SSRG International Journal of Civil Engineering
ISSN: 2348-8352/ https://doi.org/10.14445/23488352/IJCE-V11I7P102

Volume 11 Issue 7, 20-27, July 2024

Original Article

Multinomial Logit Model for Transportation Mode Choice: A Comprehensive Python Implementation with Scikit-Learn and Stats Models

Priya Hirave[1], Vidula Sohoni[2]

[1,2]Bharati Vidyapeeth (Deemed to be University) College of Engineering Pune, Maharashtra, India.

[1]Corresponding Author : pphirave@bvucoep.edu.in

Received: 05 May 2024 Revised: 05 June 2024 Accepted: 04 July 2024 Published: 31 July 2024

Abstract - *This paper examines the dynamics of transportation mode choices in response to evolving urban landscapes, utilizing the Multinomial Logit (MNL) model with the Python programming language. The aim is to provide valuable insights for researchers and practitioners across diverse contexts. As cities undergo continual transformations, understanding the connection between these changes and transportation choices becomes crucial. The adaptable Multinomial Logit model, implemented through Python, enhances the study's versatility and applicability. To fortify the research, two significant tools are strategically employed: scikit-learn for machine learning capabilities and stats models for comprehensive statistical analysis. The combined use of these tools seeks to unravel the nuanced dynamics of transportation mode choices. The paper meticulously details the methodology, offering a transparent exploration of the intricate relationships between urban evolution and transportation decisions.*

Keywords - *Evolving urban landscapes, Transportation choices, Multinomial logit model, Python programming, Scikit-learn, Stats models, Machine learning, Statistical analysis, City transformations, Modeling techniques, Valuable insights.*

1. Introduction

As cities undergo continuous expansion, the efficient allocation of transportation resources emerges as a critical imperative for sustaining growth, fostering economic development, and ensuring the overall well-being of citizens. Urban sprawl brings with it an array of challenges, including increased demands on transportation infrastructure and services. Efficiently managing these resources becomes paramount to addressing the ensuing complexities and maintaining the functionality of urban areas.

The term "efficient allocation of transportation resources" emphasizes the strategic management of various aspects of transportation, ranging from traffic flow and infrastructure usage to the optimization of public transit systems. It underscores the importance of minimizing congestion, reducing travel times, and enhancing the overall effectiveness of transportation networks. By achieving these goals, cities can mitigate the adverse environmental and economic impacts associated with inefficient transportation systems.

Sustainability stands out as a key focus in this context. As cities expand, there is a growing need to adopt environmentally conscious transportation practices. This involves promoting eco-friendly modes of transportation, reducing carbon emissions, and creating transportation systems that align with long-term ecological sustainability. Such sustainable transportation strategies are integral to ensuring that urban growth does not compromise environmental health but compromise the well-being of future generations.

Furthermore, the efficient allocation of transportation resources is intricately linked to economic development. An optimally functioning transportation system facilitates the movement of goods, services, and people, thereby supporting businesses, industries, and trade. Well-planned transportation infrastructure enhances connectivity, providing easier access to job opportunities, markets, and other economic hubs. This, in turn, contributes significantly to the economic vibrancy of expanding urban centres.

In the pursuit of the well-being of citizens, the efficient allocation of transportation resources plays a crucial role. A well-designed transportation system directly influences the quality of life for individuals within urban areas. It affects accessibility to essential services, healthcare, education, and recreational facilities. Moreover, by minimizing commute times and reducing stress associated with transportation, a thoughtful and efficient transportation system contributes positively to the mental and physical well-being of citizens.

Within the field of transportation planning, the statement introduces the concept of mode choice modelling as an essential tool. This modelling technique involves predicting and understanding individuals' preferences for different modes of transportation based on various factors such as travel time, cost, convenience, and personal preferences. Mode choice modelling serves as a foundational element for urban planners, providing valuable insights to inform decisions related to the development and optimization of transportation infrastructure. It enables planners to anticipate how people will choose between different transportation options, including cars, public transit, cycling, or walking, thus contributing to more informed and sustainable urban development.

2. Literature Review

The focus of this literature review is to explore the dynamics of travel demand and modal split analysis for work-related trips within urban settings, with a specific emphasis on the case of Pune City [1]. By synthesizing existing research findings, theoretical frameworks, and empirical studies, this literature review aims to provide a comprehensive understanding of the factors influencing commuters' mode choice behavior, the implications of transportation infrastructure developments, and the potential strategies to promote sustainable transportation options.

Current studies on transportation mode choice often overlook the specific needs and perceptions of users regarding work trips in metro cities. There is a need for a focused evaluation of how individuals choose their mode of transportation for commuting to work. Understanding these preferences is crucial for successfully implementing a new transport system that encourages the shift from private to public transport [2]. The novelty of this research is in evaluating mode choice for work trips in a metro city, aiming to shift commuters from private to public transport through user-centric insights.

3. Multimodal Transport Characteristics

Multimodal transport systems are characterized by their integration, efficiency, and convenience. They offer seamless connectivity between different modes of transport, such as buses, trains, and bicycles, allowing for easy transfers and unified ticketing systems that simplify the travel process. These systems are designed to optimize routes and schedules, minimizing travel and wait times while ensuring wide coverage, including access to underserved areas [6]. Flexibility is a key feature, providing multiple transport options to suit various needs and preferences. Sustainability is also a major focus, promoting eco-friendly modes like public transport, cycling, and walking. Multimodal transport systems are cost-effective, employing integrated pricing strategies to keep travel affordable. High safety standards, reliable and punctual service, and an emphasis on user experience through enhanced comfort and amenities further define these systems, making them a practical and attractive choice for urban commuters [7].

4. Modal Shift

The research focuses on the dynamics of Pune city's transportation landscape, particularly with the impending introduction of the Pune Metro project. Anticipation is high regarding the Metro's potential to mitigate the city's traffic woes. However, its success hinges not only on its infrastructure but also on understanding the preferences and behaviors of potential users. To this end, a comprehensive assessment of the existing public transportation systems is imperative. Factors such as travel duration, fare affordability, waiting times, and daily ridership need to be thoroughly analyzed to identify areas for enhancement. Given that the Metro project is still underway, there remains uncertainty regarding its efficacy. Therefore, conducting a Stated Preference Survey, particularly along proposed Metro corridors, can provide invaluable insights into passenger perceptions and preferences. Such data can be pivotal in shaping the integration of the Metro into Pune's transportation ecosystem. Achieving a shift from private vehicles to public transport demands that the latter offers distinct advantages, including faster travel times, seamless connectivity, reduced stress, and affordability. Establishing an Integrated Public Transport system that caters to diverse demand levels is essential for facilitating this transition effectively. Ultimately, the research aims to offer actionable insights to optimize the Metro's performance and alleviate traffic congestion in Pune and other metropolitan areas.

5. Stated Preference Survey

The effectiveness of policies and their subsequent success are intricately tied to user awareness and adherence. If users are knowledgeable about a policy and align their behavior accordingly, it is easier to predict its success post-implementation. However, when users are unaware of a policy, its success relies on their perception of it and their willingness to comply. This perception is shaped by various factors, including users' level of knowledge about the policy, their current behavior patterns, and their demographic characteristics. These factors collectively influence users' attitudes towards the policy [3].

The theory of reasoned action posits that individuals are rational decision-makers who systematically utilize available information to guide their behavior. In this framework, understanding how attitudes towards a policy translate into actual behavior are crucial. To achieve this understanding, it is essential to employ a multidisciplinary approach that combines insights from psychology and sociology within specific situational contexts [4].

These situational contexts, often referred to as scenario conditions, play a significant role in shaping users' perceptions and responses to policies [8].

The materials and methods section should contain sufficient detail so that all procedures can be repeated. It may be divided into headed subsections if several methods are described.

6. Data Sampling

Statistics is a field dedicated to collecting, analysing, and interpreting data to derive valuable insights. It serves as a vital tool in decision-making, particularly when faced with uncertainty. By quantifying and assessing the level of uncertainty associated with measured data, statistics empowers decision-makers to make informed choices. Typically, data is collected through sampling—a process of gathering observations from a larger population that is impractical to observe entirely. These observations pertain to various attributes, such as income, for each member of the population. Through statistical inference, conclusions can be drawn about population parameters, such as the mean value of these attributes. The design of the sample is crucial, aiming to maximize the information obtained about the population while minimizing costs. However, challenges arise in ensuring the sample is truly representative and in drawing valid conclusions from the data. Addressing these difficulties is essential for accurate decision-making and meaningful insights.

6.1. Methods of Sampling

Most widely accepted sampling methods rely on random sampling, where each unit in the population has an equal chance of being selected independently. Among these methods, two notable approaches stand out:

- Simple Random Sampling: This method, the simplest of all, involves assigning a unique identifier to each unit in the population and then randomly selecting these identifiers to form the sample. However, it may require large sample sizes to ensure sufficient data representation, especially for minority options of interest. For instance, randomly sampling households in a developing country might not provide adequate information on rare occurrences like multiple car ownership.
- Stratified Random Sampling: Here, the population is divided into homogeneous subgroups (strata) based on prior information. Simple random sampling is then conducted within each stratum using the same sampling rate. This approach ensures that the sample maintains the correct proportions of each subgroup, making it crucial for cases with small, distinct subgroups. It is also possible to stratify based on multiple variables, creating a multidimensional matrix of group cells. However, caution is needed to avoid creating too many cells, as it can lead to small sample sizes per cell. Despite its benefits, stratified sampling may not be effective when data is needed on options with low probabilities, requiring a third method.
- Choice-Based Sampling: This method, a subset of stratified sampling, stratifies the population based on the outcome of the choice process under consideration. It offers cost-effective data collection but may introduce bias into the sample, as it is not entirely random. Therefore, there is a risk of biased estimates when extrapolating from the sample to the entire population.

6.2. Sample Size Calculation

The sample size was calculated by using the Simple Random Sampling Technique.

$$\boldsymbol{n = z^2 \times p\,(1-p) \in^2 N}$$

Where, Z = Z Score,
ε = Margin of Error
N = Population Size
p = Population Proportion

The population size is 1,15,248 which is daily trips to Hinjewadi.
The confidence level is 85%,
So, Z = 1.44.
The margin of error is 5%.
Considering the 50% population proportion, the sample size is 207.

7. Multinomial Logit Model for Mode Choice

The Multinomial Logit (MNL) model is a statistical tool used in different fields like economics and marketing. It helps predict choices when people have to choose between multiple options.

The model works by measuring the satisfaction or desirability associated with each option. It assumes that the attractiveness of one option is not affected by the presence or absence of other options. The model calculates the probability of choosing one option over another based on their attractiveness [5].

The Multinomial Logit (MNL) model formulates the probability of choosing a specific alternative among a set of alternatives using a multinomial logistic function. This probability is expressed as a ratio involving the exponential of the utility associated with that alternative and the sum of exponentials across all available alternatives. The utility is modelled as a linear combination of explanatory variables and their associated coefficients. The general formula for the probability in the MNL model is:

$$P_n(i) = \frac{e^{Vin}}{\sum_{j_e c_n} v_{in}}$$

Where,
P(i) = Probability that mode-i is chosen
U(i) = Utility associated with the choice of alternative-i

7.1. MNL Using Python with Scikit-Learn and Stats Models

Python is extensively utilized for Multinomial Logit (MNL) modelling through popular libraries like Scikit-learn and Stats models. To implement MNL modelling with Scikit-Learn, first, the data needs to be prepared by creating a feature matrix `X` and a target array `y`. Then, a `LogisticRegression` model object is created with `multi_class='multinomial'` and an appropriate solver, followed by fitting the model to the data. Predictions can then be made using the fitted model.

On the other hand, Stats models require a similar data preparation step, after which an `MN Logit` model object is created with the target and feature data. The model is then fitted to the data, and a summary of the model can be printed. Finally, predictions can be generated using the fitted model. Both libraries offer powerful tools for MNL modeling in Python, with Scikit-learn being more user-friendly and suitable for machine learning tasks, while Stats models provide more detailed statistical analysis and summary output, catering to different preferences and requirements.

7.2. Model Development

7.2.1. Data Collection

The survey was meticulously planned and executed in two distinct phases: roadside interviews and an online questionnaire.

- Roadside Interviews: This initial phase involved volunteers conducting interviews at key locations in Hinjewadi Phase I, II, and III. Individuals from diverse demographics, spanning students, working professionals, retirees, and homemakers, were approached for their insights. A total of 207 commuters graciously participated in these face-to-face interviews, engaging with the volunteers and providing valuable feedback.
- Online Questionnaire: Concurrently, an online component was introduced to widen the reach of the survey. Utilizing Google Sheets, an electronic questionnaire was disseminated via email, Facebook, and WhatsApp channels. This approach aimed to capture responses from a broader audience. Ultimately, 50 participants responded to the online survey, adding another layer of data to the study.

Following the completion of both phases, the amassed data from the roadside interviews and online questionnaire was meticulously collated and analysed. In total, 207 responses were gathered, providing a comprehensive dataset for the subsequent analysis and interpretation.

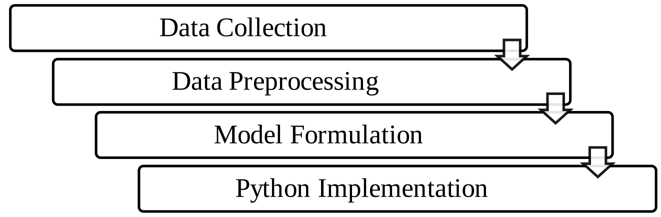


Fig. 1 Methodology of model development

7.2.2. Data Preprocessing

Data preprocessing is an essential step in preparing the dataset for analysis with the Multinomial Logit (MNL) model. This involves addressing common issues such as missing values and encoding categorical variables to ensure the dataset's suitability for modelling. Missing values are handled through techniques like imputation, where incomplete records are filled with estimated values to maintain dataset integrity. Meanwhile, categorical variables are encoded into numerical format, typically using methods like one-hot encoding, to enable the model to capture the influence of categorical variables on choice outcomes effectively. This preprocessing step is crucial for ensuring that the MNL model can appropriately analyze the dataset and produce reliable results.

Furthermore, standardizing features plays a vital role in enhancing the robustness and interpretability of the MNL model. Standardization involves scaling numerical variables to a standard scale, which facilitates the comparability of coefficients and enhances numerical stability during model estimation. By standardizing features, the model ensures that coefficients associated with each variable are directly comparable, enabling a clearer interpretation of their relative impact on choice probabilities. This step not only improves the model's performance but also contributes to its transparency and reliability, making the Multinomial Logit model a valuable tool for analyzing choice behaviours in diverse contexts.

7.2.3. Model Formulation

The Multinomial Logit (MNL) model is like a smart tool that helps us understand how people choose one thing when they have different options. Imagine you are picking a mode of transportation, like a car or bus. The MNL model looks at the chances of you choosing each option by considering how much you like or prefer each one. It is clever because it understands that you can only pick one option, even though there are many to choose from. This model is handy for figuring out why people make specific choices, whether it is about transportation or picking products.

Now, let us talk about the behind-the-scenes stuff. The presentation not only explains how to use the MNL model but also talks about the theory behind it. It dives into the math part, talking about things like the Independence of Irrelevant Alternatives (IIA). This is a bit like saying your choice between two things will not change just because more options are available. This theory helps us use the MNL model better and understand its strengths and limits when we are studying how people make choices. So, it is like having a good understanding of the theory before using this cool tool to analyze real-world decisions.

7.2.4. Python Implementation

Implementing a Multinomial Logit (MNL) model using Python involves the following steps: Utilizing the scikit-learn and stats models libraries for machine learning and statistical analysis, respectively.

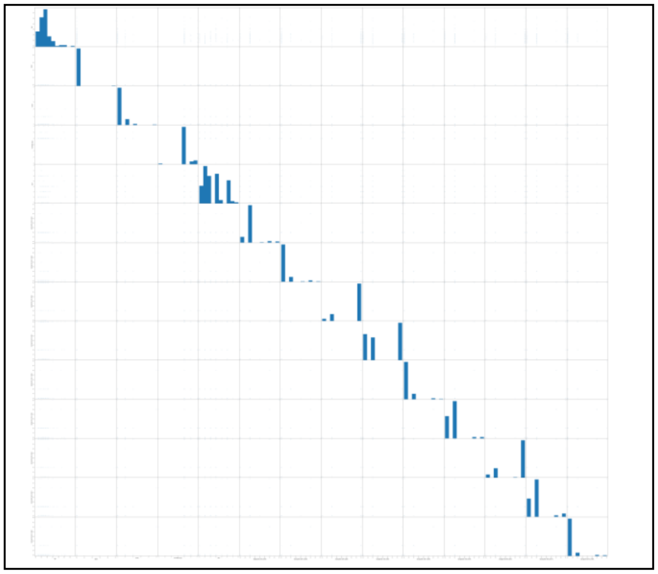
a) Import Libraries: Begin by importing the necessary libraries. For this implementation, you will need scikit-learn for machine learning tasks and stats models for detailed statistical analysis.
b) Load and Preprocess the Dataset: Load your dataset into a Pandas Data Frame and preprocess it. Handle missing values, encode categorical variables (e.g., using one-hot encoding), and standardize numerical features.
c) Train the MNL Model with scikit-learn: Use the scikit-learn Logistic Regression module for training the Multinomial Logit model.
d) Evaluate the Model: Assess the performance of the trained model using test data. You can use metrics like accuracy, precision, and recall.
e) Stats models for Detailed Analysis: For a more in-depth analysis, utilize stats models. This includes extracting and interpreting coefficients, standard errors, and p-values.
f) Interpret Results: Interpret the results obtained from the MNL model. Examine coefficients and their significance to understand the impact of different variables on the choice probabilities.

This step-by-step process provides a practical guide for implementing a Multinomial Logit model in Python using scikit-learn for machine learning tasks and stats models for detailed statistical analysis.

8. MNLogit Model in Python

The screenshots of the multinomial logit model developed in Python by using scikit-learn and the stats model are given below.

```
['Age', 'VOCR', 'VOTW', 'FOT-Week Days', 'MFIF', 'Saving time 25%-2.0 BF1', 'Saving time 25%-
1.75BF1', 'Saving time 25%-2.5BF1', 'Saving time 33%-2.5BF1', 'Saving time 33%-1.75BF1', 'Savi
ng time 33%-2.0 BF1', 'Saving time 50%-2.5BF1', 'Saving time 50%-2.0 BF1', 'Saving time 50%-1.
75BF1']
(180, 14)
(78, 14)
(180,)
(78,)
     Age  VOCR  VOTW  FOT-Week Days  MFIF  Saving time 25%-2.0 BF1  \
39    31     1     1              5   4.0                      2.0
96    31     1     1              5   4.0                      2.0
239   21     1     2              5   3.0                      1.0
123   27     1     1              5   4.0                      1.0
134   29     1     1              5   3.0                      1.0
..   ...   ...   ...            ...   ...                      ...
144   25     1     1              5   2.0                      2.0
118   24     1     3              5   6.0                      2.0
230   35     1     1              5   4.0                      2.0
189   30     1     1              5   4.0                      2.0
```



```
MNLogit Regression Results
==============================================================================
=============
Dep. Variable:                    SEX   No. Observations:
180
Model:                        MNLogit   Df Residuals:
165
Method:                           MLE   Df Model:
14
Date:                Wed, 17 Jan 2024   Pseudo R-squ.:
0.2362
Time:                        18:51:16   Log-Likelihood:
-74.692
converged:                      False   LL-Null:
-97.789
Covariance Type:            nonrobust   LLR p-value:
2.607e-05
==============================================================================
==========================
==========================
                     SEX=M       coef    std err          z      P>|
z|      [0.025      0.975]
-----------------------------------------------------------------------------
-------------------------
const                        -25.0972   1.01e+04     -0.002      0.9
98   -1.98e+04    1.97e+04
Age                            0.2964      0.087      3.398      0.0
01       0.125       0.467
VOCR                          11.4297   1.01e+04      0.001      0.9
99   -1.98e+04    1.98e+04
VOTW                           0.4227      0.745      0.567      0.5
70      -1.037       1.883
FOT-Week Days                  0.8653      0.562      1.540      0.1
24      -0.236       1.967
MFIF                          -0.3022      0.149     -2.023      0.0
43      -0.595      -0.009
Saving time 25%-2.0 BF1       -1.4421      1.046     -1.379      0.1
68      -3.492       0.608
Saving time 25%-1.75BF1        1.3151      1.140      1.154      0.2
49      -0.919       3.550
Saving time 25%-2.5BF1        -0.4032      0.295     -1.368      0.1
71      -0.981       0.174
Saving time 33%-2.5BF1         0.0472      0.262      0.180      0.8
57      -0.466       0.560
Saving time 33%-1.75BF1        1.6138      1.464      1.102      0.2
70      -1.256       4.483
Saving time 33%-2.0 BF1        0.1602      1.004      0.160      0.8
73      -1.807       2.127
Saving time 50%-2.5BF1         0.5742      0.272      2.111      0.0
35       0.041       1.107
Saving time 50%-2.0 BF1        1.5372      0.670      2.295      0.0
22       0.224       2.850
Saving time 50%-1.75BF1       -0.7807      0.840     -0.930      0.3
53      -2.427       0.865
==============================================================================
==========================
                           Results: MNLogit
=============================================================================
```

```
   Results: MNLogit
=================================================================
================
Model:                     MNLogit                 Method:
MLE
Dependent Variable:        SEX                     Pseudo R-squared:
0.236
Date:                      2024-01-17 18:51        AIC:
179.3845
No. Observations:          180                     BIC:
227.2789
Df Model:                  14                      Log-Likelihood:
-74.692
Df Residuals:              165                     LL-Null:
-97.789
Converged:                 0.0000                  LLR p-value:
2.6066e-05
No. Iterations:            35.0000                 Scale:
1.0000
-----------------------------------------------------------------
----------------
-----------------------------------------------------------------
----------------
        SEX = 0          Coef.   Std.Err.     t    P>|t|    [0.
025     0.975]
-----------------------------------------------------------------
----------------
                  const -25.0972 10086.9670 -0.0025 0.9980 -19795
.1892 19744.9948
                    Age   0.2964     0.0872  3.3982 0.0007      0
.1255     0.4674
                   VOCR  11.4297 10086.9667  0.0011 0.9991 -19758
.6618 19781.5212
                   VOTW   0.4227     0.7449  0.5675 0.5704     -1
.0373     1.8828
          FOT-Week Days   0.8653     0.5620  1.5397 0.1236     -0
.2362     1.9667
                   MFIF  -0.3022     0.1494 -2.0228 0.0431     -0
.5950    -0.0094
Saving time 25%-2.0 BF1  -1.4421     1.0458 -1.3789 0.1679     -3
.4919     0.6077
Saving time 25%-1.75BF1   1.3151     1.1400  1.1535 0.2487     -0
.9194     3.5495
 Saving time 25%-2.5BF1  -0.4032     0.2947 -1.3680 0.1713     -0
.9809     0.1745
 Saving time 33%-2.5BF1   0.0472     0.2618  0.1803 0.8569     -0
.4660     0.5604
Saving time 33%-1.75BF1   1.6138     1.4641  1.1023 0.2704     -1
.2558     4.4834
Saving time 33%-2.0 BF1   0.1602     1.0035  0.1597 0.8732     -1
.8066     2.1270
 Saving time 50%-2.5BF1   0.5742     0.2720  2.1108 0.0348      0
.0410     1.1075
Saving time 50%-2.0 BF1   1.5372     0.6698  2.2950 0.0217      0
.2244     2.8500
Saving time 50%-1.75BF1  -0.7807     0.8398 -0.9296 0.3526     -2
.4265     0.8652
=================================================================
================
```

9. Analysis and Performance Evaluation

9.1. Gender Preference Analysis

Gender preference analysis examines the tendencies and choices influenced by gender in various contexts.

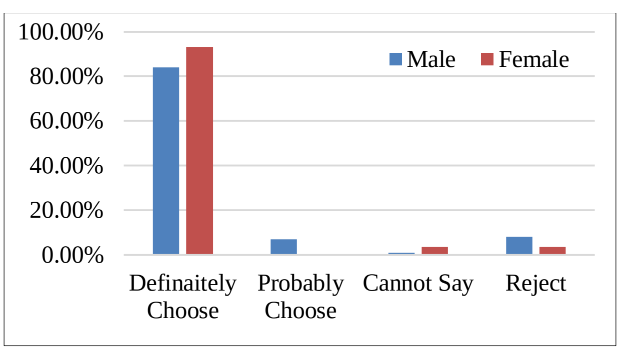


Fig. 2 Gender preference analysis

9.2. Monthly Family Income Preference Analysis

It evaluates how varying income levels influence consumer behavior and decision-making patterns.

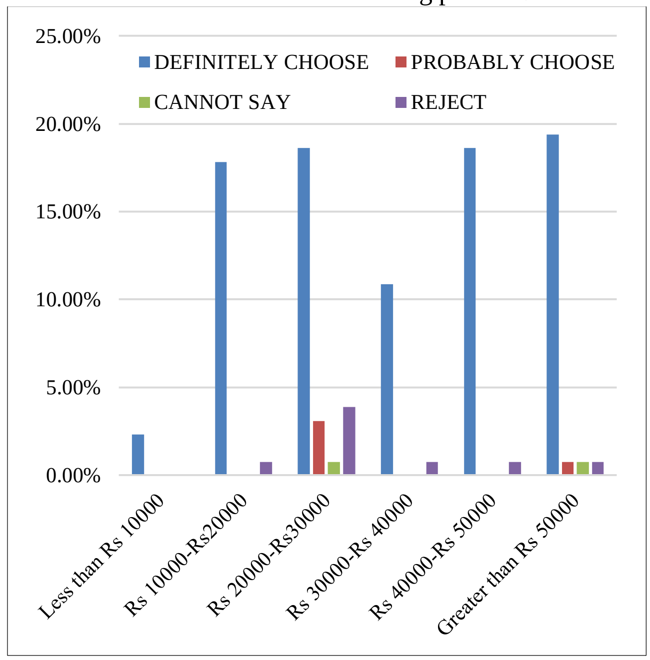


Fig. 3 Monthly family income preference analysis

9.3. Travel Slab Preference Analysis

Travel slab Preference Analysis investigates preferences and behaviors related to different travel expense categories or budget ranges.

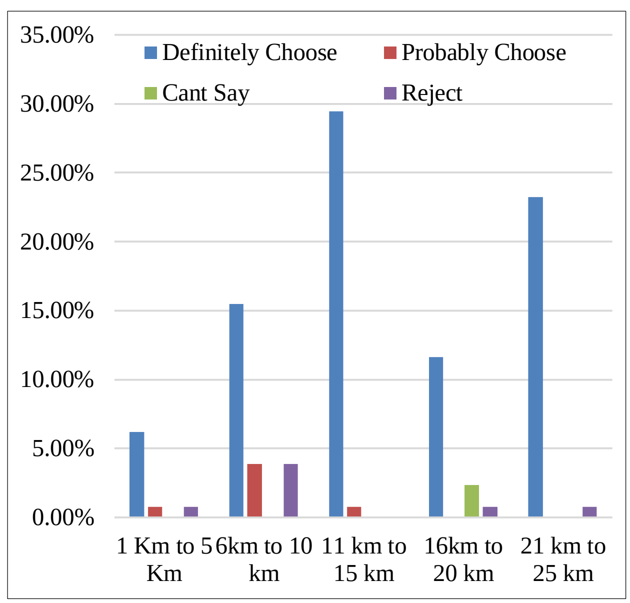


Fig. 4 Travel slab preference analysis

9.4. Age Group Preference Analysis

Age group preference analysis explores how different age demographics exhibit varying preferences and behaviors in consumer choices, services, or products.

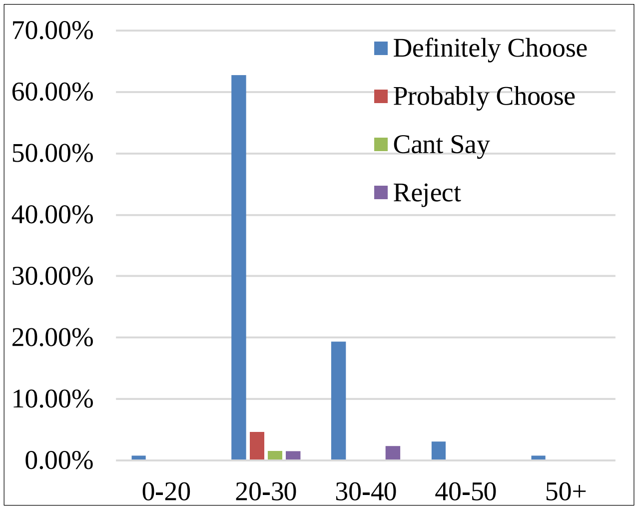


Fig. 5 Age group preference analysis

9.5. Frequency Preference Analysis

Frequency preference analysis examines patterns and tendencies in how often individuals choose or engage with specific activities, behaviors, products, or services.

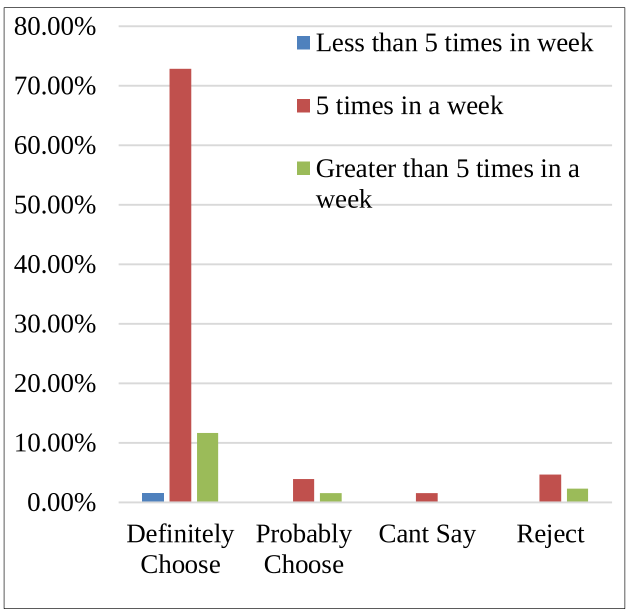


Fig. 6 Frequency preference analysis

9.6. Trip Purpose Preference Analysis

Trip purpose preference analysis investigates how individuals prioritize and choose transportation modes based on the specific purpose of their trips, such as commuting, leisure, or business travel.

9.7. Travel Time Saved Preference Analysis

Travel time saved preference analysis examines how individuals value and prioritize transportation options based on the amount of time they can potentially save compared to other modes of travel.

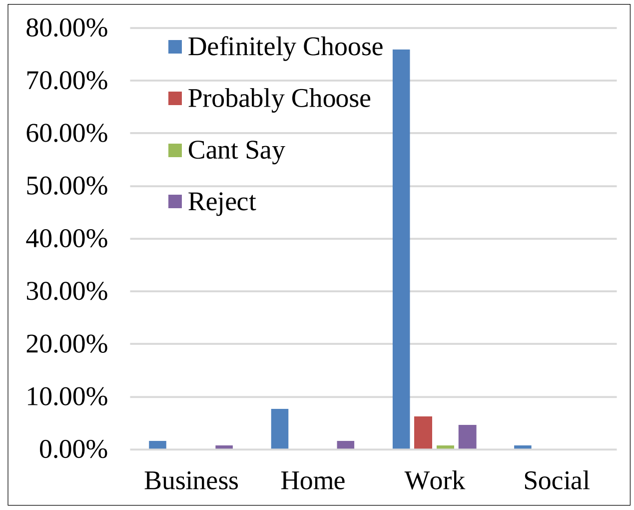


Fig. 7 Trip purpose preference analysis

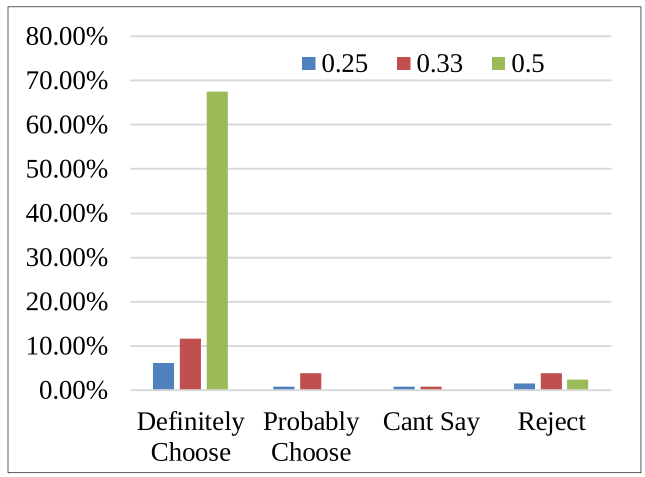


Fig. 8 Travel time saved preference analysis

9.8. Metro Fare Hike Preference Analysis

Metro fare hike preference analysis explores how commuters respond to increases in metro fare rates, examining factors such as ridership changes, mode-switching behaviors, and overall satisfaction with public transportation services.

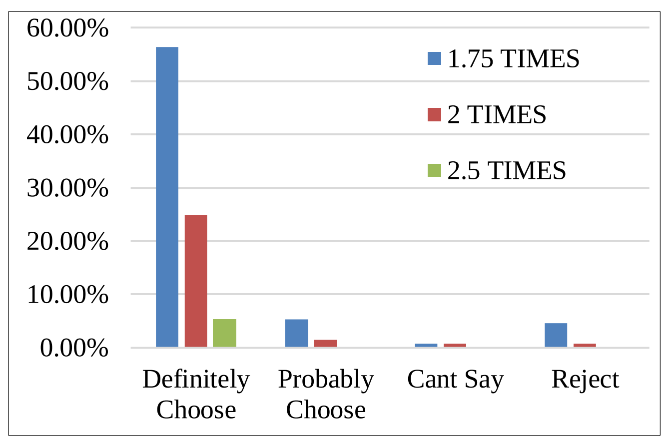


Fig. 9 Metro fare hike preference analysis

10. Interpretation and Results

Based on the provided information, it can be inferred that the Metro system, particularly the Mass Rapid Transit (MRT), will play a crucial role in facilitating modal shifts for work trips, especially for commuters traveling long distances, particularly above 10 kilometers, to locations like Hinjewadi. The following key observations and recommendations can be drawn from the study:

- Willingness to Switch to Metro: Both male and female commuters show a significant willingness to switch to the Metro system, with percentages of 84% and 93%, respectively. This indicates a strong potential demand for Metro services among the commuting population.
- Distance Preference and Time Savings: Commuters exhibit a preference for the Metro system for longer distances, with higher percentages of willingness to switch observed for distances exceeding 10 kilometers. Additionally, the study highlights the importance of time savings in commuters' mode choice decisions, with a significant proportion (83%) expressing a preference for the Metro if travel time is reduced by 50%.
- Peak Hour Frequency: Recognizing the importance of timing in commuters' travel decisions, the study emphasizes the need for a higher frequency of Metro services during peak hours, particularly for work-related trips. This aligns with the observation that a majority (87%) of commuters traveling for work would definitely choose the Metro.
- Age Group Preferences: The study identifies specific age groups that exhibit a higher propensity to prefer the Metro system for their work trips, with percentages ranging from 85% to 100%. This demographic insight can inform targeted marketing and promotional efforts to encourage Metro usage among these age cohorts.
- Modal Split for Short Distances: For commuters traveling shorter distances, particularly below 10 kilometers, the study suggests the provision of modal split options such as buses, non-motorized transport, bicycles, and walking. This underscores the importance of a comprehensive and integrated multi-modal transportation system to cater to diverse commuting needs.
- Policy Recommendations: To facilitate the successful implementation of the Metro system and promote the modal shift from private vehicles to public transport, the study advocates for policies aimed at minimizing travel costs, enhancing accessibility for lower-income groups, and improving the efficiency of the multi-modal transport network. Such policies can help maximize the Metro system's utility and accessibility, thereby encouraging its widespread adoption and usage.

11. Conclusion

The findings of the study indicate a strong potential for the Metro system, particularly the Mass Rapid Transit, to serve as a major mode of transport for work trips, especially for commuters traveling longer distances to locations like Hinjewadi. By addressing key factors such as time savings, frequency of services, demographic preferences, and policy considerations, stakeholders can effectively harness the transformative potential of the Metro system and realize its role in facilitating sustainable and efficient urban mobility.

References

[1] Kaveh Shafiee et al., "Modeling and Simulation of Vehicular Networks," *DIVANet '11: Proceedings of the First ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, Miami Florida USA, pp. 77-86, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[2] Maja M. Popović, and Jadranka J. Jović, "Concept of Expert System for Modal Split in Transportation Planning," *Yugoslav Journal of Operations Research*, vol. 16, no. 1, pp. 107-124, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[3] Li Jie et al., "Calibration of a Micro Simulation Program for a Chinese City," *Procedia - Social and Behavioral Sciences*, vol. 20, pp. 263-272, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[4] Michael P. Hunter et al., "An Investigation of Real-Time Dynamic Data Driven Transportation Simulation," *Proceedings of the 2006 Winter Simulation Conference*, Monterey, CA, USA, pp. 1414-1421, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[5] Muhammad Moazzam Ishaque, and Robert B. Noland, "Pedestrian and Vehicle Flow Calibration in Multimodal Traffic Microsimulation," *Journal of Transportation Engineering*, vol. 135, no. 6, pp. 338-348, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[6] Li-Jun Tiana, and Hai-Jun Huang, "Modeling the Modal Split and Trip Scheduling with Commuters' Uncertainty Expectation," *European Journal of Operational Research*, vol. 244, no. 3, pp. 815-822, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[7] Gordana Štefančić, Ninoslav Nogić, and Drago Bago, "Models in Planning Urban Public Passenger Transport," *Promet-Traffic&Transportation*, vol. 19, no. 4, pp. 259-262, 2007. [Google Scholar] [Publisher Link]

[8] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification," *IEEE Std*, vol. 12, no. 11, pp. 260-280, 1997. [Google Scholar]