ElleVet Customer Analysis

Prepared by: Anna Driscoll, Hunter Hornay, Anh (Nguyen) Mazur, Trang Tran

Northeastern University – ALY6980 – Fall 2024

Executive Summary

- The objective is to provide insights into:

1. What are the common characteristics of our best customers – both subscribers and 1X buyers
    1. What can we learn from their behavior and interaction with our site, product, and offers that can help us extend LTV and create a template for customer acquisition strategy
    2. If we plot these customers within the quadrants of a 4-box in which frequency of purchase runs on the x axis and LTV on the y axis, what do we learn and what can we action?
2. What common characteristics of veterinarians and clinics correlate with consistent purchase of ElleVet at scale, and what does that information offer us in predicting lifecycle and LTV behavior
    1. If we plot vets within the quadrants of a 4-box in which frequency of purchase runs on the x axis and LTV on the y axis, what do we learn and what can we action?
3. Predict the impact of of more aggressive subscription offers on revenue, acquisition, and LTV.
4. How can we use data to understand, deconstruct, and improve the connection between vets and consumers: we know that 70% of consumers identify a veterinarian's recommendation as the reason for trying ElleVet. But currently we do not have levers in use to expedite the time it takes vets and consumers to find one another. There are five cohorts that exist in a geography:
    1. There are ß vets who carry ElleVet and recommend it
    2. There are Z vets who do not carry ElleVet BUT recommend it.
    3. There are ∂ existing customers in the geography who currently use ElleVet
    4. There are µ vets who have never carried ElleVet/never heard of it
    5. There are X pet owners who have never tried ElleVet/never heard of it

How do we use geotagging, data science, analytics, to identify members of these cohorts, and facilitate interaction so that:

ß+Z+∂ => X+µ, where =>is a factor applied to bring these cohorts into contact with one another and stimulate adoption and acceptance.

What are those force factors that can be applied? How do we apply them?

5. Determine the relative LTV worth of customers based on how they first purchased:
    1. Are customers who buy 1X and then return to start a subscription more valuable than direct subscribers
    2. Are subscribers more valuable than serial non-subscribers
    3. Force rank all customer types on a weighted score of LTV and average lifespan

Data

Data Source	• Provided by ElleVet
Data Summary	• **EVS_sept24Data**: 6,389 customer transactions (sample of orders spanning 2017–2024) containing detailed order information, along with appended historical customer metrics such as lifetime orders and last order date. • **Roux Active Vet File:** A workbook of veterinary-related data, including transaction records from 2022–2024, a "Top 100 Vets" list, and additional details on veterinary clinic activity.
Data Limitations	• Lack of complete historical data (limits reliable grouping of segments based on total $ spent and order/subscription cadence) • Outlier and potential data inaccuracies • Limited geographical data for geotagging / regional segmentation • Historical subscription offers to leverage a/b testing • Missing data and N/A

Methodology

Tools Used

Python: Customer makeup insights, RMF analysis, DTC vs vet cross analysis, cluster analysis, geospatial analysis
R: Quadrant analysis, subscription offer aggressiveness analysis
Power BI: For data preprocessing and merging

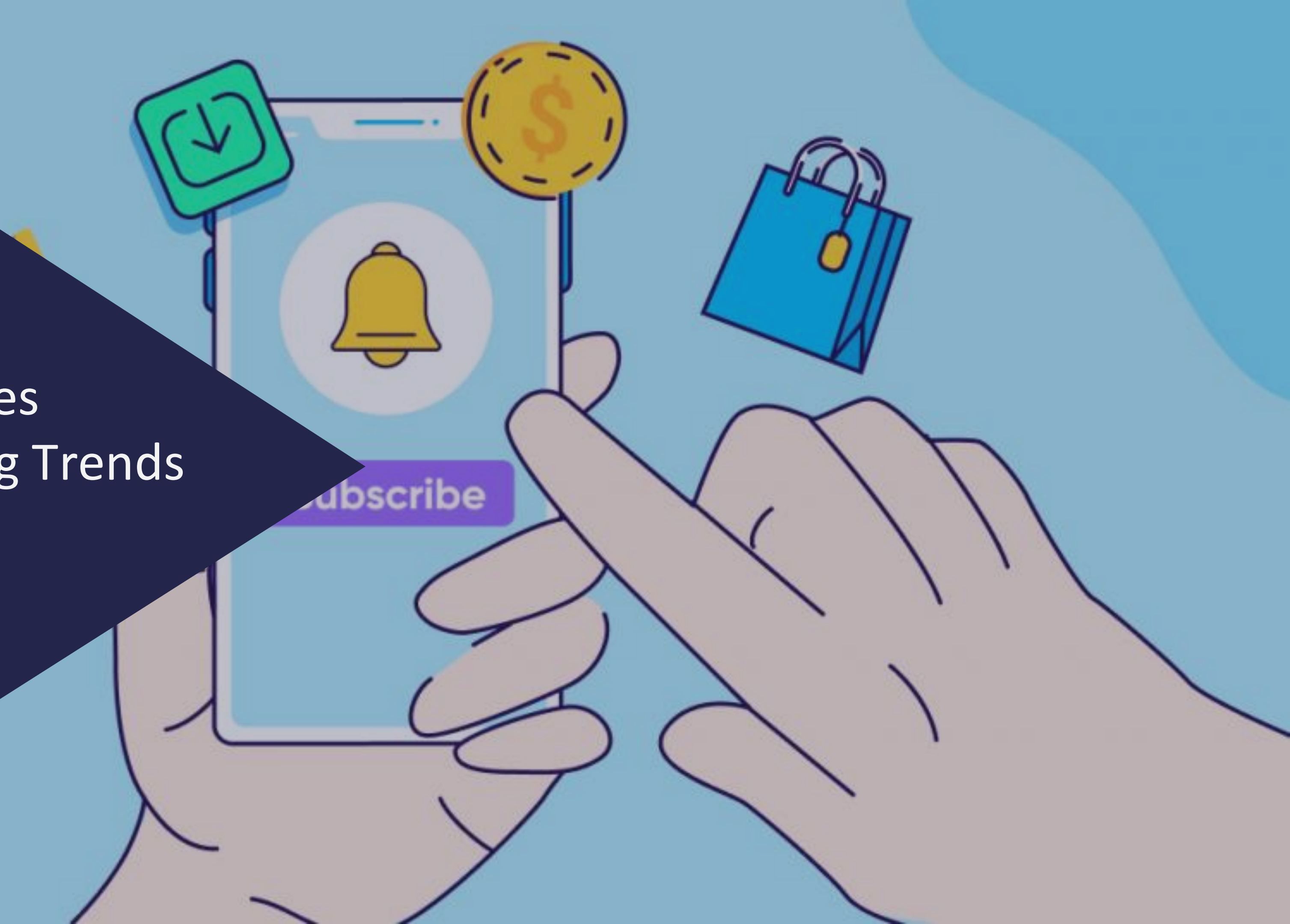
Data Transformations

Data cleaning, handling missing values, variables computed / calculated (LTV, Segment groups, etc.)

Analysis

Customer Segmentation, RFM Analysis, Random Forest, Gradient Boosting, Retention Analysis,Quadrant Analysis, Clustering methods

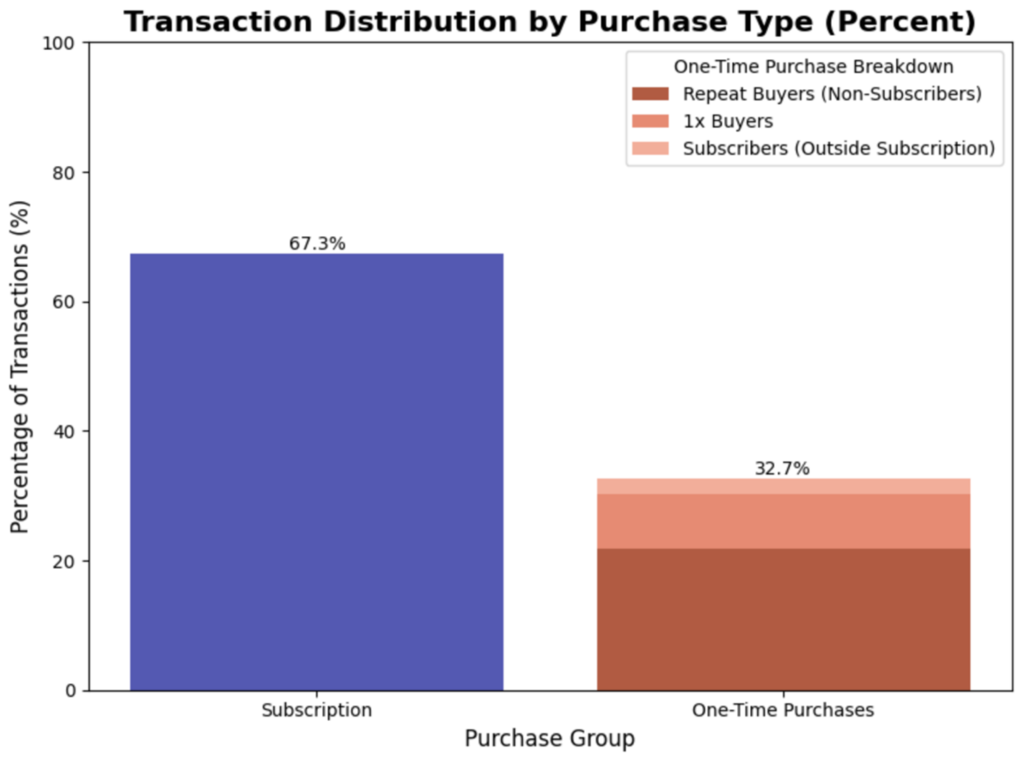
Customer Types and Purchasing Trends



Breakdown of Direct-to-Consumer Transactions

Subscribers dominate the transaction data, representing over two-thirds of direct-to-consumer orders.

Among one-time purchases, Repeat Buyers (non-subscribers) account for the largest share, followed by 1x Buyers and occasional purchases from Subscribers outside of their subscription.

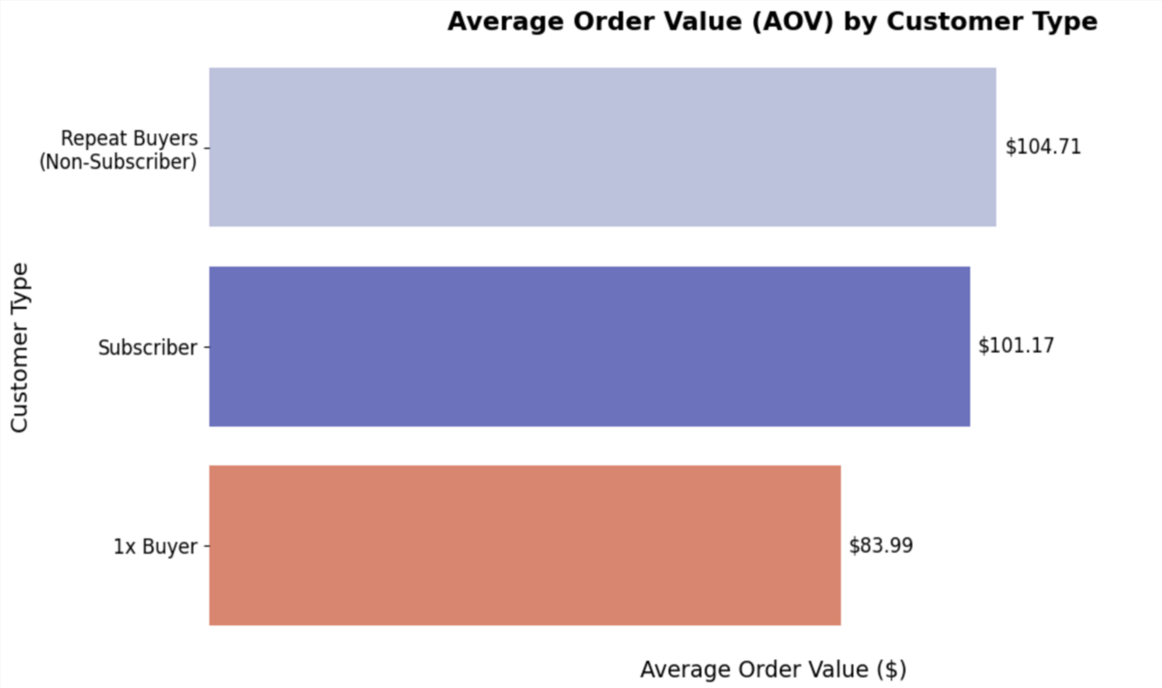


See slide notes for filtering criteria and total sample size

Note: Due to limited historical data, transaction distributions and breakouts may not fully reflect long-term trends or purchasing behaviors

Average Order Value (AOV) by Customer Type

Subscribers and Repeat Buyers (Non-Subscribers) show comparable AOV, while 1x Buyers exhibit a notably lower average per transaction.



See slide notes for for details on group definitions and AOV calculation methodology

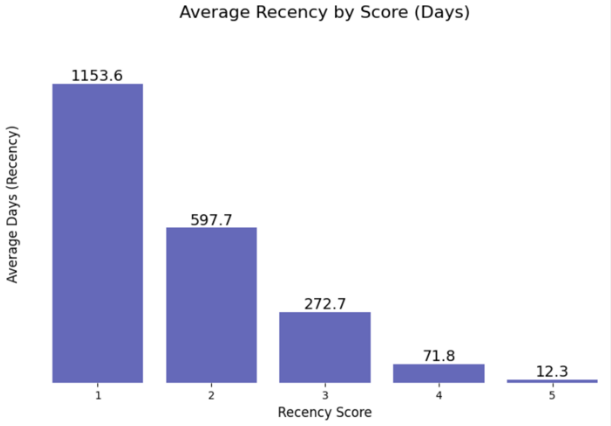
Note: Due to limited historical data, AOVs may not fully reflect long-term customer behaviors and should be interpreted with caution, especially across segments

RFM Analysis: Overview & Insights

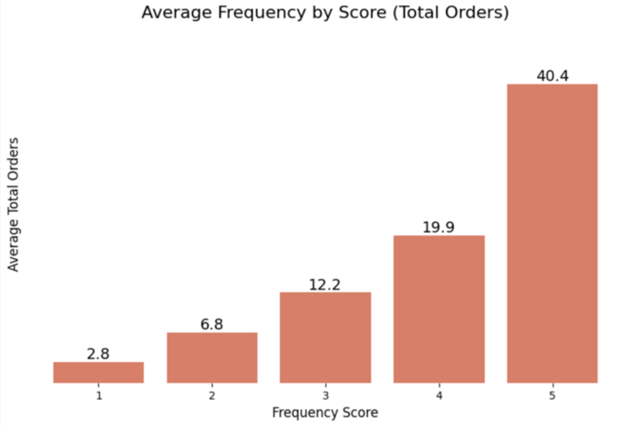
RFM Metric Overview: How Customers Are Grouped

Customers are scored from 1 to 5 for Recency, Frequency, and Monetary Value based on quintile splits, with higher scores indicating better performance in each dimension.

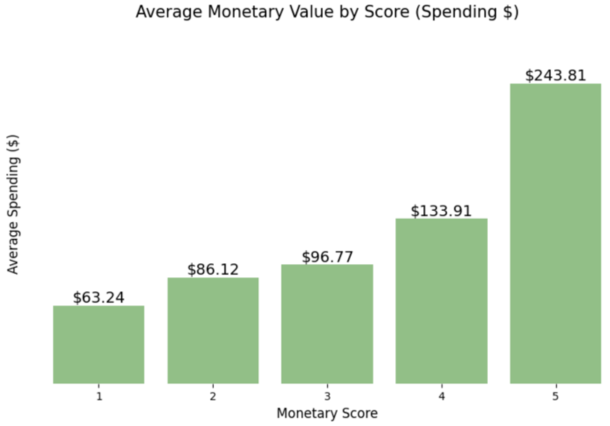
Recency



Frequency



Monetary Value*



Note: Due to the limited historical data, the Monetary Value* does fully capture the total spending of customers, limiting its reliability for RFM analysis in its current form

RFM Score Profiles of Customer Groups

VIP Customers

Highly engaged customers with significant spending and a high volume of total orders

Score	Recency Range (days)	Frequency Range (transactions)	Monetary Range ($)
5	0 to 19	29 to 527	$171.17 to $1173.35
4	20 to 68	19 to 28	$109.40 to $170.97
3	69 to 211	12 to 18	$90.84 to $109.35
2	212 to 444	6 to 11	$81.46 to $90.83
1	445 to 728	2 to 5	$0.28 to $81.41

Frequent Low Spenders

Engaged customers with high volume of total orders, but low order value

**May be re-classified (e.g., as VIPs) if analyzed on full historical data*

Score	Recency Range (days)	Frequency Range (transactions)	Monetary Range ($)
5	0 to 19	29 to 527	$171.17 to $1173.35
4	20 to 68	19 to 28	$109.40 to $170.97
3	69 to 211	12 to 18	$90.84 to $109.35
2	212 to 444	6 to 11	$81.46 to $90.83
1	445 to 728	2 to 5	$0.28 to $81.41

Infrequent High-Spenders

High order value, low volume of total orders. May include new customer

Score	Recency Range (days)	Frequency Range (transactions)	Monetary Range ($)
5	0 to 19	29 to 527	$171.17 to $1173.35
4	20 to 68	19 to 28	$109.40 to $170.97
3	69 to 211	12 to 18	$90.84 to $109.35
2	212 to 444	6 to 11	$81.46 to $90.83
1	445 to 728	2 to 5	$0.28 to $81.41

At Risk Customers

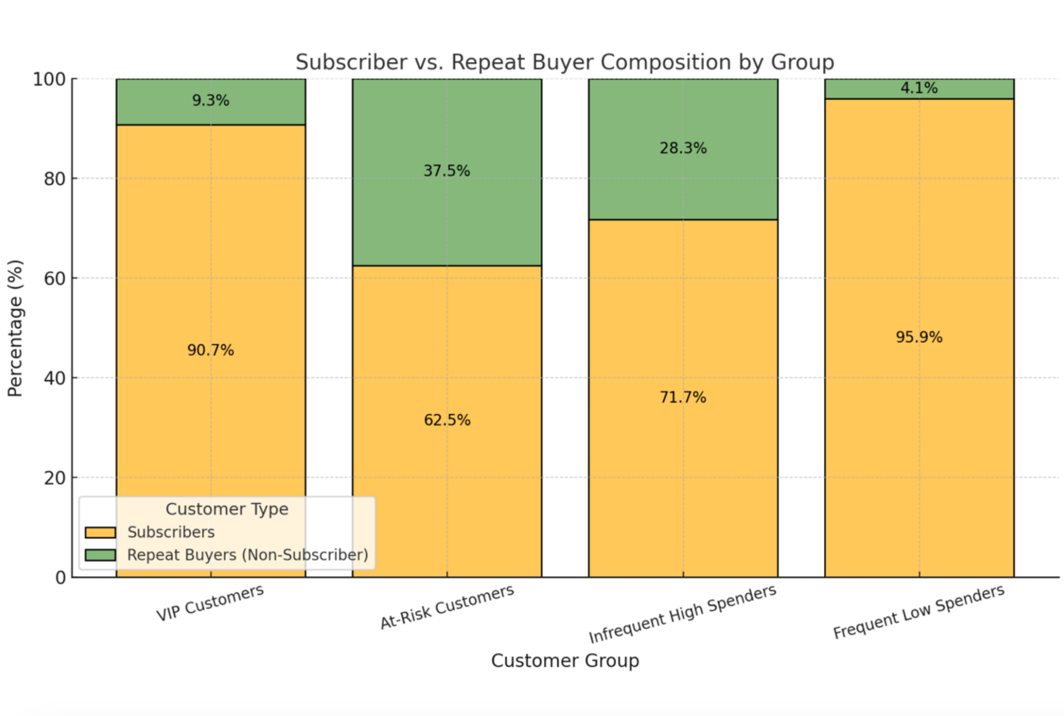
Disengaged customers (212 days since last order) with average to high spending

Score	Recency Range (days)	Frequency Range (transactions)	Monetary Range ($)
5	0 to 19	29 to 527	$171.17 to $1173.35
4	20 to 68	19 to 28	$109.40 to $170.97
3	69 to 211	12 to 18	$90.84 to $109.35
2	212 to 444	6 to 11	$81.46 to $90.83
1	445 to 728	2 to 5	$0.28 to $81.41

Monetary values are based on a limited sample and may not fully reflect historical customer behavior. Future analysis with broader historical data could refine thresholds and improve segmentation accuracy.

Subscriber vs. Non-Subscriber Distribution Across Customer Groups*

Subscribers dominate most groups, but non-subscribers make up a larger share of Infrequent High Spenders and At-Risk Customers—highlighting potential to convert/retain them.



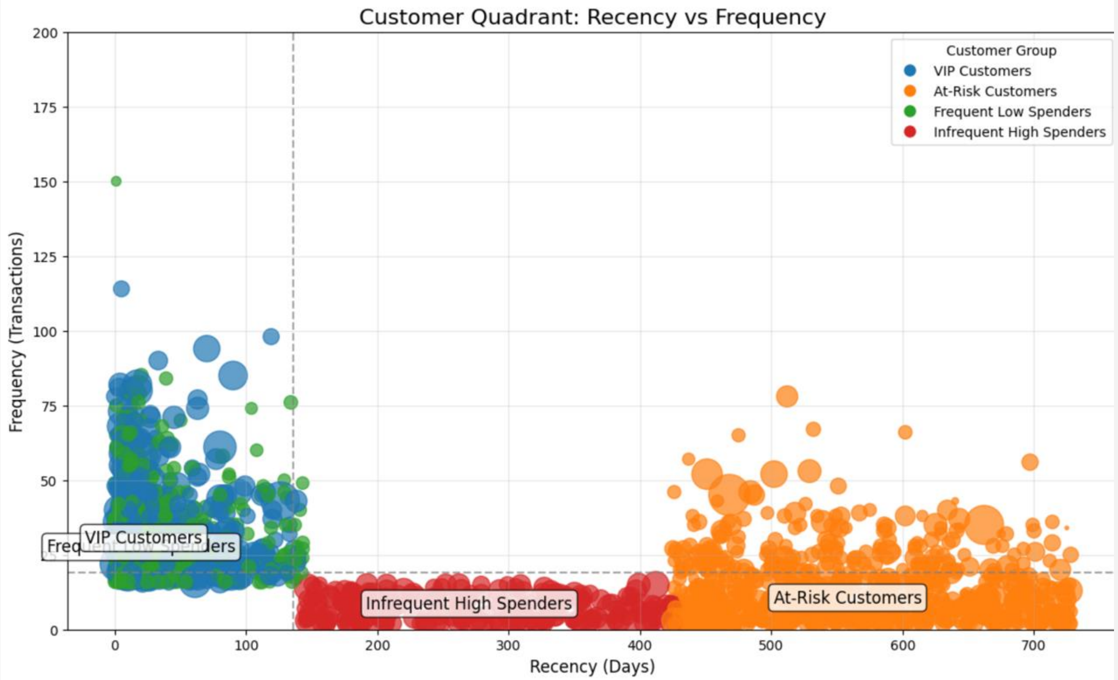
Note: Due to the limited historical data, these groups are defined based on unreliable "spend" data (i.e., total spend per customer is not accounted for)

RMF Customer Group Quadrant View

Subscribers Dominate High-Value Segments:
VIPSs – mostly subscribers - exhibit consistent purchase behavior, underscoring their role in driving revenue. Future strategies could focus on personalized rewards programs to enhance loyalty and maximize lifetime value

One-Time Buyers Exhibit Higher Churn Risk:
One-time buyers show greater variability in purchase recency and order amounts, indicating a higher potential for churn. Targeted re-engagement campaigns, such as win-back offers and upselling opportunities, could help convert them into recurring customers or subscribers.

Limited Historical Data Restricts RFM Accuracy:
The analysis is likely skewed due to limited historical data availability. This restricts the ability to confidently classify customer segments and identify trends. Expanding historical datasets should be a priority to refine and revalidate the segmentation model.



Quadrant Analysis

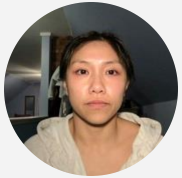
Q2. What common characteristics of veterinarians and clinics correlate with consistent purchase of ElleVet at scale, and what does that information offer us in predicting lifecycle and LTV behavior

If we plot vets within the quadrants of a 4-box in which frequency of purchase runs on the x axis and LTV on the y axis, what do we learn and what can we action?

Email: Nguyen.anh21@northeastern.edu

Follow-Up Analysis

- Latest (2023 – 2024) versus Previous (2022 – 2023)
- LTV Definition
  - Average Order Value * Total Orders
- Expanded to 100 top clinics



20 Top LTV Clinics (Latest)

Note: Analysis was done for Top 100 but visual is top 20 only.

High Frequency, Low LTV

Characteristics:

Frequent customers, but with lower average order value or spending.

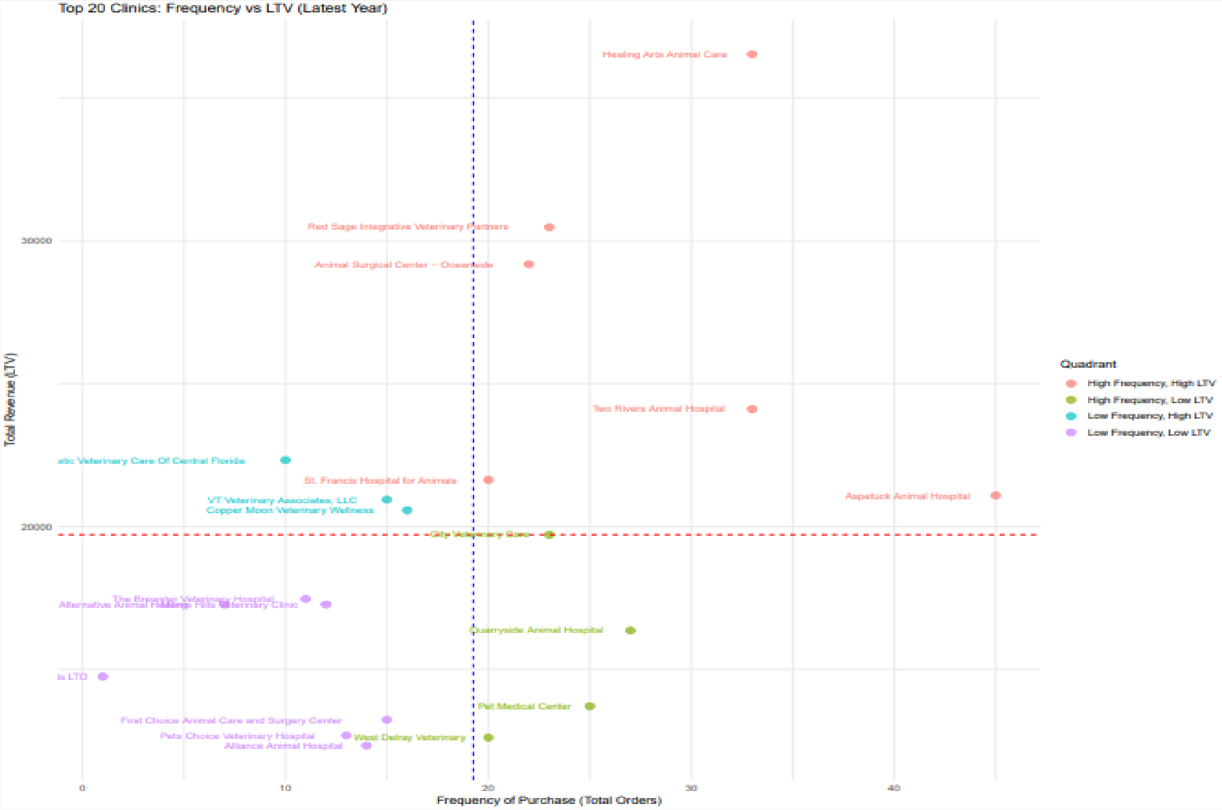
Potential for revenue growth through higher-value purchases.

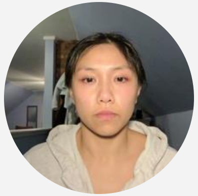
Low Frequency, Low LTV

Characteristics:

Clinics with minimal engagement and low revenue contribution.

Represents a low-priority group but could offer opportunities for activation.





High Frequency, High LTV

Characteristics:

Loyal clinics with frequent orders and high revenue generation.

Represent ElleVet's top-performing customers.

Low Frequency, High LTV

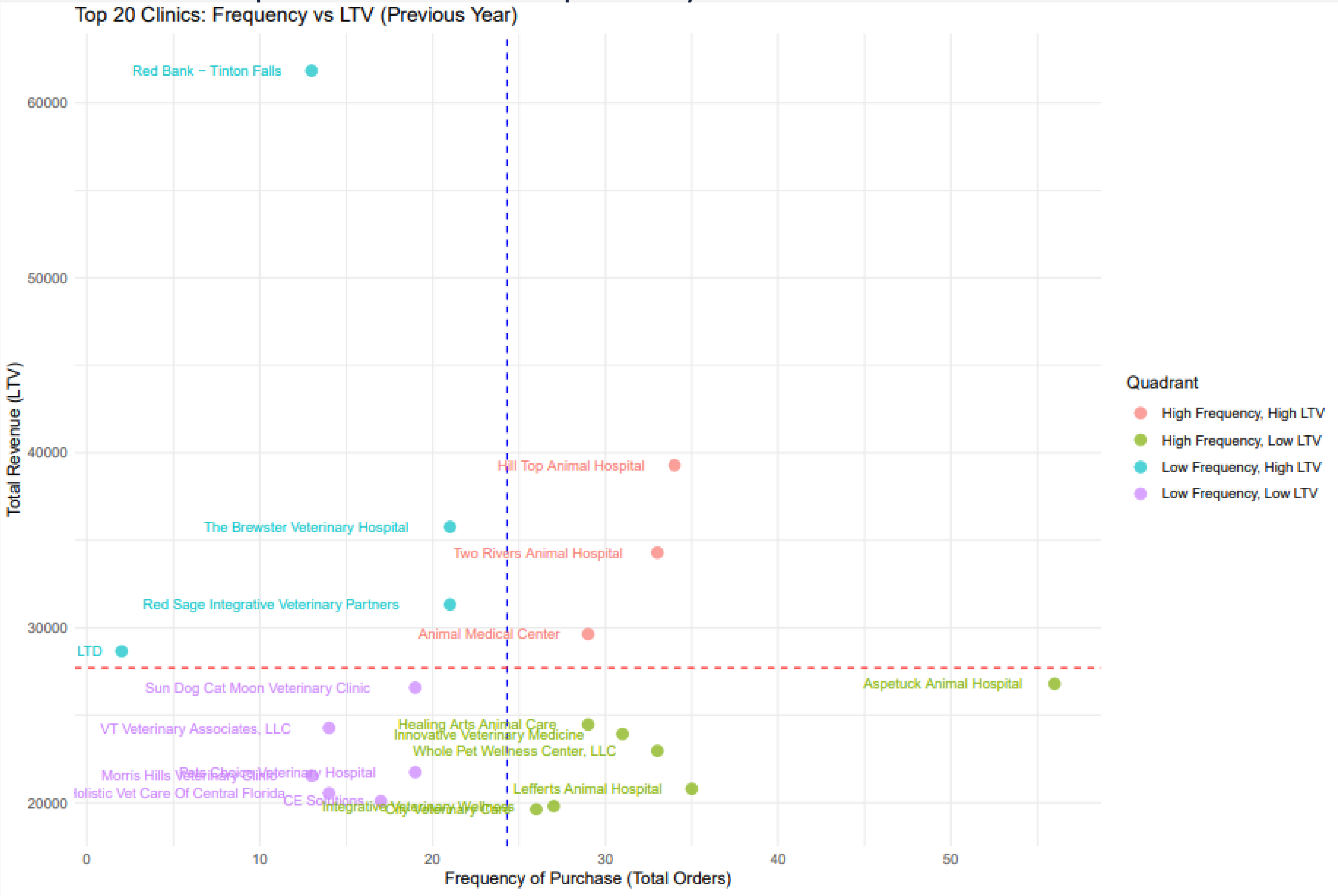
Characteristics:

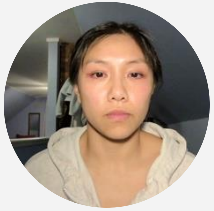
Clinics with high-value orders but infrequent purchasing.

Could become high-frequency customers with targeted engagement

20 Top LTV Clinics (Previous)

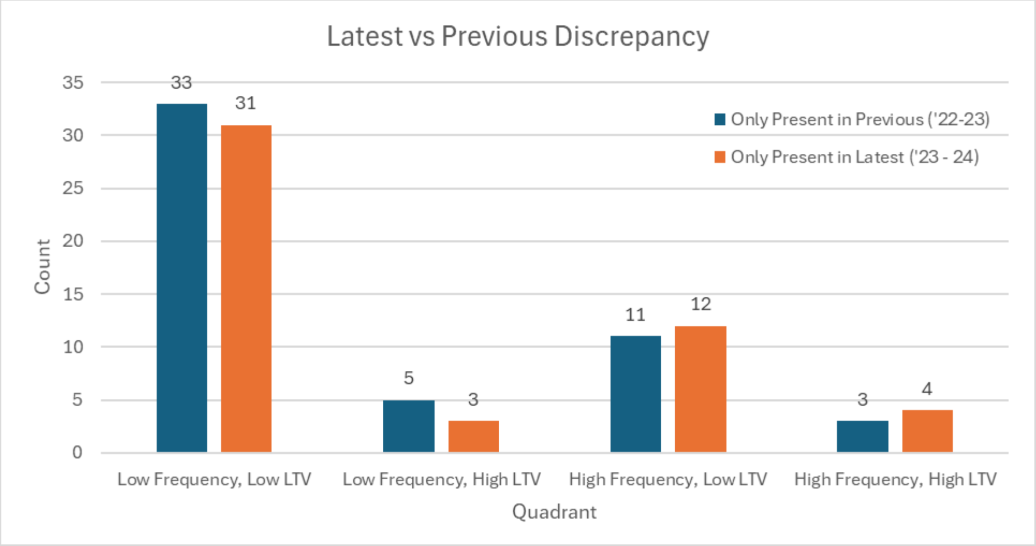
Note: Analysis was done for Top 100 but visual is top 20 only.





Summary

- 50 clinics were not present in latest data set and 50 were not present in previous data set.



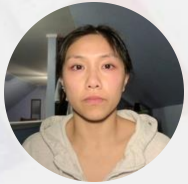
Description	Only Present in Previous ('22-23)	Only Present in Latest ('23 - 24)
Low Frequency, Low LTV	Low-value and infrequent customers, suggesting they might have stopped purchasing due to lack of engagement	The new addition of these low-value, infrequent customers might reflect an expansion of the customer base but not necessarily high-quality
Low Frequency, High LTV	While having high-value orders, were infrequent. Their absence in the latest data suggests a potential missed opportunity to convert them into more frequent buyers.	These clinics represent promising new customers, as they bring high-value orders despite low frequency.
High Frequency, Low LTV	While having high-value orders, were infrequent. Their absence in the latest data suggests a potential missed opportunity to convert them into more frequent buyers.	The addition of these clinics is valuable in terms of consistency, even if their per-order revenue is lower.
High Frequency, High LTV	The loss of these clinics is significant, as they were both frequent and high-value customers (Hill top Animal Hospital, Animal Medical Center, CE	These are the most valuable additions, as they represent high-value, consistent customers.

Summary Continued

# (out of 50)	% Clinics	Previous ('22-23)	Latest ('23-'24)
25	50%	No change	
3	6%	High Frequency, High LTV	High Frequency, Low LTV
2	4%	High Frequency, High LTV	Low Frequency, High LTV
2	4%	High Frequency, High LTV	Low Frequency, Low LTV
3	6%	High Frequency, Low LTV	High Frequency, High LTV
3	6%	High Frequency, Low LTV	Low Frequency, Low LTV
2	4%	Low Frequency, High LTV	High Frequency, High LTV
1	2%	Low Frequency, High LTV	High Frequency, Low LTV
5	10%	Low Frequency, High LTV	Low Frequency, Low LTV
1	2%	Low Frequency, Low LTV	High Frequency, High LTV
1	2%	Low Frequency, Low LTV	High Frequency, Low LTV

Implications:

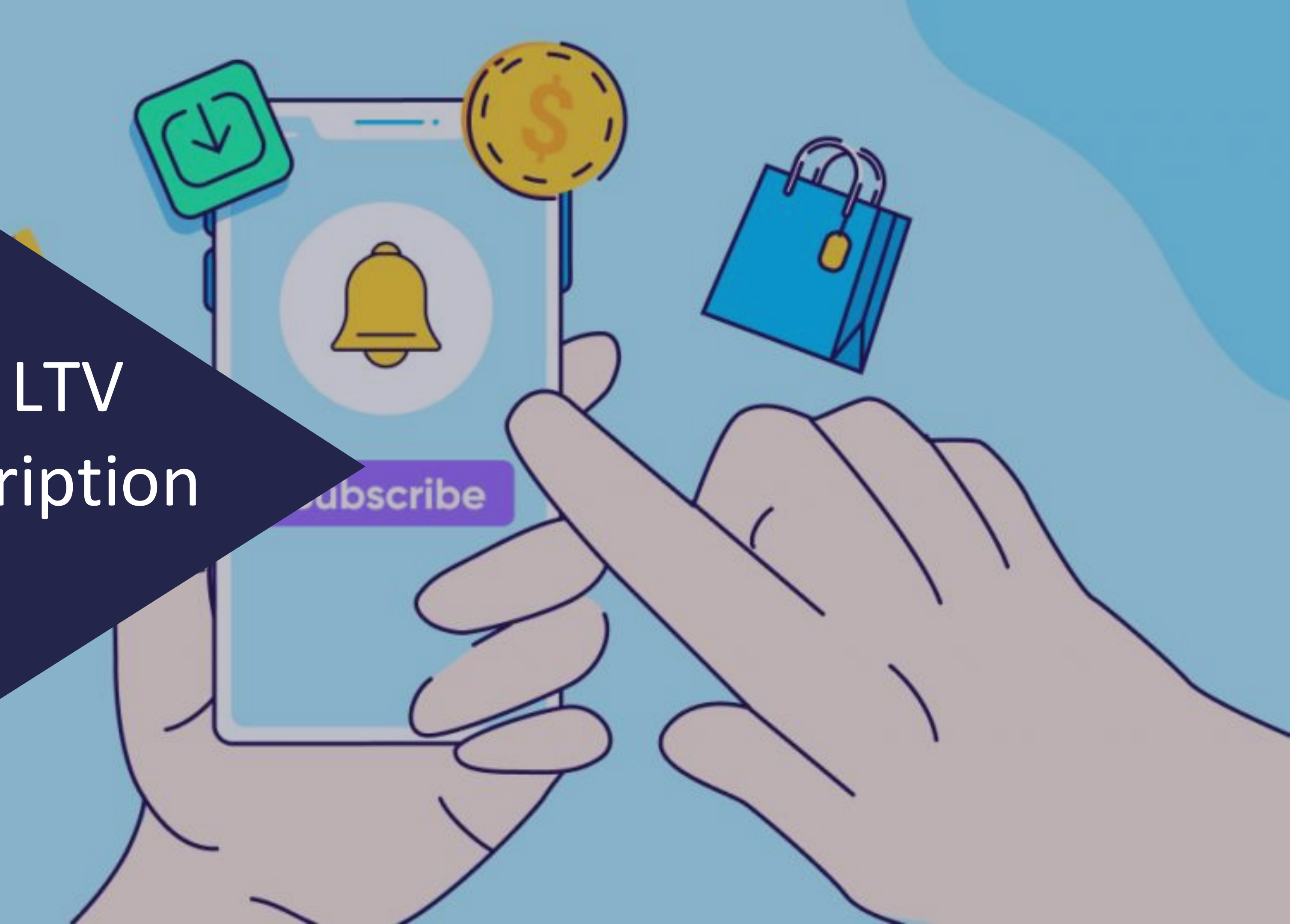
- Retention Concerns: The downward shifts, particularly among high-frequency, high-LTV clinics, highlight the need for strategies to sustain engagement and value.
- Upside Opportunities: Positive shifts from low-frequency and low-LTV clinics suggest growth potential with targeted interventions.
- Focus Areas: Efforts should be concentrated on maintaining high-performing clinics while continuing to nurture and convert low-performing ones.



Marketing Strategies for Customer Segmentation Using Quadrant Analysis

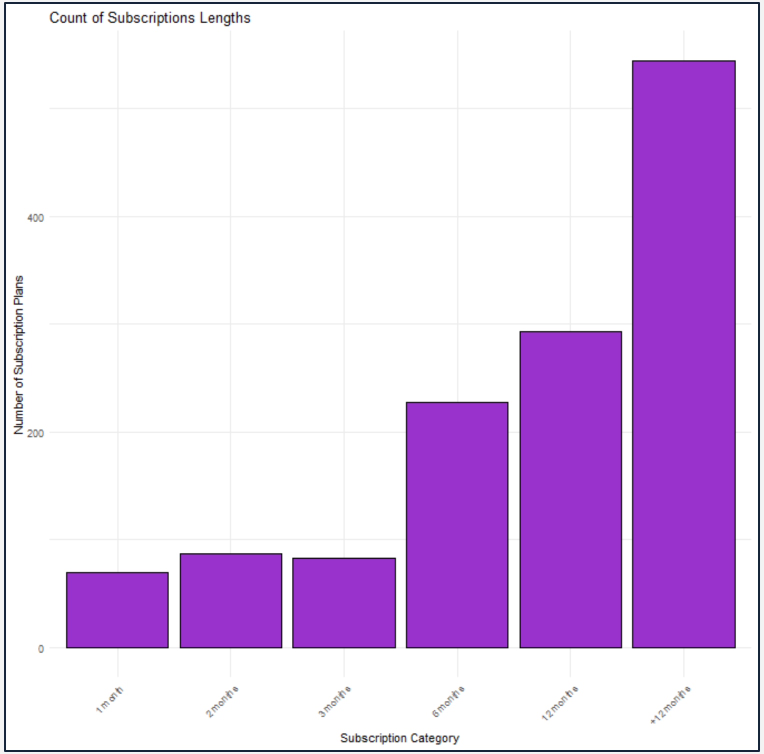
High Frequency, Low LTV **Upselling Opportunities:** Recommend premium products or larger order sizes. **Bundling Products:** Encourage purchases of bundled packages to increase order value. **Educational Content:** Provide insights on the benefits of using ElleVet's higher-value offerings. **Volume Discounts:** Incentivize larger orders with discounts for bulk purchases. **Behavior-Based Promotions:** Use data to suggest complementary products based on their purchase history.	**High Frequency, High LTV** **Loyalty Programs:** Exclusive rewards for high-frequency, high-spending clinics (e.g., volume discounts, VIP tiers). **Personalized Service:** Assign dedicated account managers to provide priority support. **Cross-Selling/Upselling:** Introduce complementary products or premium packages. **Co-Marketing Opportunities:** Collaborate on educational content, events, or promotions that spotlight their success with ElleVet products. **Early Access:** Offer early access to new products or trials as a token of appreciation.
Low Frequency, Low LTV **Awareness Campaigns:** Educate these clinics on ElleVet's value proposition through webinars, email series, or introductory offers. **Entry-Level Products:** Promote cost-effective or trial-sized products to lower barriers to entry. **Retention Focus:** Provide small incentives for continued engagement, such as discounts on second orders. **Surveys & Feedback:** Identify barriers to engagement and address specific concerns. **Referral Programs:** Encourage clinics to refer other practices for mutual benefits.	**Low Frequency, High LTV** **Re-Engagement Campaigns:** Send timely reminders or "We miss you!" offers to encourage repeat purchases. **Subscription Models:** Introduce a subscription plan for auto-replenishment. **Seasonal Promotions:** Target these clinics with campaigns tied to their typical purchasing cycles. **Consultative Selling:** Provide customized support to highlight product benefits and boost usage frequency. **Customer Stories:** Share testimonials or success stories to inspire similar engagement.

Predicting LTV and Subscription Retention



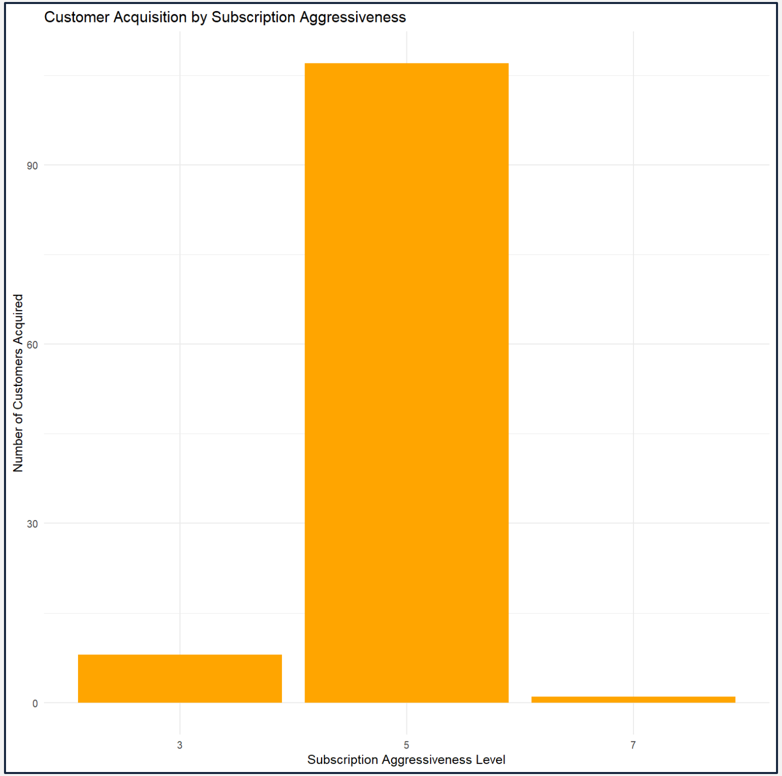
Q3: Subscription Length and Aggressiveness Analysis

- Subscriptions were categorized by length based off the subscription start and end dates to illustrate customer preferences for short-term and long-term subscriptions.
- The distribution of subscription lengths reveals a preference for longer-term subscriptions, with a significant portion of customers committing to 12-month and 12+ month subscriptions.
- This shows that customers are willing to commit to long-term plans, which offers to focus on longer-term subscriptions with subscription offers to maximize customer retention.



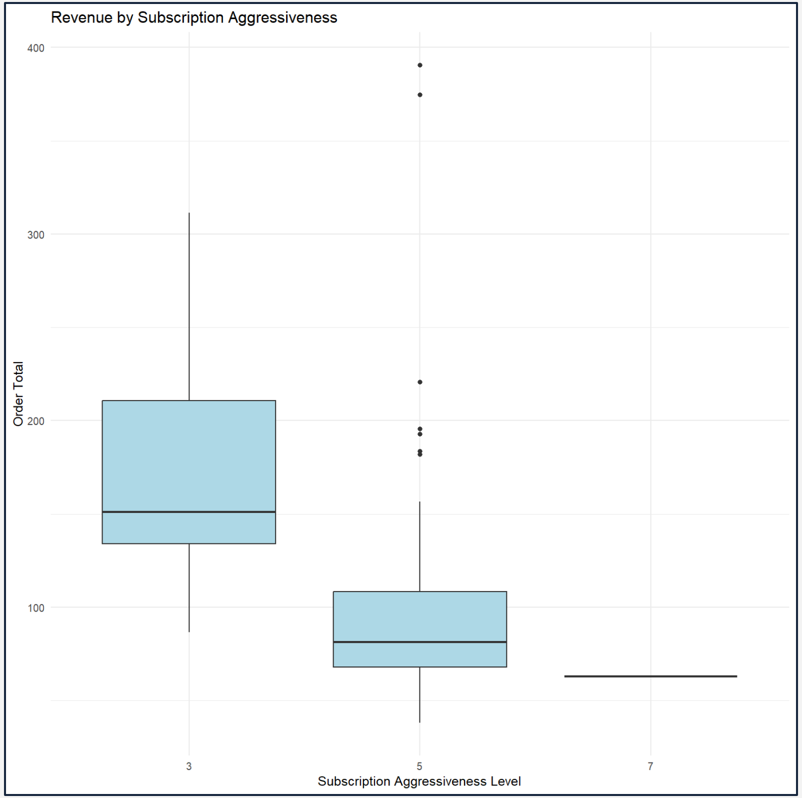
Q3: Subscription Length and Aggressiveness Analysis

- Subscription aggressiveness scoring was developed based on the discount levels and subscription lengths to classify subscription aggressiveness. If the discount is less than 10% then the subscription aggressiveness is set to 1, at least 10% it is set to 3, at least 20% is 5, and above 20% is 7.
- The chart indicates that the majority of customers were acquired at moderate levels of subscription aggressiveness, level 5. Very few customers were acquired at the highest levels of aggressiveness, suggesting that extreme discounts might not be as effective for customer acquisition.

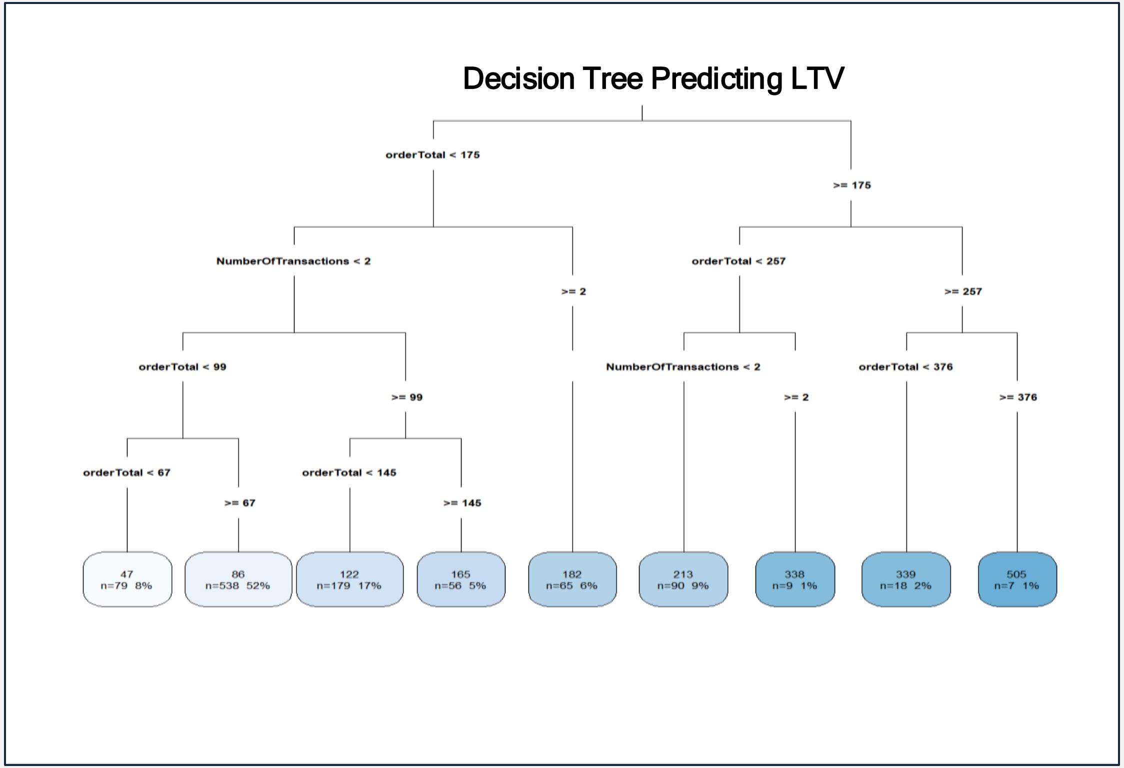


Q3: Impact of Subscription Aggressiveness on Revenue

- The chart here shows the distribution of Order Total across different levels of Subscription Aggressiveness
- The revenue across different levels of subscription aggressiveness shows a consistent revenue range across all levels, with no significant spike for more aggressive discount offers, implying that aggressive discounting doesn't necessarily lead to higher revenue per order, and lower discount levels might be just as effective in generating revenue.



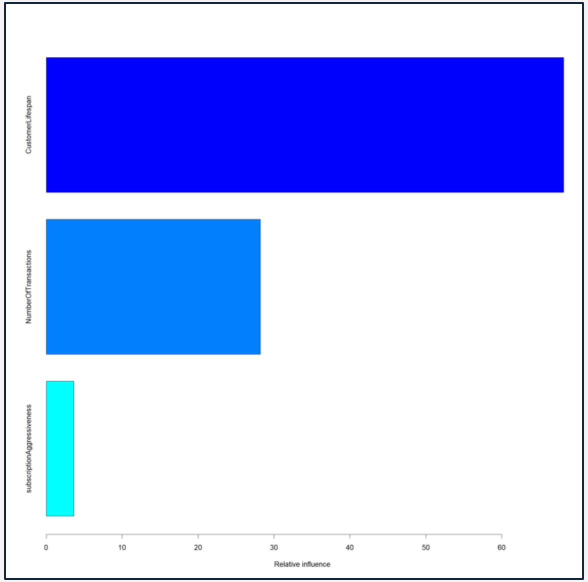
Q3: Modeling for LTV Prediction

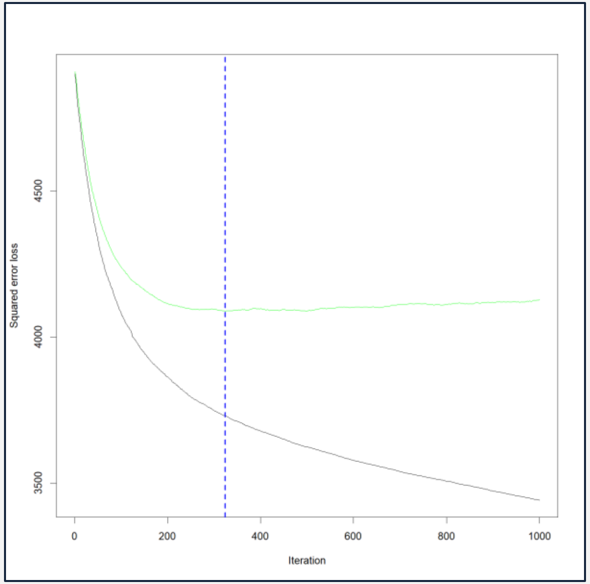


- These splits are based on value predictors that separate low and high LTV customers.
- The tree shows Number of transactions to be a strong predictor of LTV and long-term customers to have higher LTV. Customers with higher order totals have the highest predicted LTV, highlighting the influence of larger purchases.
- Providing incentives to increase transaction frequency and average order values could boost LTV for customers identified in the lower LTV segments.

Q3: LTV Influence of Predictors

- Customer Lifespan is the most significant predictor of customer value, suggesting longer relationships associate higher LTV. Number of transactions are another important feature of predicting LTV, emphasizing the value of frequent purchases.
- Subscription aggressiveness is less influential, though still shows that aggressive discounts or incentives contribute somewhat to LTV and may be more relevant for short-term.

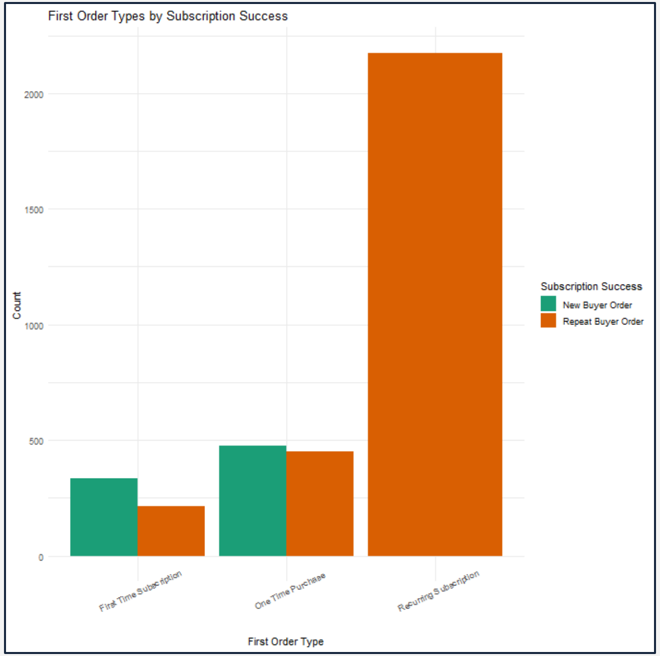


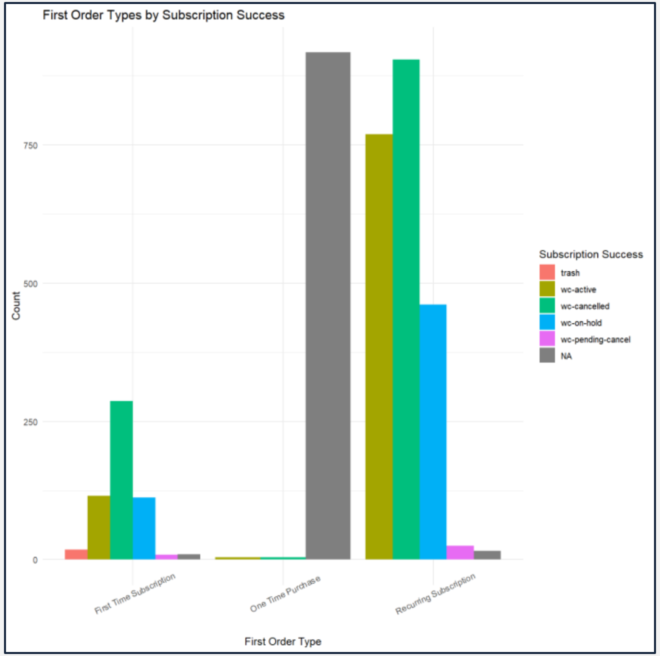


Subscription Success Distribution

- One time purchase customers are almost evenly split between new buyer and repeat buyer, suggesting a significant opportunity to convert one time purchase customers either through incentives or subscription options. First time subscriptions have overall low counts, possibly indicating either customer preference testing, or lack of promoting first time subscription orders.

- The "on-hold" subscription statuses are prominent for both subscribers, possibly showing temporary churn, and could be potential resubscribes through targeting on hold subscribers.



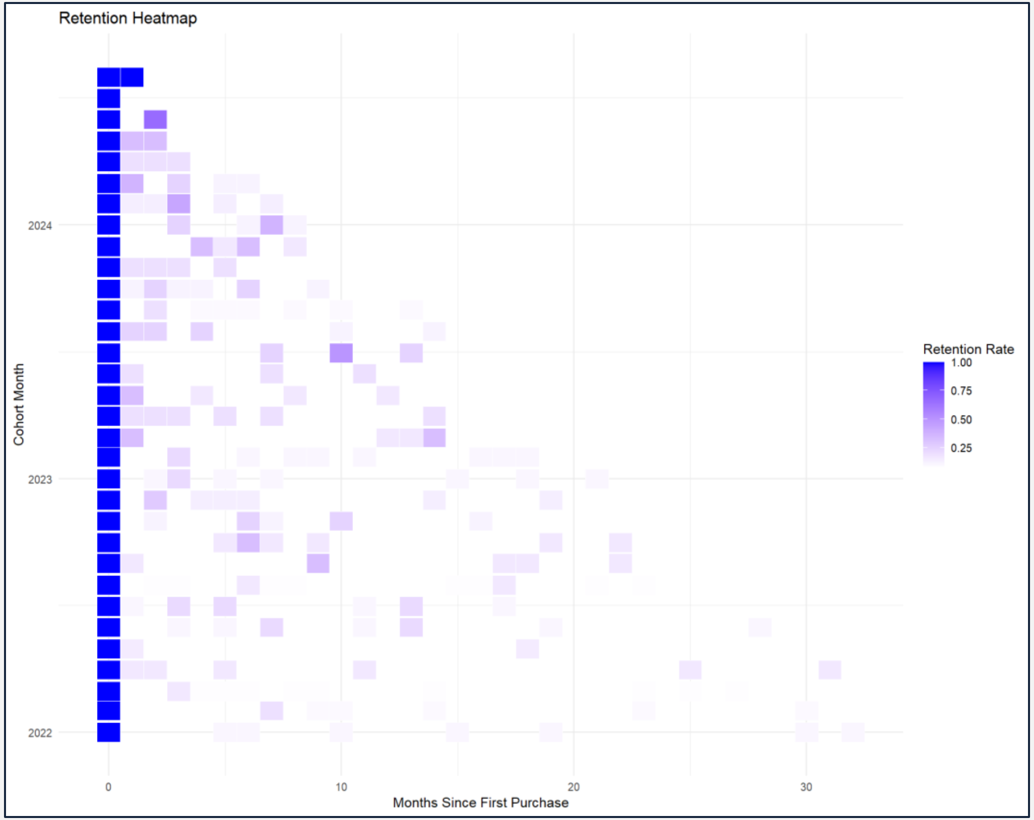


Customer Retention Heatmap

- Retention was analyzed by grouping customers based on the month of their first purchase, and monthly activity tracked. The retention rate is calculated as the proportion of customers retained each month.

- Retention rates decrease over time across all cohorts, with a more noticeable decline after the first 6-12 months of activity; though, 2024 shows to retain customers better during initial months compared to past years.

- Implement targeted re-engagement within the first 6 months to improve long-term retention.

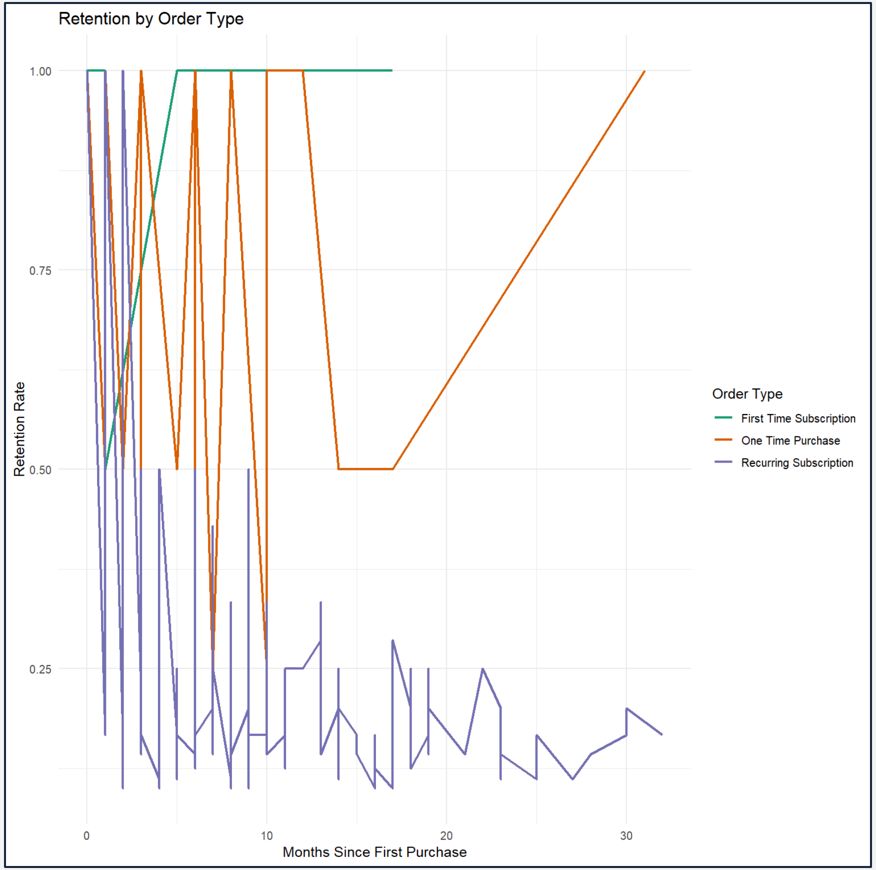


Retention Rate by Subscription Status

- Retention rates were calculated by taking the number of active customers in a month / the total customers for that order type.

- One Time purchases show sporadic return behavior, either customers return, or they don't. First time subscription rates start high, but decline after the initial months, maybe after the initial subscription period. Recurring subscriptions show high volatility over time.

- The stability of first time subscriptions in early months suggests a promising customer segment for upselling or encouraging long-term subscriptions.



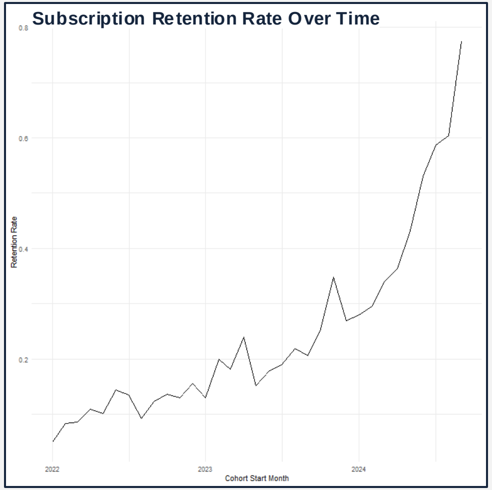
Customer Retention Rates Over Time

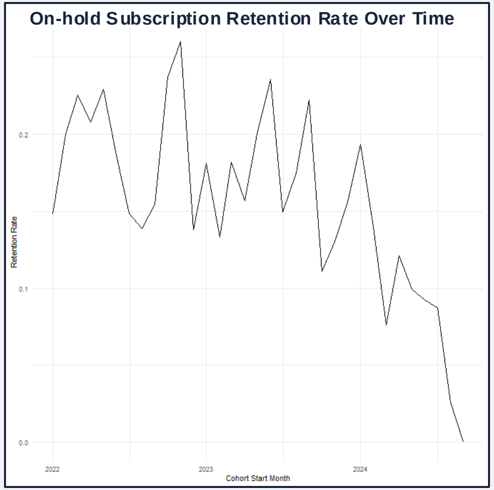
- Retention rates for subscribers have been increasing steadily, reaching close to 80% in recent months. The upward trend indicates recent efforts towards subscription values or user experience.

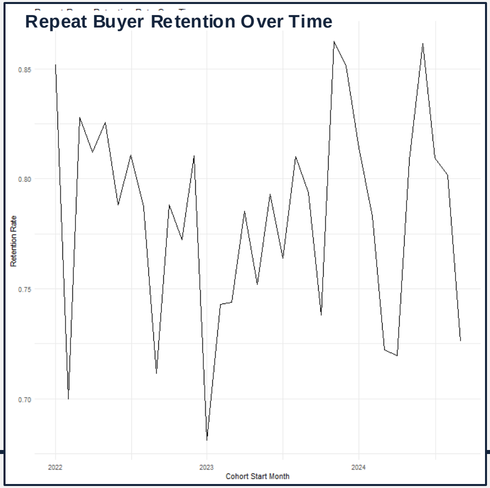
- Retention rates for on-hold subscriptions have been declining steadily over time, nearing 0% in recent months. This could indicate that once customers put their subscriptions on hold, they are less likely to reactivate

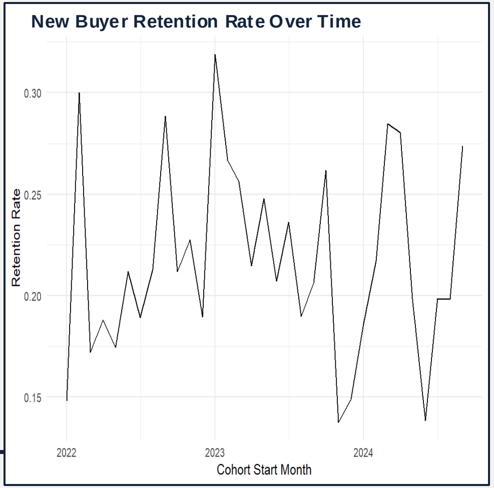
- Repeat buyers maintain a relatively high retention rate, hovering around 75%-85%, with slight dips and peaks over time; though, the generally high rates show relative customer loyalty in subscribers.

- Retention rates for new buyers are consistently low (below 30%) with significant fluctuations over time.



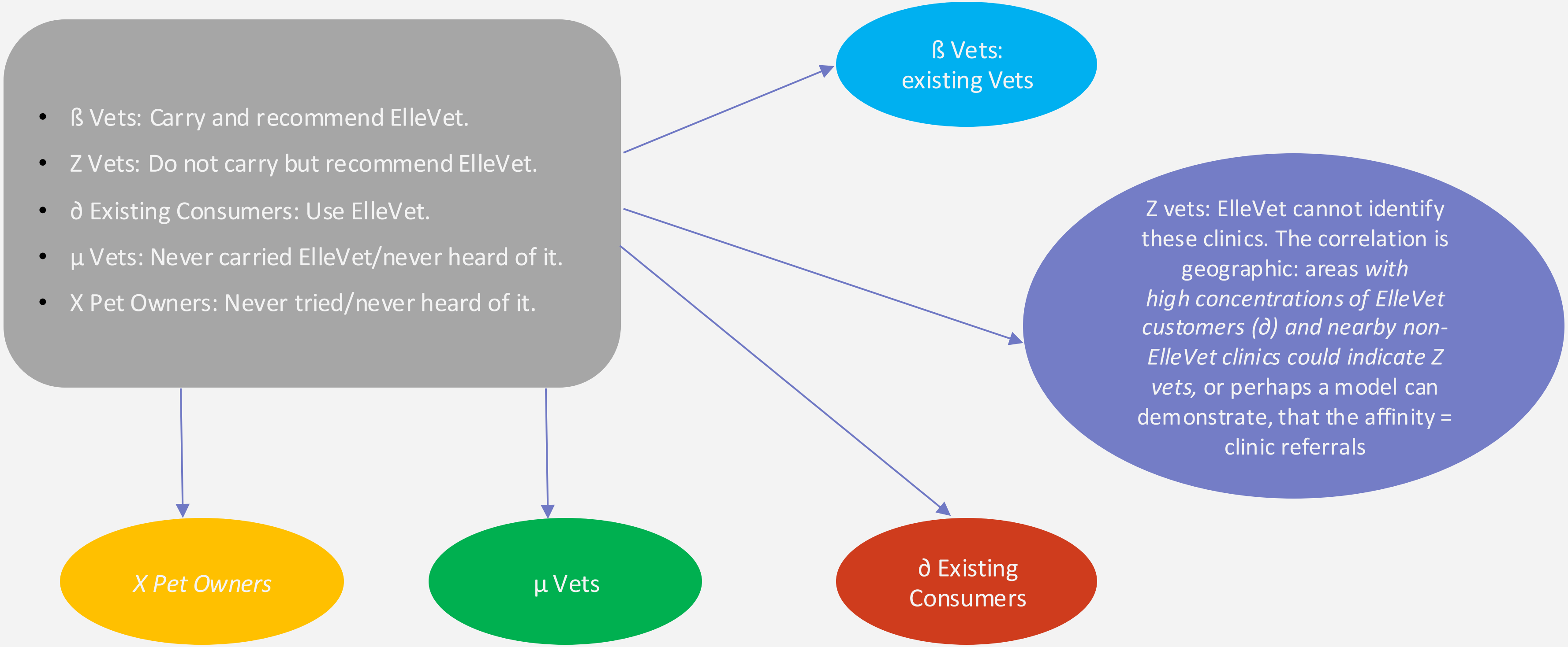






Connection between Consumers and Vet Clinics

Q4. Cross-Analysis of Consumers and Vets



Q4. Cross-Analysis of Consumers and Vets

Preprocess, impute missing values and merge two datasets: "All transactions" and "Vets' transactions"

	Customer ID	Customer Role	Address	City	State	Zipcode	Country
0	100004	Consumer	2860 Merchant Ct	Waldorf	MD	20603	United States
1	1001	Consumer	1152 Llagas Rd	MORGAN HILL	CA	95037	United States
2	1002	Vet	154 ROUTE 17K	Falls Church	NY	12550	United States
3	10023	Consumer	107 BERRY PATCH LN	DOTHAN	AL	36301-6047	United States
4	10024	Consumer	3794 SEMINOLE RD	FOREST HILL	WV	24935	United States
...	...	...	...	...	...	...	...
7083	wwh126@gmail.com	Consumer	41 Halpin Av	STATEN ISLAND	NY	10312	United States
7084	yaisaha@gmail.com	Consumer	14503 Studebaker Rd	Norwalk	CA	90650	United States
7085	yanaherbalbeauty@gmail.com	Consumer	300 Gorge Road	Cliffside Park	NJ	07010	United States
7086	youcancontactdiana@gmail.com	Consumer	557 Atlantic Avenue Apt. 3G	BROOKLYN	NY	11216	United States
7087	zoozlikar@gmail.com	Consumer	13944 S SHORE DR	TRUCKEE	CA	96161	United States

7088 rows × 7 columns

The merged dataset includes 7088 unique IDs of Consumers and Vets

Q4. Cross-Analysis of Consumers and Vets

Generate geolocation features 'latitude' and 'longitude' by Geopy - Nominatim

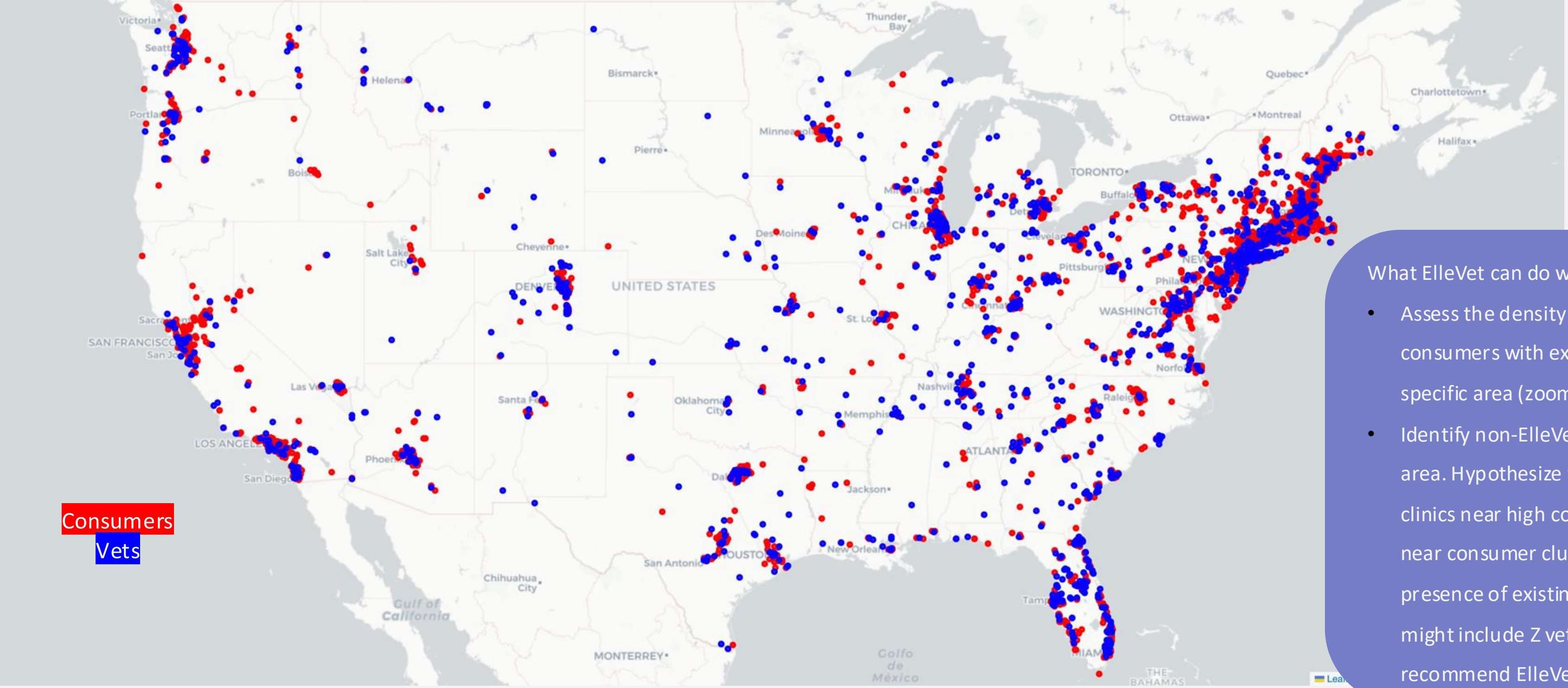
	Customer ID	Customer Role	Address	City	State	Zipcode	Country	latitude	longitude
1	1001	Consumer	1152 Llagas Rd	MORGAN HILL	CA	95037	United States	37.133084	-121.679663
4	10024	Consumer	3794 SEMINOLE RD	FOREST HILL	WV	24935	United States	37.572540	-80.795297
5	100242	Consumer	2009 AUSTRIAN WAY	COLORADO SPRINGS	CO	80919	United States	38.932074	-104.862528
7	10031	Consumer	10325 JUNCTION HILL DR	LAS VEGAS	NV	89134	United States	36.213829	-115.321555
9	100395	Consumer	531 SAVANNAH RD	LADSON	SC	29456	United States	33.023528	-80.093853
...	...	...	...	...	...	...	...	...	...
7082	wsantiago126@gmail.com	Consumer	200 Broad Street	STAMFORD	CT	06901	United States	41.055403	-73.534527
7083	wwh126@gmail.com	Consumer	41 Halpin Av	STATEN ISLAND	NY	10312	United States	40.553929	-74.186100
7084	yaisaha@gmail.com	Consumer	14503 Studebaker Rd	Norwalk	CA	90650	United States	33.899683	-118.100136
7085	yanaherbalbeauty@gmail.com	Consumer	300 Gorge Road	Cliffside Park	NJ	07010	United States	40.813744	-73.990066
7087	zoozlikar@gmail.com	Consumer	13944 S SHORE DR	TRUCKEE	CA	96161	United States	39.319252	-120.257972

4584 rows × 9 columns

```
Customer Role
Consumer    3236
Vet         1348
Name: count, dtype: int64
```

The final geolocation dataset includes 4584 records after filtering out nulls from columns

Q4. Geospatial Analysis of Consumers and Vets

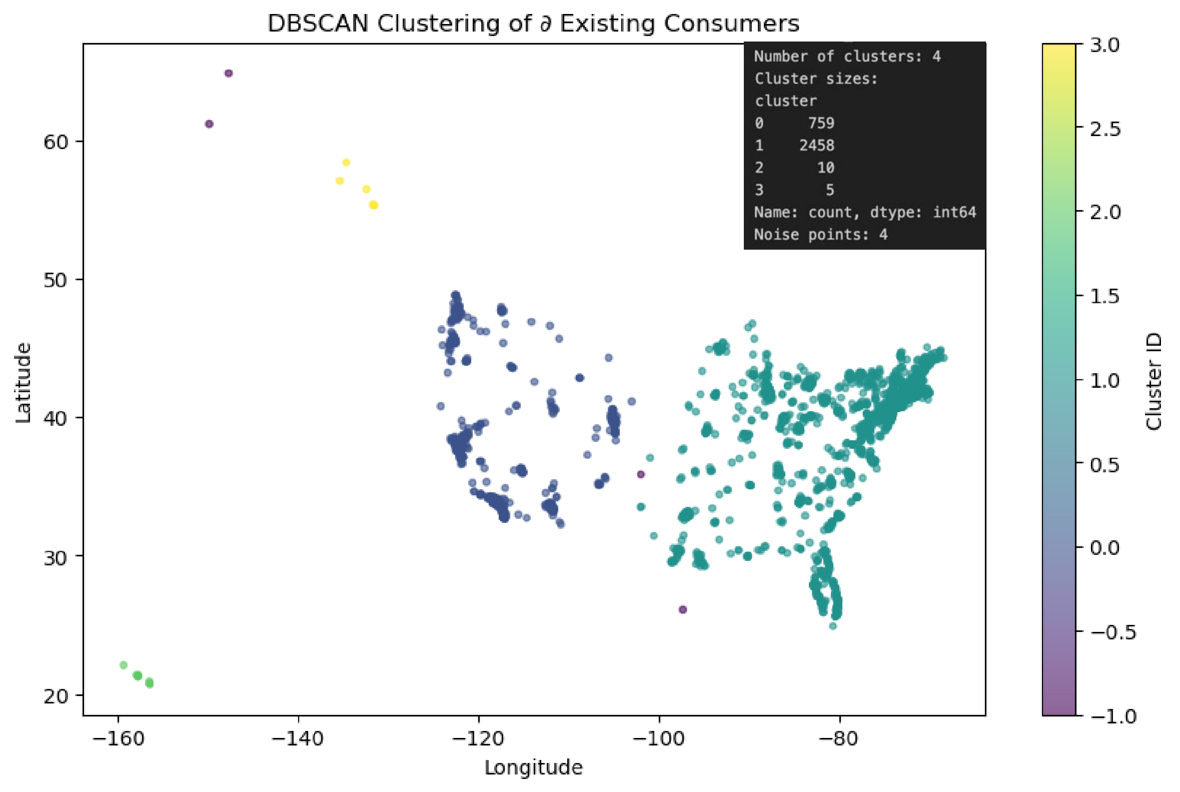


What ElleVet can do with this map:

- Assess the density of existing consumers with existing vets in each specific area (zoom in).
- Identify non-ElleVet clinics in each area. Hypothesize that non-ElleVet clinics near high consumer density or near consumer clusters w/o the presence of existing Ellevet clinics might include Z vets who recommend ElleVet but not carry it.

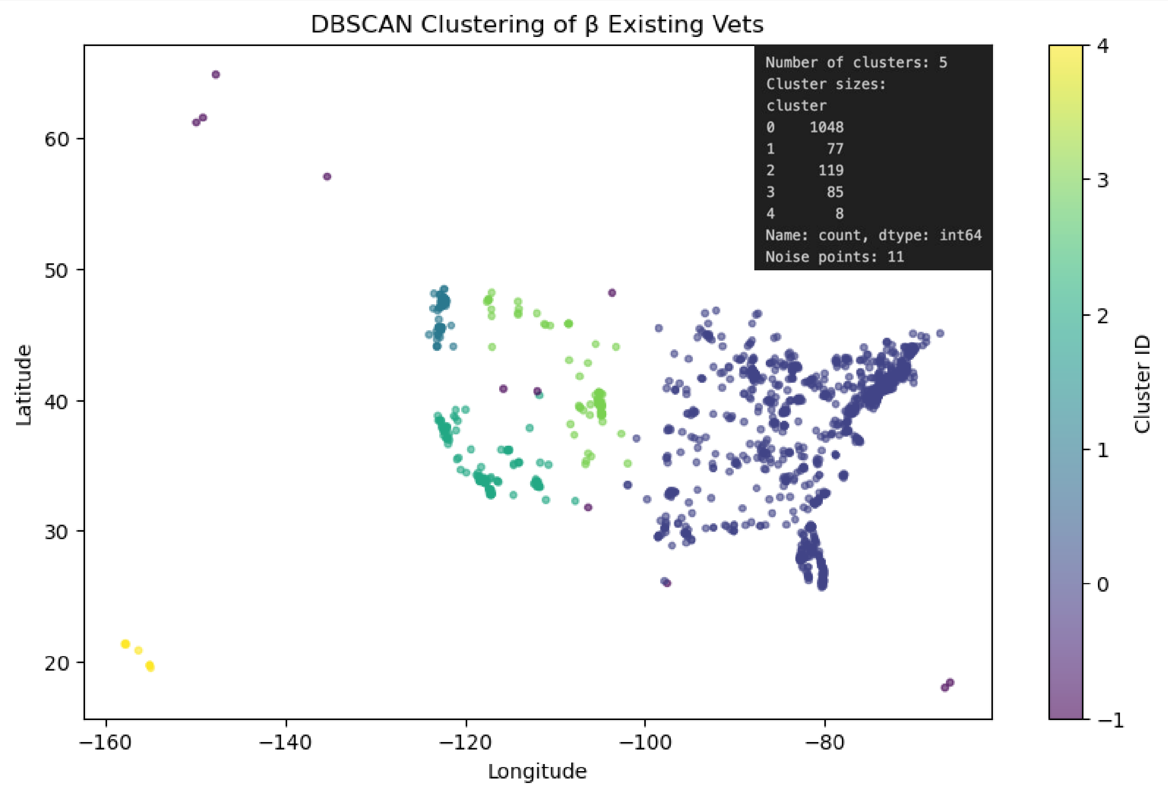
Interactive Map created by Folium library in Python

Q4. Cluster Analysis of Existing Consumers



- Cluster 1 (2,458 consumers): A critical region due to its large size. If this cluster does not have sufficient β clinics nearby, it represents a significant opportunity for outreach or marketing.
- Cluster 0 (759 consumers): A region with moderate potential. Worth considering for targeting if β clinic coverage is inadequate.
- Clusters 2 and 3 (10 and 5 consumers): These small clusters might indicate emerging or niche markets. Investigate whether these areas could grow with focused efforts or if they are less critical.
- Noise Points: These represent isolated customers, who are of lower priority for outreach.

Q4. Cluster Analysis of Existing Vets



- Cluster 0 (Major Cluster – 1048 vets): significantly larger than the others, indicating a region with a high concentration of vet clinics. This area could be a stronghold for ElleVet or a target for reinforcing the supply chain.
- Clusters 1–3 (Medium Clusters): represent smaller but still noteworthy groupings of vet clinics. These regions could benefit from marketing efforts to solidify ElleVet's presence or gather insights on customer behavior.
- Cluster 4 (Small Cluster): A small and isolated group that may indicate a niche region with limited but focused potential.
- Noise Points: The 11 noise points could represent vets in areas with outliers geographically.

Q4. Predictive modeling concepts

Build a Propensity Model to Infer Z Vet Clinics (who recommend ElleVet but do not carry it)

1. **Feature Engineering for the model:**
   - Customer Density: calculated consumer density at the zip code level
   - Distance to β Clinics: created a demo for calculating the distance from a clinic to Nearest β Clinic
   - Demographic and Socioeconomic Data: incorporated Population and Household Income; no data for Pet Ownership rate by zip code was found.
2. **Train the model:**
   - We need a training set in which 'is_z_vet' is the target column, along with geolocation features, engineered features, and other demographic and socioeconomic variables for non-ElleVet clinic dataset.
   - Use classification algorithms to train the model predicting the likelihood that a clinic in a region is a Z vet based on customer presence and proximity to β clinics.
3. **Manual Validation**
   - Select a few regions where the model predicts the presence of Z vets and manually investigate (via outreach or online research) to verify if the local clinics are recommending ElleVet without carrying it.
   - Use feedback from validation efforts to improve the model's accuracy and incorporate any newly acquired data to refine predictions.

Recommendations – Data Enrichment

- Survey and Feedback Mechanism: Use direct customer surveys or digital feedback loops (emails, post-purchase surveys) to ask ∂ customers how they heard about ElleVet.

- X Pet Owners (Never tried/never heard of): Estimate the population of pet owners (X) within each geography and cross-reference this with data from ∂ customers and β vet locations to identify under-penetrated areas.

- µ Vet (Unaware Vets): Identify clinics that are not in proximity to either β vets or ∂ customers' clusters => Can be categorized as µ Vets.

Recommendations - Interaction (β + Z + ∂ => X + μ)

1. Localized Marketing and Geofencing (Visibility):

- Run localized marketing campaigns targeting pet owners (X) in regions with high concentrations of ∂ customers and nearby β clinics => Create visibility for ElleVet in the areas most likely to yield new customers.
- For μ vets, target them with similar campaigns introducing them to ElleVet.

2. Incentivized Vet-Consumer Referrals (Trust and Incentive):

- Referral Programs: Offer incentives for β and Z vets to refer new X customers.
- Encourage existing customers (∂) to refer friends (X) or leave online reviews to influence local pet owners.

3. Educational Outreach (Knowledge):

- Develop educational materials for μ vets through webinars, free trials, or CE credits on ElleVet's efficacy => Transition them from μ to β or Z.
- For X pet owners, run campaigns that educate them on the benefits of ElleVet, targeting areas where ∂ customers are dense but ElleVet awareness is low.

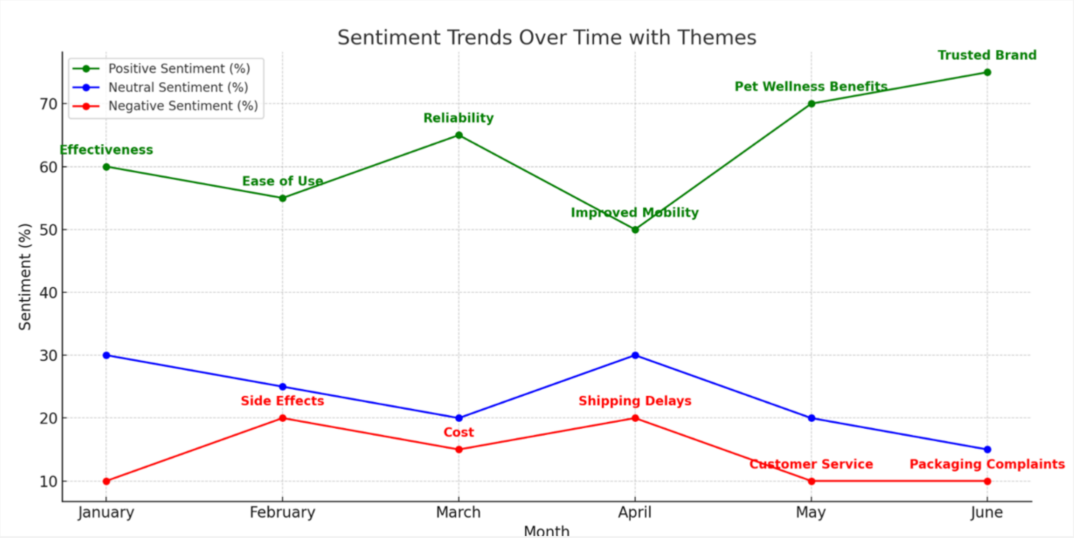
Future Project Proposals

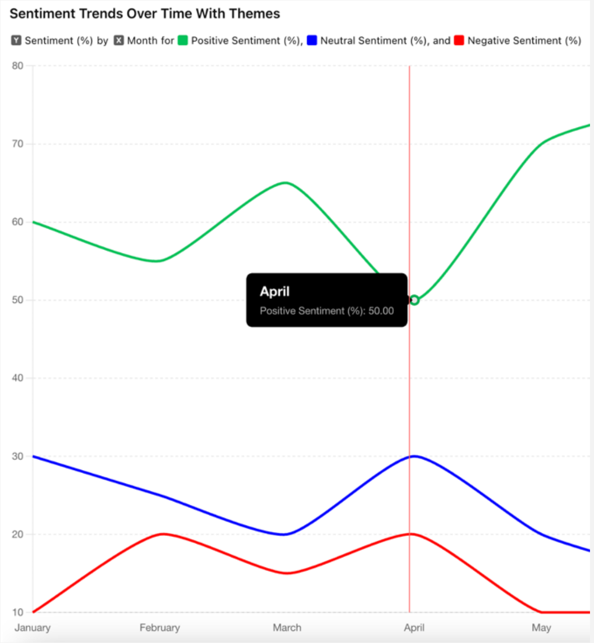
Sentiment and Feedback Analysis Proposal (Anna)

- **Objective:** Leverage sentiment and thematic analysis to understand customer and veterinarian perceptions of ElleVet's products. Drive actionable insights to refine marketing, product strategies, and engagement efforts.
- **Approach:**
  - Use machine learning (e.g., XGBoost) and thematic analysis (e.g., LDA) to classify sentiment and extract themes from customer reviews, social media mentions, and vet survey responses.
  - Integrate sentiment scores, thematic insights, and engagement metrics into a centralized dashboard for real-time insights.
- **Requirements:**
  - **Data Collection:** Aggregate customer reviews, social media mentions, and annual vet surveys (segmented by recommenders, non-recommenders, and unaware vets).
  - **Tools:** Use survey platforms (e.g., Qualtrics), social monitoring tools (e.g., Sprout Social), and secure dashboard integration.
- **Benefits:** Uncover drivers of satisfaction, identify adoption barriers, and enhance customer retention strategies through informed decision-making.

Sentiment Feedback Analysis Proposal (Cont'd)

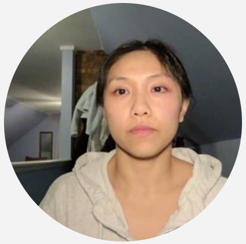
Proposed Dashboard Examples: Trending Sentiment & Thematic Insights





Ellevet Product Recommendation System (Anh)

- Proposal Overview:
  - Objective: Develop a predictive recommendation system using machine learning to personalize product suggestions for ElleVet’s customers.
  - Expected Benefits:
    - Enhanced customer engagement and retention
    - Increased sales through personalized recommendations
    - Data-driven insights for marketing and product development
- Literature
  - Success Companies:
    - Netflix, Amazon, Spotify
- Key Data Requirements:
  - Historical purchase data (sales records, product demand)
  - Customer interaction data (clicks, search queries, browsing time)
  - Demographic and psychographic profiles (age, income, preferences)
  - Customer feedback (reviews, surveys)



ElleVet Pet Medical Record Database System Proposal

- This project proposes the development of a digital pet medical record database system hosted by ElleVet Sciences. The system aims to streamline veterinarian care by providing a centralized platform for storing, accessing, and analyzing pet health data.
- **Objective**:
  - Create a cloud-based pet medical record database for veterinarians and pet owners to access, manage, and analyze pet health information. Provide tools to track health trends, manage treatments, and improve communication between veterinarians and pet owners.
- **Features**:
  - Interfaces tailored for veterinarian and pet owner use and accessibility.
  - Pet profiles with demographics, vaccination records, and treatment logs.
  - Visual health tracking.
- **Data & Use:**
  - Descriptive, predictive, and comparative analytics to track treatment outcomes and identify trends.
  - Enable insights into product performance and the identification of high-risk pets for early intervention.
- **Benefits**:
  - Strengthen ElleVet's partnerships with veterinary clinics.
  - Gather data to improve ElleVet product efficacy and inform product development.
  - Enhance the value of ElleVet products through tangible health outcomes.

Predictive Analysis of Customer Churn: Spatial and ML Approaches

- **Objective:** Improve customer lifetime value (LTV) by identifying predictors of customer loyalty and churn.
- **Data Sources:**
  - Quantitative Data:
    - Customer Demographics
    - Customer Behavioral Data
    - Spatial Data
    - Historical Churn Data
  - Qualitative Data:
    - Customer Feedback or Reviews
    - Exit Survey Data
  - Derived Features
    - Geospatial clusters based on customer density
    - Sentiment scores from textual reviews
- **Data Analysis Proposal**
  - EDA (Exploratory Data Analysis): Create dashboards displaying geographic clusters of customers, churn risk percentages, and purchase patterns.
  - Predictive Modeling:
    - Machine learning models such as stochastic gradient boosting, XGBoost, and random forests will identify churn predictors.
    - Ensemble learning methods will combine multiple models for improved accuracy.
- **Literature Review**

Appendix

References

Ariyanto, M. (2022). Quadrant analysis: One simple method to prioritize and take action for business strategy. Medium. Retrieved October 20, 2024, from https://mvchid.medium.com/quadrant-analysis-one-of-simple-method-to-set-up-prioritize-and-action-for-business-strategy-70bdc567a597

Dolnicar, S. (2003). "Using Cluster Analysis for Market Segmentation—Typical Misconceptions, Established Methodological Weaknesses, and Some Recommendations for Improvement." Australasian Journal of Market Research, 11(2), 5-12. https://ro.uow.edu.au/commpapers/139/

No author. (n.d.). Customer Lifetime Value-Experience Matrix. Priority Matrix. Retrieved October 30, 2024 from https://sync.appfluence.com/templates/customer-lifetime-value-experience-matrix/#:~:text=The%20top%2Dright%20quadrant%20includes,effective%20solutions%20or%20minimal%20engagement.

Statsig. (n.d.). How to calculate lifetime value (LTV). Statsig. Retrieved October 14, 2024, from https://www.statsig.com/perspectives/how-to-calculate-lifetime-value-(ltv)Links to an external site.

Survature. (n.d.). What is quadrant analysis? Survature Help Center. Retrieved October 20, 2024, from https://help.survature.com/what-is-quadrant-analysis/