Data Analytics Using R

Project 1 - Predicting Public Opinion on a Proposed Casino in Toronto

Students
Huyen Thi Thu Pham

Contents

1 Introduction

Toronto's City Council is considering the establishment of a casino in the community and seeks to understand public opinion on this matter. To gather insights, a survey was conducted to capture the views and opinions of residents. The survey comprised 11 questions aimed at:

- Understand residents' attitudes toward a proposed casino
- Assessing its alignment with Toronto's image
- Gathering opinions on specific proposals and locations

The survey included a mix of open-ended questions and multiple-choice questions, allowing respondents the flexibility to express their views fully.

The original dataset has 17766 observations and 92 variables, this report will only focus on Q1_A, Q3_ responses, and Gender to provide insights into public sentiment toward the casino proposal.

This report conduct multinominal logistic regression and PCA to predict a general feeling of respondents towards a new casino based on their answers from a survey data collected by City of Toronto. I also run EFA to find the main driving factors in popullation opinions.

2 Relationship between responses to Q1_A and various factors captured in Q3_

Before conducting data analysis, the dataset was preprocessed to ensure it was ready for modeling. The detailed data processing steps are provided in Appendix A.

2.1 Data Exploration

To gain a deeper understanding of the data in Q1_A and Q3_, I will begin by examining the distribution of responses in Q1_A to capture the overall sentiment of participants toward the proposed casino. Following this, I will analyze the responses in Q3_ to identify the key factors that shape these sentiments.

From Figure 1 reveals that a significant majority (66.2%) of respondents are strongly opposed to the casino. In contrast, the 21.2% who are strongly in favor. The percentage of neutral or mixed responses just 3.1% of respondents, while those somewhat opposed and somewhat in favour is 5% and 4.5% respectively.

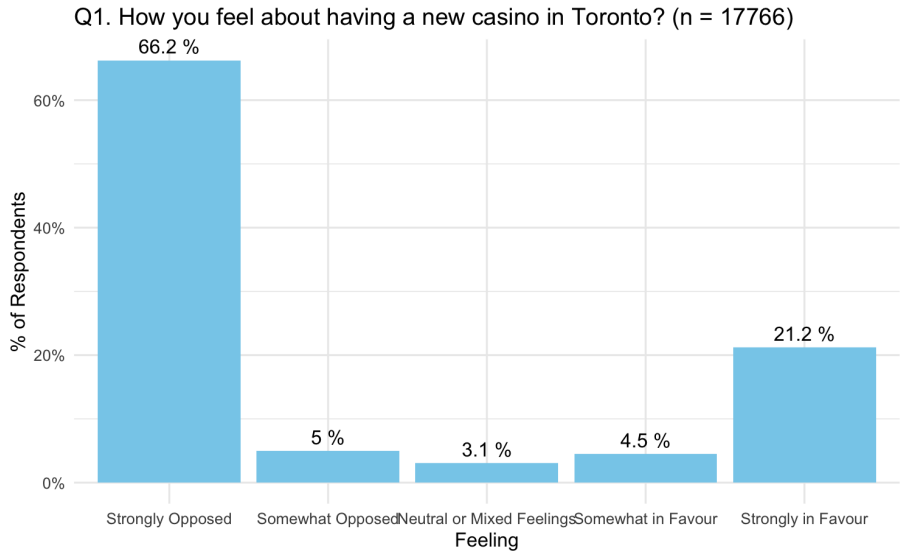

Figure 1: Feeling toward a proposed casino in Toronto

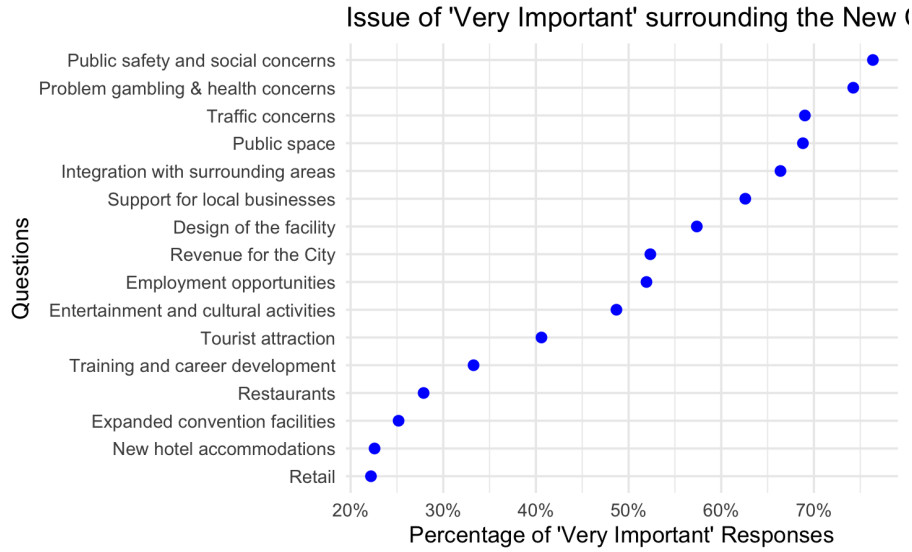

Figure 2: Percentage of 'Very Important' responses vs. each question

Questions Q3_ examined respondents' perceptions of various social and environmental concerns related to the casino, without offering a 'No Casino' option, to capture opinions on managing the casino's potential impacts. Figure 2 illustrates that issues such as safety, health, addiction, traffic, public space, and integration with surrounding areas are top priorities, with a high percentage of respondents rating these concerns as 'Very Important'. Addressing these concerns is critical for any future development plans to gain broader public support.

2.2 Multinomial Logistic Regression Model

The goal of this analysis was to predict public sentiment toward the proposed casino (Q1_A) using responses to questions Q3_A through Q3_P. Since Q1_A is a categorical variable with multiple response levels, a multinomial logistic regression model was employed, treating Q1_A as the dependent variable and the 16 responses (Q3_A to Q3_P) as independent predictors.

This model provides insights into how various factors influence the likelihood of respondents expressing different levels of sentiment, ranging from strong opposition to strong favorability. The detailed technical specifications and methodology are outlined in Appendix B.

Key Findings:

- The analysis identified Q3_N (Tourist Attraction), Q3_L (Revenue for the City), Q3_B (Employment Opportunities), and Q3_F (Hotel Accommodations) as the most significant predictors of strong support for the casino, with Q3_N as the most influential.
- The model demonstrated low misclassification rates for extreme sentiment categories such as "Strongly Opposed" and "Strongly in Favor". However, the misclassification was higher for intermediate categories like "Neutral" or "Mixed Feelings".
- Overall, the model shows good predictive power with clear patterns across predictors.

2.3 Principal Component Analysis (PCA) for Dimensionality Reduction

In Section 2.2, public sentiment toward the proposed new casino in Toronto was assessed using a regression model based on responses to survey questions Q3_A to Q3_P, incorporating 16 predictor variables. While comprehensive, the inclusion of all variables increased the model's complexity, potentially affecting its performance. To mitigate this, Principal Component Analysis (PCA) was introduced in this section to reduce dimensionality and evaluate whether this reduction could improve prediction accuracy.

First, categorical data were converted to numeric format, making the data suitable for PCA. Since all data were on the same scale, PCA was appropriate for reducing dimensions without requiring additional scaling.

PCA was performed by calculating eigenvalues to determine the importance of each principal component. Components with larger eigenvalues indicating higher importance.

After testing with various principal components and analyzing the scree plot (details

in Appendix C), two principal components were retained.

These two principal components captures approximately 61.87% of the original variance, significantly reducing the dataset's complexity while preserving most of the information. Principal component 1 is the most important one, accounting for 42.64% of information in the data. Principal component 2 explains 19.23% of the variance.

Detailed technical specifications are provided in Appendix C.

Key Findings:

- The PCA-based model with PC1 and PC2 is effective in predicting Q1_A, explaining about 61.87% of its variance, and has the same performance as the model using the original data.
- Using PCA reduces the dimensionality of the model without sacrificing any predictive power. Both models are statistically significant, and either approach would be appropriate depending on the context (e.g., interpretability vs. dimensionality reduction).

2.4 Exploratory Factor Analysis (EFA)

To explore the underlying structure of public opinions regarding a new casino in Toronto, I performed Exploratory Factor Analysis (EFA) on responses to survey questions Q3_A to Q3_P. The aim was to identify the main driving factors influencing the population's perceptions.

Tests was conducts and it is concluded that the data were suitable for EFA. Details technical specifications are provide in Appendix D.

The factor analysis performed using Maximum Likelihood (ML) extraction and Varimax rotation reveals several key insights. Here's insights from the results:

Main Driving Factors

1. **ML1 - Economic and Development:** This factor is driven by high loadings on variables such as

   - Q3_N: ("Tourist attraction"): 0.80
   - Q3_L ("Revenue for the City") : 0.76
   - Q3_C: ("Entertainment and cultural activities"): 0.76
   - Q3_B ("Employment opportunities"): 0.75
   - Q3_J ("Restaurants"): 0.73
   - Q3_F ("New hotel accommodations"): 0.71
   - Q3_D ("Expanded convention facilities"): 0.70
   - Q3_P: ("Training and career development"): 0.67
   - Q3_K ("Retail"): 0.67
   - Q3_A ("Design of the facility"): 0.64
   - Q3_M ("Support for local businesses"): 0.62
   - Q3_E: ("Integration with surrounding areas"): 0.53

   High loadings on items like "Tourist attraction", "Revenue for the City", "Entertainment and cultural activities", "Employment opportunities" and "Support for local businesses" suggest that Economic & Development factors are key drivers of opinions. It explains 37% of the variance in the data, making it the most significant factor.

2. **ML2 - Social and Public Concerns:** The second factor is characterized by loading on

   - Q3_E: ("Integration with surrounding areas"): 0.47
   - Q3_H ("Public safety and social concerns"): 0.92
   - Q3_G ("Problem gambling & health concerns"): 0.90
   - Q3_O ("Traffic concerns"): 0.68
   - Q3_I ("Public space"): 0.67

   Respondents scoring high on this factor reflect a more cautious focus on the potential negative social and health consequences of the new casino are likely to oppose the casino due to fears of increased crime, gambling addiction, and negative social impacts. It explains 20% of the variance.

The two factors together explain 57% of the total variance as show in Table 9.

This is meaningless comment. They are uncorrelated as you used orthogonal rotation. If you would try oblique rotation and then see significant correlation, that would be different story. I discussed that in class.

They appear to be distinct. ML1 and ML2 (0.01) are nearly uncorrelated, indicating they are distinct factors measuring separate constructs which is shown in Table 1

Table 1: Factor Correlations

	ML1	**ML2**
ML1	1.00	0.01
ML2	0.01	1.00

Base on the results, the factor analysis suggests that two factors are sufficient to capture the main dimensions of the survey data, with ML1 focusing on economic and development opportunities, and ML2 on social and public concerns.

2.5 Factors driven by Gender

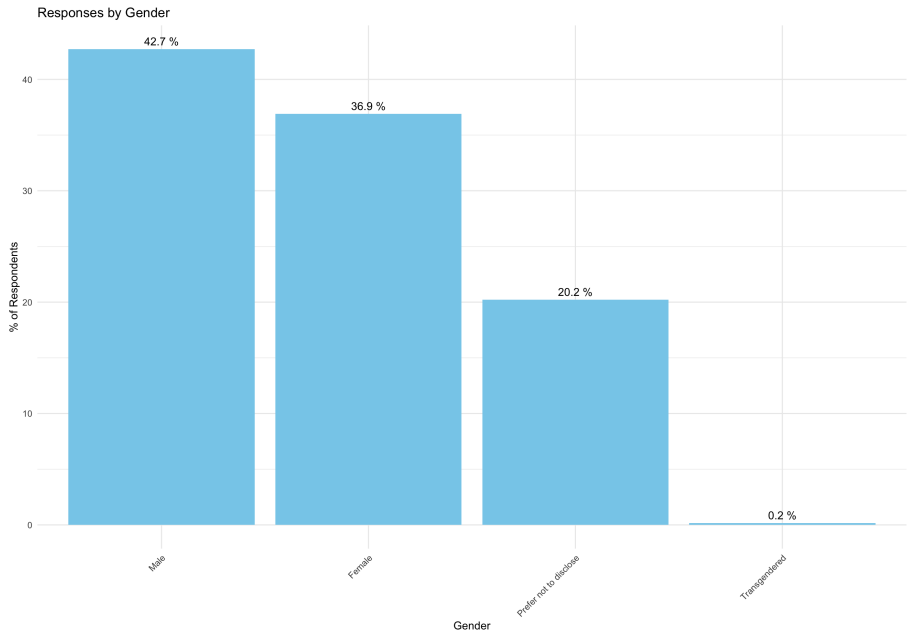

Figure 3: Distribution of Gender

This distribution indicates that the majority of the survey participants are either male or female with larger proportion of male (42.7%) compared to female (36.9%), with a small percentage transgendered and another 20.3% preferring not to disclose their gender.

To determine if there are differences in factors driven by gender, a t-test was conducted to analyze the key factors associated with gender differences. Detailed results are provided in Appendix E. Based on the analysis, the following conclusions were drawn:

- Males are more likely to be driven by factors related to economic and development opportunities offered by the casino, showing a generally more positive view.
- Females are significantly more concerned about social and public safety issues, showing higher caution towards the potential negative impacts on community health and safety.

These differences suggest that gender plays a critical role in shaping public opinion towards the new casino, with men prioritizing economic and development opportunities and women emphasizing social and public concerns. Understanding these differences can help tailor communication and address the specific concerns of each gender group more effectively.

3 Conclusion

The analysis aimed to understand public sentiment towards a proposed new casino in Toronto by examining survey responses through various statistical techniques. The use of multinomial logistic regression provided insights into how specific factors influence public opinion, revealing that respondents are strongly divided, with clear distinctions between those in favor and those opposed. The model performed well in predicting extreme sentiments but struggled with intermediate categories, highlighting areas for improvement in capturing more significant opinions.

To enhance model performance, Principal Component Analysis (PCA) was employed to reduce dimensionality and improve interpretability. The PCA-based regression model demonstrated same performance compared to the original model, with dimensionality reduction, PCA was effective in capturing the core elements of public sentiment.

Exploratory Factor Analysis (EFA) further identified two key factors driving public perceptions: economic and development opportunities, and social and public concerns. The analysis highlighted gender differences, with males prioritizing economic benefits and infrastructure development, while females were more concerned about social and public safety issues.

Overall, the findings underscore the complexity of public opinion on the new casino, shaped by diverse and sometimes conflicting factors. These insights provide valuable guidance for stakeholders to address specific concerns and better communicate the potential benefits and drawbacks of the casino to the public.

4 Appendices

Appendix A Data Processing

The data set has 17766 observations and 92 variables, primarily representing survey responses related to the proposal of building a new casino in Toronto. Answers were on a 5 point scale:

1. Strongly in Favour
2. Somewhat in Favour
3. Neutral or Mixed Feelings
4. Somewhat Opposed
5. Strongly Opposed

The objective of this analysis is to determine whether the City of Toronto should proceed with the construction of a new casino, based on public sentiment as captured in the survey responses. Specifically, I will focus on question Q1_A: "Please indicate on the scale below how you feel about having a new casino in Toronto," as a predictor of overall public opinion. Additionally, responses to questions Q3_A through Q3_P will be used to assess the influence of these factors on respondents' attitudes, which rate the importance of various factors related to the new casino on a 4-point scale below:

1. Very Important
2. Somewhat Important
3. Not Important At All
4. Unsure

The factors include:

- Design of the facility (Q3_A)
- Employment opportunities (Q3_B)
- Entertainment and cultural activities (Q3_C)
- Expanded convention facilities (Q3_D)
- Integration with surrounding areas (Q3_E)
- New hotel accommodations (Q3_F)
- Problem gambling & health concerns (Q3_G)
- Public safety and social concerns (Q3_H)
- Public space (Q3_I)

- Restaurants (Q3_J)
- Retail (Q3_K)
- Revenue for the City (Q3_L)
- Support for local businesses (Q3_M)
- Tourist attraction (Q3_N)
- Traffic concerns (Q3_O)
- Training and career development (Q3_P)

I also use Gender as as covariates for further data analysis after PCA/EFA.

The missing values was checked and result as table 2

Table 2: Missing Values Data

Q1_A	**Q3_A**	**Q3_B**	**Q3_C**	**Q3_D**	**Q3_E**	**Q3_F**	**Q3_G**	**Q3_H**
1	3307	3329	3415	3562	3291	3606	2567	2594
Q3_I	**Q3_J**	**Q3_K**	**Q3_L**	**Q3_M**	**Q3_N**	**Q3_O**	**Q3_P**	**Gender**
3023	3511	3610	3354	3324	3520	2899	3671	2892

I apply Mode Imputation technique to handle missing values for categorical variables:

1. Q1_A: Missing values were replaced with the most frequently occurring category, "Strong Opposed," as shown in Figure 1
2. Q3_A to Q3_P: Missing responses likely indicate uncertainty or lack of opinion about the factors. Imputing these values with "Unsure" appropriately captures this uncertainty
3. Age: Missing values were replaced with "Prefer not to disclose" to reflect respondents' privacy preferences and reluctance to provide their age

Below are variables were dropped:

- Q1_B1 to Q1_B3: Reasons for the rating in Q1_A (text responses)
- Q2_A and Q2_B: Fit of the new casino with the image of Toronto and reasons (text responses)
- Q3_Comments: Comments supporting selections in Q3_A to Q3_P (text responses)
- Q4_A to Q10: Additional questions and comments (text responses)
- Demographic Variables: Postal Code, GroupName, DateCreated

Table 3: Summary Statistics for Selected Variables

Variable	Not Imp.	Somewhat Imp.	Very Imp.	Unsure
Design of the facility (Q3_A)	3519	3212	10222	813
Employment opportunities (Q3_B)	3817	3998	9078	873
Entertainment (Q3_C)	4595	3807	8471	893
Convention facilities (Q3_D)	7699	4516	4318	1233
Integration (Q3_E)	2576	2535	11949	706
New hotel accommodations (Q3_F)	8545	4200	3853	1168
Gambling & health (Q3_G)	1747	1990	13610	419
Public safety & social (Q3_H)	1531	1947	13941	347
Public space (Q3_I)	1970	2670	12550	576
Restaurants (Q3_J)	7444	4578	4783	961
Retail (Q3_K)	8761	4005	3893	1107
Revenue for the City (Q3_L)	4161	3659	8989	957
Local businesses (Q3_M)	2827	3085	11102	752
Tourist attraction (Q3_N)	6651	3350	6719	1046
Traffic concerns (Q3_O)	1953	2766	12591	456
Training & career (Q3_P)	6106	4186	5952	1522

Summary Statistics:

Based on the provided summary statistics:

- **Not Important:** Responses in questions Q3_D, Q3_F, Q3_J, Q3_K, and Q3_P have a higher rate of "Not Important".
- **Very Importance:** Responses in all other questions have a higher rate of "Very Important".

Appendix B Multinominal logistic regression

1. **Coefficients:** The coefficients reveal the impact of each predictor on sentiment outcomes, Figure 4 present the coefficients of the model.

proper name

- Q3_N.L and Q3_B.L has a consistently large positive effect across all categories , highest value, 2.4786 for "Strongly in Favour", implying these predictors have strong influence on favourability.
- Q3_G.L has negative coefficients, especially for "Strongly in Favour" (-2.33), suggesting it decreases the likelihood of "Strong in Favour".

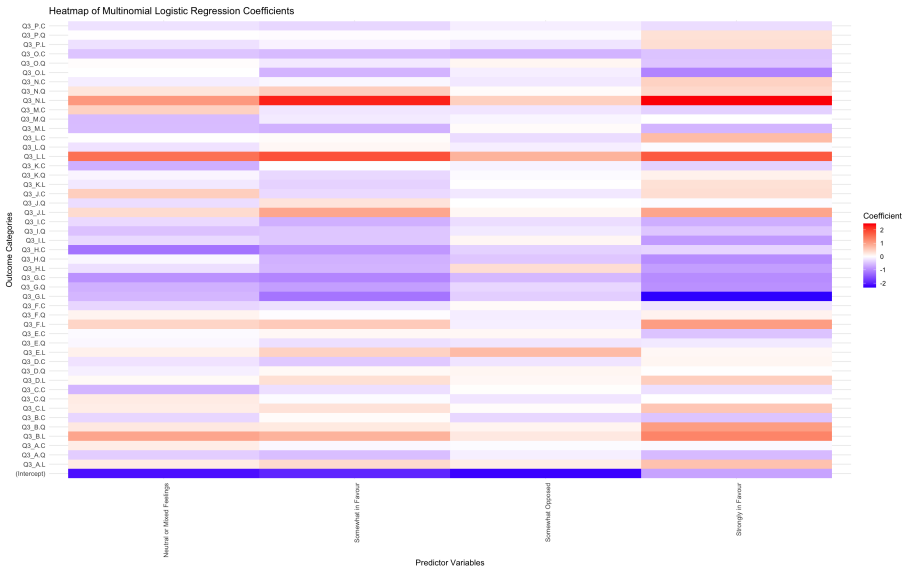

Figure 4: Heatmap of Multinominal Logistic Regression Coefficients

2. **Standard Errors:** Larger standard errors suggest more uncertainty in the estimate for that variable (e.g., Q3_L.C or Q3_M.C in "Strongly in Favour" at 0.236 or 0.2296).
3. The Residual Deviance: 19105.97
4. AIC: 19497.97
5. The predicted probabilities:

   From Table 4:

   - Observation 1 has a 97.37% probability of being "Strongly Opposed."
   - Observation 2 has an 89.18% chance of being "Strongly in Favour." and so on.
   -

   From Table 5:

Table 4: Predicted Probabilities of Sentiment Categories

	Strong Opp.	Somewhat Opp.	Neutral	Somewhat Fav.	Strong Fav.
1	0.9737	0.0174	0.0081	0.0005	0.0003
2	0.0020	0.0024	0.0136	0.0902	0.8919
3	0.0005	0.0003	0.0029	0.0317	0.9646
4	0.5781	0.2264	0.1106	0.0565	0.0284
5	0.2354	0.1188	0.0705	0.1677	0.4076
6	0.9892	0.0064	0.0038	0.0004	0.0001

Table 5: Confusion Matrix for Multinomial Model

			Actual		
Predict	**Strong Opp.**	**Some. Opp.**	**Neutral**	**Some. Fav.**	**Strong Fav.**
Strong Opp.	11255	786	337	243	343
Somewhat Opp.	7	4	5	0	1
Neutral	14	4	13	11	2
Somewhat Fav.	17	11	32	56	42
Strong Fav.	475	82	159	487	3380

- Out of 11,255 actual "Strongly Opposed" observations, 11,255 were correctly predicted, with minimal misclassification into other categories.
- Misclassifications are more frequent in categories like "Somewhat Opposed" and "Neutral or Mixed Feelings," where small sample sizes exist.

Appendix C Principal Components Analysis

Experiment tests results to choose number of Principal Components

PCA reduces the dataset's complexity while preserving as much variability as possible. By calculating the eigenvalues of the training dataset, components were scaled according to their importance, with larger eigenvalues indicating higher importance. As shown in Table 6, the first six components capture approximately 80% of the variance in the data.

Table 6: Importance of Components

	PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**
St. dev.	2.6922	1.8081	1.05015	0.83198	0.77343	0.74621
Pro. of Var.	0.4264	0.1923	0.06487	0.04072	0.03519	0.03275
Cum. Pro.	0.4264	0.6187	0.68353	0.72424	0.75943	0.79219

From the table, the data retain 61.87% of the original information after reducing the dimensionality by 87.5% (from 16 variables to 2).

The scree plot (Figure 5) suggests an optimal number of components to retain, as indicated by the "elbow" at the third principal component, suggesting that additional components contribute less to explaining variance.

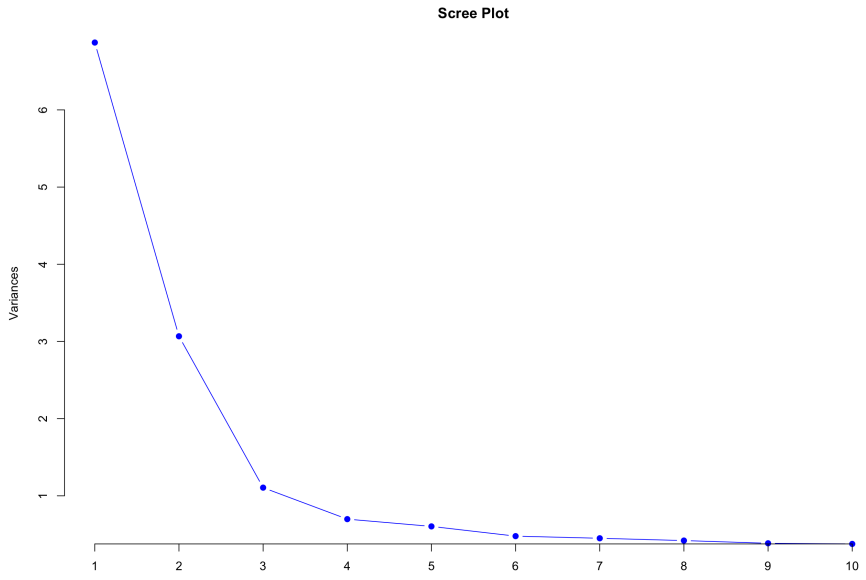

Figure 5: Screeplot

Kaiser's Criterion recommends retaining three principal components, while Jolliffe's Criterion suggests retaining between two and three.

Jolliffe's never suggest less than Kaiser's

After running models with both two and three components, retaining two principal components was determined to be the optimal choice. This choice balances capturing a significant amount of variance (approximately 61.87%) while reducing the dataset's complexity and proportion of large deviations.

Table 7: Loadings of Variables on Principal Components 1 and 2

Variable	**PC1**	**PC2**
Q1_A	0.2302	-0.2567
Q3_A	0.2687	0.0876
Q3_B	0.2975	0.0102
Q3_C	0.2960	-0.0109
Q3_D	0.2602	-0.1294
Q3_E	0.2494	0.2105
Q3_F	0.2581	-0.1856
Q3_G	0.0652	0.4718
Q3_H	0.0861	0.4692
Q3_I	0.1623	0.3641
Q3_J	0.2698	-0.1678
Q3_K	0.2465	-0.1768
Q3_L	0.2945	-0.0325
Q3_M	0.2720	0.1452
Q3_N	0.2875	-0.1747
Q3_O	0.1269	0.3783
Q3_P	0.2666	-0.0160

- PC1 is driven by a broader range of questions (Q3_B (Employment opportunities):0.2975, Q3_C (Entertainment):0.2960, Q3_L(Revenue for the City): 0.2945, Q3_N (Tourist attraction):0.2875), indicating these questions may represent a common underlying factor (Economic and development opportunities)
- PC2 is driven more strongly by specific questions (Q3_G (Gambling and health concern):0.4718, Q3_H (Public safety abd social concerns):0.4692, Q3_O (traffic concerns) :0.3782), possibly reflecting a different factor of the data (Safety and public concern).

Linear Regression on PCA Scores

1. Intercept (2.09): The average value of Q1_A is 2.09 when PC1 and PC2 are 0.
2. PC1 coefficient (0.383):For every unit increase in PC1, the predicted value of Q1_A increases by 0.383, holding PC2 constant.
3. PC2 coefficient (-0.427): For every unit increase in PC2, the predicted value of Q1_A decreases by 0.427, holding PC1 constant.

The high t-values and very low p-values (all ¡ 2e-16) for PC1 and PC2 suggest that these two principal components are both significant predictors of Q1_A.

Residual Analysis:

The residuals range from -2.98 to 4.56, with a median of -0.02, suggesting that the model is reasonably good at predicting Q1_A, with some errors in the predictions, this is understandable.

Model Fit:

- R-squared (0.5996): The model explains around 59.96% of the variance in Q1_A using the first two principal components (PC1 and PC2). This is a moderately strong fit, indicating that the PCA-based model captures a significant amount of the variability in the data.
- F-statistic (1.33e+04)and p-value (¡ 2.2e-16), so model is statistically significant.

A regression analysis using the selected principal components was compared with the original data model (Table 8).

Table 8: Comparison of PCA Model and Original Model

	PCA Model	**Original Model**
Adjusted R-squared	0.5996	0.5996
AIC Value	52279.76	52279.76

The PCA model is effective as the original model at explaining the variability in respondents' general feelings towards a new casino, as indicated by the same adjusted R-squared value and a AIC score.

Appendix D Exploratory Factor Analysis

Before conducting EFA, missing values were imputed, and the reliability of the dataset was assessed using Cronbach's alpha, which yielded a value of 0.906. All correlation coefficients were above 0.3, indicating internal consistency and suitability for factor analysis. The correlation matrix 6 showed two groups of variables with high positive correlations, suggesting underlying factor structures.

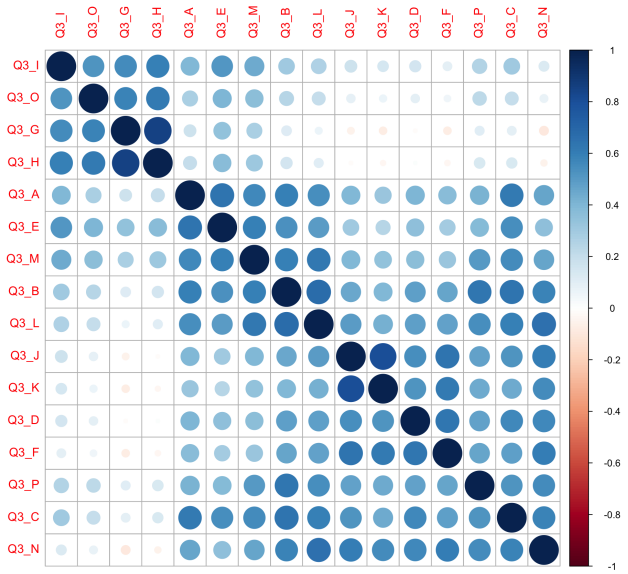

Figure 6: Correlation matrix

The determinant of the correlation matrix was 0.00002762, confirming that the data were not singular. Bartlett's Test of Sphericity resulted in a highly significant p-value, further supporting the appropriateness of factor analysis. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was greater than 0.75 for all variables, overall MSA = 0.92, indicating that the data were suitable for EFA.

The factor analysis was conducted using the maximum likelihood (ML) method with varimax rotation, which helps to identify the underlying structure among the items (Q3_A to Q3_P). The analysis extracted two factors (ML1, ML2) based on the correlation matrix.

- ML1 has high loading for items such as "Tourist attraction" (Q3_N), "Revenue for the City" (Q3_L), "Entertainment and cultural activities" (Q3_C), and "Employment opportunities" (Q3_B). These items suggest that ML1 represents factors related to Economic and Development.
- ML2 has high loadings for items like "Public safety and social concerns" (Q3_H), "Problem gambling & health concerns" (Q3_G), "Traffic concerns"

Table 9: Factor Analysis Using ML Method

Item	**ML1**	**ML2**	**h2**	**u2**	**com**
Q3_N	0.80		0.64	0.36	1.0
Q3_L	0.76		0.61	0.39	1.1
Q3_C	0.76		0.61	0.39	1.1
Q3_B	0.75		0.62	0.38	1.2
Q3_J	0.73		0.54	0.46	1.0
Q3_F	0.71		0.51	0.49	1.0
Q3_D	0.70		0.49	0.51	1.0
Q3_P	0.67		0.49	0.51	1.1
Q3_K	0.67		0.46	0.54	1.0
Q3_A	0.64		0.49	0.51	1.4
Q3_M	0.62		0.55	0.45	1.7
Q3_E	0.53	0.47	0.50	0.50	2.0
Q3_H		0.92	0.85	0.15	1.0
Q3_G		0.90	0.81	0.19	1.0
Q3_O		0.68	0.48	0.52	1.1
Q3_I		0.67	0.50	0.50	1.2
SS Loadings	**5.96**	**3.17**			
Proportion Var	**0.37**	**0.20**			
Cumulative Var	**0.37**	**0.57**			
Proportion Explained	**0.65**	**0.35**			
Cumulative Proportion	**0.65**	**1.00**			

(Q3_O) and "Public space" (Q3_I). These seem to align with Social and Public concern-related factors.

- Communality (h2): between 0.46 to 0.85, which is resonably. Items such as Q3_H (0.85), Q3_G (0.81), and Q3_N (0.64) show high communalities, explain a large portion of their variance.
- Uniqueness (u2): Q3_H with lowest value 0.15 is the most of the item's variance is captured by the model.
- Complexity (com) is mostly equal 1 or close to it.

Model Fit

The model has a Tucker Lewis Index (TLI) of 0.817, and the Root Mean Square Error of Approximation (RMSEA) is 0.126, indicating a good fit. The Root Mean Square of the Residuals (RMSR) is 0.06 indicates that the residuals are small.

All communality is greater than 0.45, which suggest that the extracted factors explain most of the variance in the variables being analysed.

Proportion of large deviations: Approximately 29% (below 50%) of the correlations matrix, which is expected.

Appendix E Factors driven by Gender

1. **Economic & Development Opportunities:**
   - **t-test Results:** The t-value is 10.722 with a p-value of $< 2.2e-16$, which is highly significant, indicating a statistically significant difference between males and females in how they perceive economic opportunities related to the new casino.
   - **Mean Scores:**
     - Mean in group Male: 0.1199
     - Mean in group Female: -0.0548
   - **Confidence Interval:** The 95% confidence interval for the difference in means is (0.1428, 0.2067).

   Males have positive, higher mean, suggesting that male may prioritize or have a stronger perception of economic benefits compared to females, possibly influenced by a greater interest in the economic promises associated with the casino.

2. **Social and Public Concerns:**
   - **t-test Results:** The t-value is -10.806 with a p-value $< 2.2e-16$, indicating a very significant difference between the genders.

- **Mean Scores:**
  - Mean in group Male: -0.03068
  - Mean in group Female: 0.13904
- **Confidence Interval:** The 95% confidence interval for the difference in means is (-0.2005, -0.1389).

Males, with a negative mean score, appear less concerned about Social and Public Concerns, highlighting a gender divide in the perception of social risks associated with the casino. Women may be more protective of community well-being and cautious about potential negative social outcomes.