**KPMG Insights: Transactions, Demographics, and CLV**

**Task Explanation**

**Task 1: Prepare the datasets for analysis by cleaning and correcting any inconsistencies.**

**Solution:**

1. ***For customer address data:***

* Created a blank workbook and used Power Query to import our dataset.
* Removed duplicate, blank rows, and those too with errors.
* Checked the data type for each column which was correct.
* State column primarily had 3 states only – New South Wales, Victoria and Queensland but some cells had their state code instead, like NSW for New South Wales, VIC for Victoria, and QLD for Queensland, so replaced these abbreviations with proper state name.
* For the address column, used trim and clean in format option for non-printable values and capitalized first letter of each word for consistency.

1. ***For customer demographic data:***

* Imported customer demographic sheet and removed duplicate, blank and error rows. Checked data type of every column.
* Selected whole table and replaced null values with N/A to maintain consistency.
* Trimmed and cleaned Default column for non-printable values but the column had invalid and unnecessary data, removing the column from our table will be a better option.
* Gender column has abnormalities – F, Femal, Female described Females while M and Male were for Males, and U for Unknown. We used transform > replace values to have only 3 entries – Male, Female, Unknown (make sure to select Match entire cell content to change only required values)
* Spelling of Agriculture corrected from Argiculture in Job\_Industry\_Category column.

1. ***For Transactions data:***

* Imported Transactions sheet and removed duplicate, blank and error rows. Checked data type of every column. Product\_First\_Sold\_Date column had wrong data type, changed to Date. Changed Data type of List\_Price and Standard\_Cost to Currency.
* Changed null values in Brand, Product\_Line, Product\_Class, and Product\_Size to N/A. Also capitalized first letter of each value for consistency.

1. ***For New customer list:***

* Imported new\_customer\_list sheet and removed duplicate, blank and error rows. Checked data type of every column.
* Capitalized first letter of every word in First name and Last name columns (for consistency, in case)
* Gender column had Male, Female, and U as entries. Changed U to Unknown for consistency.
* Replaced Null values in Job Title and Job Industry Category with N/A for easy understanding. Corrected spelling of Agriculture from Argiculture.
* Trimmed and Cleaned Address column, capitalized first letter of each word for consistency.
* State column has abbreviations – NSW, VIC, and QLD. For consistency in database, replaced abbreviations with their state names – New South Wales, Victoria, and Queensland respectively.
* Columns after Property Valuation has irrelevant and unnecessary data except Rank and Value, not to be included in our filtered dataset.

**Task 2: Segment customers based on demographic and transaction data to identify key customer groups.**

**Solution:**

1. ***Segmentation by Wealth Segment:***

* Created a Pivot Table using Customer Demographics data.
* Added Wealth segment column in rows section and Customer id and Tenure in Values section.
* Set Customer id value settings to Count and Tenure to Average with 2 decimal values.

1. ***Segmentation by Gender:***

* Created a Pivot Table using Customer Demographics data.
* Added Gender column to rows, Customer Id to Values (set value field as Count), Past 3 years bike related purchases to Values (set value field as average).
* Formatted table for better readability.

1. ***Segmentation by Job Category:***

* Created a Pivot Table using Customer Demographics data.
* Added Job Industry Category column to rows, also added Job title column to rows as subcategory for these industries for better breakdown and analysis.
* Added Customer id column to Value (set value field as Count), and Wealth segment to Value (set value as %age of Grand Total).
* This shows Wealth distribution within Job Industry and the Job titles.

After this I used Charts to prepare a dashboard for this analysis for better readability and understanding.

**Task 3: Analyze transaction data to identify trends and patterns.**

**Solution: The Slicer in Dashboard can be used to check monthly orders per Brand, Sales and Average List Price per Product Line.**

1. ***Sales Trend Analysis:***

* Created Pivot table with Month column and Transaction\_ID to get total orders per month. To get Earnings per month, created a pivot table with Month column and List Price.
* Used Bar charts to represent data visually and analyzed seasonal trends and drops and added my conclusion in an adjacent text box.

1. ***Product Performance Analysis:***

* Created Pivot table with Brands column in Rows section and List Price in Values to get total sales per brand. Created Line Chart to depict this visually.
* Similarly created a Pivot Table with Product Line in Rows and List Price in Values twice (as sum and as average). Used Column chart to depict this.

1. ***Customer Purchase Behavior:***

* Created a Pivot Table with Customer ID in Rows and List Price in Values to get total transaction value of customers. Then Sorted Transaction value column in descending order and selected top 10 customers. Right Click and selected “Keep only selected” option. Represented this using Bar chart.
* To get average number of purchases per customer, Create Pivot Table with Customer ID in Rows and Transaction ID in Values (as count). Select a new cell and write formula =AVERAGE(B25:B518) to get 5.72 as answer.

**Task 4: Analyze the new customer dataset to provide insights into potential new customer behavior and value.**

**Solution:**

1. ***New Customer Demographics:***

* Created a Pivot Table with Job Category in Rows section, Wealth\_Segment in Columns and First Name in Values as count as we don’t have Customer ID for them. Represented this using a Stacked bar chart.
* To calculate average past 3 year bike-related purchases, we can use Average() function directly on the column. But I thought of representing it gender-wise too. So, I created a Pivot table with Gender in Rows and Past 3-year bike-related purchase in Values as Average. It gave me average Gender wise and overall, too. Represented it using bar chart.

1. ***New Customer Location Analysis:***

* To get state-wise distribution of new customers, I created a Pivot Table with State in Rows and First Name in Values as Count to get total customer per state. Then represented it using Filled Map chart.
* To analyze correlation between property\_valuation and wealth\_segment columns, I used Line chart instead of Scatter Plot as the data was big. Created a Pivot Table with Property\_valuation in Rows and Wealth\_segment in Columns with First Name in Values as Count to get distribution of customer as per these 2 fields. Then created a Line chart.

1. ***Potential Revenue from New Customers:***

* To get Potential Revenue from customers based on past 3 years bike-related purchases and value, We can Add a new column “Potential Revenue” in Query Editor with the formula =Past\_3year\_bike\_related\_purchases \* Value. We get Potential Revenue of each customer in NewCust sheet while total potential revenue stands at ₹**43980.91003**.

**Task 5: Calculate and analyze the customer lifetime value to identify the most valuable customers.**

**Solution:**

1. ***CLV Calculation:***

**CLV formula:** Average Purchase Value \* Purchase Frequency \* Customer Lifespan

* Firstly, create a Excel Table with Customer Id and Total Transaction made by each of them and Total Revenue generated.

**Average Purchase Value =** Total Revenue/ Total Transactions by the customer

**Purchase Frequency =** Number of transactions made in a period of time, here we have data for 2017 only, so Total Transaction is our Purchase Frequency.

**Customer Lifespan =** Tenure

* Add Average Purchase Value column with the formula, we have Purchase Frequency already and import Tenure as Customer Lifespan.
* Add a Column “CLV” and use the formula to get CLV for each customer. Customer 5034 is not in our dataset and hence CLV and other data can not be extracted for the same.

1. ***Segment CLV Analysis:***

* For data analysis, Import columns like Gender, Wealth\_Segment, Job Category, Owns\_Car etc. using Customer ID.
* Created a dashboard for analysing Average CLV as per various criterias.

**Task 6: Summarize findings and provide actionable recommendations for business strategies.**

**Solution:**

***Summary of Key Insights:***

**1. Customer Segmentation (Demographics):**

* ***Customer Location:***Over 50% of customers reside in New South Wales, whether old or new.
* ***Average Age:*** 48.2 years with a wide range from 22 to 92 years. Several Users had their DOB missing.
* ***Bike Purchases:***On average, customers made approximately 49 bike-related purchases over the past three years.
* ***Tenure:***Customers have an average tenure of 10.7 years, indicating a relatively long-standing relationship with the company.

**2. Transaction** **Analysis**:

* ***Average Transaction Value:*** The average list price is approximately ₹1107.83, with a standard cost of ₹556.05.
* ***Online Orders:*** About 50% of the transactions are made online.
* ***Product Range:*** A wide range of products is sold, with product IDs ranging from 0 to 100. There are 6 Brands with 4 Product Lines, 3 Classes and 3 Sizes.

**3. New Customer Analysis:**

* ***Customer Location:*** Similar to the Customers dataset, over 50% of New Customers are clustered in New South Wales State.
* ***Tenure:*** New customers have an average tenure of around 11.4 years, suggesting some might have been recently reactivated.

**4. CLV (Customer Lifetime Value) Analysis:**

* ***Average CLV:*** The average Customer Lifetime Value is approximately $66,079.
* ***Total Revenue:*** The average revenue generated per customer is around $6341, with some high-value customers contributing significantly more.
* ***Average Purchase Value:*** Customers spend about $1106.86 per transaction on average.

***Recommendations:***

**1. Marketing Strategies:**

* ***Target High-Value Customers:*** Focus on marketing strategies that specifically target high-value customer segments identified through the CLV analysis. These could include loyalty programs, personalized offers, and exclusive deals to retain these customers.
* ***Reactivation Campaigns:*** Given the long tenure of some new customers, consider running reactivation campaigns targeting past customers who have not made a purchase in recent years.

**2. Business Expansion:**

* ***Geographic Targeting:*** Analyze the locations with higher customer density (based on postcode data) to identify potential areas for business expansion. Opening new stores or focusing marketing efforts in these regions could lead to higher customer acquisition.
* ***Online vs. Offline Strategy:*** Since 50% of transactions are made online, consider enhancing the online shopping experience, perhaps by offering localized promotions or improving delivery options in high-density areas.

**3. Product Offering Improvements:**

* ***Popular Products:*** Focus on the most popular product categories (based on product IDs with the highest sales) to ensure they are well-stocked and promoted.
* ***Price Optimization:*** Given the wide range in product pricing and costs, consider a price optimization strategy to maximize profit margins on high-demand products.

***Additional Recommendations***

**1. Customer Experience Enhancement:**

* ***Personalized Customer Service:*** Implement a personalized customer service approach, especially for high-value customers. This could include dedicated customer support, personalized maintenance reminders, or special service packages that enhance the ownership experience.
* ***Customer Feedback Loop:*** Create a feedback loop where customers can easily provide insights on their purchasing experience and product satisfaction.

**2. Marketing and Promotions:**

* ***Seasonal Promotions:*** Analyze transaction data to identify any seasonal trends in bicycle purchases. Use this information to design targeted marketing campaigns and promotions during peak seasons, such as the summer or holiday periods.
* ***Bundling and Upselling:*** Consider bundling bicycles with complementary products like helmets, accessories, or maintenance kits.

**3. Product Development and Diversification:**

* ***Focus on Popular Models:*** Consider expanding the range or offering enhanced versions of popular bicycles.
* ***Sustainable and Eco-Friendly Products:*** Given the growing trend towards sustainability, consider including eco-friendly bicycles, such as those made from recycled materials or electric bikes. Promote these products to environmentally-conscious customers.