

Assignment Brief: Hackathon on Customer Segmentation

Data Dictionary

Data Provided:

Raw_data: This Excel file dataset included the following sheets of data:

- **Transactions_data:** This dataset included the transactions data of the customers across all the different states in Australia.
- **NewCustomerList:** This dataset included the new customers who visited the automobile bike company recently.
- **CustomerDemographic:** This dataset included entire details of the Customer Demographics.
- **CustomerAddress:** This dataset included the address of the Customers.

Transactions_data

Column Name	Description
transaction_id	Unique identifier for each transaction
product_id	Identifier for the product involved in the transaction
customer_id	Identifier for the customer involved in the transaction
transaction_date	Date when the transaction occurred
online_order	Indicator of whether the order was placed online (TRUE for yes, FALSE for no)
order_status	Current status of the order (e.g., 'Approved', 'Pending', 'Cancelled')
brand	Brand of the product
product_line	Line of the product (e.g., 'Mountain', 'Road', 'Standard', 'Touring')
product_class	Classification of the product based on quality (e.g., 'Low', 'Medium', 'High')
product_size	Size of the product (e.g., 'Small', 'Medium', 'Large')
list_price	Listed price of the product at the time of the transaction
standard_cost	Standard cost of the product to the company
product_first_sold_date	Date when the product was first sold

NewCustomerList

Column Name	Description
first_name	The customer's first name
last_name	The customer's last name
gender	The customer's gender
past_3_years_bike_related_purchases	Number of bike-related purchases in the past three years
DOB	Date of birth of the customer
job_title	Job title of the customer
job_industry_category	Industry category of the customer's job
wealth_segment	Wealth segment classification of the customer

deceased_indicator	Indicator if the customer is deceased ('Y' for yes, 'N' for no)
owns_car	Indicator if the customer owns a car ('Y' for yes, 'N' for no)
tenure	Length of time the customer has been with the company
address	Customer's residential address
postcode	Postal code of the customer's address
state	State of the customer's residence
country	Country of the customer's residence
property_valuation	Valuation of the customer's property
Rank	Rank of the customer based on some criteria (e.g., 'Wealth segment', 'Property valuation')
Value	Value score assigned to the customer (used in ranking)

CustomerDemographic

Column Name	Description
customer_id	Unique identifier for the customer
first_name	The customer's first name
last_name	The customer's last name
gender	The customer's gender
past_3_years_bike_related_purchases	Number of bike-related purchases by the customer in the past three years
DOB	Date of birth of the customer
job_title	Job title of the customer
job_industry_category	Industry category of the customer's job
wealth_segment	Classification of the customer's wealth segment
deceased_indicator	Indicator if the customer is deceased ('Y' for yes, 'N' for no)
default	Information on default payment method or other default settings ('unstructured data')
owns_car	Indicator if the customer owns a car ('Y' for yes, 'N' for no)
tenure	Length of time the customer has been associated with the company

CustomerAddress

Column Name	Description
customer_id	Unique identifier for the customer
address	Residential address of the customer
postcode	Postal code associated with the customer's address
state	State where the customer resides
country	Country of the customer's residence
property_valuation	Valuation of the customer's property