Client Requests

1. The operations team wants to ensure the integrity of the events data by removing duplicates. Check for and remove duplicate rows in the events dataframe based on store_id, campaign_id, and product_code. How many duplicate rows were removed?
2. How many cities have more than 5 stores?
3. The sales team has identified missing values in the quantity_sold(before_promo) data. Estimate these values using the median quantity sold before the promotion. How many missing values were filled, and what is the median used for imputation?
4. Identify the product category with the lowest base price before the promotion.
5. What is the total quantity sold after the promotion for the BOGOF promo type during the Diwali campaign?
6. Which store recorded the highest quantity sold after the promotion during the Diwali campaign?
7. Understand which campaigns had the most successful outcomes. Compare the total quantities sold before and after the promotions for the Sankranti and Diwali campaigns. Which campaign saw a greater increase in sales?
8. Which product recorded the highest Incremental Revenue Percentage (IR%) during the Sankranti campaign? What is the IR% for this product?
9. Which store in Visakhapatnam recorded the lowest Incremental Sold Units Percentage (ISU%) during the Diwali campaign? What is the ISU% for that store?
10. Which promo type had both a negative Incremental Revenue Percentage (IR%) and Incremental Sold Units Percentage (ISU%) during the Sankranti campaign?

Key Metrics:

- **IR% (Incremental Revenue):** IR% measures the percentage change in revenue after a promotion compared to the revenue before the promotion. It helps assess how effective a promotion was in driving revenue growth.

- **ISU% (Incremental Sold Units):** ISU% calculates the percentage change in the number of units sold after a promotion compared to the units sold before the promotion. It indicates the effectiveness of a promotion in boosting sales volume.
