UNIVERSITY OF CONNECTICUT

OPIM 5671- DATA MINING AND BUSINESS INTELLIGENCE

Prof. Sudip Bhattacharjee

Text Mining Project Report

“McDigest: Mining McDonald's Food Reviews”

Group 8

Satakshi Deshmukh

Hema Pradhiksha Dhanapal

Rohit Kamineni

Sri Kavya Reddy Narra

Sravya Machavarapu

Table of Contents

SECTION 1: Executive Summary

SECTION 2: Project Introduction

SECTION 3: Data Analysis, Cleaning & Preparation

SECTION 4: Modeling

SECTION 5: Business Recommendations and Conclusion

SECTION 6: References

SECTION 1: Executive Summary

Our objective was to utilize SAS Enterprise Miner to derive actionable business insights from McDonald's reviews, enabling the company to target specific areas for improvement. This project utilized a dataset comprising over 33,000 anonymous reviews of McDonald's outlets across the United States, sourced from Google reviews, for a comprehensive analysis of customer feedback.

The dataset was then subjected to a variety of data mining techniques such as parsing, clustering, regression, and decision trees. These techniques were useful in understanding category-wise aspects that are hindering McDonald's business growth.

The project demonstrated the efficacy of SAS Enterprise Miner in extracting actionable insights from large customer review datasets. The findings from the analysis could effectively guide various business decisions, ultimately enhancing overall customer satisfaction.

SECTION 2: Project Introduction

In the competitive realm of the fast-food industry, the essence of customer satisfaction and understanding their feedback is paramount for a giant like McDonald's to maintain and boost customer loyalty and operational efficiency. The project, centered on text mining, aims to meticulously analyze over 33,000 anonymous reviews from McDonald's stores across the United States, gathered from Google reviews. By harnessing advanced data analysis techniques, including geographic assessments, extracting categories from textual reviews, and predictive modeling, this initiative seeks to delve into the depths of customer feedback. The primary goal is to extract meaningful insights regarding customer sentiment and discern factors that influence customer satisfaction. Such a comprehensive analysis is designed to spotlight potential areas for operational enhancements, refine customer service strategies, and identify successful attributes of stores that can be replicated across the franchise network.

The significance of this endeavor lies in its capacity to leverage text mining to transform vast amounts of unstructured customer feedback into actionable insights. This approach not only aids in identifying common grievances from customers but also in understanding how these perceptions might vary across different geographic locations. Such insights are valuable for strategic decision-making, enabling McDonald's to fine-tune its operations, tailor marketing strategies more effectively, and ultimately, enrich the customer dining experience. By focusing on the feedback provided by customers, McDonald's can align its service offerings more closely with customer expectations, leading to enhanced brand loyalty, operational efficiencies, and improved market positioning. This project represents a strategic endeavor to utilize text mining technologies to transform raw data into a strategic asset, thereby assisting McDonald's in staying ahead in the fiercely competitive fast-food industry.

BUSINESS PROBLEM

The business problem focuses on the need for McDonald's to deeply understand customer experiences, preferences, and points of satisfaction and dissatisfaction across its franchise network in the United States. In the highly competitive fast-food industry, maintaining customer loyalty and operational excellence is essential for sustained success.

The core challenge lies in extracting meaningful insights from a vast pool of unstructured text data to identify key themes, sentiment trends, and predictive factors that can influence customer satisfaction. These insights are critical for McDonald's to pinpoint specific areas for improvement in service, operations, and overall customer experience. Additionally, understanding how customer feedback varies by location can help McDonald's tailor its strategies to meet regional preferences and address timely concerns, ensuring a more personalized and responsive approach to customer service.

Finally, based on the insights gained from the comprehensive feedback analysis and the understanding of regional customer preferences and pain points, it's important to develop tailored marketing and operational strategies. So, to summarize, the three business challenges can be articulated as follows:

- Comprehensive Feedback Analysis
- Understanding Regional Customer Preferences and Pain Points
- Tailored Marketing and Operational Strategies

By solving this business problem, McDonald's hopes to enhance its operational strategies, improve customer satisfaction, and replicate successful practices across its stores, ultimately leading to increased brand loyalty, operational efficiency, and market competitiveness.

DATA

Data Source

Our group thoroughly investigated numerous sources to find datasets spanning a wide range of fields. Through our deliberations, we assessed each dataset's advantages and drawbacks, integrating a variety of perspectives to achieve a collective agreement. Following a thorough evaluation, we selected the McDonald's Store reviews dataset for our project. Embracing a collaborative and comprehensive decision-making process enabled us to select a dataset that aligned with the needs and aspirations of all team members.

(The dataset can be accessed here.)

Data Description

Initially, the dataset contained over 33,000 records with 10 columns.

The data set provided valuable insights into customer experiences and opinions about various McDonald's locations nationwide. The dataset includes information such as store names, categories, addresses, geographic coordinates, review ratings, review texts, and timestamps.

The Key features are :

reviewer_id: Unique identifier for each reviewer (anonymized)

store_name: Name of the McDonald's store

category: Category or type of the store

store_address: Address of the store

latitude: Latitude coordinate of the store's location

longitude: Longitude coordinate of the store's location

rating_count: Number of ratings/reviews for the store

review_time: Timestamp of the review

review: Textual content of the review

rating: Rating provided by the reviewer

Project Goal

Our goal from analyzing the McDonald's reviews dataset is to gain deep insights into customer satisfaction and operational areas that can be optimized for a better dining experience. By examining the unique identifiers, store characteristics, and geographical data, we can understand patterns of customer preference and performance variations across different locations. Analyzing textual reviews alongside ratings allows us to identify common themes of customer praise or dissatisfaction, enabling targeted improvements in service quality, menu options, and store environment.

SECTION 3: Data Analysis, Cleaning & Preparation

Data Analysis

The project aimed to delve into a large collection of customer feedback, seeking vital insights into their sentiments and perspectives.

We have observed that the reviews originated from various states, indicating a wide geographic spread of feedback. Additionally, it became clear that the reviews could generally be categorized into distinct themes such as service quality, food quality, and others.

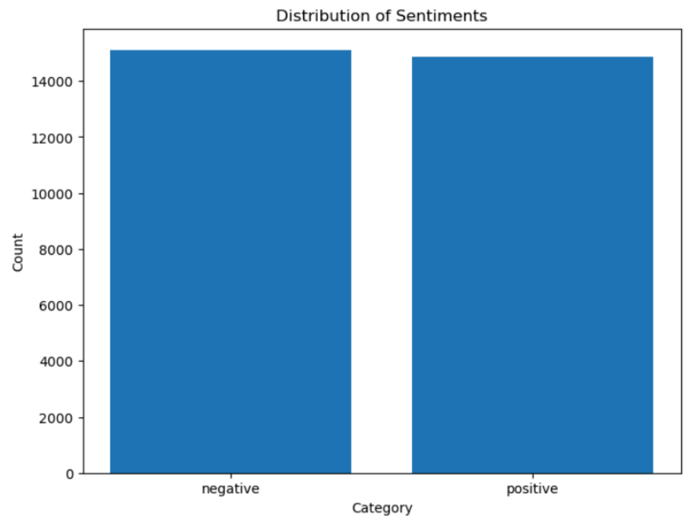
Moreover, a deeper dive into the textual content of the reviews provided insights into recurring customer concerns and praises, such as wait times, cleanliness of the premises, staff friendliness, and menu diversity.

This initial phase of data analysis not only highlighted the diverse sources and categories of feedback but also underscored the importance of a nuanced approach to understanding customer experiences across different dimensions. By aggregating and analyzing this rich dataset, we aim to uncover actionable insights that can drive targeted improvements in service and food quality, tailored to regional preferences and temporal factors.

Data Cleaning & Preparation

The preliminary data purification was conducted using JMP software, where entries with missing values and nonsensical text were removed. After this initial cleaning, the dataset comprised approximately 29,969 entries.

To distinguish between positive and negative feedback, the rating column was split into two separate groups - reviews rated between 1 and 3 stars were classified as negative, while those with a 4 or 5-star rating were deemed positive in **SAS Studio**. This method was selected due to its outstanding ability to identify the elements affecting reviews positively and negatively.



The above image depicts the distribution of sentiments (positive and negative) columns which we have created based on ratings.

Upon completing the initial review, we decided to conduct a **regional analysis** across the US, enabling the company to concentrate on insights specific to different regions. To achieve this, we grouped the data by state using Excel & SAS, organizing the information to reflect region-specific trends.

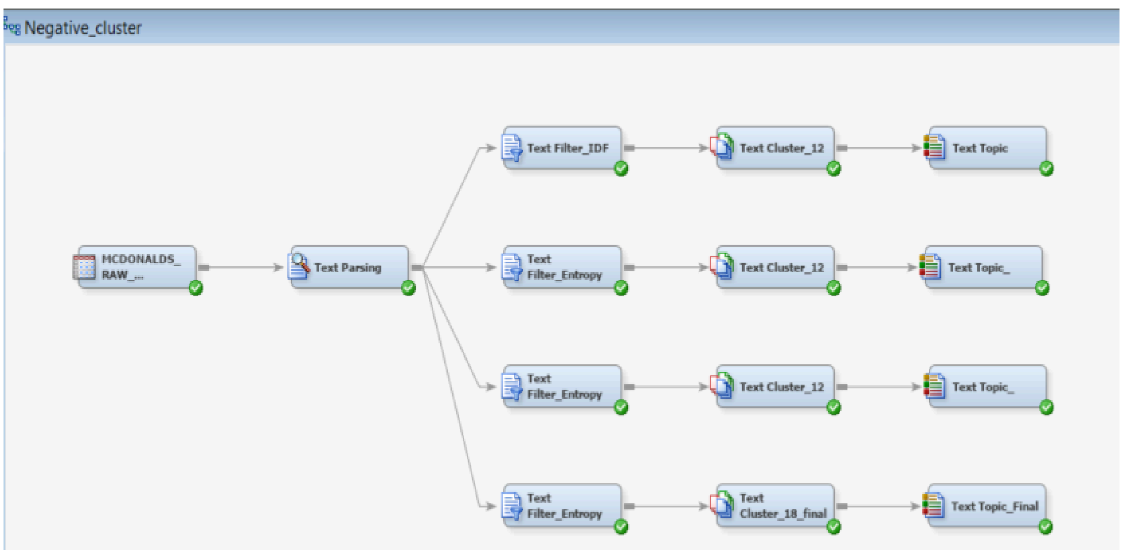
In essence, our approach to data cleaning and organization played a pivotal role in ensuring our dataset was tidy and suitable for subsequent analysis.

SECTION 4: Modeling

Initially, our objective was to separately examine the positive and negative feedback reviews before venturing into regional modeling. This step was crucial to gain a comprehensive understanding of the primary factors contributing distinctly to both negative and positive reviews. Consequently, utilizing the partitioned data from SAS Studio, categorized solely based on the nature of the reviews as positive or negative—we proceeded with the modeling phase. This methodical approach ensured that we could accurately identify and address the specific aspects influencing customer experiences across different spectrums before delving into region-specific analysis.

Negative Feedback Modeling

MODELING DIAGRAM OF SAS ENTERPRISE MINER

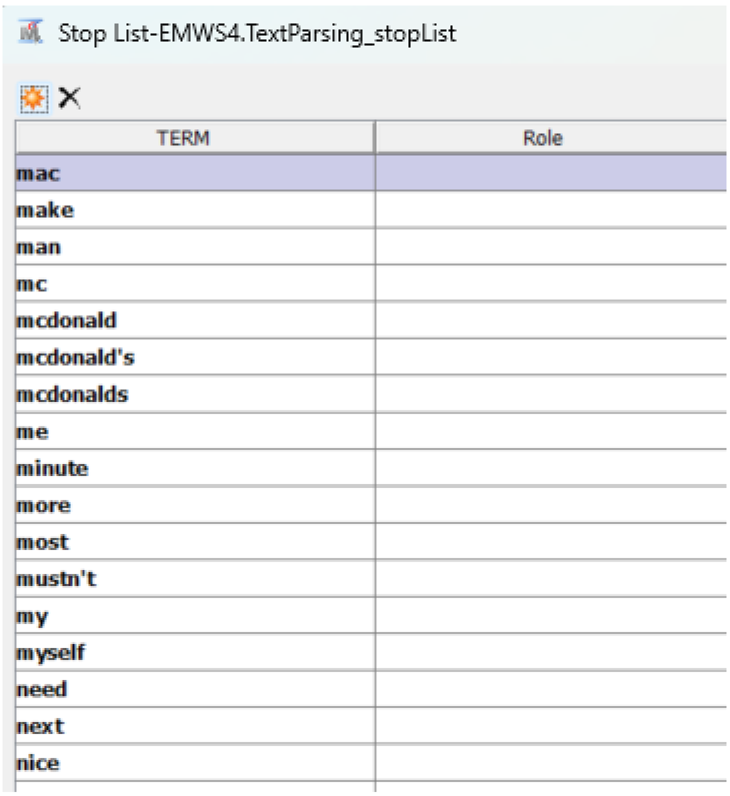


Our modeling started with the selection of the **data source** titled- “MCDONALDS_RAW_NEGATIVE”

Initially, the **Text Parsing node** was executed using its default settings:

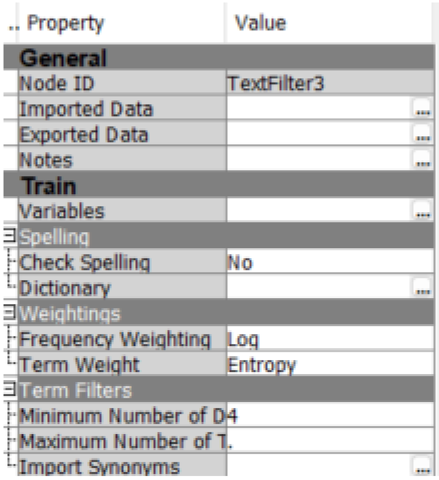
Property	Value
General	
Node ID	TextParsing2
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Parse	
Parse Variable	
Language	English

Then we had redundant words like McDonald’s, good, bad, worse, and service that prevented better clusters and topics from being formed. Hence, further, we started to create a stop list for the text parsing node, followed by a progression to the **Text Filter** node.

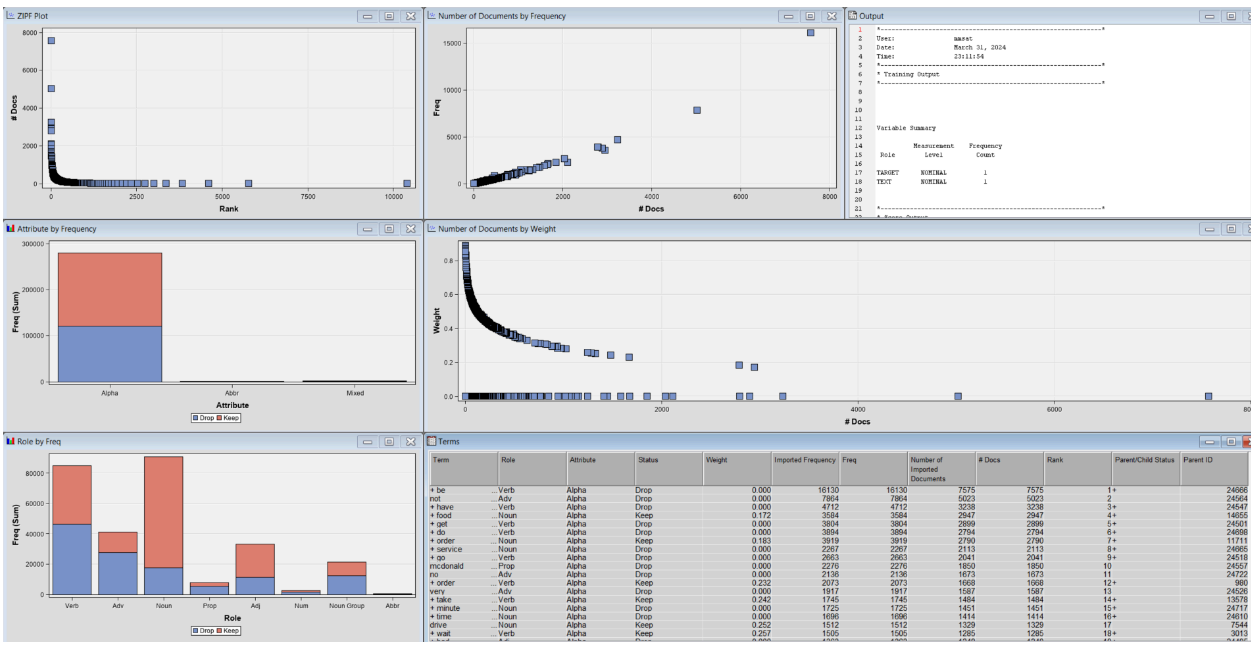


Subsequently, adjustments were made to set the frequency weights to logarithmic, while also exploring various term weighting methods, including Entropy and Inverse

Document Frequency (IDF), to optimize our analysis. Then we were able to get better results using Log and Entropy for all of our models.



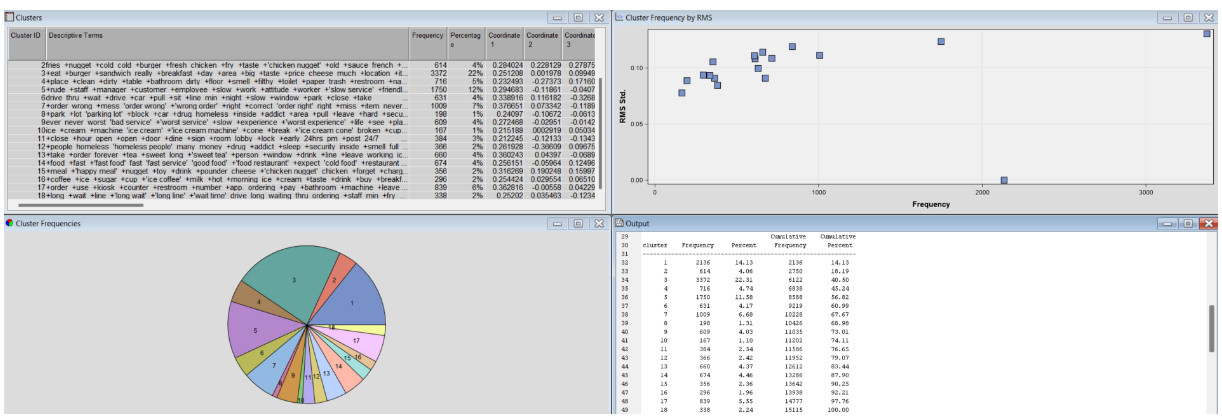
Property	Value
General	
Node ID	TextFilter3
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Spelling	
Check Spelling	No
Dictionary	...
Weightings	
Frequency Weighting	Log
Term Weight	Entropy
Term Filters	
Minimum Number of D	4
Maximum Number of T	.
Import Synonyms	...



Subsequently, a **Text Clustering node** was executed utilizing its default parameters. We further refined our approach by adjusting the Singular Value Decomposition (SVD) resolution to low, medium, and high settings while keeping the SVD dimensions at their default values. Our experimentation revealed that the optimal Text Cluster results were achieved when the SVD resolution was set to low.

Following that, we have also explored varying the number of clusters to enhance the precision of our topic-wise analysis.

Property	Value
General	
Node ID	TextCluster2
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Transform	
SVD Resolution	Low
Max SVD Dimensions	100
Cluster	
Exact or Maximum Numb	Exact
Number of Clusters	18
Cluster Algorithm	Expectation-Maximizatio
Descriptive Terms	15



Subsequently, we executed the **Text Topic Node**, adjusting the number of multi-term topics to 21.

Property	Value
General	
Node ID	TextTopic4
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
User Topics	...
Term Topics	
Number of Single-term T	0
Learned Topics	
Number of Multi-term T	21
Correlated Topics	No

Following the execution of the Text Topic Node, we evaluated the words and categorized them into similar topic names, such as Cleanliness, Cold/State

Food, etc., as illustrated in the image below.

Additionally, we adjusted the Term Cutoff settings, which aided in removing redundant and irrelevant words from the identified topics

File Edit

Topics

Topic	Category	Term Cutoff	Document Cutoff	Number of Terms	# Docs
Cleanliness	User	**0.054**	**0.0646666667**	25	2404
Cold/Stale Food	User	**0.10175**	**0.072**	14	3286
Disappointed/Neutral	User	**0.017**	**0.07925**	196	2346
Drive Through	User	**0.024**	**0.0685714286**	141	2582
Employee Attitude	User	**0.0185**	**0.0715**	211	2052
Machine Malfunction	User	**0.0625**	**0.0635**	20	947
Meal Combo	User	**0.0213333333**	**0.0686666667**	187	1596
Orders	User	**0.022125**	**0.0755**	160	2213
Security	User	**0.018**	**0.069**	70	1088

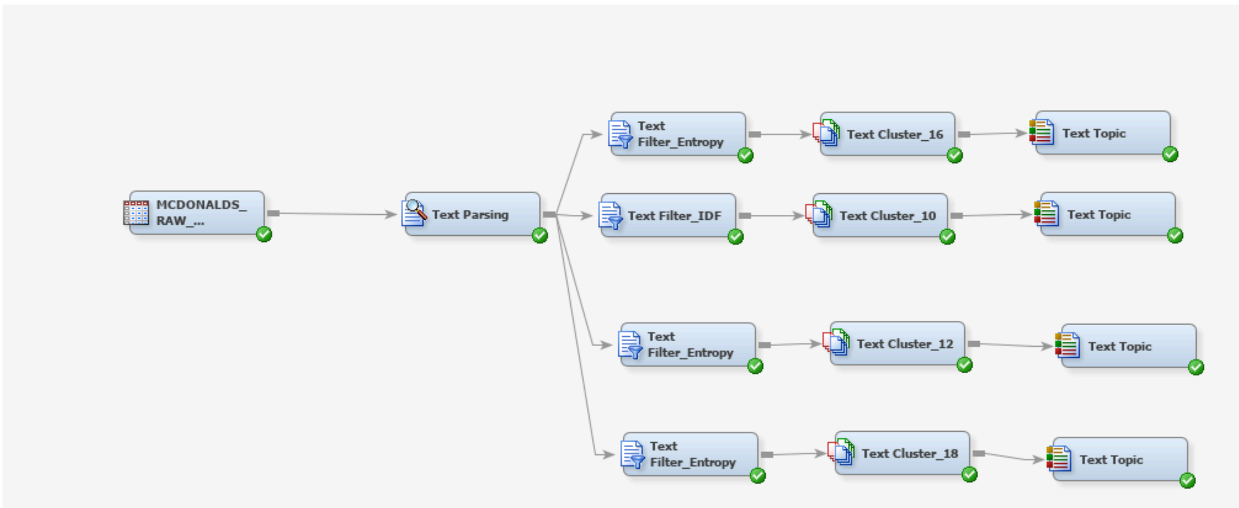
Furthermore, we noticed that certain similar words appeared across multiple topics. To achieve more focused and specific topic categorization, we added these redundant terms to the stop list in the Text Cluster node.

Insights

In our analysis, we have identified that clusters 3 and 15 collectively contribute to a topic concerning meal combos, accounting for 24% of the feedback. Furthermore, clusters 14 and 2 are indicative of issues related to cold or stale food. Clusters 10 and 16 highlight problems with machine malfunctions. Beyond these specific issues, there are additional concerns to note. Notably, security concerns have arisen due to the presence of homeless individuals, as evidenced by complaints of drug smells, and issues in the restrooms and parking lot, which are represented in cluster 12. Moreover, cluster 4, accounting for 4% of the feedback, delineates issues related to the cleanliness of the restrooms, with descriptions of filthy conditions, unpleasant odors, and dirty floors littered with paper.

Positive Feedback Modeling

MODELING DIAGRAM OF SAS ENTERPRISE MINER



Our modeling started with the selection of the **data source** - "MCDONALDS_RAW_POSITIVE"

In this stage, we replicated the process previously described for Negative Feedback modeling up to the Text Topic node.

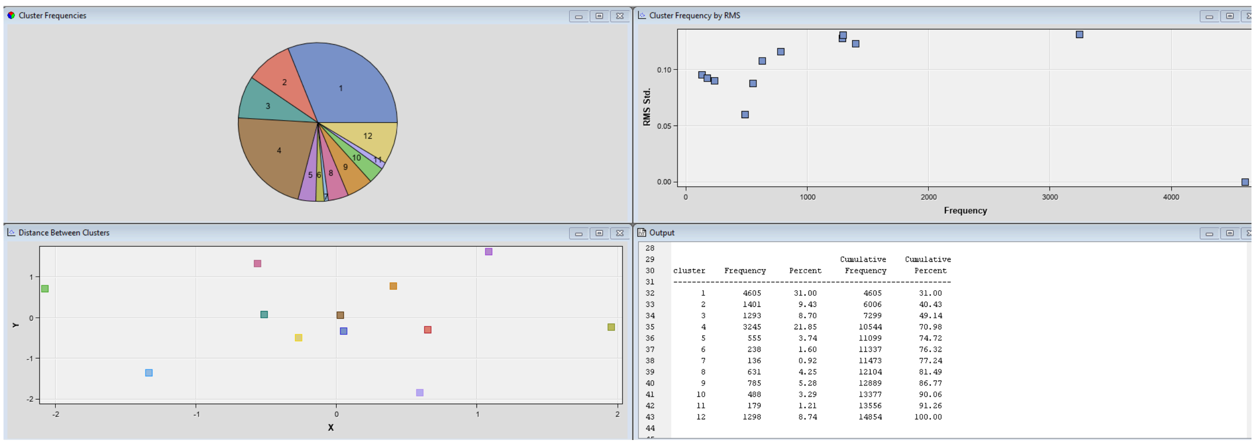
Topic	Category	Term Cutoff	Document Cutoff	Number of Terms	# Docs
+friendly,+employee,friendly service,+friendly staff,+friendly employee	Multiple	0.022	0.084	27	1163
+fresh,+fresh food,+hot,+food,+location	Multiple	0.022	0.078	28	646
+place,good place,+great place,clean,clean place	Multiple	0.022	0.081	19	1080
food,+restaurant,+location,ï,+quick service	Multiple	0.023	0.061	45	533
+good service,+clean,food,+location,+great food	Multiple	0.021	0.072	7	260
+fast service,+clean,+friendly staff,+clean,+morning	Multiple	0.021	0.073	8	312
+great service,+employee,+food,+great food,+great crew	Multiple	0.021	0.068	6	223
good food,+food,+quick service,+price,+clean	Multiple	0.021	0.075	17	450
+clean,+restaurant,clean place,+bathroom,clean restaurant	Multiple	0.022	0.079	32	905
+drive,thru,+drive,+line,+long	Multiple	0.023	0.07	50	659
+coffee,good coffee,great coffee,ice,ï	Multiple	0.023	0.06	46	425
+order,+take,+order,+correct,+kiosk	Multiple	0.024	0.071	98	1157
+eat,+place,+bite,+meal,+cheap	Multiple	0.023	0.064	51	713
+staff,+friendly staff,+quick service,+location,helpful	Multiple	0.022	0.077	34	996
+food,+great food,+fast food,fresh,+expect	Multiple	0.023	0.077	49	1469
fries,+burger,french,french fries,+nugget	Multiple	0.024	0.061	90	647
+hot,hot food,+quick service,+food,+area	Multiple	0.023	0.071	47	662
people,+quick service,+work,really,homeless people	Multiple	0.024	0.061	72	745
+clean,+employee,+location,helpful,+table	Multiple	0.023	0.063	70	734
+breakfast,+quick service,+location,+morning,good breakfast	Multiple	0.023	0.059	48	427
+area,play,+pizza,+location,+play area	Multiple	0.025	0.065	120	1087

We got the above topics which we were able to rename into the following major categories:

Topics

Topic	Category	Term Cutoff	Document Cutoff	Number of Terms	# Docs
Ambience	User	0.0223333333	0.0743333333	92	2083
Cheap Meals & Breakfast	User	0.023	0.0615	78	993
Friendly Staff	User	0.0215	0.0755	42	1673
Good/Fresh food	User	0.0228	0.0706	134	1385
Others	User	0.023	0.06625	160	2156
Play Arena	User	0.025	0.065	120	1087
Quick Drive through	User	0.023	0.07	50	659

Results



Cluster ID	Descriptive Terms
1	'excellent atention' atention
3	+eat +restaurant +price also ï +menu +cheap +park clean +restroom ever +experience +lot convenient never
5	+order +hot +take +fresh +kiosk +manager +table +correct +bring right +use +cashier +window accurate +pay
4	food +location +'good service' +hour +open +late +night 'excellent service' +quality attention +clean +'quick service' +day efficient +dine
11	+busy +pizza little +pretty +pasta +arcade +visit +large +world cool upstairs orlando +wait different +floor
15	+food 'good food' +hot +'fast food' +'great food' 'hot food' 'good fast food' +fresh +'fresh food' fresh +'fast service' 'delicious food' 'cheap food' delicious +tasty
7	+place +area play 'good place' +'great place' +'play area' 'clean place' 'play place' 'nice place' clean favorite fun +play +clean +cool
17	fries +hot +fresh +nugget +big chicken french +sandwich +chicken 'french fries' +'chicken nugget' +'chicken sandwich' 'fresh fries' +shake crispy
6	+friendly +staff +'friendly staff' +clean 'friendly service' 'very friendly staff' 'very fast service' 'fast friendly service' 'good staff' professional helpful 'great staff' +'quick service' +'fast service' atmosphere
18	really +burger +new delicious +expect well +yummy +pounder +favorite bacon +cheese mcd 'quarter pounder' different +taste
14	+drive thru +line +long +'long line' 'fast drive' +move +wait +window +lane +car accurate +open quickly +hour
12	+employee +'great service' +manager courteous +'friendly employee' +'great employee' +job +customer 'great job' +kind +crew polite +keep +friendly +smile
8	+clean +'fast service' +'clean bathroom' +'good service' respectful 'clean store' +neat +'clean restroom' +'good product' +bathroom +highway +fancy modern professional +mistake
9	people +work 'homeless people' 'good people' 'friendly people' 'great people' 'nice people' many 'a lot of' +lot security money +run +worker +kind
16	+coffee 'good coffee' 'great coffee' 'ice coffee' 'iced coffee' +ice +cup +cream ice sugar vanilla +sausage bacon +taste +biscuit
13	+breakfast +egg +sausage +biscuit +pancake 'good breakfast' 'quick breakfast' +'great breakfast' +'breakfast sandwich' mcmuffin hash +sandwich +bagel +cheese +brown
2	+meal +'quick meal' +enjoy +'happy meal' +toy 'good meal' value +bus +son +deal +price +stop +'great place' +look +year
10	ice cream 'ice cream' +cone +machine +'ice cream cone' 'ice cream machine' +break vanilla +cream +pie +'apple pie' +coke +apple +soda

Insights

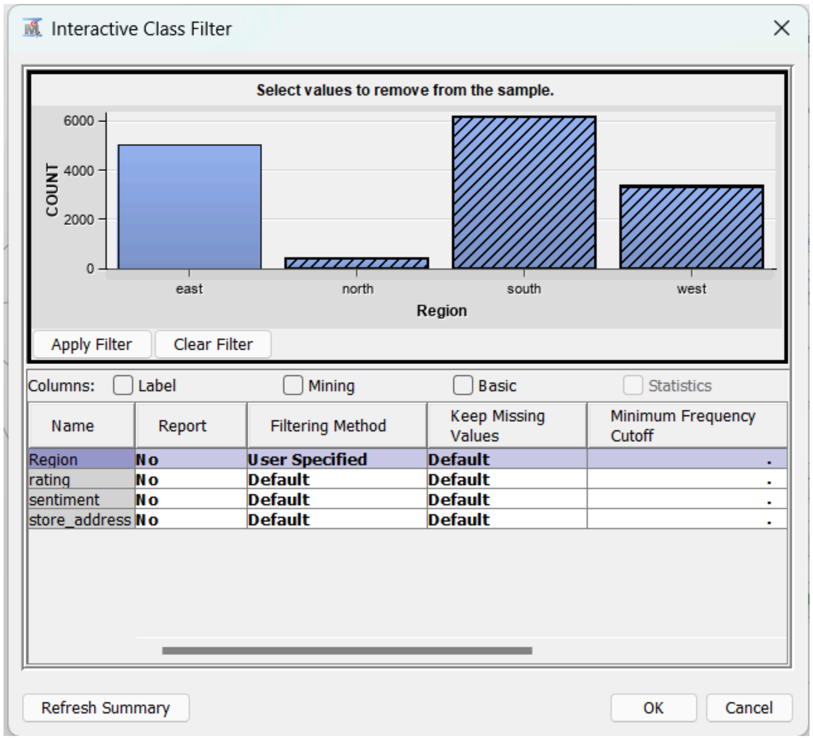
In our analysis of customer feedback, the clustering technique has provided insightful categorizations that align with specific aspects of our offerings, leading to the identification of distinct topics that are particularly well-received by our customers. For instance, Cluster 11 vividly demonstrates a strong customer satisfaction associated with our pasta and pizza offerings, suggesting these menu items resonate well with our patrons' preferences. Similarly, Cluster 18 garners specific positive feedback for the Quarter Pounder with Cheese, indicating this classic burger remains a beloved choice among our customer base. Furthermore, Cluster 7 sheds light on the play arena as a significant aspect of customer appreciation, highlighting its value as a family-friendly feature that enhances the dining experience. These examples underscore the effectiveness of clustering in pinpointing areas of success within our operations, thereby guiding us toward maintaining and enhancing these key aspects that contribute positively to customer satisfaction.

Region-wise Negative Feedback Modeling

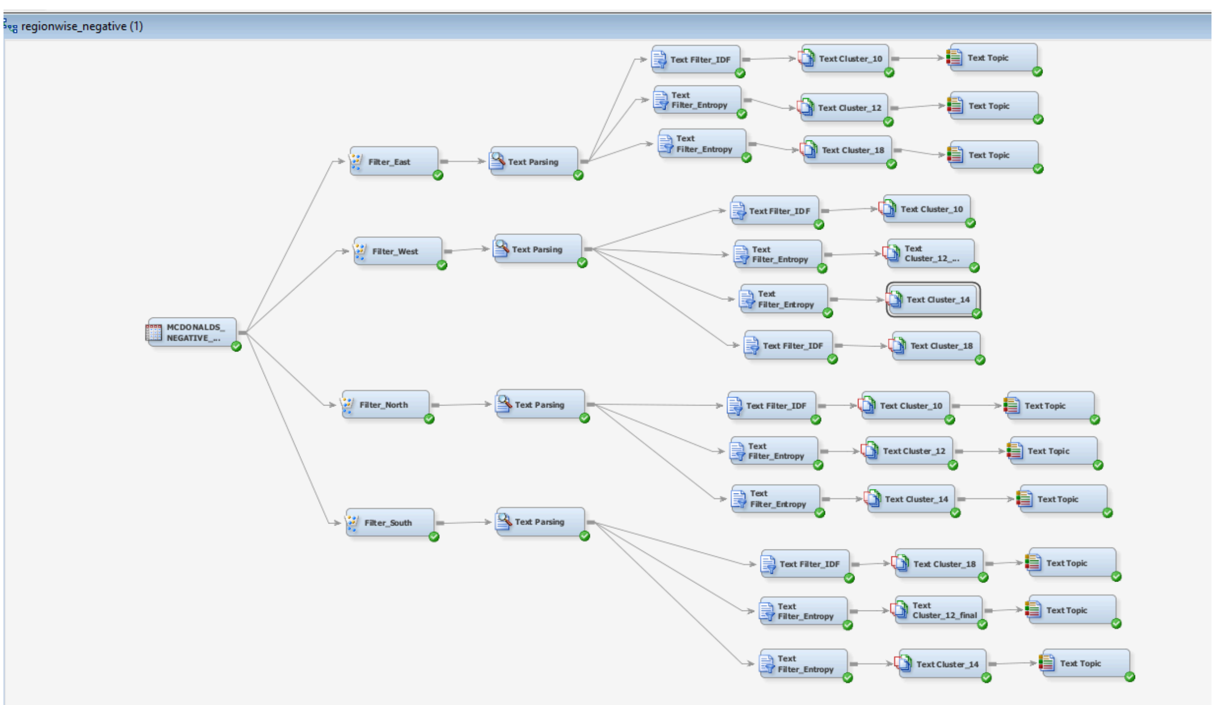
After evaluating both positive and negative reviews, we gained a thorough understanding that certain factors played a significant role in contributing to the overall negative feedback.

Next, we aimed to discern the regional factors impacting both positive and negative reviews. To achieve this, we segmented the data by region, utilizing the Filter node after importing the dataset.

Below is an image showcasing the Filter node configuration:



MODELING DIAGRAM OF SAS ENTERPRISE MINER

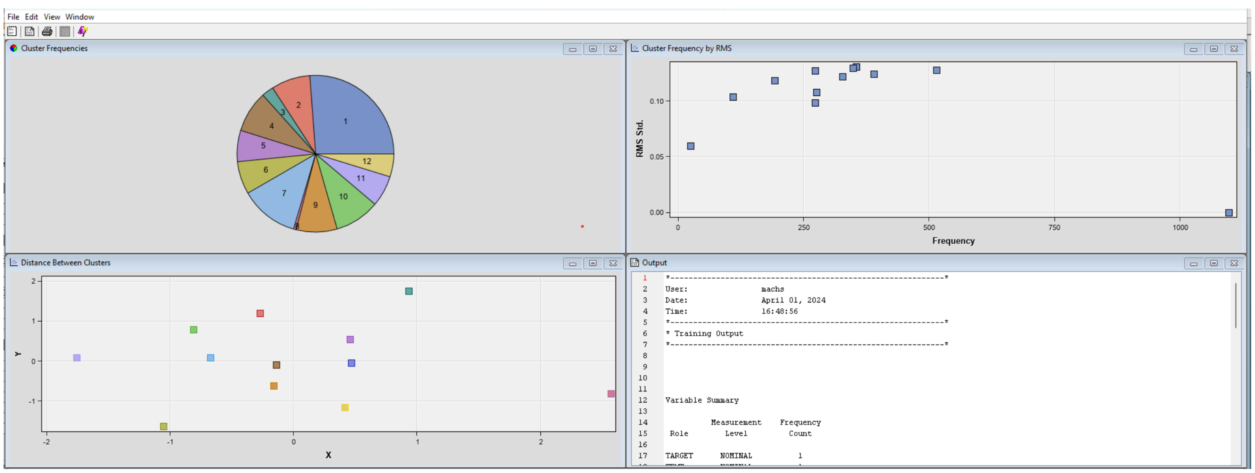


We have our modeling process by establishing the **data source** - “MCDONALDS_NEGATIVE_REGION”

Here we have followed a similar process for each region as mentioned earlier in Negative Feedback modeling till the Text topic node.

We noted that setting the number of clusters to 12 yielded the most effective categorization in terms of clustering.

Eastern Region Results

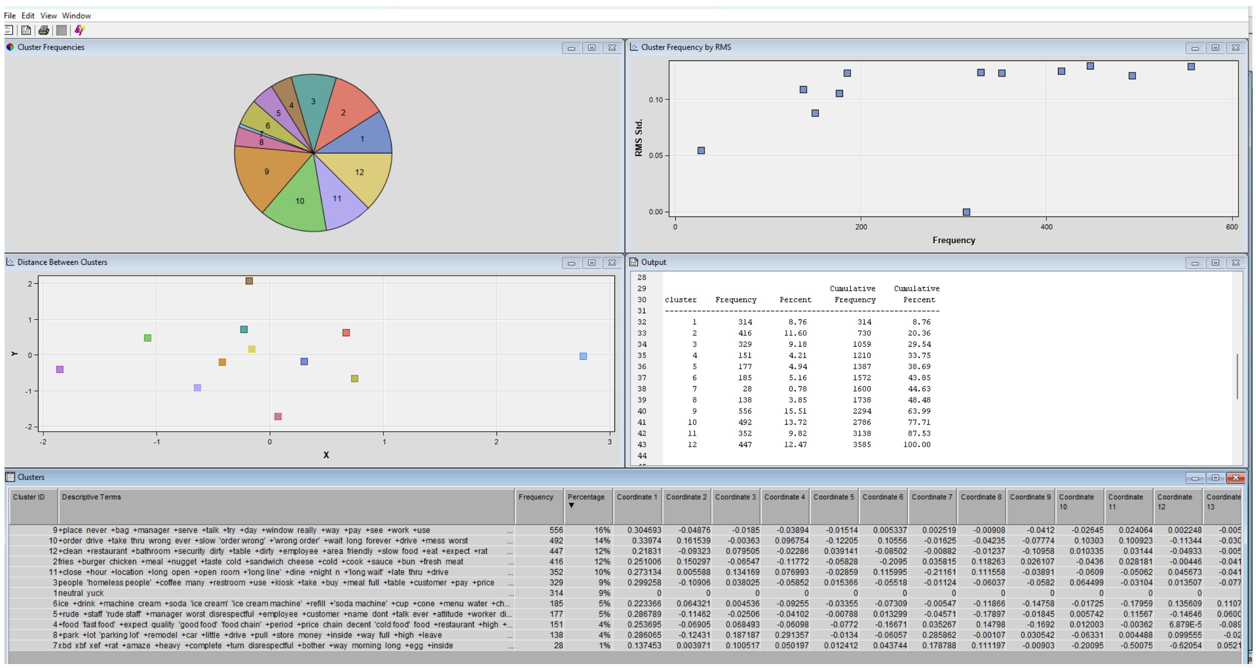


Cluster ID	Descriptive Terms	Frequency	Percentage ▼
1	neutral disapointed ...	1097	26%
7	+people drive homeless +eat thru 'homeless people' many +drug inside +line security +wait +sit +car window ...	516	12%
10	+place +clean +dirty +table dirty +bathroom 'bad service' +avoid +floor pretty 'at all' really +crowd +recommend +area ...	391	9%
4	never +coffee +egg breakfast +see +way +bag n cheese +suck ever +put +drink +stop +morning ...	355	8%
9	+meal ice +sandwich +nugget chicken cream +tea +'chicken nugget' +sauce +chicken +order +big +old food +piece ...	349	8%
2	+rude +staff +employee +manager +attitude +experience worst +speak +nasty +young unprofessional ever +visit +friendly +...	328	8%
11	+order wrong +mess 'order wrong' 'wrong order' +correct +take mobile +wait +item +keep right +miss +speak +receive ...	276	7%
5	+customer +use +bathroom +restroom +kiosk +pay +card +deal +register t +lock security also +buy +morning ...	274	7%
6	+food +fast +'fast food' 'fast service' fast 'good food' +'food place' slow cold +quick +expect +busy waiting +taste +restaurant ...	273	7%
12	fries +burger +salt fresh +pounder +fry double french cold +cold +cheeseburger +large cheese +hot +look ...	193	5%
3	+hour +close +slow open +open 'slow service' 24hrs 24/7 hrs +door usual +'terrible service' +advertise +inform +early ...	110	3%
8	xbf xef xbd sick +chance away +egg +become stuff 'fast service' +handle +piece +bite +hot thru ...	25	1%

Insights

The primary dissatisfaction among customers in the East relates to cleanliness and security, accounting for 26% of feedback with 1,104 mentions. Clusters 7,10 and 5 specifically highlight these concerns.

Western Region Results

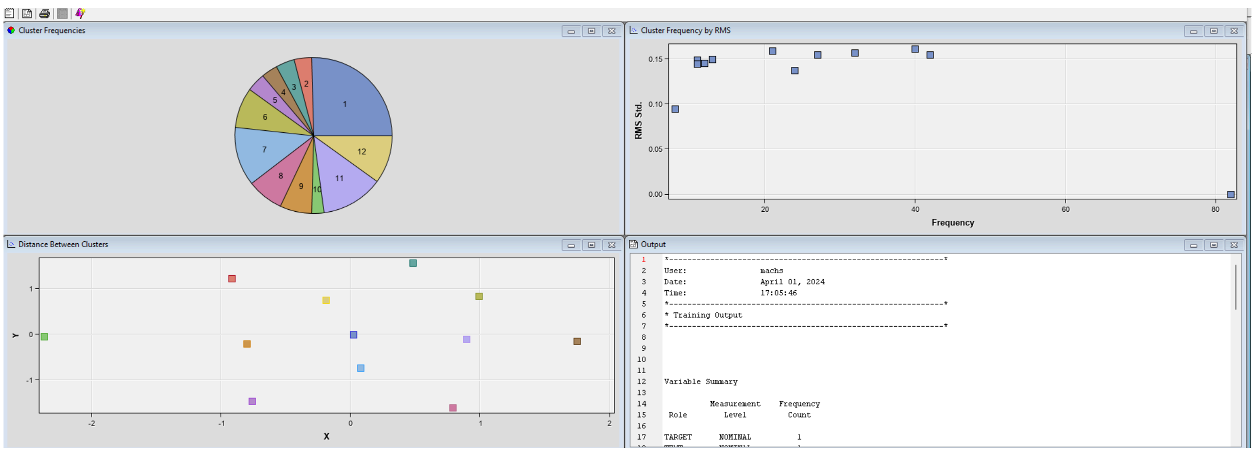


Cluster ID	Descriptive Terms	Frequency	Percentage ▼
9	+place never +bag +manager +serve +talk +try +day +window really +way +pay +see ...	556	16%
10	+order drive +take thru wrong ever +slow 'order wrong' +'wrong order' +wait long foreve...	492	14%
12	+clean +restaurant +bathroom +security dirty +table +dirty +employee +area friendly +sl...	447	12%
2	fries +burger chicken +meal +nugget +taste cold +sandwich cheese +cold +cook +sauc...	416	12%
11	+close +hour +location +long open +open room +'long line' +dine +night n +'long wait' ...	352	10%
3	people 'homeless people' +coffee many +restroom +use +kiosk +take +buy +meal full +...	329	9%
1	neutral yuck ...	314	9%
6	ice +drink +machine cream +soda 'ice cream' 'ice cream machine' +refill +'soda machine' ...	185	5%
5	+rude +staff 'rude staff' +manager worst disrespectful +employee +customer +name dont...	177	5%
4	+food 'fast food' +expect quality 'good food' 'food chain' +period +price chain decent 'col...	151	4%
8	+park +lot 'parking lot' +remodel +car +little +drive +pull +store money +inside +way ful...	138	4%
7	xbd xbf xef +rat +amaze +heavy +complete +turn disrespectful +bother +way morning l...	28	1%

Insights

In the West, the most significant issues are with the drive-through service, as identified in clusters 9 and 10, which have the highest frequency. This is followed by concerns about cleanliness (12%, cluster 12) and meal quality (12%, cluster 2).

Northern Region Results



Cluster ID	Descriptive Terms	Frequency	Percentage ▼
1	neutral ...	82	25%
11	quality long food also +take +line staff +nugget hair +long bag +customer hot +pay 'fast food' ...	42	13%
7	dirty +run place +expect +table available +eat +place +clean slow people person drive area +wait ...	40	12%
12	+order +correct tea 'at all' +refund cheese home ice ridiculous +mess never +charge wrong window +call ...	32	10%
6	+friendly cold +fry +fresh fries +worker chicken +sandwich fish today old +attitude wrong +look breakfast ...	27	8%
8	+slow thru drive slow area inside +'happy meal' +serve +throw +use mcd hot 'fast food' +clean +hour ...	24	7%
9	oak park management unprofessional rude location manager +fix +pull coming open +see madison phone +day ...	21	7%
3	+hard +start +taste +leave mcd money really +purchase +throw +worker cheese fish inside +charge +clean ...	13	4%
5	+app. 'pick up' +park lot +close many store +hour +receipt way maybe +call 'at all' +place +pull ...	12	4%
2	coffee n +refuse ni +throw coming lesson rather +morning new sorry +day +learn breakfast maybe ...	11	3%
4	sometimes +hamburger ketchup +long +wait +attitude +fix +pull +serve +suck money +morning little +line problem ...	11	3%
10	okay +meal +'happy meal' +use store +take experience way many put +taste +forget something fries +order ...	8	2%

Insights

The North region presents a mostly neutral perspective. However, specific issues are noted, including poor management at Oak Park. Unavailability of tables leading customers to overcrowded drive-throughs.

Southern Region Results

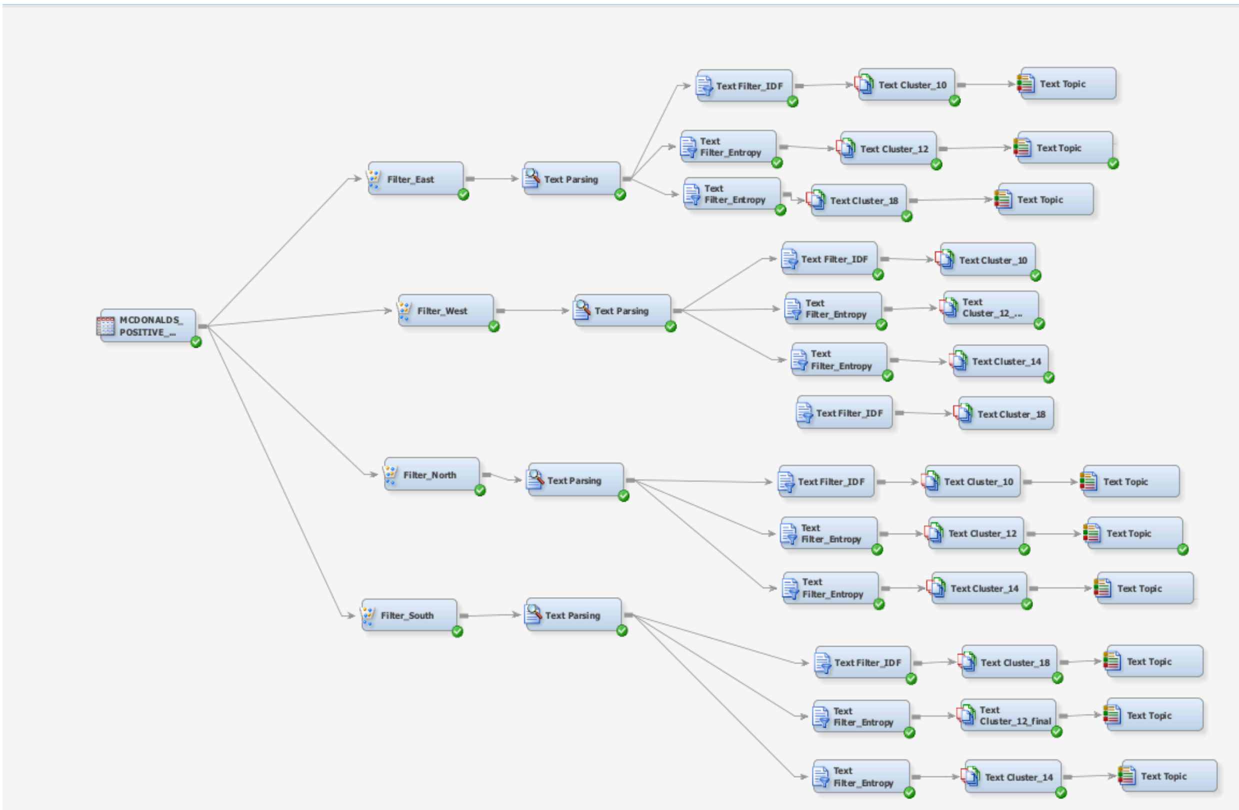
Cluster ID	Descriptive Terms		Frequency	Percentage ▼
6	+order drive thru +manager +meal +window +car +sauce +forget +drink +nugget +pay +drive +leave +attitude	...	1298	18%
1	neutral disapointed	...	739	11%
10	+kiosk n ice +machine +cream +use +counter t +cup +table +drink +person +pay +take +see	...	676	10%
11	+food fries +burger cold +nugget +sandwich +cold chicken +fresh +taste +old +hot +'chicken sandwich' +chicken +dry	...	671	10%
7	+area +clean +table +bathroom +dirty dirty play +floor +smell 'play area' +filthy +dine +expect really +friendly	...	657	9%
2	+order +close +take +hour wrong open +door +open 'order wrong' +dine +sign right forever +mess +night	...	648	9%
9	+wait +long +line +breakfast +park +'long wait' +lot +egg +work +morning +way +menu far +day +late	...	621	9%
3	+food +place +eat food +big +restaurant going slow +miss +expect +recommend +try +dirty people +care	...	501	7%
8	+slow +employee +customer +'slow service' +'rude employee' +rude +speak english absolutely people waiting +put +work water +restaurant	...	392	6%
12	+staff +rude slow +friendly xbd xbf xef people unprofessional busy +care +visit +customer +manager management	...	355	5%
4	ever never worst +'worst service' +life far +experience +see +slow +place +recommend +manager +dirty management +attitude	...	310	4%
5	+coffee +sugar +ice +cup water okay right +morning +cream +taste +drink +leave +hot +put ice	...	152	2%

Insights

The South faces the largest number of complaints about drive-through service, followed by issues with faulty machines (12%, cluster 10) and cold food (10%, cluster 11), equally impacting customer satisfaction.

Region-wise Positive Feedback Modeling

MODELING DIAGRAM OF SAS ENTERPRISE MINER

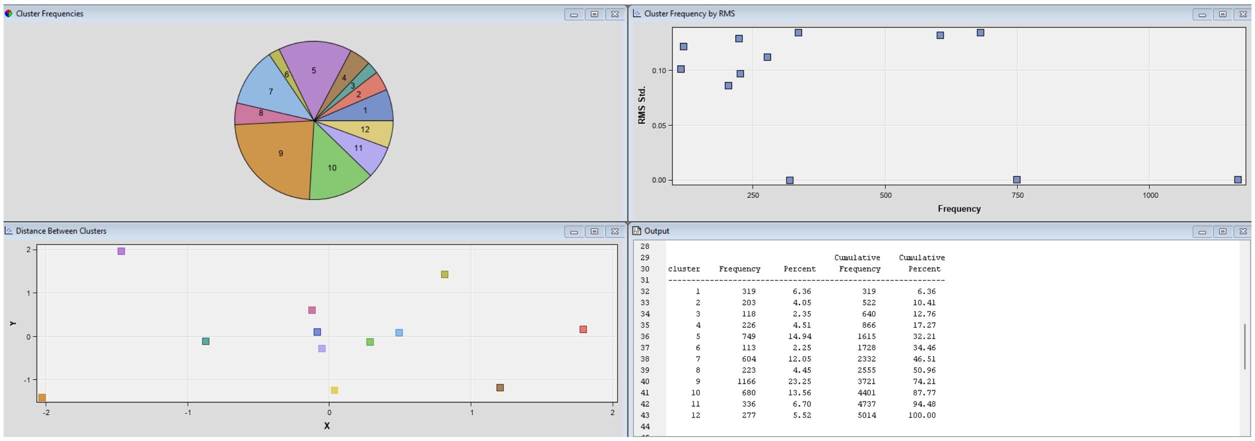


We have our modeling process by establishing the **data source** - "MCDONALDS_POSITIVE_REGION"

Here we have followed a similar process for each region as mentioned earlier in Negative Feedback modeling till the Text topic node.

We noted that setting the number of clusters to 12 yielded the most effective categorization in terms of clustering.

Eastern Region Results

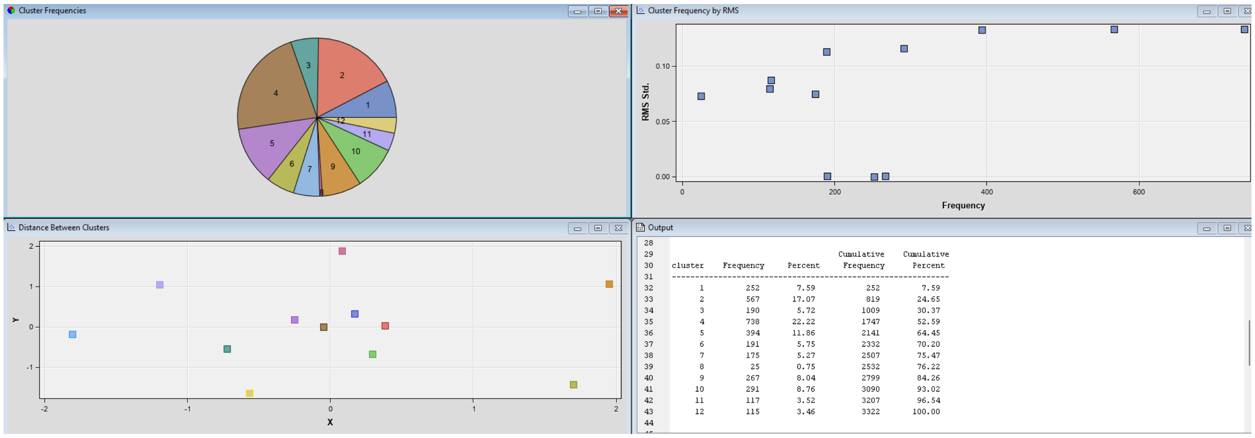


Cluster ID	Descriptive Terms		Frequency	Percentage ▼
9	excellent	...	1166	23%
5	good	...	749	15%
10	+location fries +price +big +spot +area pretty convenient really clean +customer +restaurant +employee +taste much	...	680	14%
7	+order people +hot +take xbd xbf xef +wait +fresh +place +kiosk +line right n +food	...	604	12%
11	drive thru +experience +night +late 'quick service' d +dine +tasty +snack efficient +renovate +busy +amaze +expect	...	336	7%
1		...	319	6%
12	+clean +friendly +staff 'friendly staff' 'clean place' 'friendly service' environment professional +restroom modern +place +bathroom inside 'quick service' +'great service'	...	277	6%
4	+good +place +'good service' 'good place' 'great place' 'good attention' 'nice place' 'good stuff' attention 'very good service' beautiful downtown quiet +snack +safe	...	226	5%
8	+coffee +eat 'good coffee' ice +menu +apple cream well +pie +dollar +'apple pie' +meal something favorite +new	...	223	4%
2	+food +good 'good food' +'fast food' 'great food' 'good fast food' 'delicious food' delicious +'best food' fresh amazing +fresh accessible +price +worker	...	203	4%
3	+'great service' food +breakfast 'excellent service' +chicken +egg 'breakfast sandwich' +sandwich 'great time' juice pizza bacon attention cheese 'chicken sandwich'	...	118	2%
6	+'fast service' +clean accessible convenience border +sit 'delicious food' +organize consistent polite +option efficient +look 'good attention' +friend	...	113	2%

Insights

East with most locations out of all in the dataset, contributed equally for all the topics identified in the positive analysis. However, within cluster ID 8, a specific mention of apple pie is worth mentioning.

Western Region Results

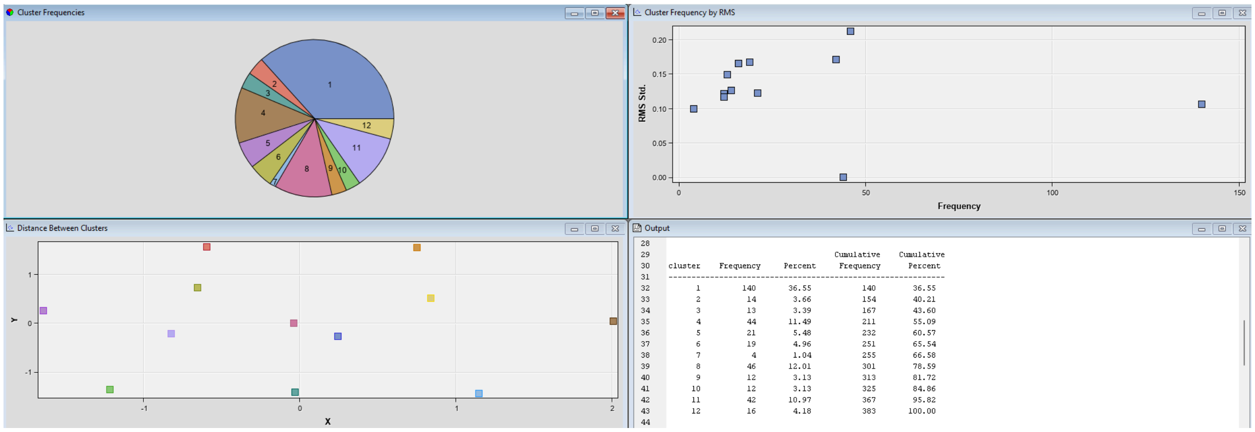


Cluster ID	Descriptive Terms	Frequency	Percentage ▼
4	+order drive thru +take +line +long +hour +wait +employee +use +open food +dine +drive +look ...	738	22%
2	+coffee +breakfast +menu ice cream chicken +burger +ice really n +nugget fries +'fast service' +new tasty ...	567	17%
5	+location never also santa monica +see +keep beach pier +help +try +right people +work +day ...	394	12%
10	+food +hot +fresh +'fast food' 'hot food' fries 'great food' fresh cold 'quick service' +expect french tasty several +'good price' ...	291	9%
9	excellent ...	267	8%
1	...	252	8%
6	good ...	191	6%
3	+clean 'great service' +'clean restroom' people 'homeless people' +lot +park +run inside +friendly +machine +restroom security +area +soda ...	190	6%
7	+good +'good service' 'good food' 'good fast food' 'very good service' +food 'best food' +'good price' 'good quality' +price quality +'quality food' +congratulation +strip +location ...	175	5%
11	+place 'great place' 'good place' 'nice place' 'clean place' +good clean 'quick meal' +dessert +meal +dude +bring okay +crew +clean ...	117	4%
12	+friendly +staff +'friendly staff' 'friendly service' 'very fast service' 'great staff' atmosphere 'timely manner' manner helpful timely dirty +'apple pie' fairly professional ...	115	3%
8	xbd xbf xef +crew several +reopen +remove +advertise +return +healthy +renovate +change dirty +offer +customer ...	25	1%

Insights

The majority contribution is focusing on drive-through, and it being quick and fast however we had to check out parking, as it portrays a negative sentiment even in positive reviews.

Northern Region Results

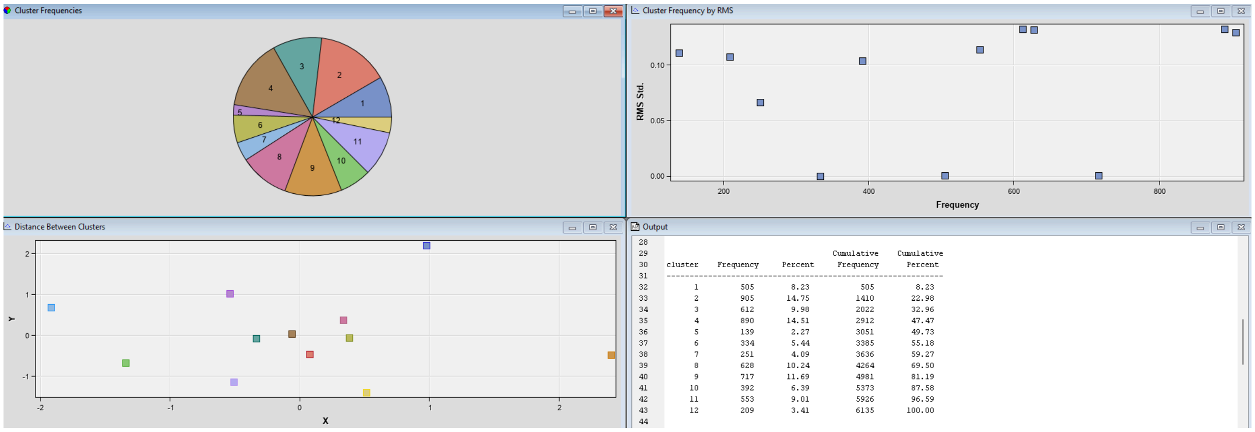


Cluster ID	Descriptive Terms	Frequency	Percentage ▼
1	excellent ...	140	37%
8	+location wait never +restaurant +meal park +clean oak +employee also convenient tasty +long day +problem ...	46	12%
4	good ...	44	11%
11	+good 'hot food' food hot 'good service' 'good food' 'fast service' 'great service' +slow people quickly sauce fresh excellent ...	42	11%
5	delicious fresh fries hot 'great service' also tasty courteous +employee food staff +order place friendly +good ...	21	5%
6	clean +enjoy fresh food good friendly convenient quickly sauce courteous fries staff excellent ...	19	5%
12	pleasant friendly staff 'fast service' +wait +enjoy +clean drive +order place good food ...	16	4%
2	place really clean +problem courteous +meal staff +good +order friendly food good ...	14	4%
3	+line big +order sauce +long efficient drive +problem oak +wait park +restaurant +eat staff friendly ...	13	3%
9	+eat +sandwich breakfast +enjoy +wait day place +long food ...	12	3%
10	thru drive +slow people quickly efficient day +long +order friendly ...	12	3%
7	sweet tea +order +employee day fries +good excellent ...	4	1%

Insights

The North region presents a mostly neutral perspective. However, specific issues are noted, including poor management at Oak Park. Unavailability of tables leading customers to overcrowded drive-throughs.

Southern Region Results



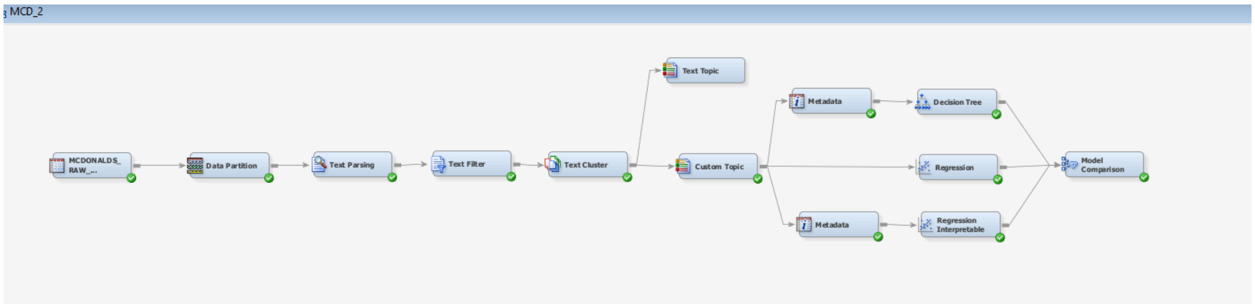
Cluster ID	Descriptive Terms	Frequency	Percentage ▼
2	+order +clean 'fast service' +kiosk +use +new +store ordering +restaurant +correct +pretty quickly +table +bring right ...	905	15%
4	+pizza +area play +visit +arcade +pasta +menu +big +large +see +floor +'play area' fun +world +'great service' ...	890	15%
9	excellent ...	717	12%
8	+eat +breakfast +coffee +meal +cheap mickey +egg +sweet night +year +taste +nugget +price +drink +morning ...	628	10%
3	fries xbd xbf xef +manager ice +customer cream +burger crew +employee +job +smile +morning well ...	612	10%
11	+friendly +staff +place +clean 'friendly staff' helpful 'good place' +'great place' clean favorite 'friendly service' +old 'good attention' +restaurant +worker ...	553	9%
1	good ...	505	8%
10	+food +fresh +hot 'fast food' +'fresh food' +'great food' 'hot food' fresh +tasty 'friendly service' +'great service' +taste +good really +price ...	392	6%
6	...	334	5%
7	+good +food +'good service' 'good food' 'very good service' 'good fast service' 'always good service' 'good attention' +pretty attention affordable value +product 'fast service' +deal ...	251	4%
12	+drive thru +line +long +'long line' 'fast drive' +wait +window +car +move long +correct pretty quickly busy ...	209	3%
5	people +work really +'quick service' 'good people' 'friendly people' hard 'great people' +hard atmosphere +help many +worker busy ever ...	139	2%

Insights

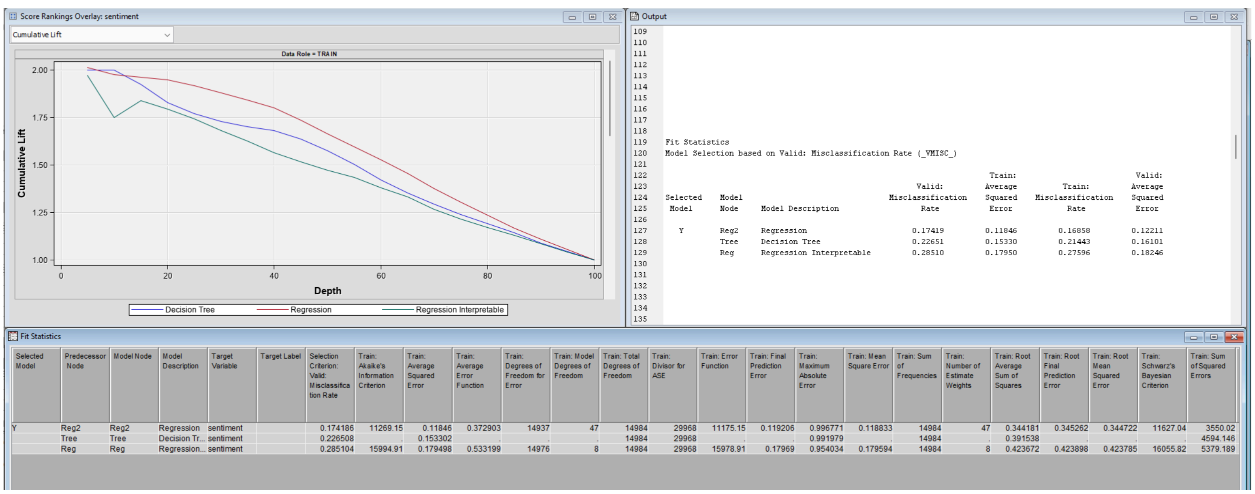
The South faces the largest number of complaints about drive-through service, followed by issues with faulty machines (12%, cluster 10) and cold food (10%, cluster 11), equally impacting customer satisfaction.

Predictive Sentiment Analysis (Supervised Model):

MODELING DIAGRAM OF SAS ENTERPRISE MINER



Results



We have connected all the nodes up to the text topic similarly. We created 4 custom topics: Food, Ambience, Service, Others.

We observed that the Regression model performed the best with a misclassification rate of ~ 0.17.

The maximum likelihood estimation tables in the results of the Regression (Interpretable) model showed us that all the topics apart from 'Others' are significant to the prediction.

SECTION 5: Business Recommendations and Conclusion

Eastern Region

Cleanliness and Security: Implement daily cleaning schedules for bathrooms and dining areas, with additional staff assigned during peak hours. Install security cameras and increase security patrols to deter drug-related activities. Collaborate with local organizations to address homelessness issues around premises, potentially through outreach programs or partnerships.

Western Region

Parking and Drive-Through Efficiency: Explore options for expanding parking facilities or optimizing existing space to accommodate more vehicles. Extend drive-through service hours to operate 24/7 as advertised, ensuring adequate staffing levels for overnight shifts. Design a more efficient order management system to prioritize drive-through orders based on necessity.

Northern Region

Food Quality and Order Accuracy: Provide additional training for management teams, particularly in locations like Oak Park, to improve operational efficiency and ensure consistency in service standards. Conduct regular equipment maintenance checks to prevent ice cream machine malfunctions. Enhance quality control measures to address complaints about cold or stale food and inaccuracies in orders. Introduce mobile ordering options to reduce wait times and enhance order accuracy.

Southern Region

Employee Training and Machine Maintenance: Develop a comprehensive employee training program focusing on customer service skills, with a specific emphasis on drive-through interactions. Implement regular customer service training programs to improve interactions and address issues of rudeness. Implement performance metrics to monitor employee attitudes and provide feedback for improvement. Establish a preventive maintenance schedule for machines to reduce instances of malfunctions, potentially involving more frequent servicing or equipment upgrades.

CONCLUSION

In conclusion, our in-depth analysis of customer feedback across diverse regions has provided invaluable insights into the unique challenges and opportunities faced by McDonald's. By understanding regional preferences and pain points, we are empowered to tailor our marketing and operational strategies effectively.

Today, we've uncovered specific areas for improvement in cleanliness, security, parking, drive-through efficiency, food quality, employee attitude, and machine maintenance. These findings serve as a roadmap for action, guiding us toward enhancing the overall customer experience and driving business growth.

Moving forward, we commit to prioritizing regional-specific initiatives aimed at addressing customer concerns and delivering exceptional service consistently across all McDonald's locations. By harnessing the power of data and customer feedback, we will continue to innovate, adapt, and exceed expectations, ensuring McDonald's remains the preferred choice for customers nationwide.

SECTION 6: References

- OPIM 5671 – Course Materials on Data Mining and Business Intelligence
- https://www.ibm.com/topics/unsupervised-learning
- https://towardsdatascience.com/a-friendly-introduction-to-text-clustering-fa996bcefd04