

# ALY 6980 CAPSTONE PROJECT

Group 3

**Pei-Yu Jheng, Shreeya Ambre, Sean Mclean, Theodore Smith**



A magnifying glass is positioned over a bar chart. The chart has three groups of bars labeled Q2, Q3, and Q4. Each group contains two bars, one blue and one green. The blue bars are consistently taller than the green bars across all quarters. The magnifying glass is centered over the Q3 bars, making them the most prominent.

# Introduction

---

- **Data Cleaning & Integration:** Ensure accurate and consistent datasets.
- **Streamlit Dashboard:** Track customer interactions efficiently.
- **Growth Optimization:** Prioritize high-potential opportunities.

# BUSINESS QUESTION

Utilizing past data, how can Big Sky improve leads and transactions by building stronger client relationships?

# Tools and Data To Use

---



R AND PYTHON FOR DATA CLEANUP



INTERACTIVE DASHBOARD FOR  
ASSESSMENT AND EDITS



COMPARISON OF DATASETS FOR INDUSTRY  
RESEARCH

# Data Cleaning Process

---

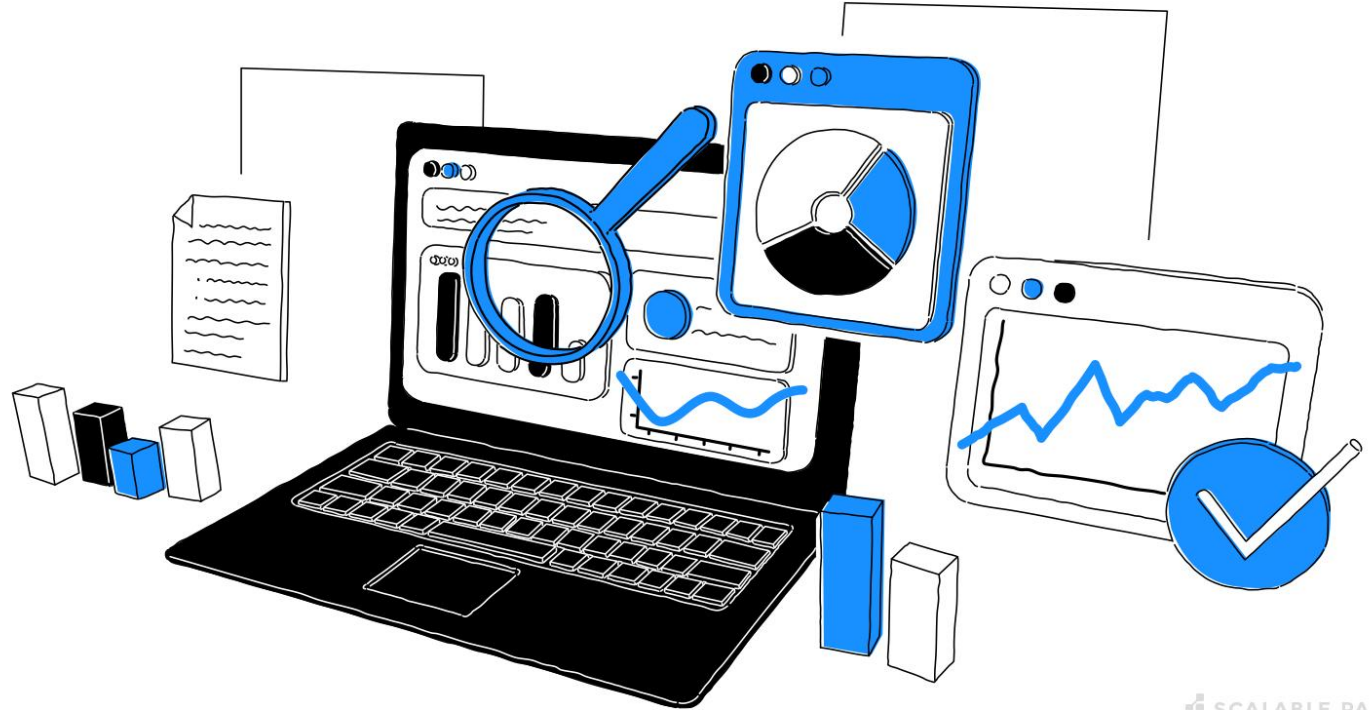
- Statistical Overview: Print table stats and identify missing values.
- Merging: Combine datasets and ensure smooth integration.
- Column Name Standardization: Convert to lowercase, replace spaces/dots with underscores.
- Duplicate Columns: Identify and distinguish duplicates.
- Unnecessary Columns: Remove irrelevant columns (e.g., id, "last enriched time").
- Finding Missing Values



# Exploratory Data Analysis

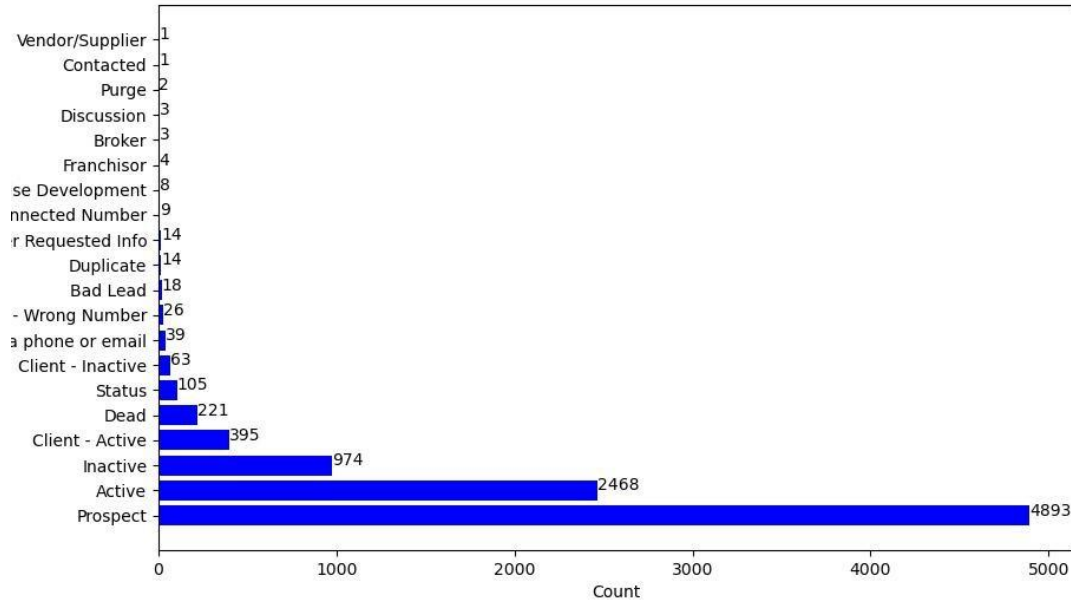
---

- **Identify Trends & Patterns:**  
Focus on key attributes to target potential clients.
- **External Datasets Comparison:** Evaluate similar external datasets for added insights.
- **Variable Relationships:**  
Examine correlations for business insights and addressing concerns.

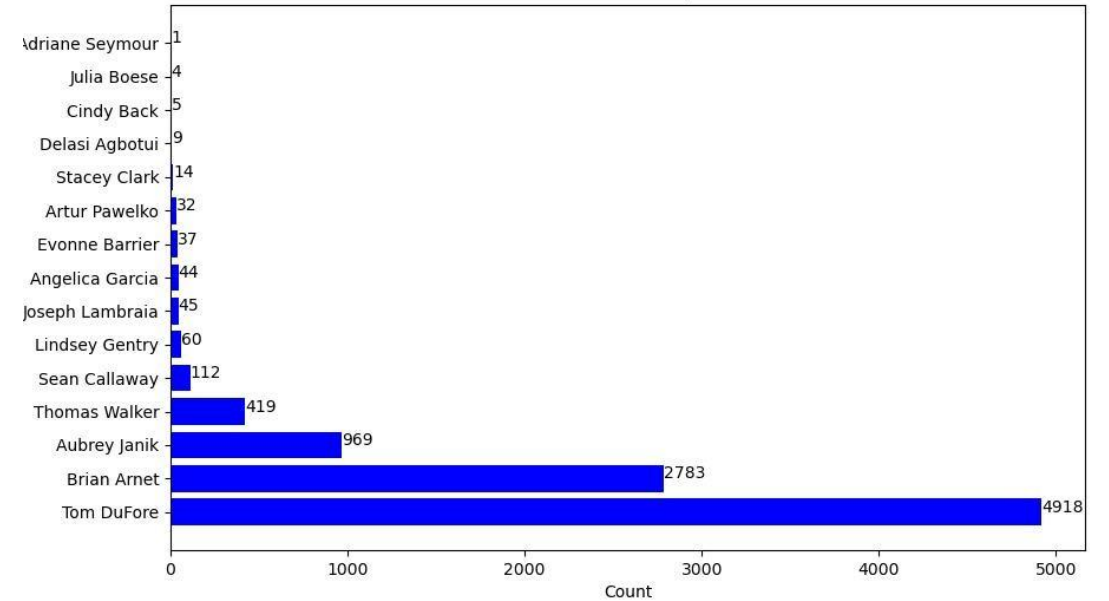


# EXPLORATORY DATA ANALYSIS (CONT.) - CLIENT ACCOUNTS

Count of Status

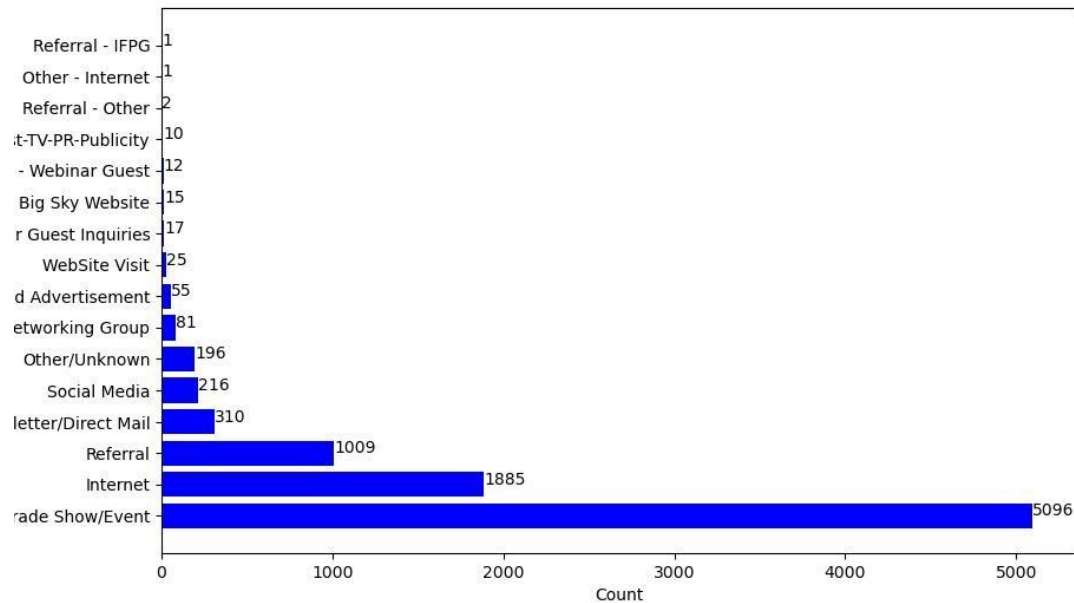


Count of Accounts by Owner

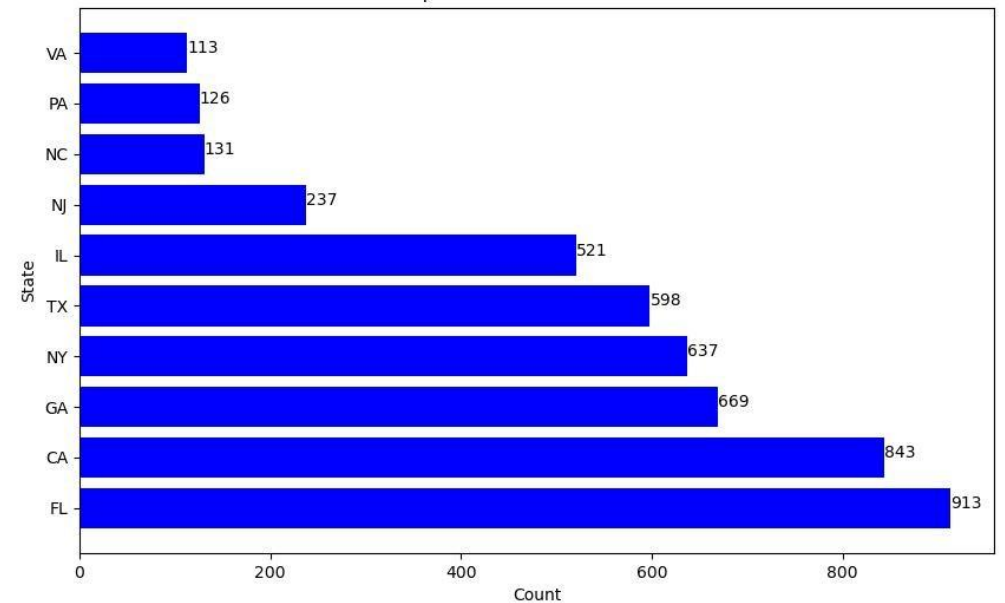


# EXPLORATORY DATA ANALYSIS (CONT.) - LEADS AND CONTACTS

Count of Lead Sources



Top 10 State Contact Count





# Key Questions & Issues to Address

---

- **Focus Areas for Client Acquisition:**
- Our analysis will highlight key areas to focus on for landing more clients.
- **Inquiry-to-Callback Time:**
- What is the average, minimum, and maximum number of days from inquiry to callback?
- **Business Relationship Stages:**
- Count the number of “probabilities” in the dataset, which may represent stages in business relationship development.
- **State/Region Impact:**
- Do certain states or regions influence the number of inquiries and probabilities in the dataset?



A decorative graphic on the left side of the slide, featuring a dark blue background with a dense cluster of overlapping circles in various colors including purple, green, blue, orange, and pink. The circles vary in size and opacity, creating a bokeh-like effect.

# Final Words/What To Expect

---

- Construction of customer management accounts.
- Data collection methods to gain important metrics.
- Creation of Interactive dashboard on data insights.
- Plan and solutions for long-term growth.

# Concluding Thoughts

---

- Opportunity to improve internal process.
- The implementation of data governance.
- Forward-focus plan for Big Sky.
- The rise of new leads.

