

ABSTRACT

Company Background – Leimberg Information Services, or LISI, is a niche content provider for Estate and Tax attorneys, CPAs, and financial professionals and companies. LISI offers newsletters, webinars, podcasts, courses, and other content to its customers. Historically, LISI has been run like a hobby. It is now under new ownership, and they want to expand profitably and sustainably. Due to poor data collection, and no real organizational structure, they didn't know where to start.

Overview of Solutions – We propose three methods to drive LISI's growth:

- 1) Develop an extensive podcast platform.
- 2) Expand the reach of their newly created product The Institute.
- 3) Build a networking platform based on the expertise of LISI's subscribers.

INTRODUCTION

Exploratory Analysis was first carried out on multiple datasets to gain a better understanding of LISI's customers. The findings were used to further analysis and form recommendations for growth.

All Contacts - The dataset consisting of a growing email list of relevant professionals, including newsletter subscribers and non-subscribers. However, the current email cadence is not effectively using this feature to showcase LISI's benefits. The dataset contains over 80,000 records with over 50% of customers who have opted out, and a significant proportion of opt-out reasons are unsubstantial or negative.

Subscribers - This dataset with newsletter subscribers was used to analyze growth and locations. LISI had a large spike from 2018 to 2021 but then fewer people joined in 2022 causing the spike to lessen. For the location analysis, 7785 unique subscribers were identified based on subscriber email found all around the United States. The top states include California, New York, and Florida with many other customers in-between the different coasts showing a strong foundation.

Webinars and Other Products - Data contains information other products and services had over 2000 items offered that were purchased globally. After the United States, LISI has a growing international presence. Overall, the top items purchased were LISI's Tool Access, PP Loan Tracker, and Bob Keebler Webinars.

Limitations of Current Data - The Team was able to quickly gather insights for basic analysis but if LISI wants to truly connect with their subscribers, LISI needs to emphasize better data collection and quality. The following are some apparent flaws with the current datasets that hinder decision making and growth.

- 1) Excessive free-form fields reveal lack of data controls.
- 2) Redundant data provides no new information.
- 3) The current data fields only generate surface-level analytics.

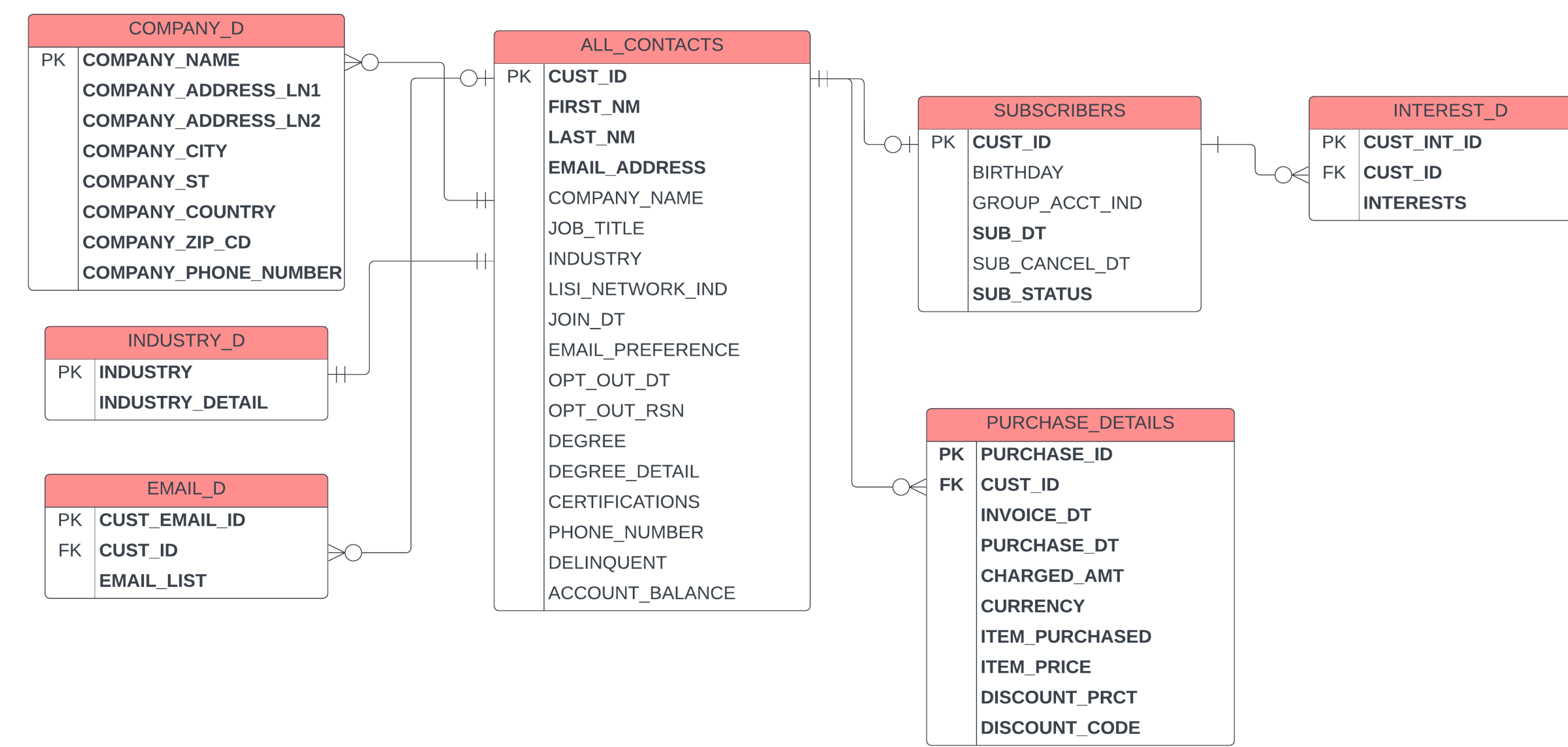
METHODS AND MATERIALS

For further analysis, we identified more methods and materials to aid LISI in better data collection and data analysis.

- LucidCharts and Excel – Data Dictionaries and Entity Relationship Diagram (ERD)
- Tableau/Tableau Prep and Excel - Data Cleaning and Visualization
- ChatGPT – Google Sheets and Excel Extension for Text Generation
- Python – Unsupervised K-means Clustering Algorithm with Term Frequency Inverse Document Frequency (TF-IDF) Vectorizer for Industry Analysis

RESULTS

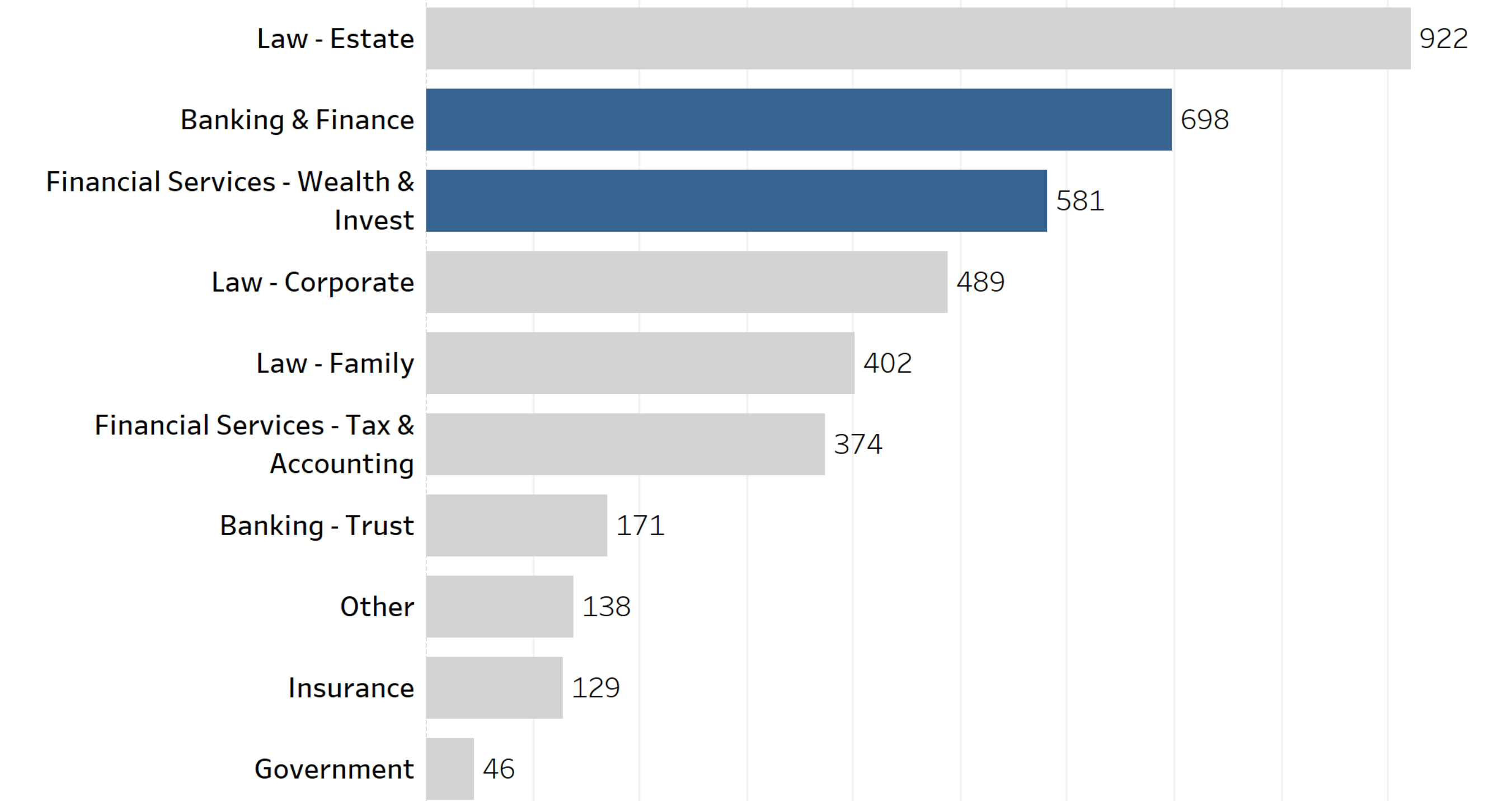
Data Modeling and Design - Since data collection can be very costly if done wrong, we summarized all current data into data dictionaries so that we can note major flaws such as excessive redundancy, poor usability, and a lack of integrity. To help mitigate, we created new dictionaries and an ERD of potential future-state datasets containing a richer combination of data elements. The ERD below shows a starting point for LISI's future data consisting of three main tables containing unique IDs for appropriate table joins, and four sub-tables containing more detailed information regarding more granular customer information. All tables in this recommended ERD have IDs and keys that assist in proper joins to allow for ease of reporting and analytics.



Industry Analysis - LISI appeals to other industries complementary to estate and tax law which can be areas of growth. However, only company name is available which can be difficult to analyze. Therefore, categories needed to be created to be able to assign each customer to an industry based on the company name. After generating text descriptions of the company names with ChatGPT, we carried out NLP methods of preprocessing the text, running it through a TF-IDF vectorizer, and feeding the text into a K-means clustering algorithm, k=10. We utilized the vectorizer to extract features from the clusters of the top 10 terms prevalent to help name the clusters and see the industry groupings. Sub industries were also identified from the clusters to note specializations of the customers within their industry. This information and machine learning process will allow LISI to continue to analyze customer characteristics and build other products and services as they continue to grow and collect more data.

Top Terms In Clusters 0-3			
Cluster 0: Law - Estate	Cluster 1: Fin Serv - Wealth & Invest	Cluster 2: Banking - Trust	Cluster 3: Insurance
law	financial	trust	insurance
legal	management	management	life
firm	wealth	wealth	mutual
specializing	investment	services	financial
estate	services	company	northwestern
areas	planning	financial	services
business	company	institutions	businesses
providing	individuals	families	company
litigation	families	provides	annuities
clients	specializes	individuals	individuals

Companies in Each Cluster



DISCUSSION & CONCLUSIONS

Revise Current Data Structure for Better Data Governance- Data is representative of credible decision making. Without strong data governance, LISI puts themselves at risk for privacy breaches, data misuse, higher data management costs, poor data quality, and, ultimately, a loss in competitive advantage.

Use Customer Surveys to Reveal Future Growth Opportunities - The attitude of how people react and interact with products/services, as well as those who choose to do so, can very quickly reveal future growth opportunities. Use surveys to assess current operating performance and plan for future improvement.

Better Customer Email Communications - To avoid displeasing customers, LISI should enhance the email delivery and subscription system, particularly for customers who do not opt for any subscription options and receive all emails. Additionally, it is important to format and accurately record useful client information during data collection, based on our future data recommendations, which can facilitate efficient response to client requests.

Expand The Institute - Prioritize program materials in areas generating the most interest based on customer dataset analysis and industry analysis. Consider offering certificates or credentials globally to individuals interested in the LISI Institute to expand the company's reach. Provide professional continuing education credits such as CLE (Law) and CFP (Finance) credits to retain existing clients and potentially attract those who have discontinued their participation in the program.

LISI Listens – Podcasts - The foundation of LISI Listens is existing content, but we expect that experts will continue to contribute to LISI's knowledge base. LISI can convert old webinar files to audio form and potentially different languages to provide more convenient content to attract more professionals. Content can also be guided by the top industries among customers. More importantly, podcasts will benefit from more data collection following the ERD model and survey results for deeper analysis to be done.

LISI Network – Internal Network of LISI Professionals - New contact sharing service for individual and company customers to request or be matched to freelance work, job positions, expert witness testimony, and more with LISI answering inquiries with applicable candidates. The ERD model and data governance and collection is needed to filter on specific characteristics for matches LISI does not currently have such as industry, degree, and certificates. Methods and materials from the industry analysis will help maintain updated information for this service that will provide another revenue stream for LISI.