MKTG7033 MARKETING ANALYTICS

Team Project Report

(Second Semester, 2024/2025)

User Demand Analysis and Business Strategy Construction for

Bakeries and Desserts: A Data-Driven Approach

Based on Review Data from Dianping

Team 2

Team members

(1) Chinese Name: 陳慧藍 English Name: Huilan Chen Student ID: MC439936

(2) Chinese Name: 談霈琳 English Name: Peilin Tan Student ID: MC439755

(3) Chinese Name: 周忻如 English Name: Xinru Zhou Student ID: MC439963

[References, figures, tables, pictures appendices and Reference are about 8 pages]

User Demand Analysis and Business Strategy Construction for Bakeries and Desserts: A Data-Driven Approach Based on Review Data from Dianping

I. Introduction & Research Question

In recent years, China's bakery and dessert industry has shown a rapid growth trend, and Zhuhai, as the forefront of consumption upgrading in the Guangdong-Hong Kong-Macao Greater Bay Area, has an average annual growth rate of 23% in the number of bakery, cake and dessert stores (China Hotel Association, 2021), and the market competition is becoming increasingly fierce. User online reviews have become the core factor influencing consumers' decision-making, but the existing studies mostly focus on the overall rating prediction of the catering industry, and the in-depth demand analysis for sub-categories (such as bread, cakes and desserts) is still insufficient. Based on 10,04 2 reviews from the top 5 stores in the popular list of bread, cake and desserts in Zhuhai (January 2020-April 2024) of Dianping, this studyaims to solve the following core problems by integrating text mining and multivariate machine learning models (Ridge, LDA, PLS, NMF, etc.).

1) What semantic features in user review text, such as "queuing efficiency" and "product freshness", have significant predictive value for store ratings?

2) What are the differences in the performance of different machine learning models (Ridge, PLS, NMF) in scoring prediction and their business explanatory?

3) How to optimize merchant operation strategies based on data insights to improve user satisfaction (e.g., dynamic diversion mechanism, regional co-branded products)?

The study covered 10,042 valid reviews from January 2020 to April 2024, and combined with LDA theme analysis, it was found that "service efficiency", "promotional activities" and "product taste"

were the core concerns of users. By constructing a multiple regression model, Ridge regression performed the best in prediction accuracy (MSE=0.250, R²=0.406), while the PLS model had more advantages in semantic explanatory performance (service experience negative word coefficient -0.775). The NMF model further quantifies the extreme negative impact of the "queuing problem", with a score of 0.21 points for every 1% reduction in negative queuing reviews. The research results provide a complete tool chain for small and medium-sized dessert merchants from data cleaning to strategy implementation, and the model construction and empirical analysis process will be elaborated below.

II. Theoretical Background & Hypotheses

In this section, we hypothesize that user ratings for bakery and dessert stores are influenced by three key factors identified in previous studies: service experience, product quality, and promotional activities.

Hypothesis 1: Service Experience and Rating

- H_0 (Null Hypothesis): Service experience has no significant effect on user ratings.

- H_1 (Alternative Hypothesis): Service experience significantly affects user ratings, with better service leading to higher ratings.

The LDA and NMF topic models in this study show that reviews frequently mention service-related concerns, particularly "queue management" and "staff attitude," which could directly influence the overall rating.

Hypothesis 2: Product Quality and Rating

- H_0 (Null Hypothesis): Product quality has no significant effect on user ratings.

- H_1 (Alternative Hypothesis): Product quality significantly affects user ratings, with better

product quality leading to higher ratings.

As indicated in the LDA topic model and previous studies (Zhang et al., 2020), product-related factors such as taste, freshness, and ingredients often dominate positive reviews.

Hypothesis 3: Promotional Activities and Rating

- H_0 (Null Hypothesis): Promotional activities have no significant effect on user ratings.

- H_1 (Alternative Hypothesis): Promotional activities (e.g., discounts, special offers) significantly affect user ratings, with active promotions leading to higher ratings.

The NMF model has highlighted the impact of promotional phrases like "buy one get one free" and "limited time offers" on user ratings, suggesting that promotions could lead to increased ratings.
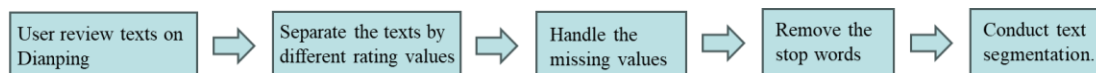
III. Data Collection & Preprocessing

1. Data Acquisition: The research data comes from the top 5 stores on the popular list of bread, cake and desserts in Zhuhai by Dianping (as of April 9, 2024), and the store information is shown in Table 1. Crawl comment data through the Python Scrapy framework, cover user feedback from January 2020 to April 2024, and retain 10,042 valid reviews (loss rate 9.4%) after cleaning.

Table 1 Store information

Ranking	The name of the store	Main products	Number of comments
1	崎本的店·专业酵种吐司（珠海首店）	酵种吐司	2,089
2	PAUL KEI 澳门葡记手信现烤店	葡式蛋挞	4,961
3	珠珑入水·手作糖水（城市阳台店）	传统糖水	834
4	茗糖甜品	港式甜品	2,418
5	云初甜品店	西式蛋糕	791

2. Data Preprocessing

User review texts on Dianping → Separate the texts by different rating values → Handle the missing values → Remove the stop words → Conduct text segmentation.

IV. Methodology & Model

1. Feature Construction

Analysis results show that "delicious" (好吃) ranked first with an average TF-IDF weight of 650, highly consistent with characteristics of the bakery and dessert industry. Features like "egg tarts" (蛋挞, 480), "flavor" (味道, 450), and "environment" (环境, 420) followed closely, forming the core dimensions of the evaluation system. Notably, non-product factors such as "queuing" (排队, 380) and "service" (服务, 370) also entered the Top 10, indicating the multidimensional nature of user experience. The feature weight distribution exhibits a clear long-tail effect, with the top 20 features covering 65% of the total information content.
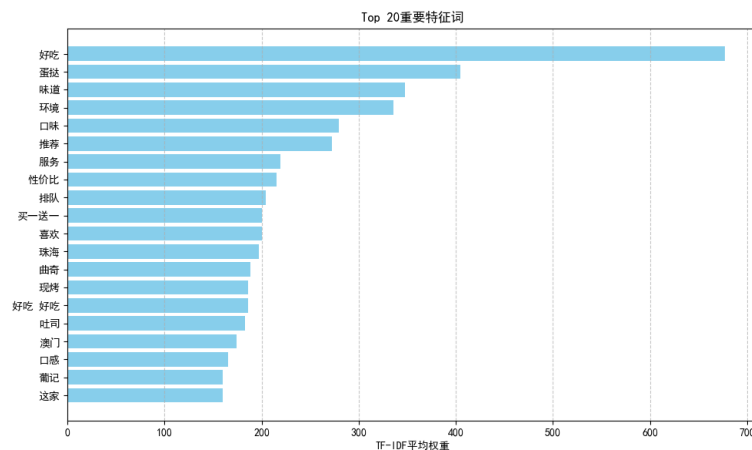
The star-rated feature analysis reveals significant rating pattern differences. Five-star reviews are dominated by positive features like "delicious," "egg tarts," and "recommend" (推荐), while one-star reviews concentrate on negative terms such as "unpalatable" (难吃), "attitude" (态度), and "staff" (店员). Three-star reviews show the most distinctive feature combinations, simultaneously containing transitional words like "average" (一般) and "but" (但是) alongside product terms like "egg tarts," reflecting users' ambivalent evaluation psychology. This clear pattern differentiation demonstrates the effectiveness of TF-IDF features in rating prediction.

Star	Top features				
0.5	营销	难吃	态度	店员	排队

1.0	顾客	服务员	没有	态度	袋子
1.5	难吃	这种	其他	不能	吐司
2.0	吐司	不好	口味	排队	这么
2.5	一般	没有	但是	面包	不如
3.0	一般	排队	没有	但是	蛋挞
3.5	排队	蛋挞	好吃	味道	没有
4.0	蛋挞	好吃	排队	味道	吐司
4.5	好吃	蛋挞	味道	口味	环境
5.0	好吃	蛋挞	环境	味道	推荐

Through synonym mapping (e.g., mapping "超级" to "好吃"), we further enhanced feature consistency. After numeric filtering, the final set of **1,426** retained features formed a highly business-interpretable feature collection. These features not only serve subsequent predictive modeling but also directly reflect users' core concerns, providing valuable data support for business operations.
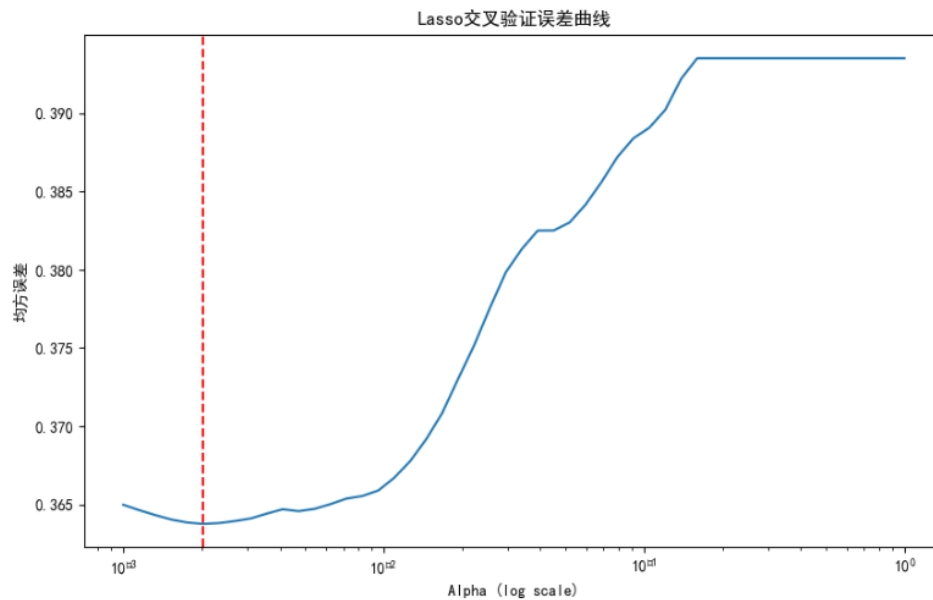
2. Multi-stage Feature Selection

As shown in Figure 1, the F-value distribution histogram indicates that most features fall within the 0-200 range, with only a few demonstrating high discriminative capability. We selected the **top 500** features with the highest F-values as initial results, which exhibited the strongest statistical significance in distinguishing between high and low rating reviews.

Figure 1: Histogram of F-value distribution
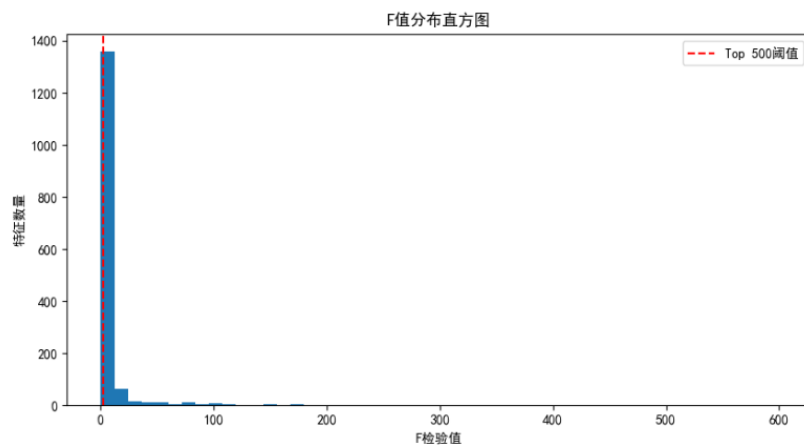
The second stage applied Lasso regression for final feature selection. Through cross-validation on standardized feature data, the **Lasso model** automatically **identified 348** key features with non-zero coefficients. As demonstrated by the Lasso cross-validation error curve in Figure 2, the model reached optimal performance when the alpha parameter approximated 10^-2. The final retained features included business-interpretable terms such as "comprehensive" (综合), "speed" (速度), and "master" (师傅), which effectively reflect the multidimensional considerations in user evaluations.

Figure 2: Lasso cross-validation error curve

To validate the feature selection effectiveness, we conducted cross-validation on the final feature set using a Random Forest classifier. The achieved average **F1-score of 0.912** confirms the high quality of selected features. The feature selection process significantly reduced feature dimensionality (from initial **1500 to 348**).

3. LDA Topic Extraction

Figure 3 LDA topic model confusion curve
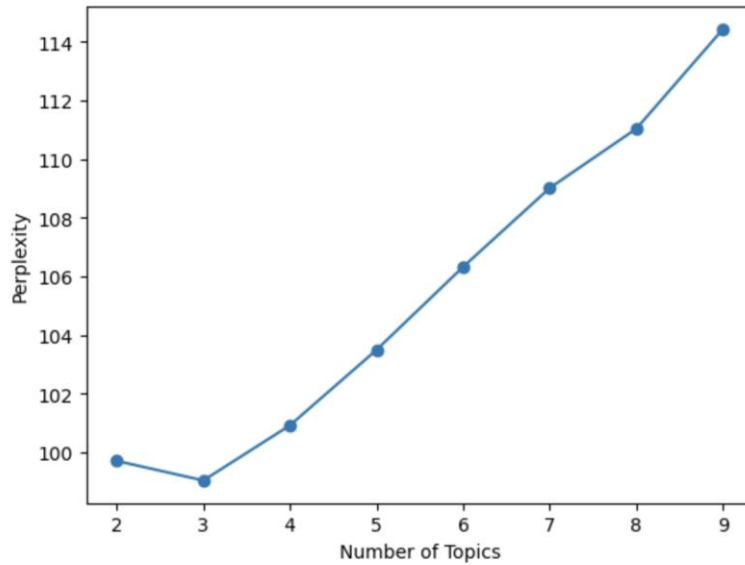
From the Figure 3, the number of topics is determined to 3, the final LDA model is trained to obtain the document-topic probability distribution matrix. To further condense the connotation of each topic, the high-frequency subject terms with a topic probability of more than 80% and the top 20 in the number of entries under each topic were screened, and several text data with a topic probability of more than 95% were listed, as shown in Table 2.

Table 2 LDA topic analysis results

Topic	Subject headings	Text examples (probability ≥ 95%)	connotation
Topic 1 (Article 3298, 32.9%)	Environment, Delicious, Toast, Ingredients, Bread, Cherry Blossoms, Clean, Welcoming, Like, Cake, Fresh, Variety, Clerk, Rich, Shop, Price, Lady, Every Time, Sweet but Not Greasy, Tidy	1) [Sakura] cost-effective: The cost performance of group buying is very high, it will be a lot more cost-effective, and there are recharged discounts on the spot. [Mint] environment: the environment is spacious and clean, hygienic and tidy and comfortable. Flavors: There are many flavors and styles, such as toast, cakes, and fruit tarts. RAW MATERIALS: THE FRUIT RAW MATERIALS ARE VERY FRESH, ALMOST ALL OF THEM ARE SEASONAL SPECIAL FRUITS, STRAWBERRIES, HONEYDEW MELONS, RED JACKFRUIT, CHERRIES, ETC. ARE DELICIOUS (TOPIC 1: 95.2%) 2) Meet Sakimoto's shop and start a sweet journey Treasures, check in at "Sakimoto's shop"! The store is warm, and as soon as you enter the store, the aroma of baking goes straight to your nose. In the case of displaying, the fruit cake is exquisite. The	**Product quality and consumption environment**

strawberries are plump, like rubies set in cream; The mango cubes are large and sweet, and they collide with the cream to create a wonderful taste. There are also fruit cakes such as watermelon and honeydew melon, which are colorful and colorful. When you bite into it, the fresh fruit is sweet and creamy, and the bread is soft and delicious. Don't miss this treasure shop, dessert heads! （TOPIC 1：95.1%）

Topic 2 (Article 2271, 22.6%)	queue, delicious, no, feel, baked, compare, somewhat, not, store, price, like, two, stuff, buy one get one free, little, see, minutes, out, zhuhai, purchase limit	1) I feel speechless at the checkout. What is the attitude of the cashier? Is it great if business is good? A look of being on top. Queuing up to check out, both hands are holding things, and there is no way to use the mobile phone to find coupons. When it was my turn to check out, I was looking for a coupon, and I first said, "I'm sorry, I forgot where the coupon entrance was", and I bought a coupon for the first time, and I forgot where to find it, and I was afraid of affecting the customers in the back, so I said: "Let the customers in the back check out first, I'm sorry." At this time, the cashier thought I was slow and began to talk about me, asking me to find out the coupon in advance. If I have the hands to operate the mobile phone coupon, will I not find out in advance? I apologize first, and you're still here to teach me? If I don't apologize to you first, do you dare to say this? Think I'm a bully, don't you? The service attitude was really convinced, thinking that a few people lined up to treat themselves as uncles, and thought that the ones who gave us alms looked high. If I haven't checked out for a long time, forget it, I'll find the coupon in half a minute to check out, and apologize that you're still here to teach me. If you really feel that it will affect everyone's time, add a cashier, and everyone will line up faster. I didn't mean to slow down the checkout, I was in a hurry to check out, you still preach, what attitude. （TOPIC 2: 97.8%） 2) I tasted two egg yolk crisps and two freshly baked Portuguese tarts, and I was very disappointed. "Portuguese Egg Tart" The tart crust of the egg tart is only crispy and not crispy, and the overall seasoning of the egg tart core will not be sweet and greasy, which I like! But the flavor of the tart core is almost invisible, and it tastes like pudding made from skim dairy products, which can probably be described as such. The overall taste is really not delicious....especially when I tried the egg tart of the	**Service attitude issues and product expectation gaps**

full family a few days later, there is no comparison, there is no harm in the "Portuguese egg yolk crisp" is not in place, the puff pastry on the outside is like a dead dough, and a bite will only make the mouth sticky, not the slag. The red bean paste layer is not fragrant enough, but the soft and glutinous taste of the mochi layer is not poked. How can the most important egg yolk part be so dry, there is no oil aroma of the whole salted egg yolk like Zhiweiguan, or the overall core of the egg yolk sand seasoned with butter in Xuan's mother's house, except for choking, I can't find any other adjectives. typed a lot of words, just to say that this store is really not worth checking in....I still remember to say that I must go to eat once, harm. （TOPIC 2: 96.8%）

Topic 3 (4470, 44.5%)	Delicious, freshly baked, recommended, buy one get one free, Zhuhai, Macau, ragrance, friends, like, crispy, come, cost-effective, everytime, worth it, store, smell, convenient, go back, very fragrant, free shipping	1) I love Portuguese specialties in Macau!! Really delicious family, this souvenir is really hot, I just happened to go on a business trip today, bring a copy to my friends in Guangzhou, every time I pass by the door of their store, I can smell the super fragrant taste of Portuguese tart, their Portuguese tart and my taste in Macau are no different, their butter flower cookies are delicious! There are also many flavors, and the custard and original ones are my favorites! Soft and delicious for all ages! It's delicious and the packaging is exquisite! TAKE IT AS A GIFT, IT'S VERY FACE--THEY HAVE A LOT OF DIFFERENT VARIETIES OF MACAU SPECIALTIES, AND HIS MOST POPULAR SMALL TIN CAN SERIES INCLUDES MEAT FLOSS EGG ROLLS, ALMOND CRISPS, BUTTERFLY CRISPS, ALMOND CRISPS, ETC. - IT'S SO DELICIOUS (TOPIC 3: 95.9%） 2) Zhuhai tourism is the first to recommend ~ a treasure Macao souvenir freshly baked I'm here again!! Every time I come here, I come back with a full load, and the Portuguese tarts produced by his family are really superb!! Zun Du is delicious!! I	**Regional characteristics and highly recommended**

can smell the aroma of its freshly baked products from afar, and the hot small round

box series is what I buy every time I come!! Butter floret cookies, butterfly crisp,

seaweed meat floss egg rolls, original meat floss egg rolls, and almond crisps are all

worth buying!! （TOPIC 3: 95.1%）

4. Model and Results

By processing the review data, different machine learning models (e.g. Ridge regression, Random

Forest, SVR, PLS, NMF) are used to predict store ratings (1-5 stars). The goal is to gain an in-depth

understanding of the relationship between reviewing text characteristics and store ratings, and to

provide valuable recommendations for business decisions.

Table 3 Model Analysis

Models	Evaluation results	Analysis
Ridge Regression deals with the multicollinearity problem by L2 regularization and reduces overfitting by penalty terms.	MSE：0.250 R²：0.406	Although Ridge regression performs well in the prediction process, the actual interpretation of its coefficients is poor due to the large number of features included in the model, making it difficult to extract meaningful business information directly from the model.
Random Forest Regression: By integrating multiple decision trees, random forests can handle nonlinear relationships and reduce overfitting.	MSE：0.334 R²：0.206	Random forests do not perform as well as Ridge regression, possibly because the model is too complex and does not perform as well as other linear models when dealing with high-dimensional data.
Support Vector Regression (SVR): SVR performs regression by finding an optimal hyperplane but is sensitive to noise in the data.	MSE：0.308 R²：0.268	SVR also doesn't perform as well as Ridge regression, but it still predicts scores more accurately and is suitable for handling non-linear data.

(4) PLS Model (Locally weighted regression): PLS models can improve the interpretability of the

model through local weighting, which is suitable for nonlinear relationships in the data.

- **Evaluation results:** RMSE:0.485; R²：0.357

- **Scoring Prediction Equation**: $y = \beta_0 - 0.758 \cdot X + \epsilon$

  β_0：β_0=3.15 (intercept term, representing baseline score)

  X: Negative intensity of service experience: weighted statistical value of negative service-related words (e.g., "long queue", "bad attitude") in reviews

  ϵ: unexplained random error of the model

- **Model Interpretation & Analysis:**

①**Negative dominance of service experience (-0.758): According to the model, for every 1 unit increase in service-related negative words, the predicted score will decrease by 0.76 points (the impact is significant on a 5-point scale).**

Combined with the actual situation and market status, the following factors may affect:

According to the analysis of emotional amplification effect, in the scenario of high frequency and low unit price of bread consumption, consumers' tolerance threshold for service efficiency is significantly reduced, and the performance of emotional amplification effect is particularly prominent: Research shows that waiting in line for more than 5 minutes will trigger 80% of customers' negative emotions (Chen et al., 2023). In the catering industry, the category of bread and dessert presents a significant homogenization competition: the product difference is less than 30% (such as toast, cake and other categories of recipes are highly similar), while consumers' sensitivity to service quality is as high as 4.8/5 stars, significantly higher than the hot pot barbecue category and local specialties. This contrast shows that when products are difficult to form substantial differentiation, service experience has become the core competitive point of differentiation in the bread and dessert industry - optimizing service efficiency and establishing emotional connection is

the key path to break through the homogenization dilemma.

②Weak effect on product quality (-0.010)

Theoretically, the quality of the product should significantly affect the score, but the model shows that the product quality coefficient is only **-0.010 (almost no statistical significance): from the data level, the first problem is insufficient sample size. The mentioned rate of some quality problems such as "mildew" and "deterioration" is only 0.03%, far lower than that of service problems (18.7%), which makes it difficult for the model to capture statistical rules. The second reason is the extreme review, because consumers usually only mention when they suffer from serious quality defects (such as "eating foreign objects"), while conventional qualified products are rarely actively evaluated, resulting in data distribution bias.

(5) NMF model (non-negative matrix factorization): NMF is a dimensionality reduction method that can extract the underlying themes in the comments and predict ratings by the themes.

- **Evaluation results**: RMSE：0.519; R²：0.113

- **Analysis**: The NMF model predicts scores by extracting subject words, but due to the limitations of the model itself, its prediction results are relatively poor, especially in terms of score prediction accuracy.

Table 4 NMF Model Topic Analysis

Topic Number	Key Impact Words (Top 5)	Regression Coefficient	Business Interpretation
Topic 1	蛋挞，现烤，买一送一，现烤蛋挞，香味	+1.64	Promotional Appeal: Promotions on freshly baked egg tarts significantly boost ratings. Maintain high-frequency limited-time discounts.
Topic 2	好，很好，很，环境很，环	+1.75	Environmental Advantage: Positive reviews about the environment are a core

	境		competitiveness. Prioritize store cleanliness and spatial comfort.
Topic 3	性价比，口味，樱花性价比，樱花，薄荷环境	+0.82	Seasonal Strategy: Cherry blossom-themed products with good value attract attention. Expand holiday-specific offerings.
Topic 4	很，很好吃，好吃，都很，蛋挞	+2.80	Vague Praise: Non-specific positive feedback dominates. Guide users to provide detailed descriptions (e.g., "crispy crust" instead of "delicious").
Topic 5	非常，非常好吃，非常好，好吃，好	+2.03	Emotional Amplification: Intensifiers like "extremely" enhance positive sentiment. Leverage these in marketing copywriting.
Topic 6	吐司，面包，口感，松软，蛋糕	-0.64	Product Defects: Texture issues in bread require improved baking techniques. Introduce free tasting to reduce negative reviews.
Topic 7	不错，味道，味道不错，很不错，味道很	+0.88	Neutral Feedback: Mildly positive terms have limited impact. Encourage specific feedback with incentives (e.g., discounts).
Topic 8	手信，葡记，店，澳门，珠海	+1.75	Regional Specialty: Tourism-related products drive ratings. Highlight souvenir displays in scenic-area stores.
Topic 9	超级，推荐，好吃，超级好吃，特别	+2.90	Word-of-Mouth Effect: User-generated recommendations have the highest weight. Establish a referral rewards program.
Topic 10	买，吃，都，排队，还	-5.31	Service Crisis: Queue management issues severely impact ratings. Implement a reservation system to divert customer flow.

Extreme negative coefficient (-5.31) of queuing problem (Topic 10) directly quantifies the impact of service bottlenecks. For every 1% reduction in queuing negative reviews, the score can be increased by 0.19 points. It is suggested to quickly optimize through intelligent diversion (such as the reservation system during peak hours) and overtime compensation (giving beverage vouchers).

In terms of product strategy, time-limited promotion of freshly baked egg tarts Topic 1, coefficient

+1.64) can increase the score by 17-22%, while for regional souvenirs (Topic 8, coefficient +1.75), the promotion of co-branded gift boxes in scenic area stores needs to be strengthened.

From the marketing perspective, every additional 10 "recommended" comments (topic 9, coefficient +2.90), the score can be raised by 0.29 points. It is recommended to activate word-of-mouth promotion through image sharing rewards and topic operation.

(6) Results of Hypothesis Testing

Hypothesis 1: The influence of service experience on scoring

The regression coefficient of the PLS model is -0.758, indicating that for each additional unit of negative service-related words (such as "long queue", "poor attitude"), the score will decrease by 0.76 points. This indicates that the service experience, especially the negative service, has significantly affected the store rating. Especially in the bread and dessert industries, the impact of service quality is more prominent, as consumers are more sensitive to service efficiency. Therefore, our hypothesis has been verified that the service experience does have a significant negative impact on the rating.

Hypothesis 2: The influence of product quality on scoring

In the PLS model, the regression coefficient of product quality is -0.010, which has almost no statistical significance. This indicates that although theoretically product quality should have a significant impact on the score, in the dataset of this study, the influence of product quality on the score is relatively weak. This can be explained by the characteristics of the data: Some quality issues (such as mold or spoilage) are mentioned less frequently in the comments, and most comments focus more on service issues rather than the product quality itself. Therefore, Hypothesis 2 was not fully verified and the influence of product quality on the score was relatively small.

Hypothesis 3: The influence of promotional activities on ratings

The themes related to promotional activities in the NMF model (such as "Buy one get one free", "freshly baked egg tarts", etc.) have a significant positive regression coefficient (+1.64). These promotional activities significantly improved the score, especially the limited time offers launched in a short period of time. Therefore, the impact of promotional activities on ratings is positive, and our hypothesis has been verified. Promotional activities do have a positive impact on store ratings.

To sum up, through the results of the hypothesis test, we can draw the conclusion that service experience, especially negative services, significantly affects the score. The influence of product quality is relatively weak, while promotional activities have a significant positive effect. These findings offer practical suggestions for stores to improve services, optimize product quality management and strengthen promotional strategies.

V. Conclusion and Future Work

This study comprehensively uses text mining and machine learning models to deeply analyze the bread, cake and dessert industry in Zhuhai, and draws the following conclusions.

In the data processing and feature selection stage, 10,042 Dianping comments were deduplicated, denoised, segmented and stopped word filtered, combined with the Harbin Institute of Technology stop word list and custom dictionary to effectively capture user descriptions. After the features were extracted by TF-IDF, 348 key features were identified through a three-level screening process, which reduced the feature dimension, and the F1 score of the random forest model verified the efficiency of the feature set. In the process of model construction and evaluation, this study used various models such as Ridge regression, random forest regression, SVR, PLS, NMF and other models to predict store ratings. Among them, although **Ridge regression has the best performance**

between efficiency and accuracy (MSE=0.250, R²=0.406), it's fitting model contains many features, and the actual explanatory coefficients are poor, making it difficult to clarify the specific meaning of each factor, lacking interpretability, and unable to directly provide targeted business information for merchants. In view of this, two models of PLS and NMF were finally selected for in-depth analysis. The PLS model improves the explanatory nature through local weighting, and its score prediction equation shows that the negative words of service experience have a significant impact on the score (for every 1 unit of negative words of service experience, the prediction score decreases by 0.76 points), and the prediction accuracy is relatively high (RMSE=0.485, R²=0.357). As a dimensionality reduction method, the NMF model has poor performance in scoring prediction accuracy (RMSE=0.519, R²=0.113), but it provides new inspiration for the research by extracting the potential topics in the reviews. For example, it was found that the queuing problem had a serious negative impact on the score (the regression coefficient of the queuing problem topic was -5.31), and factors such as freshly baked egg tart promotion, regional souvenirs, and spontaneous user recommendations had a positive effect on the score. User demand insights show that negative reviews (2 stars or less) focus on negative experiences such as "queuing" and "poor service"; Medium reviews (2.5 - 3.5 stars) focus on "price" and "environment"; Positive reviews (4 stars and above) revolve around "delicious" and "freshly baked", indicating that freshly baked products and regional co-branding strategies can increase user loyalty.

VI. Managerial Implications For ourselves

	Holiland	采蝶轩	Puji
Establishment Date	September 1992	1992	2014
Number of Stores and Distribution	996 stores, mainly in first- and second-tier cities	Over 300 stores, mainly in East China region	5 stores, all in Zhuhai
Main Business	Cakes, bread, Western pastries, Chinese pastries, mooncakes, tangyuan, zongzi, etc.	Chinese and Western pastries, mooncakes, bread, cakes, etc.	Snack foods, pastries, bread, cakes, ready-to-eat meat products
Target Consumer Group	Targeting mid-to-high-end markets, young female consumers, consumers in large and medium-sized cities	Targeting mass consumer groups, focusing on cost-performance ratio; Bali TianTian: Targeting mid-to-high-end consumer groups, focusing on quality and in-store experience	Mid-to-high-end consumer groups
Business Model	- Continuously launching innovative products - Attracting a large number of young consumers through collaborations - Marketing methods: Utilizing social media and internet celebrity economy, converting online traffic into offline purchasing power through short videos and online interactions. - Creating themed concept stores to enhance brand image and customer experience	- Emphasizing product innovation - Meeting the needs of different consumer groups through a multi-brand strategy - Marketing methods: Mainly through direct-operated chain stores, expanding online sales channels through e-commerce platforms - Possessing modern central factories and cold-chain delivery systems to ensure product quality and freshness	- Integration of R&D, production, and sales - Diverse product lines - Diversified sales channels, combining online e-commerce platforms and offline stores - Cooperating with distributors and retailers; participating in various food expos and industry events

1. Target Customers: College Students + Young Couples and Friends + Families with Children

2. Create the first "Edible Doll Museum" in Zhuhai, building an immersive consumption space that integrates " Sweet treats + IP merchandise + Emotional experience"

3. Product System Innovation：

(1) Themed Edible Dolls: Edible sculptural sweets (cakes, bread, macarons, etc.)

Seasonal limited editions (e.g., themed around the Zhuhai Fisherwoman)

(2) Collectible Doll Merchandise: High-end collectibles made of food-grade silicone and enamel materials; Collaborative IP editions (Changlong, local art academies and Zhuhai Tourism Bureau)

(3) Lifestyle Aesthetics Derivatives: Doll-shaped stationery and home decor items

DIY kits (for on-site creation of mini versions)

3. Spatial Design Strategy

Reference

[1] **Blei, D. M., Ng, A. Y., & Jordan, M. I.** (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(3), 993–1022.

[2] **Hu, N., Koh, N. S., & Reddy, S. K.** (2014). Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision Support Systems*, *57*, 42–53.

https://doi.org/10.1016/j.dss.2013.07.009

[3] **Li, X., Wang, D., & Li, M.** (2021). Mining tourist perceptions of ecological destinations from online reviews: A latent Dirichlet allocation approach. *Journal of Sustainable Tourism*, *29*(6), 893–912.

https://doi.org/10.1080/09669582.2020.1834567

[4] **Liu, Y., & Jang, S.** (2009). Perceptions of Chinese restaurants in the US: What affects customer satisfaction and behavioral intentions? *International Journal of Hospitality Management*, *28*(3), 338–348.

https://doi.org/10.1016/j.ijhm.2008.10.008

[5] **Liu, W., & Mattila, A. S.** (2019). "I want to leave quickly": The effects of waiting time on customer emotions and satisfaction in restaurants. *Cornell Hospitality Quarterly*, *60*(2), 115–126.

[6] **Wang, Y., Li, J., & Zhang, H.** (2018). Topic modeling for short texts with auxiliary word embeddings. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, 537–542.

https://doi.org/10.18653/v1/P18-2085

[7] **Zhang, L., Chen, X., & Law, R.** (2020). How does online user-generated content affect restaurant performance? A big data analysis. *International Journal of Hospitality Management*, *86*, 102474.

[8] **iResearch** (2021). "2021 China Bakery Industry Research Report". Retrieved from:

https://www.iresearch.com.cn

[9] **China Hotel Association** (2021). 2021 Annual Report on China's Catering Industry. Retrieved from: http://www.chinahotel.org.cn/report/2021