Customers' Surveys Analysis

Survey analysis is the process of turning the raw material of your survey data into insights and answers you can use to improve things for your business. It's an essential part of doing survey-based research. Fig 1 is the simplest example of analysis, "mostly people associated family with actual meaning of life". This is how we extract insights from raw data.

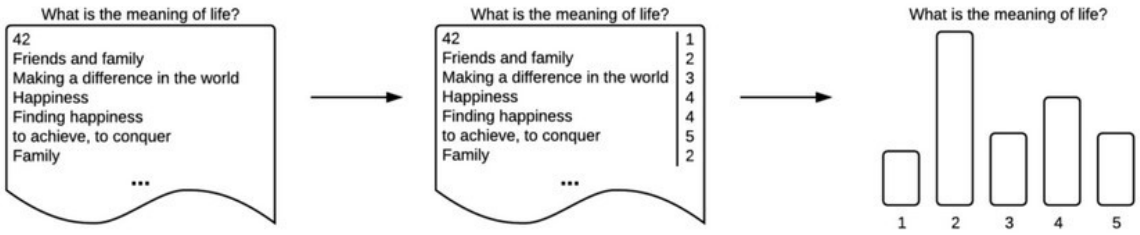

Figure 1: Simple Survey Analysis Example

Types of Questions in Interviews/ Surveys

- **True/ False/ Fixed Options**
  - Questions such as what is your gender, color, degree etc. are the fixed options questions. These questions are important for grouping and segregation such as 70% of interviewees were male and 30% were female. Or Most of the interviewees were of Masters in Computer Science etc.
  - Analysis of such questions is simple, just need Pandas Knowledge and PowerBI for storytelling.
- **Quantitative**
  - Some questions might have different values, but they are quantitative in nature such as age of interviewee, experience in years, salary etc. These variables are can also be treated as fixed option but with different visuals.
  - Analysis of these are also straight forward.
- **Close Ended**
  - Some questions are close ended, such as responses having some format such as "love to read books", "sometimes play games". Now these are phrases, we need to extract the intent from such sentences such as in first phrases, LOVE is about the object BOOK.
  - Data pre-processing techniques of NLP (POS Tags, NER), Python functions
- **Open Ended**
  - These have no structures, interviewees are open to explain. No restriction on length, sentences etc. Questions such as explain this phenomenon.
  - Advance Data pre-processing techniques of NLP (POS Tags, NER), Python functions
  - Topic Modeling required
  - Machine learning techniques such as text classification, text clustering etc.

What a survey analysis should include?

- Take a look at your top survey questions.
- Determine sample size.
- Use cross tabulation to filter your results.
- Benchmarking, trending, and comparative data.
- Crunch the numbers.
- Draw conclusions/ suggestions/ recommendations.

Is survey analysis qualitative or quantitative?

- A survey can be qualitative, quantitative or mix methods. If your survey involves a questionnaire with scalable answers then it is a quantitative survey. If your survey has descriptive questions with in-depth answers then it is a qualitative survey. If your survey has both of them then it is a mixed-method survey.

Proposed Methodology

From a high level perspective (as shown in fig 2), we have customers' interviews in different domains such as retail, MFG, utilities etc. We have to apply following steps on the data:-

1) **Data Pre-processing/ Normalization**
   a. We have different surveys/ interviews like year wise customers' interviews, department wise interviews, designation wise interviews etc. We need all these interviews in some common form. There can be smaller change in column names such as GENDER in one file, but Gender in other file and Male/ Female in some other files. We need all the data to have in common structure for analysis. We can also perform some pre-processing such as removal of unnecessary text.
2) **Data Merging Pipe Line**
   a. One all the files/ interviews became same in structure, we will combine them. The combine form will be then used for data analysis.
3) **Analysis**
   a. We will apply data analysis including both basic and advance analysis as per the requirements and person who will use the system. We will different visualization charts for different variables. Frequent questions, frequent terms, most frequent responses, group wise categories, gender wise groups etc.
4) **Reporting**
   a. The last step include reporting, which is storytelling, insight extraction, recommendation and suggestions in some user friendly way.

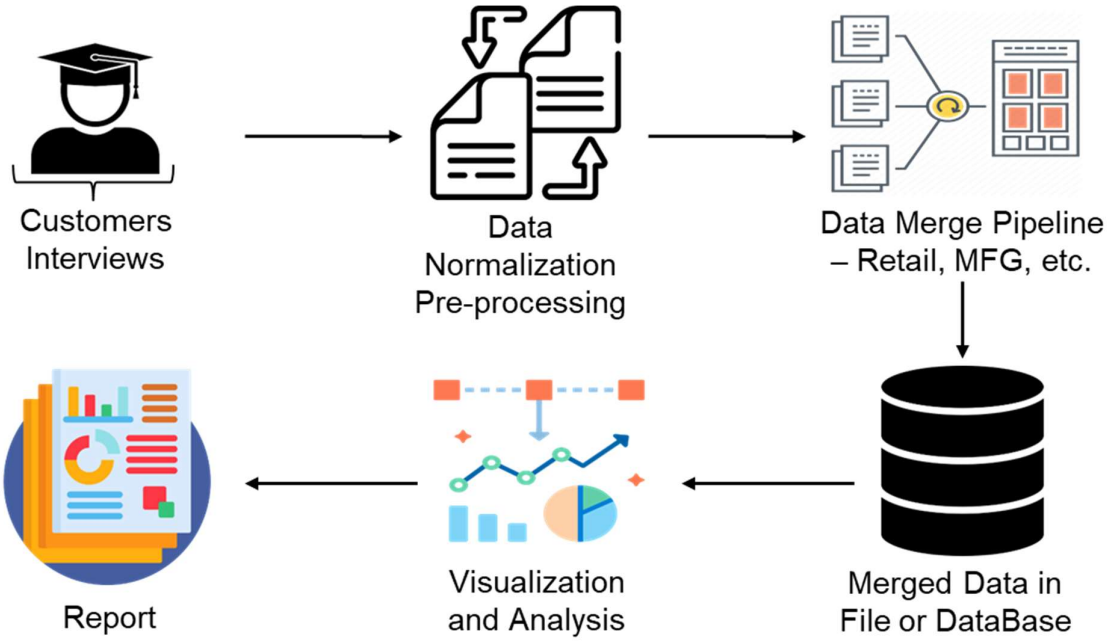

Figure 2: Proposed Methodology High Level Perspective

Fig 2 shows the detailed/ advanced analysis of the interviews. We will leverage the NLP techniques such as N-Grams (most frequent terms and most important terms). We will do following type of analysis.

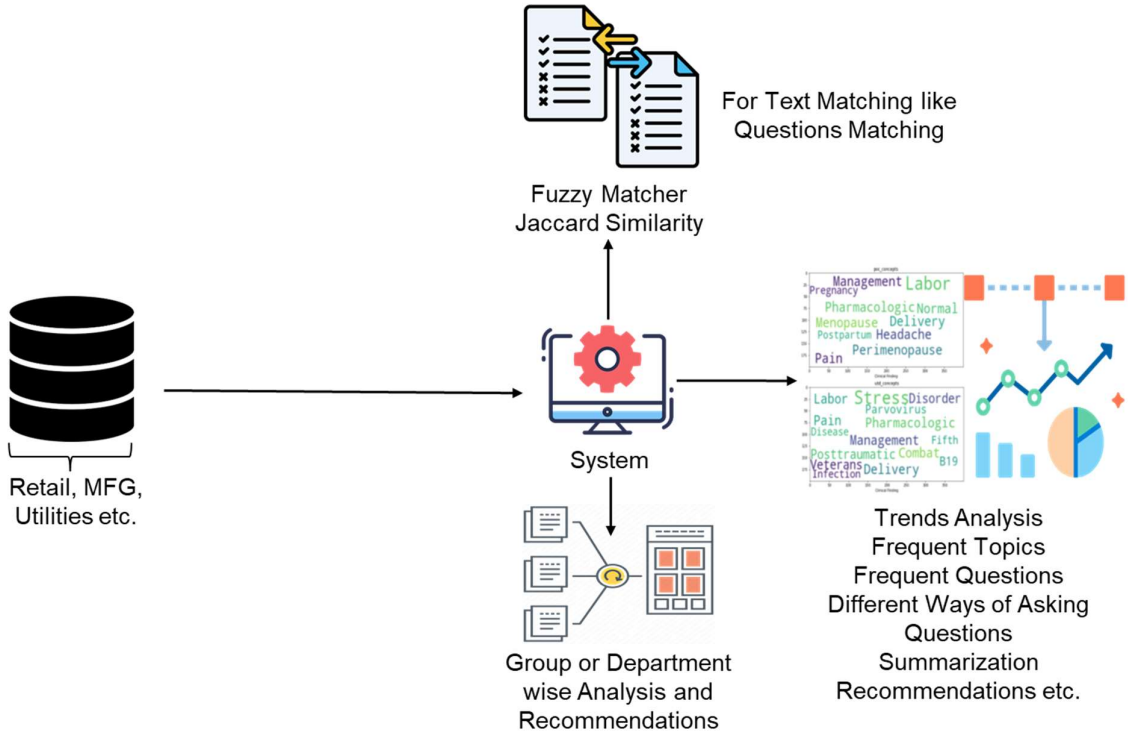

Figure 3: Detailed/ Advanced Analysis of Interviews

- **Group Wise Thought Analysis.** Age wise responses, gender wise responses, department wise responses, designation wise responses using word cloud.
- **Frequent Terms/ Important Terms.** Most frequent terms discussed overall, most discussed terms group wise using N-Grams/ Frequency Counter.
- **Topic Modeling using Machine Learning**. Use of automated machine learning/ statistical based approach for extraction of topics.
- **Text Clustering**. Grouping based on responses using unsupervised machine learning technique such as K-Means clustering.
- **Sentiment Analysis**. Do sentiment analysis of customers' responses.
- **User Friendly Dashboard**. Last but not the least, we need to show everything in a user friendly dashboard as shown in fig 4 (Power BI Dashboard).

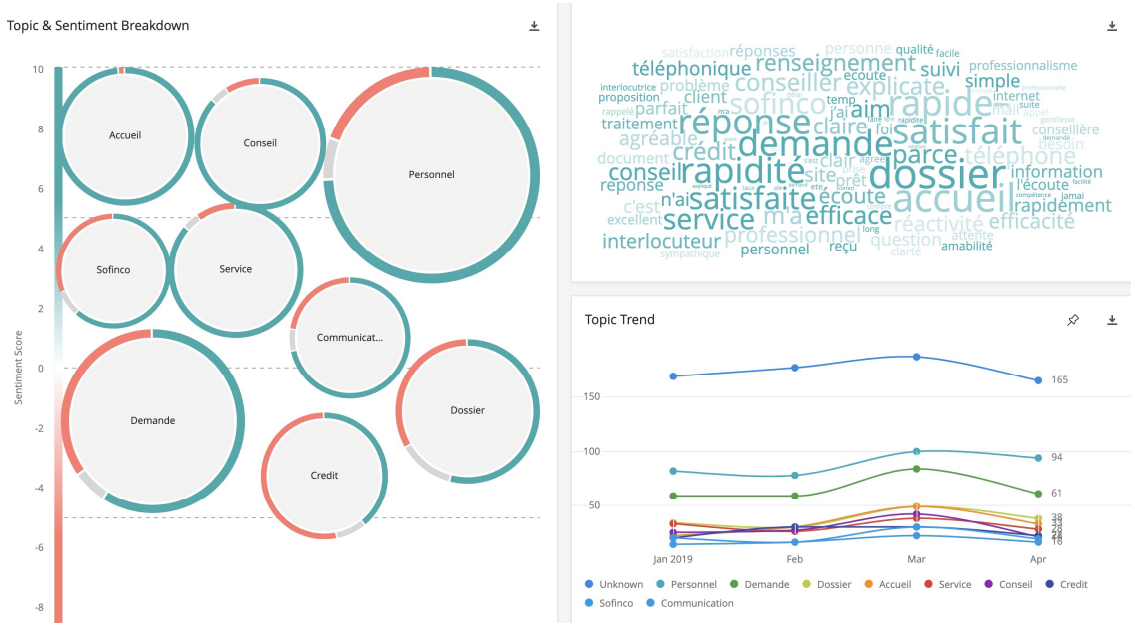

Figure 4: Power BI Dashboard Example

Technical Requirements

- **Python → (Python, Pandas library, Numpy manipulation, matplotlib, seaborn)**
- **Visualization → (Power BI, matplotlib, seaborn)**
- **Machine Learning → (K-Means Clustering and LDA for topic modeling)**
- **NLP → (NLTK, N-Grams, Spacy, NER)**
- **Sentiment Analysis → (VADER/ might need some customization)**