

train.csv: Train data containing the coupons offered to the given customers under the 18 campaigns

| Variable | Definition |
|-------------------|---|
| id | Unique id for coupon customer impression |
| campaign_id | Unique id for a discount campaign |
| coupon_id | Unique id for a discount coupon |
| customer_id | Unique id for a customer |
| redemption_status | (target) (0 - Coupon not redeemed, 1 - Coupon redeemed) |

campaign_data.csv: Campaign information for each of the 28 campaigns

| Variable | Definition |
|---------------|-----------------------------------|
| campaign_id | Unique id for a discount campaign |
| campaign_type | Anonymised Campaign Type (X/Y) |
| start_date | Campaign Start Date |
| end_date | Campaign End Date |

coupon_item_mapping.csv: Mapping of coupon and items valid for discount under that coupon

| Variable | Definition |
|-----------|--|
| coupon_id | Unique id for a discount coupon (no order) |
| item_id | Unique id for items for which given coupon is valid (no order) |

customer_demographics.csv: Customer demographic information for some customers

| Variable | Definition |
|----------------|---|
| customer_id | Unique id for a customer |
| age_range | Age range of customer family in years |
| marital_status | Married/Single |
| rented | 0 - not rented accommodation, 1 - rented accommodation |
| family_size | Number of family members |
| no_of_children | Number of children in the family |
| income_bracket | Label Encoded Income Bracket (Higher income corresponds to higher number) |

customer_transaction_data.csv: Transaction data for all customers for duration of campaigns in the train data

| Variable | Definition |
|-----------------|--|
| date | Date of Transaction |
| customer_id | Unique id for a customer |
| item_id | Unique id for item |
| quantity | quantity of item bought |
| selling_price | Sales value of the transaction |
| other_discount | Discount from other sources such as manufacturer coupon/loyalty card |
| coupon_discount | Discount availed from retailer coupon |

item_data.csv: Item information for each item sold by the retailer

| Variable | Definition |
|------------|--------------------------------|
| item_id | Unique id for item |
| brand | Unique id for item brand |
| brand_type | Brand Type (local/Established) |
| category | Item Category |

test.csv: Contains the coupon customer combination for which redemption status is to be predicted

| Variable | Definition |
|-------------|--|
| id | Unique id for coupon customer impression |
| campaign_id | Unique id for a discount campaign |
| coupon_id | Unique id for a discount coupon |
| customer_id | Unique id for a customer |

To summarise the entire process:

- Customers receive coupons under various campaigns and may choose to redeem it.
- They can redeem the given coupon for any valid product for that coupon as per coupon item mapping within the duration between campaign start date and end date
- Next, the customer will redeem the coupon for an item at the retailer store and that will reflect in the transaction table in the column coupon_discount.