 100   |
| Sunglasses          | P02       | Summer    | 50    |
| Half Sleeve T-shirt | P03       | Summer    | 200   |
| Capri               | P04       | Summer    | 350   |
| Saree               | P05       | Evergreen | 400   |
| Earrings            | P06       | Evergreen | 30    |
| Kurta               | P07       | Evergreen | 150   |
| Ethnic Shoes        | P08       | Evergreen | 300   |
| Winter Cap          | P09       | Winter    | 150   |
| Sweatshirt          | P10       | Winter    | 250   |
| Long Sleeve T-shirt | P11       | Winter    | 300   |
| Jeans               | P12       | Winter    | 600   |
| Thermocoat          | P13       | Winter    | 270   |
| Sherwani            | P14       | Wedding   | 2000  |
| Lahenga             | P15       | Wedding   | 4000  |

# SHOPPING DATA

- This table contains the OrderID, CustomerID, date of purchase, ProductID, Product, Quantity purchased, Unit Price, and Revenue respectively
- The data was grouped by the OrderID and CustomerID in patches of color as seen which indicates items bought in pairs by the same customer and these pairs are frequently purchased as indicated by patches of the same color

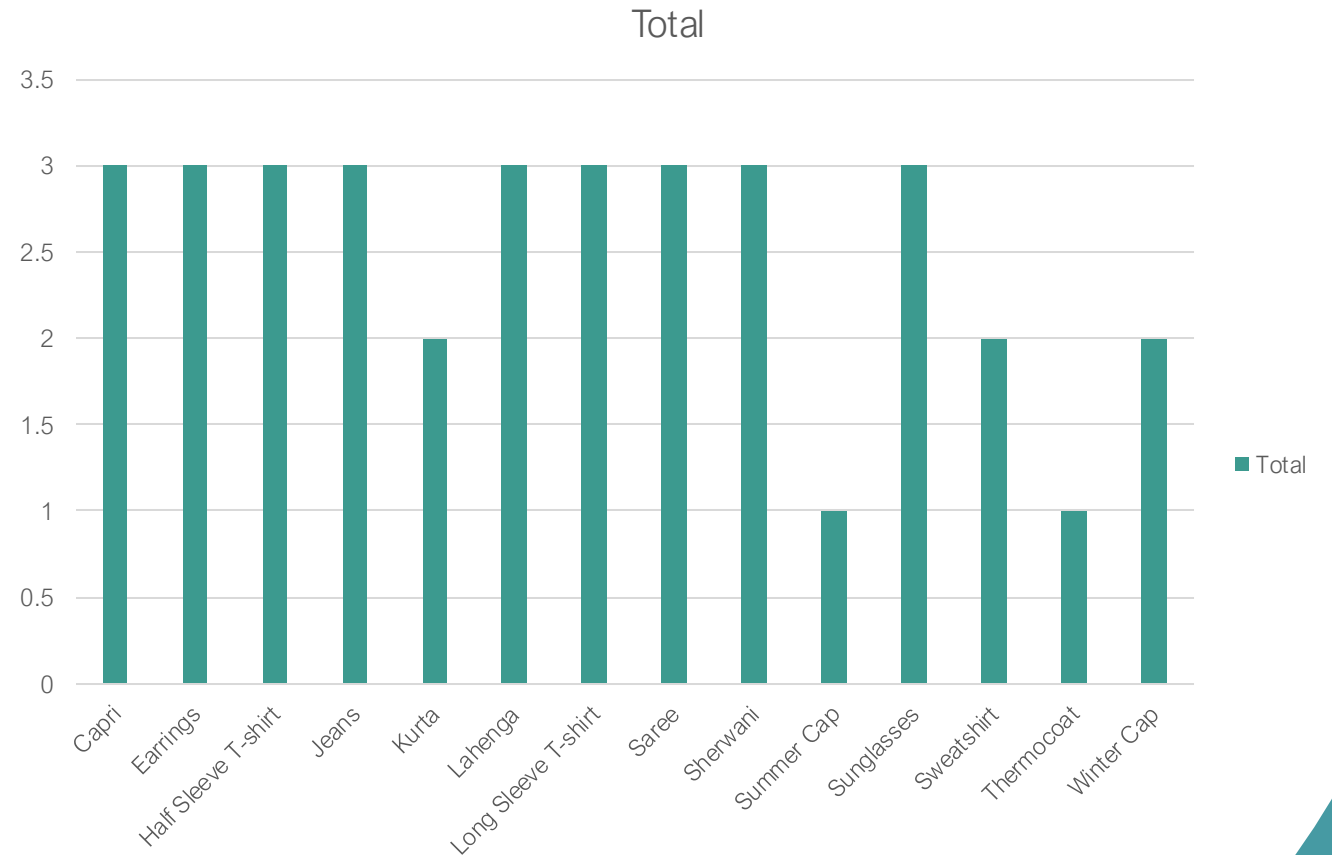
|      |     |            |     |                     |   |      |      |  |  |  |
|------|-----|------------|-----|---------------------|---|------|------|--|--|--|
| 1001 | 101 | 27/06/2023 | P02 | Sunglasses          | 1 | 50   | 50   |  |  |  |
| 1002 | 102 | 28/07/2023 | P07 | Kurta               | 1 | 150  | 150  |  |  |  |
| 1003 | 103 | 29/07/2023 | P03 | Half Sleeve T-shirt | 1 | 200  | 200  |  |  |  |
| 1003 | 103 | 29/07/2023 | P04 | Capri               | 2 | 350  | 700  |  |  |  |
| 1004 | 104 | 31/08/2023 | P05 | Saree               | 1 | 400  | 400  |  |  |  |
| 1004 | 104 | 31/08/2023 | P06 | Earrings            | 1 | 30   | 30   |  |  |  |
| 1005 | 105 | 01/09/2023 | P03 | Half Sleeve T-shirt | 1 | 200  | 200  |  |  |  |
| 1005 | 105 | 01/09/2023 | P04 | Capri               | 2 | 350  | 700  |  |  |  |
| 1006 | 105 | 02/09/2023 | P07 | Kurta               | 2 | 150  | 300  |  |  |  |
| 1007 | 104 | 07/09/2023 | P05 | Saree               | 2 | 400  | 800  |  |  |  |
| 1007 | 104 | 07/09/2023 | P06 | Earrings            | 1 | 30   | 30   |  |  |  |
| 1008 | 106 | 05/10/2023 | P15 | Lahenga             | 1 | 4000 | 4000 |  |  |  |
| 1009 | 107 | 16/10/2023 | P14 | Sherwani            | 1 | 2000 | 2000 |  |  |  |
| 1010 | 108 | 26/10/2023 | P15 | Lahenga             | 1 | 4000 | 4000 |  |  |  |
| 1011 | 103 | 27/10/2023 | P03 | Half Sleeve T-shirt | 1 | 200  | 200  |  |  |  |
| 1011 | 103 | 29/07/2023 | P04 | Capri               | 2 | 350  | 700  |  |  |  |
| 1012 | 109 | 27/10/2023 | P14 | Sherwani            | 1 | 2000 | 2000 |  |  |  |
| 1013 | 110 | 28/10/2023 | P15 | Lahenga             | 1 | 4000 | 4000 |  |  |  |
| 1014 | 111 | 29/10/2023 | P14 | Sherwani            | 1 | 2000 | 2000 |  |  |  |
| 1015 | 101 | 01/11/2023 | P10 | Sweatshirt          | 2 | 250  | 500  |  |  |  |
| 1016 | 103 | 02/11/2023 | P11 | Long Sleeve T-shirt | 1 | 300  | 300  |  |  |  |
| 1016 | 103 | 02/11/2023 | P12 | Jeans               | 1 | 600  | 600  |  |  |  |
| 1017 | 103 | 04/11/2023 | P13 | Thermocoat          | 1 | 270  | 270  |  |  |  |
| 1018 | 101 | 05/11/2023 | P10 | Sweatshirt          | 2 | 250  | 500  |  |  |  |
| 1019 | 105 | 06/11/2023 | P11 | Long Sleeve T-shirt | 1 | 300  | 300  |  |  |  |
| 1019 | 105 | 06/11/2023 | P12 | Jeans               | 1 | 600  | 600  |  |  |  |
| 1020 | 106 | 10/12/2023 | P05 | Saree               | 1 | 400  | 400  |  |  |  |
| 1020 | 106 | 10/12/2023 | P06 | Earrings            | 1 | 30   | 30   |  |  |  |
| 1021 | 107 | 01/01/2024 | P11 | Long Sleeve T-shirt | 1 | 300  | 300  |  |  |  |
| 1021 | 107 | 01/01/2024 | P12 | Jeans               | 1 | 600  | 600  |  |  |  |
| 1022 | 101 | 01/02/2024 | P09 | Winter Cap          | 1 | 150  | 150  |  |  |  |
| 1022 | 101 | 01/02/2024 | P02 | Sunglasses          | 1 | 50   | 50   |  |  |  |
| 1023 | 111 | 03/02/2024 | P09 | Winter Cap          | 1 | 150  | 150  |  |  |  |
| 1023 | 111 | 03/02/2024 | P02 | Sunglasses          | 1 | 50   | 50   |  |  |  |

# Insights



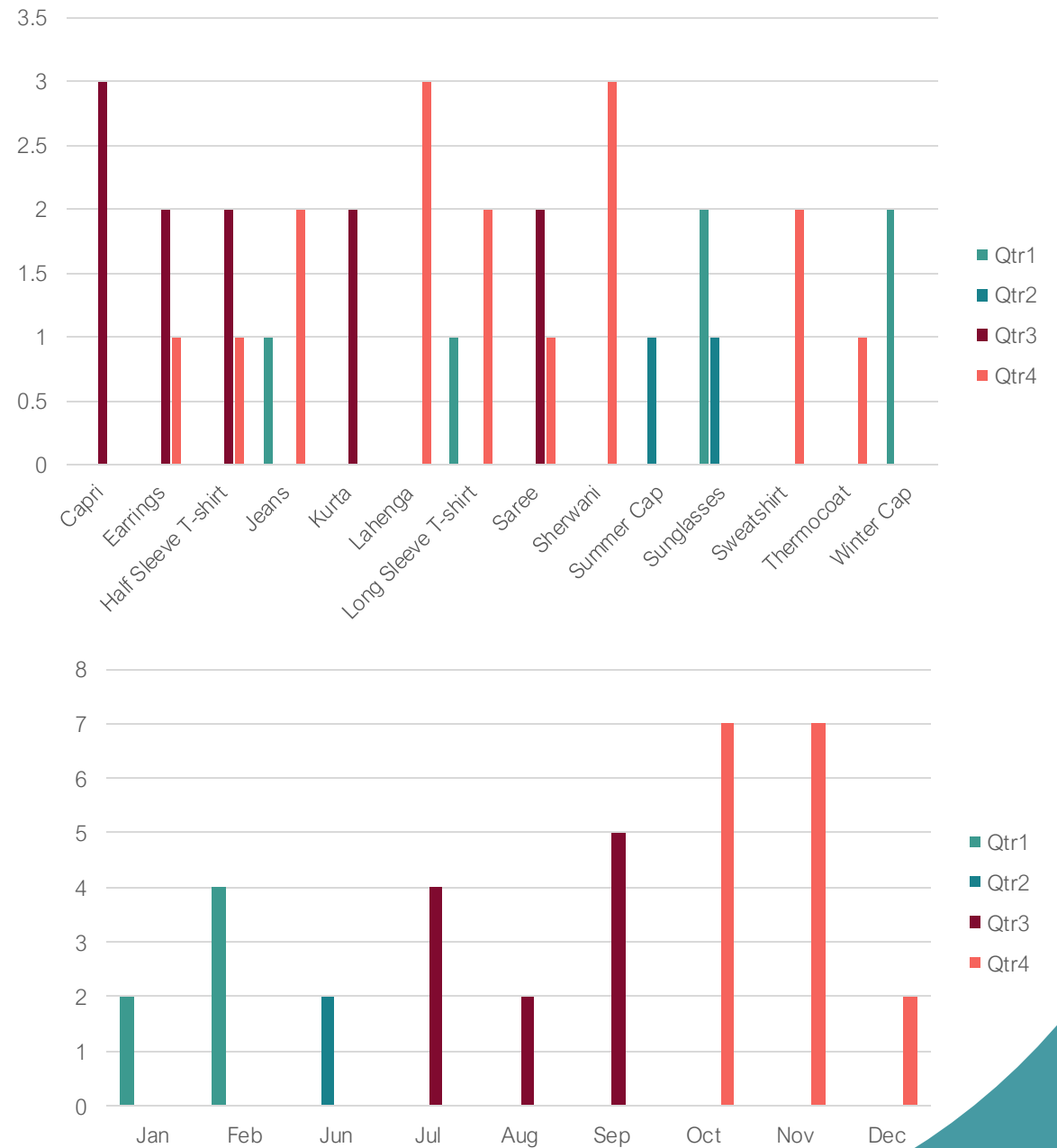
# COUNT OF EACH ITEM PURCHASED

- This chart was created using a pivot table on the given data
- On the X-axis we can see the name of each item
- On the Y-axis we can see the frequency of purchase
- From this bar chart, we can say that all items are frequently purchased except summer cap and thermo-coat, so a suggestion would be to spend more time and money on marketing them



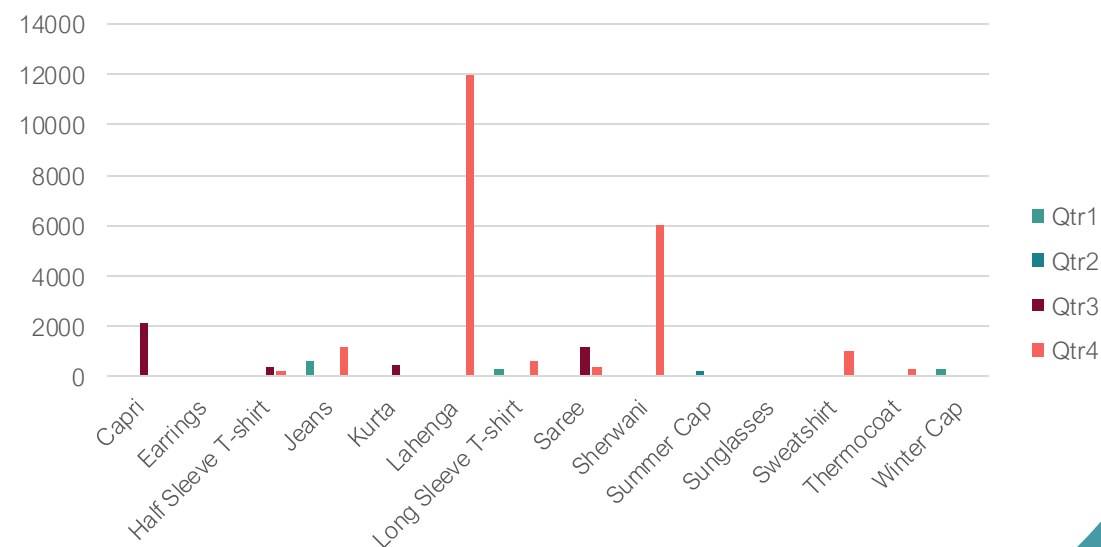
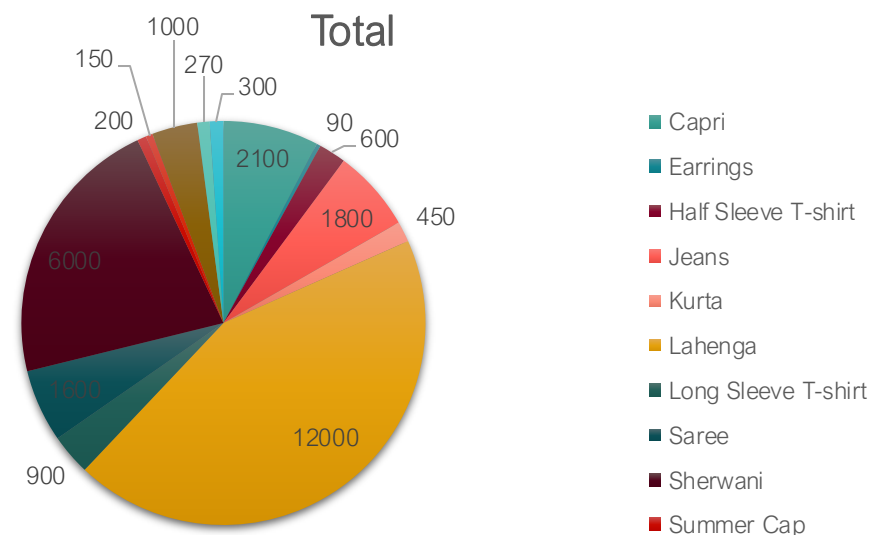
# QUARTER-WISE PURCHASES

- These charts were created using a pivot table on the given data
- From both the charts, we can see that most of the purchases are made during the 4<sup>th</sup> quarter, that is, during the months of October, November and December. This must be because the wedding season overlaps with winter. So, winter apparel and festive wear are high-in demand. Rest of the year, purchases are quite consistent except in the months of March, April, and May which indicates that demand is next to none during the summer season. A suggestion here would be to find innovative ways to market clothes during the summer season



# BUSINESS REVENUE

- These charts were created using a pivot table on the given data
- From both charts, we can see that the item *Lahenga* contributes to nearly half of the total revenue followed by *sherwani* and *earrings*. This again proves the fact that the business flourishes during the festive season and purchases are fairly consistent rest of the year





# Association Rule Mining



# What is an association rule?

An association rule represents the pattern/co-occurrence of two item sets by using an if-then condition. For example, a rule (Apple)  $\rightarrow$  (Banana) means “IF Apple is in a transaction, THEN Banana is also in that transaction”. This is a concept that works for problems like the Market Basket Analysis in which we predict which items are most probably going to be sold together. Let’s dive into the concept, shall we?

# Antecedent and Consequent

These are important concepts related to the if-then structure of association rules. Here's a brief explanation to add to the existing content:

- Antecedent (If part): The item or itemset that appears on the left side of the association rule. It's the "if" part of the rule.
- Consequent (Then part): The item or itemset that appears on the right side of the association rule. It's the "then" part of the rule.

For example, in the rule  $X \rightarrow Y$ :

- X is the antecedent
- Y is the consequent

This rule suggests that if X occurs in a transaction, then Y is likely to occur as well. In the context of market basket analysis, if a customer buys item X, they are likely to buy item Y too.

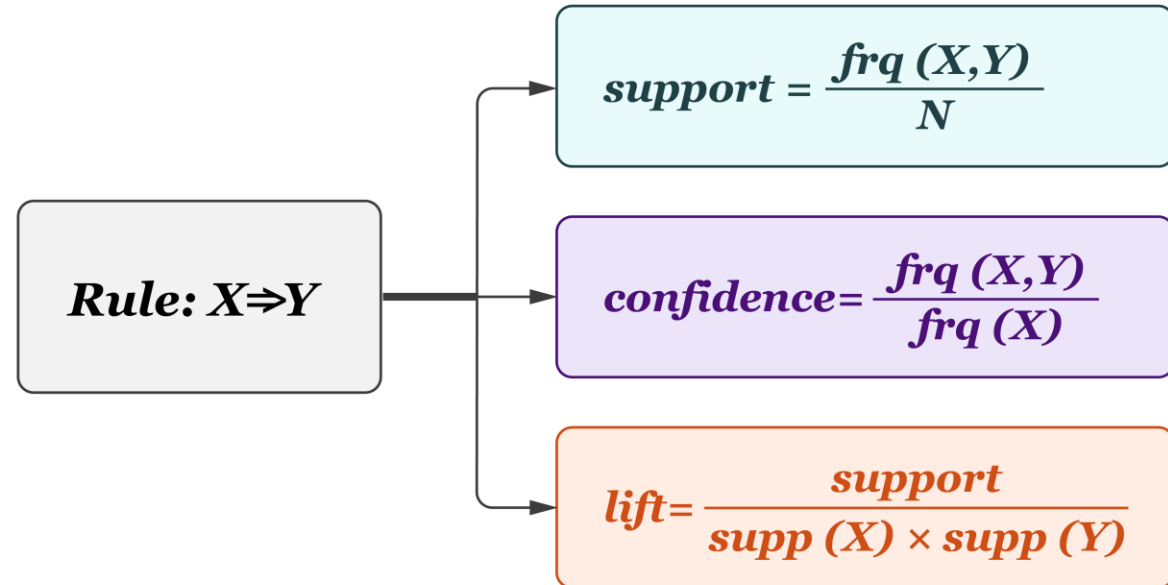
# Support, Confidence and Lift

Association Rule Mining consists of three metrics. These metrics help identify relationships between items in a dataset, such as products frequently purchased together:

- **Support:** Support measures how frequently an item (a group of items) appears in the dataset. It indicates how common an item is
- **Confidence:** Confidence measures the likelihood that item Y is purchased when item X is purchased. It's the ratio of the number of transactions that include both X and Y to the number of transactions that include X alone
- **Lift:** Lift measures how much more likely two items are to be bought together than expected if they were independent. It compares the observed co-occurrence of X and Y to what would be expected if X and Y were independent

# HIGHLIGHTS

- *Support* indicates how common an item is in the dataset.
- *Confidence* shows the likelihood of one item being purchased when another is purchased.
- *Lift* reveals the strength of the association between two items, adjusting for how common they are individually.



# USING ASSOCIATION RULE ON OUR DATA

| Product             | Support single | Items paired               | Frequency | Support   | Confidence | Lift |
|---------------------|----------------|----------------------------|-----------|-----------|------------|------|
| Summer Cap          | 0.027777778    | Summer Cap, Sunglasses     | 1         | 0.0833333 | 3          | 36   |
| Sunglasses          | 0.083333333    | Half Sleeve T-Shirt, Capri | 3         | 0.25      | 3          | 36   |
| Kurta               | 0.055555556    | Saree, Earrings            | 3         | 0.25      | 3          | 36   |
| Half Sleeve T-shirt | 0.083333333    | Long Sleeve T-Shirt, Jeans | 3         | 0.25      | 3          | 36   |
| Capri               | 0.083333333    | Winter Cap, Sunglasses     | 2         | 0.166667  | 2          | 24   |
| Saree               | 0.083333333    |                            |           |           |            |      |
| Earrings            | 0.083333333    |                            |           |           |            |      |
| Half Sleeve T-shirt | 0.083333333    |                            |           |           |            |      |
| Capri               | 0.083333333    |                            |           |           |            |      |
| Kurta               | 0.055555556    |                            |           |           |            |      |
| Saree               | 0.083333333    |                            |           |           |            |      |
| Earrings            | 0.083333333    |                            |           |           |            |      |
| Lahenga             | 0.083333333    |                            |           |           |            |      |
| Sherwani            | 0.083333333    |                            |           |           |            |      |
| Lahenga             | 0.083333333    |                            |           |           |            |      |
| Half Sleeve T-shirt | 0.083333333    |                            |           |           |            |      |
| Capri               | 0.083333333    |                            |           |           |            |      |
| Sherwani            | 0.083333333    |                            |           |           |            |      |
| Lahenga             | 0.083333333    |                            |           |           |            |      |
| Sherwani            | 0.083333333    |                            |           |           |            |      |
| Sweatshirt          | 0.055555556    |                            |           |           |            |      |
| Long Sleeve T-shirt | 0.083333333    |                            |           |           |            |      |
| Jeans               | 0.083333333    |                            |           |           |            |      |
| Thermocoat          | 0.027777778    |                            |           |           |            |      |
| Sweatshirt          | 0.055555556    |                            |           |           |            |      |
| Long Sleeve T-shirt | 0.083333333    |                            |           |           |            |      |
| Jeans               | 0.083333333    |                            |           |           |            |      |
| Saree               | 0.083333333    |                            |           |           |            |      |
| Earrings            | 0.083333333    |                            |           |           |            |      |
| Long Sleeve T-shirt | 0.083333333    |                            |           |           |            |      |
| Jeans               | 0.083333333    |                            |           |           |            |      |
| Winter Cap          | 0.083333333    |                            |           |           |            |      |
| Sunglasses          | 0.083333333    |                            |           |           |            |      |
| Winter Cap          | 0.083333333    |                            |           |           |            |      |
| Sunglasses          | 0.083333333    |                            |           |           |            |      |

# OBSERVATIONS AND SUGGESTIONS

1. Lift=1 : X, and Y are independent(no association)
2. Lift>1 : X and Y are positively associated(they occur together frequently)
3. Lift<1: X and Y are negatively associated(they occur together less often than expected)

As seen from the previous table, the lift values are way above 1 for the pair of items, this indicates a strong relationship between the items and a high likelihood of them being purchased together

Hence, it's suggested that these items be sold together, to positively generate revenue and maintain a successful business

# Thank you