

# Unlocking Customer Insights

## A Statistical Investigation

Tools Used : Python, Pandas, Matplotlib, Seaborn, SciPy

# *Introduction*

- **Objective:** Explore and interpret customer behavioral patterns using statistical and exploratory data analysis to uncover actionable insights.
- **Scope of Analysis:** 10,675 customer records across demographics, transactions, and engagement factors, analyzed through descriptive, correlation, and hypothesis tests.
- **Approach:** A balance of data-driven exploration and storytelling to convert analytical findings into strategic business insights.

## *Understanding the Dataset*

- **Dataset Composition:** The dataset includes 10,675 records and 12 variables capturing demographics, transactions, and behavioral attributes of U.S. customers.
- **Data Quality:** No missing values were detected across any feature; data types include 8 categorical and 4 numerical variables.
- **Key Variables:** Features include CustomerID, Name, State, Education, Gender, Age, Married, NumPets, JoinDate, TransactionDate, MonthlySpend, DaysSinceLastInteraction.
- **Relevance:** The structure allows for comprehensive univariate, bivariate, and inferential analysis to understand spending and engagement patterns.

# *Descriptive Statistics Summary*



## **Numerical Insights**

- Average Age: 49 years (SD = 18.2) → diverse customer age base
- Average Monthly Spend: \$332 (median = \$282) → right-skewed distribution
- Days Since Last Interaction: median 445 → many customers haven't engaged recently



## **Categorical Distributions**

- Gender: Balanced mix (Male 35%, Female 32%, Non-Binary 33%)
- Education: Masters (21%) most common
- Marital Status: Slightly more non-married (52%)

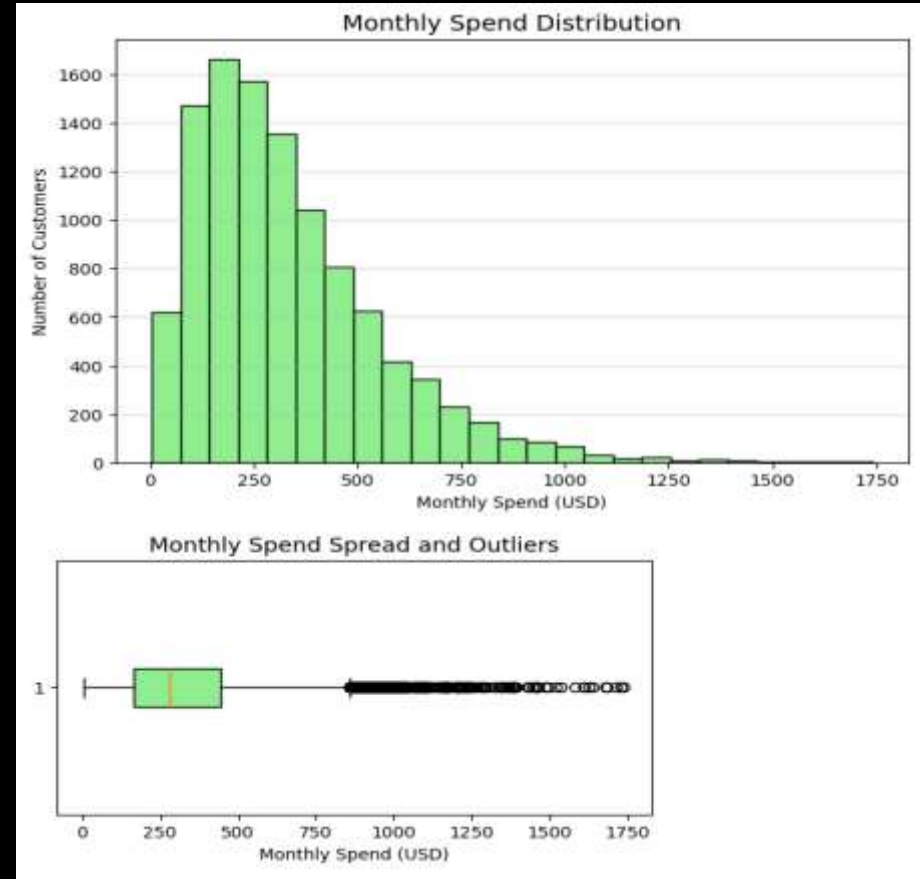


## **Statistical Characteristics**

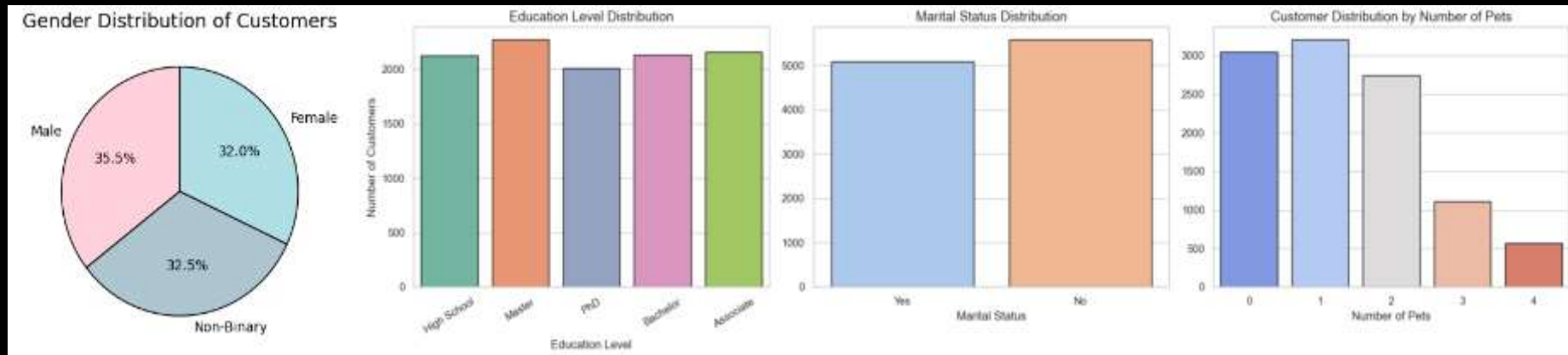
Right-skewed spend distribution suggests few premium spenders drive total revenue, while broad age dispersion ensures balanced engagement.

# Customer Age & Spending Patterns

- **Age Distribution:** The customer base spans from 18 to 80 years, with concentrations among young adults (18–25) and seniors (70–80), indicating a broad demographic reach.
- **Spending Behavior:** Monthly spend distribution is right-skewed, most customers spend under \$400, with a few high-value outliers exceeding \$1,000.
- **Interpretation:** Balanced age participation highlights inclusivity, while the presence of outlier spenders implies strong potential for premium-targeted offerings.



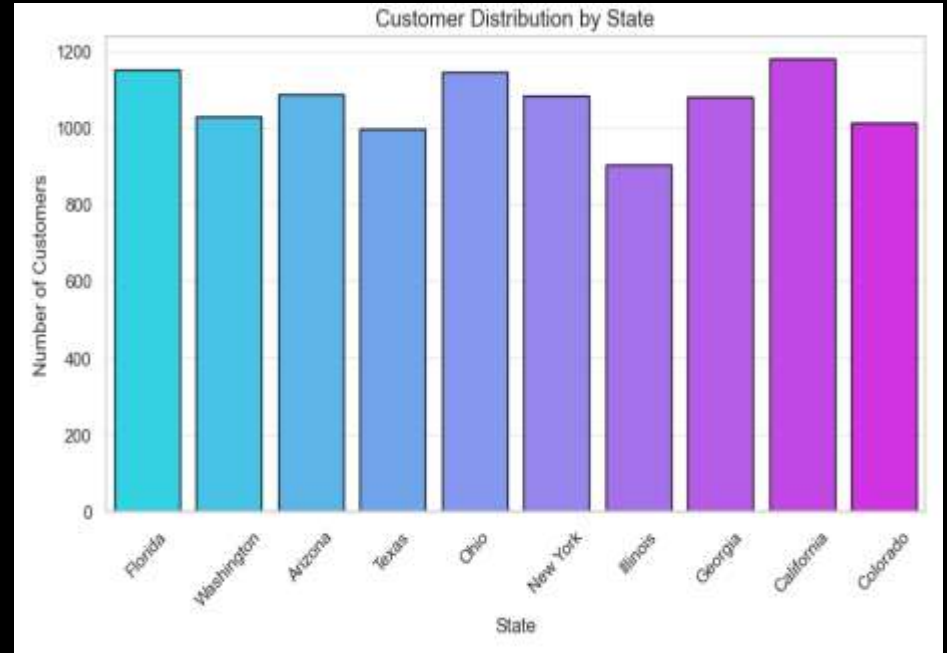
# Demographic Overview



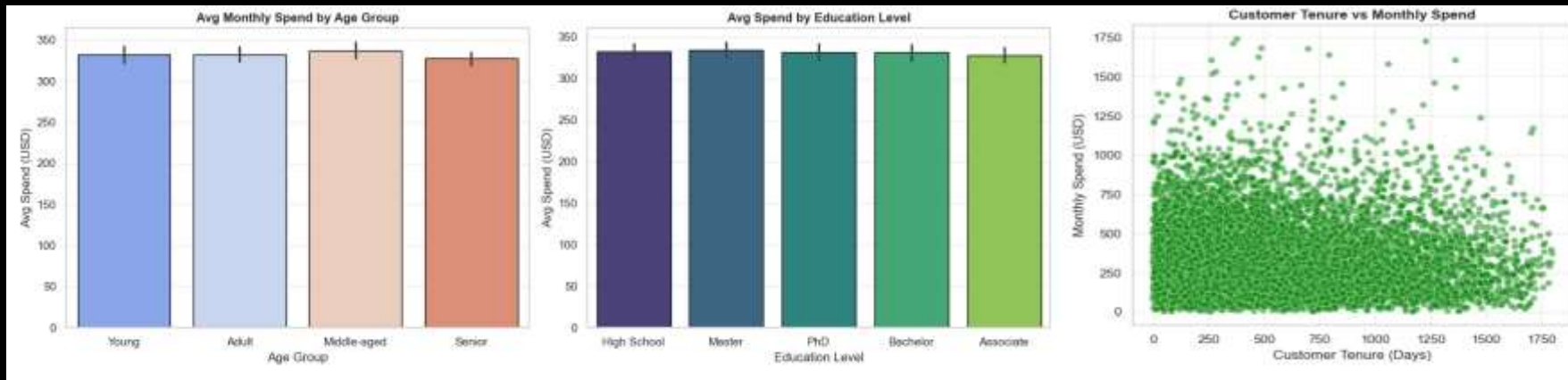
- **Gender Distribution:** Balanced representation: Male (35%), Non-Binary (33%), and Female (32%), indicating inclusivity across gender identities.
- **Education Levels:** Customers are well-educated, Masters (21%) and Associates (20%) dominate, with smaller PhD representation.
- **Marital Status & Pets:** Slightly more non-married individuals (52%), while married customers are more likely to own pets, an insight supported by Chi-Square testing.
- **Implication:** A demographically diverse and balanced customer base supports broad, lifestyle-driven marketing rather than segment-limited approaches.

## Geographic & State-Wise Insights

- **Customer Concentration:** California and Florida lead in customer count, reflecting strong brand presence and high potential revenue contribution.
- **Regional Variability:** States like Illinois and Washington show lower representation, potential targets for regional marketing and growth strategies.
- **Engagement Implications:** Balanced spending behavior across states indicates that offers and campaigns can be standardized nationally without significant regional bias.
- **Strategic Takeaway:** Focus on high-performing states for retention while launching targeted awareness campaigns in underrepresented markets.



# Linking Age, Education, and Tenure to Spending Patterns

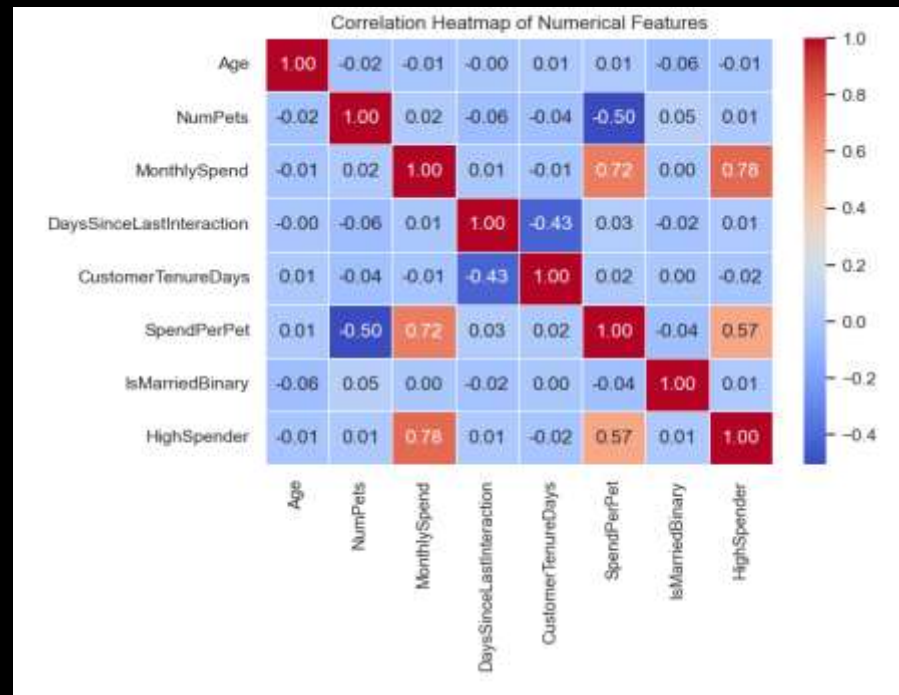


- **Spending by Age Group:** Middle-aged customers exhibit slightly higher average spending, but overall spend remains consistent across all age categories, indicating broad engagement.
- **Spending by Education Level:** Master's and PhD holders spend marginally more, though the difference is statistically insignificant, lifestyle appears to influence spend more than education.
- **Customer Tenure and Interaction:** No strong trend between customer tenure and spending. However, recent interactions correlate with higher spend, emphasizing engagement recency over longevity.
- **Strategic Implication:** Focus marketing automation on reactivation campaigns targeting customers with declining interaction frequency to sustain spend levels.



# Correlation Heatmap: Understanding Relationships Among Key Variables

- **High Positive Correlations:** Monthly Spend and High Spender (0.78), strong alignment validating spending-based segmentation. Spend per Pet also shows moderate correlation (0.57) with high-value customers.
- **Negative Correlations :** Days Since Last Interaction and Customer Tenure (-0.43), newer customers interact more frequently; engagement declines with tenure.
- **Behavioral Signals :** NumPets inversely relates to Spend per Pet (-0.50), suggesting diminishing marginal spend as pet count increases.
- **Interpretation :** Customer engagement recency is more predictive of spending than static demographics, emphasizing dynamic behavior modeling.



## *Hypothesis Testing Summary*

- **Gender vs Spending (T-Test):** P-value: 0.7345 → Fail to Reject  $H_0$  — No significant spending difference between males and females.
- **Education vs Spending (ANOVA):** P-value: 0.9224 → Fail to Reject  $H_0$  — Education level does not influence monthly spending.
- **Marital Status vs Pets (Chi-Square):** P-value: 0.000 → Reject  $H_0$  — Married customers are significantly more likely to own pets.
- **Age vs Activity & State vs Spending:** Both correlations non-significant (P-values > 0.34) — Engagement and spending uniform across age groups and states.

# *Business Insights & Strategic Recommendations*



## **Gender-Neutral Marketing**

Since spending behavior is consistent across genders, campaigns should focus on lifestyle and engagement triggers rather than demographics.



## **Pet-Oriented Offers for Married Customers**

Significant link between marital status and pet ownership , cross-sell pet or family-related products to married segments.



## **Universal Pricing & Engagement**

Education and state-wise uniformity suggest consistent pricing models and nationwide campaigns are effective.




## **Reactivation Focus**

Target customers with long inactivity periods using personalized engagement campaigns to sustain revenue.

## *Conclusion & Takeaways*

- **Behavioral Consistency:** Customer behavior remains largely consistent across gender, education, and geography, supporting unified engagement strategies.
- **Lifestyle Differentiators:** Marital status and pet ownership stand out as key lifestyle factors driving unique behavioral trends.
- **Engagement Over Demographics:** Recency of interaction is a stronger predictor of spending than age or tenure emphasizing proactive re-engagement.
- **Strategic Direction:** Focus marketing on behavioral segmentation, engagement-driven personalization, and family-oriented offers to maximize loyalty and ROI.

*Thank You !*

— Divya Shaw  
Data Analyst | Turning Data into Decisions  
 [linkedin.com/in/divya-shaw144/](https://www.linkedin.com/in/divya-shaw144/)