Global Scientific JOURNALS

GSJ: Volume 12, Issue 11, November 2024, Online: ISSN 2320-9186
www.globalscientificjournal.com

Geospacial analysis and predictive modelling for urban transportation in Akure, Nigeria

Aderinola Olumuyiwa Samson[a], *Faloyo Taye John[b], Opeyemi David[c]

[a]Civil and Environmental Engineering Department, Federal University of Technology Akure, Ondo State, Nigeria
[b]Civil and Environmental Engineering Department, Federal University of Technology Akure, Ondo State, Nigeria
[c]Civil and Environmental Engineering Department, Olusegun Agagu University of Science and Technology, Ondo State, Nigeria

Corresponding author email: help2real@yahoo.com*

A B S T R A C T

Rapid urbanization in developing cities presents unique challenges for transportation systems. This study investigates urban transportation patterns in Akure, Nigeria, using geospatial analysis and machine learning to inform sustainable planning strategies. A two-pronged approach was employed. First, geospatial analysis involved surveying 440 Akure residents (age, occupation, income, preferred transportation). Additionally, geospatial data on trip volumes within high, medium, and low-density zones was obtained. Second, machine learning models (Random Forest and XGBoost) were trained on the data to predict transportation choices. Model performance was evaluated based on classification accuracy, error metrics (mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE)), and R-squared value. Analysis revealed young adults (18-30 years old) comprised the largest demographic group (37.5%), with self-employment being most common (32.0%). The gender distribution favoured males (54.4%), and 27.7% earned 50,000-100,000 Naira monthly. Taxis were the most popular mode (30.6%), followed by motorcycles (23.5%) and personal cars (28.2%). Spatial density significantly influenced mode selection. High-density zones (4,144 weekly trips) favored taxis (31.2%) and motorcycles (23.5%) for manoeuvrability in congested areas. Low-density zones (800 weekly trips) preferred personal cars (20.75%) due to sparser populations and potentially longer distances. The medium-density zone (916 weekly trips) demonstrated a balanced mix (taxis: 21.8%, personal cars: 19.8%, motorcycles: 15.1%, tricycles: 9.5%). The Random Forest model outperformed XGBoost in predicting transportation choices (accuracy: 78% vs. 74%). It also exhibited lower error metrics (MSE: 805.76, MAE: 603.25, RMSE: 28.39) and a higher R-squared value (0.70), indicating a stronger predicted-actual value correlation. This study highlights the interplay between demographics, spatial density, and transportation preferences in Akure. Geospatial analysis and machine learning offer valuable insights for understanding urban transportation dynamics. The Random Forest model's superior performance suggests its potential as a tool for urban transportation planning in Akure. By considering zone-specific characteristics and resident needs, policymakers can develop tailored solutions for sustainable and efficient transportation systems.

Keywords: Urban transportation, Geospatial analysis, Dominant modal choice, Random forest and XGboost

GSJ: Volume 12, Issue 11, November 2024
ISSN 2320-9186

1.0 INTRODUCTION

The rapid urbanization experienced in many developing nations has ushered in a plethora of challenges, with urban transportation emerging as a paramount concern. Nowhere is this more evident than in Akure, the capital city of Nigeria's Ondo State. The convergence of rapid population growth and urban expansion has led to a heightened demand for transportation services [1]; [2]. However, In the realm of urban transportation, the concept of mode selection takes centre stage [3]. It encompasses the myriad decisions individuals make when determining how to navigate a city or urban area. This decision-making process encompasses a spectrum of transport options, from the simplicity of walking to more intricate choices such as cycling, utilizing public transit, or relying on private vehicles such as cars and motorcycles.

Economic considerations also enter the equation. The choices individuals make in transportation carry substantial financial implications, both for themselves and for the cities they inhabit. Private car ownership, for example, can become prohibitively expensive due to ongoing expenses like fuel, maintenance, insurance, and parking fees. In contrast, opting for public transit or alternative modes often proves more economically prudent. transportation modes can vary widely within a city, giving rise to disparities in social equity [4]. Vulnerable populations may face barriers in accessing reliable and affordable transportation, resulting in mobility inequities. A comprehensive grasp of mode selection can shed light on these disparities and guide strategies to address them [5]

Akure faces multifaceted challenges within its current urban transportation landscape. Daily traffic congestion persists, resulting in extended travel times, economic inefficiencies, and environmental degradation [6]. Concurrently, escalating air pollution and carbon emissions from a growing vehicular populace pose substantial health and environmental risks [7]. Furthermore, disparities in transportation access and affordability are apparent, disproportionately impacting vulnerable populations. To tackle these challenges comprehensively, this research adopts a data-driven approach. Leveraging geospatial data, travel surveys, and advanced machine learning techniques, this study scrutinize the factors influencing mode selection within Akure.

2.0 METHODOLOGY

Study Area

Akure, the capital of Ondo State in Nigeria, is located in the northern part of the state, positioned at approximately 70.151 degrees latitude North and 50.151 degrees longitude East. Covering an area of around 30.02 square kilometres, Akure had a population estimated at 403,000 based on the 2006 census. This population was divided into 175,495 (49.68%) males and 177,716 (50.32%) female.

Sampling Methods

Following the study of [8], Akure metropolis was divided into three zones according to its land use configuration. This is to make collection, sorting and extraction of data accurate. The zones are high, medium and low-density zones. Figure 1 showed the delneaton of the study area into populaton density zone.

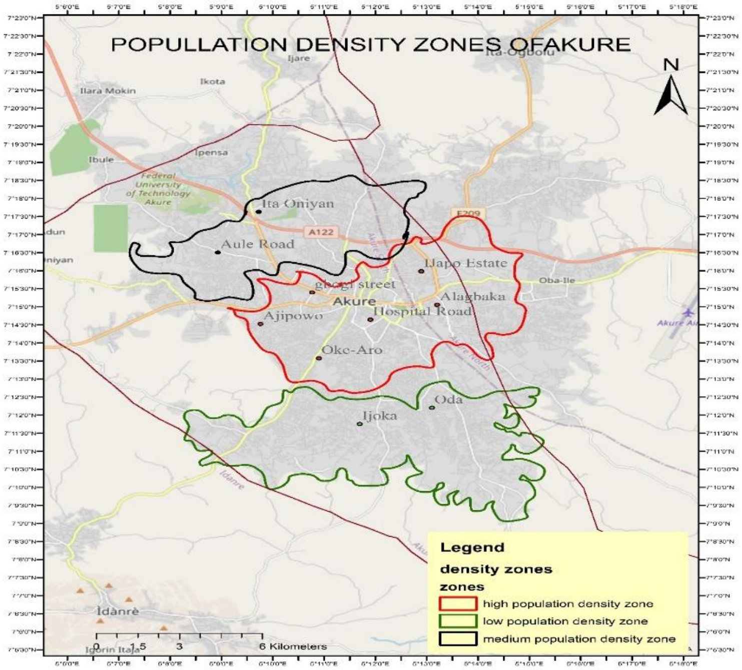

Figure 1: Delineation of study area into population density zone

Study Population

According to the National Population Commission (2006), Akure had 403,000 residents as of the 2006 census, and its annual population growth rate was 3.67 percent. Equation (3) was used to determine the estimated population for 2023 based on the population of the base year.

$$P_p = P_i (1 + r)^n \quad (1)$$

Where;

P_p is the projected population,

r is the Population growth rate (%),

n is the number of years from which projection is made, and

P_i is the base year or initial population.

Based on equation (1), the 2023 population is calculated thus,

P_p (2023) = $403.000\ (1 + 3.67/100)^{17}$

P_p (2023) = 743,717 People

The National Bureau of Statistics (2016) predicts that there are 5 people per family in Nigeria's urban areas. The research population is calculated by multiplying the expected population by .Therefore,

Number of Households = $\frac{743717}{5}$

Number of Households = 148,743 households

Sampling technique

The systematic random sampling technique, which involved choosing each item from a list one at a time, was used to carry out the travel survey. This method of sampling was easier and more affordable, and it could be effectively used even when working with enormous populations. According to [9], the right sample size for a travel survey relied on the number of residents in the area under investigation. The recommended sample size for use in this investigation was provided in Table 1 and was based on the advice of [9].

Table 1: Recommended Sample Size

Population	Suggested Sample size (%)
Below 50,000	10.0≥ x ≤20.0
50,000 – 300,000	≥2.5≤12.5
300,000 – 500,000	≥2.0≤6.7
500,000 – 1,000,000	≥1.5≤5.0
Above 1,000,000	≥1.0≤4.0

Source: Handbook of Transport modelling (Hensher and Button, 2007).

For a population of 148,743 households for the study are computed in section 3.2, the minimum number of samples according to Table 3.2 is 2.5% of the population and is computed as follows:

$$Minimum\ Sample\ Size\ =\ 0.025\ \times\ 148{,}743$$

$$Minimum\ Sample\ Size\ =\ 3{,}719$$

To accurately gauge the true breadth of travel behavior within a population, the Australian Transport and Infrastructure Council proposed a recommendation suggesting an optimal response rate of 80% or above. The estimated number of survey questionnaires that would have needed to be collected in order to reach an 80 percent response rate was determined as follows:

$$Minimum\ number\ of\ questionnaire\ =\ 0.80\ \times\ 3719$$

$$Minimum\ number\ of\ questionnaire\ =\ 2{,}975$$

3.0 Results and Discussion

Geospatial analysis and predictive modelling are used to assess the urban transportation system of Akure using the two-pronged approach geospatial analysis involved surveying 440 Akure residents (age, occupation, income, preferred transportation). Additionally, geospatial data on trip volumes within high, medium, and low-density zones was obtained. Second, machine learning models (Random Forest and XGBoost) were trained on the data to predict transportation choices.

Exploratory Data Analysis

Demographics

Figure 2 depicts the age distribution of the respondents. The majority fall within the age range of 18-30 years, comprising 165 individuals and accounting for 37.5% of the total sample. The smallest segment consists of respondents above 60 years old, with 21 individuals, accounting for 4.8% of the total sample.

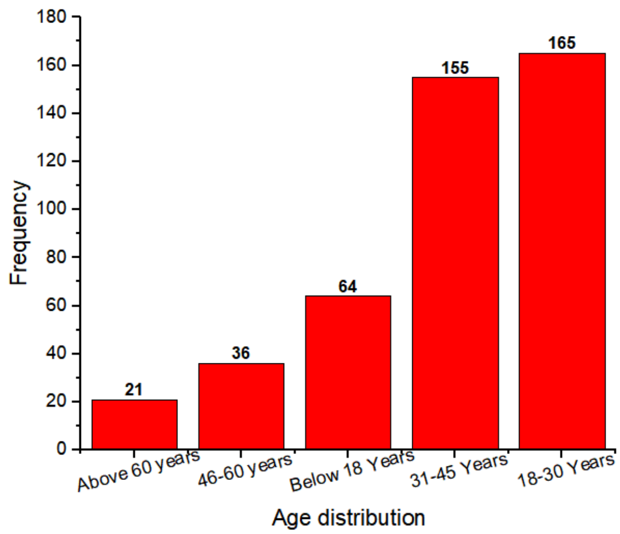

Figure 2: Distribution of Respondents Age

Examining the distribution of respondents' monthly income (Figure 3) reveals diverse economic backgrounds. The majority of respondents earn incomes falling within the range of 50,000 - 100,000 Naira, with 122 individuals, representing 27.7% of the total sample while. 44 earns between 20, 000 – 30,000 and 29 earns between 10,000 and 20,000.

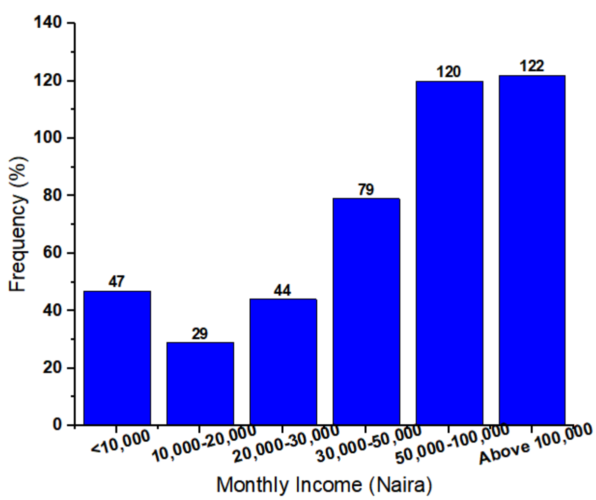

Figure 3: Distribution of Respondents Monthly Income (In Naira).

Mode preference

Daily activities within Akure show a diverse range of choices among respondents due to there preferred mode of transportation. Taxis emerge as the most popular mode, with 30.6% of respondents opting for this mode of transportation, closely followed by motorcycles (23.5%) and personal cars (28.2%). Tricycles (9.4%) and walking (2.4%) are also utilized by a portion of respondents. Interestingly, cycling accounts for 3.5% of preferred modes, while other modes of transportation collectively represent 2.4%. These results underscore the varied transportation preferences within Akure, highlighting the importance of understanding and catering to these preferences in urban planning and transportation development initiatives.

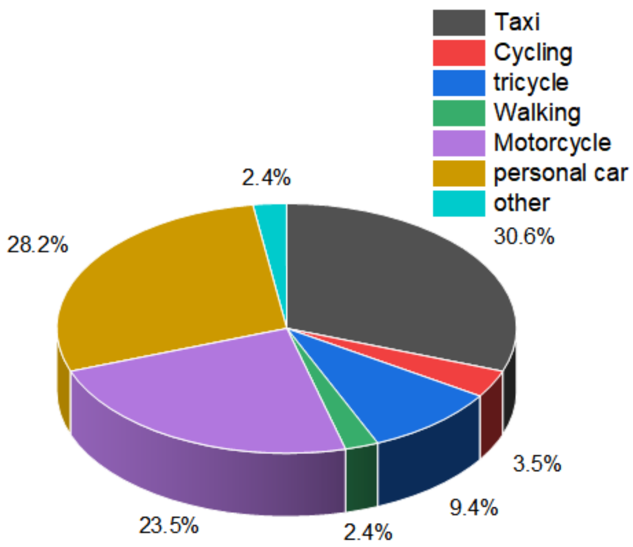

Figure 4: Daily Activities Travel within Akure

Geo-spatial analysis

High-Density Zones Analysis

In the high-density zones encompassing Ijapo Estate, Hospital Road, Ajipowo, Oba Adesida Rd, Oja Oba, Alagbaka, Akure Stadium, Gbogi Street, and Oke-Aro, a significant volume of weekly trips is observed, totalling 4,144 trips based on the provided data. Oja Oba notably stands out with 504 weekly trips, primarily due to its vibrant commercial activities and status as a major marketplace within Akure. Analyzing specific areas within these high-density zones further reveals the transportation dynamics. Alagbaka and Oke-Aro showcase a significant preference for taxis and personal cars, reflecting their mixed-use nature with residential, commercial, and institutional activities. Overall, the data underscores the complex transportation dynamics within high-density zones in Akure, with a notable reliance on taxis, motorcycles, and personal cars for daily commuting and activities. Understanding these mode preferences is crucial for urban planners and policymakers to develop sustainable transportation solutions and alleviate potential traffic congestion issues in these densely populated areas, particularly emphasizing the significance of Oja Oba as a pivotal transportation hub.

Medium-Density Zones Analysis

The medium-density zone encompassing Aule Road, Ita Oniyan, and Ijigba experiences a moderate volume of weekly trips, totaling 916 trips based on the provided data. Aule Road contributes 344 weekly trips, while Ita Oniyan adds 332 weekly trips, and Ijigba contributes 240 trips to the total count. These areas represent a middle ground in terms of population density and activity levels compared to the high-density zones. This diversity in land use contributes to the variety of transportation modes observed. Additionally, the proximity of these zones to both high-density and low-density areas influences the mode choices, with residents opting for modes that best suit their travel needs within and outside the immediate vicinity.

Low-Density Zones Analysis

The low-density zones of Oda and Ijoka see notably lower weekly trips compared to high and medium-density areas, totaling 800 trips based on the provided data. Oda contributes 444 weekly trips, while Ijoka adds 356 weekly trips to the total count. These areas represent regions with lower population density and comparatively less activity in terms of daily commuting and travel. Analyzing the prevalent modes of transportation within the low-density zones reveals a different pattern compared to higher-density areas. Due to the lower population density and less commercial activity, the reliance on public transportation modes is relatively lower. Taxis remain a common choice, constituting 8.75% of trips, indicating their availability even in low-density areas. Personal cars are more prevalent, representing 20.75% of trips, as residents may prefer private vehicles for mobility due to the less crowded nature of these zones.

Table 2 Trip distribution for Origiin-destation in Akure

Origin/destination	**Oda**	**ijoka**	**Ijapo estate**	**Hospital Rd**	**Oke-aro**	**Gbogi**	**Ajipowo**	**Aule Rd**	**Ita oni-yan**	**Alagbaka**	**Total no. of trips**
Oda	100	56	64	28	56	32	28	32	28	20	444
Ijoka	36	92	36	24	28	24	32	28	32	24	356
Ijapo Estate	32	32	100	32	32	36	36	32	36	92	460
Hospital Rd	20	68	72	104	32	32	28	24	24	28	432
Oke Aro	24	32	32	28	84	28	32	28	20	24	332
Gbogi	28	28	24	24	20	92	28	32	32	32	340
Ajipowo	20	32	32	36	36	32	96	24	24	36	368
Aule Rd	28	36	36	32	28	28	24	76	28	28	344
Ita Oniyan	16	28	24	28	32	36	32	32	68	36	332
Alagbaka	44	32	32	36	36	28	24	24	32	92	380
Oba Adesida Rd	76	52	20	32	24	36	52	56	36	52	436
Oja Oba	92	56	36	28	20	32	64	64	28	84	504
Akure stadium	60	32	20	32	36	24	36	28	32	60	360
Idanre Garage	44	36	32	36	32	28	36	32	28	28	332
Ijigba	28	24	20	28	28	32	12	16	32	20	240
Total no. of trips	648	636	580	528	524	520	560	528	480	656	

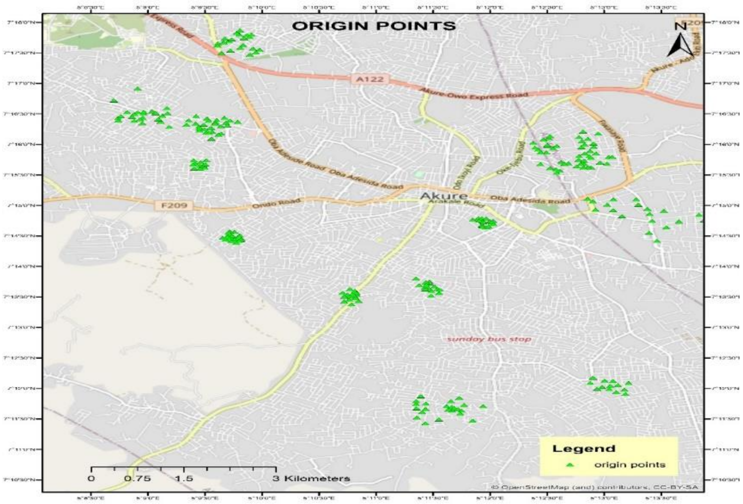

Figure 5: Origin Points

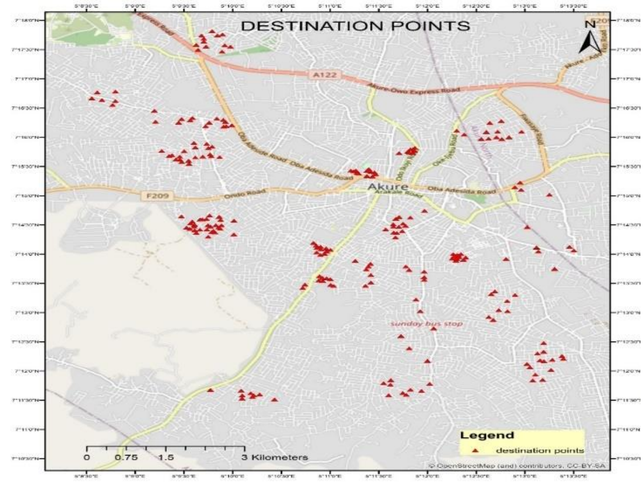

Figure 6: Destination points

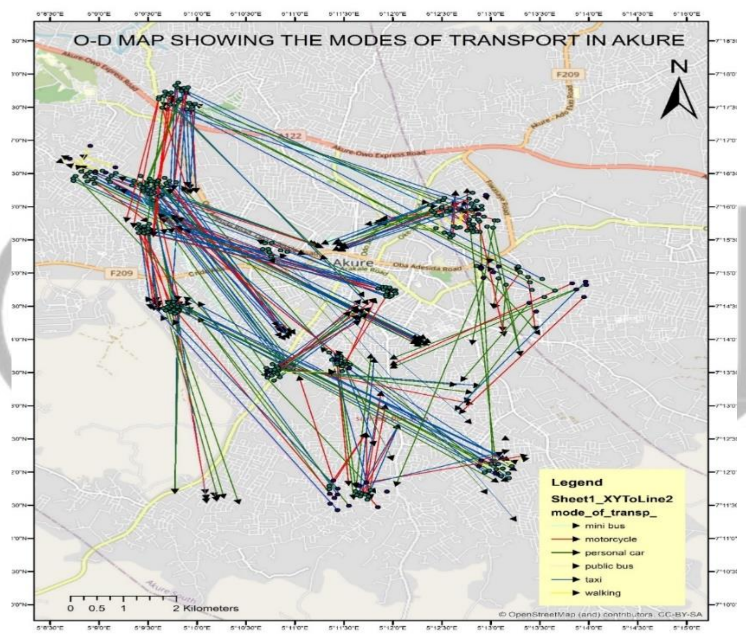

Figure 7: O-D map showing modes of transportation in Akure

Modelling

Random forest regression and was used to estimate the traffic mode data. Each model was trained, validated and tested on the mobility dataset. The dataset was randomly split into 70% train, 15% validation and 15% testing set.

Random Forest Model Classification

The Random Forest classification model's results for transportation mode prediction in Akure exhibit a generally robust performance across various metrics. Analyzing precision, recall, and F1-scores, it's evident that the model excels in accurately predicting certain transportation modes. For instance, it achieves exceptionally high precision (0.97) and recall (0.95) for identifying instances of personal car usage. This indicates a strong ability to differentiate and correctly classify trips involving personal vehicles, showcasing the model's effectiveness in capturing patterns associated with this mode of transportation. Similarly, the model demonstrates notable proficiency in predicting motorcycle usage, with a precision of 0.94 and a high recall of 0.97. This suggests that the model reliably identifies instances where motorcycles are used as a means of travel.

Table 3: Random Forest classification report

	Precision	Recall	F1-score	Support
Personal car	0.97	0.95	0.96	38
Motorcycle	0.94	0.97	0.95	32
Public bus	0.87	0.82	0.85	40
Tricycle	0.76	0.86	0.81	37
cycling	0.43	0.42	0.42	36
Taxi	0.94	0.85	0.89	39
Accuracy			0.78	258
Macro average	0.78	0.77	0.78	258
weighted average	0.78	0.78	0.78	258

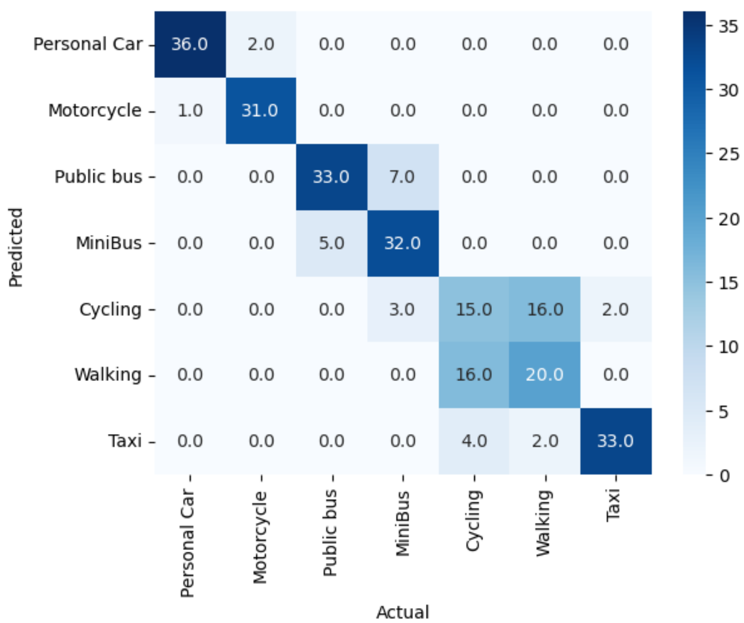

Figure 8: Random Forest confusion matrix

a) <u>Random Forest Regression Model</u>

The model evaluation report for the random forest model shows that the model achieved about 70% accuracy on the testing set. The model root mean square error as shown in table 4 is 28.39 and the mean absolute error of 603.25. The frequency of mobility as shown in figure 14 indicate that more than 45 residents' moved through a road axis using a preferred transportation mode. This increase in rate as indicated by the model output could lead to traffic congestion in the under studied route. The model equation and the coefficient of each model predictors are explained below

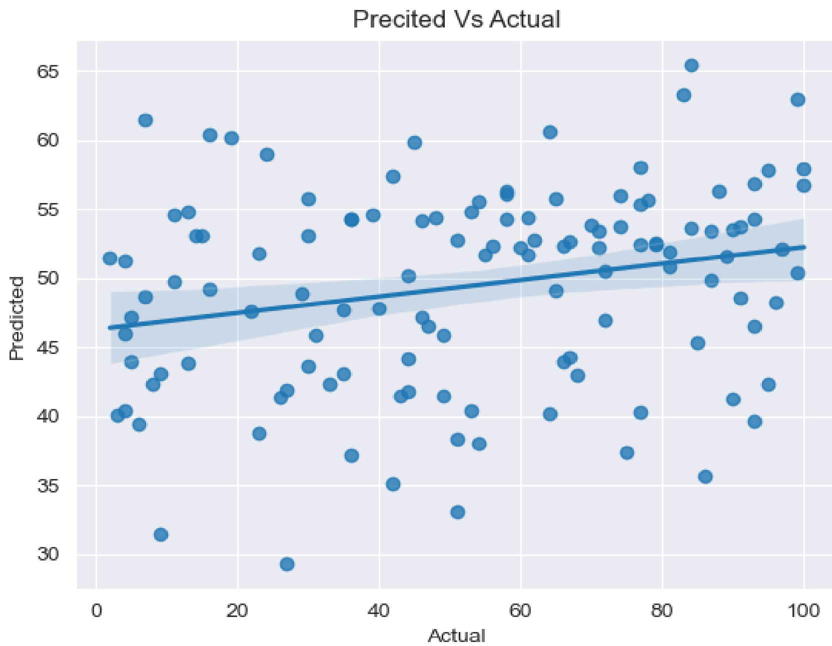

Figure 9: Prediction output of the random forest model

Table 4: Random Forest Model Evaluation Metrics

Metrics	Random Forest
Mean Square error	805.76
Mean absolute error	603.25
Root mean square	28.39
R^2	0.70

$$
\begin{aligned}
mode\ frequency \\
&= -0.031(gender) - 0.057(Owncar) + 0.066(areawithmotorableroad) \\
&+ 0.019(occupation) + 0.003(factorinfluencingmodechoice) \\
&+ 0.004(workplace) + 0.007(schooloruni) + 0.129(market) \\
&+ 0.026(hospitalorhealthcenter) + 0.003(shoppingcenterormall) \\
&- 0.049(religiousplace) - 0.031(parkorrecreationarea) \\
&- 0.108(governmentoffices) + 0.079(age) + 0.012(monthlyincome)
\end{aligned}
$$

Figure 10: Rrandom forest regression model coefficients

XGboost Classification model

The XGBoost model achieves strong F1-scores for personal car (0.94), motorcycle (0.85), and taxi (0.98), indicating robust performance in accurately predicting these transportation modes. Support represents the number of instances for each class in the dataset. Higher support values suggest more data available for training and testing the model for that particular class. In this report, support values vary across different transportation modes, with higher support for personal car, motorcycle, public bus, and walking, and lower support for cycling and tricycle. The overall accuracy of the XGBoost model is reported as 0.74, indicating that it correctly predicts the transportation mode in 74% of cases. The macro average and weighted average provide overall assessment metrics across all classes, with both averages indicating consistent performance across multiple transportation modes.

Table 5: XGboost Classification model

	Precision	Recall	F1-score	Support
Personal car	0.92	0.96	0.94	47
Cycling	0.50	0.40	0.44	5
Motorcycle	0.85	0.84	0.85	56
Public bus	0.69	0.58	0.63	62
Tricycle	0.48	0.53	0.50	19
walking	0.48	0.59	0.53	41
Taxi	0.97	0.96	0.98	28
accuracy			0.74	258
Macro average	0.70	0.69	0.70	258
Weighted average	0.75	0.74	0.74	258

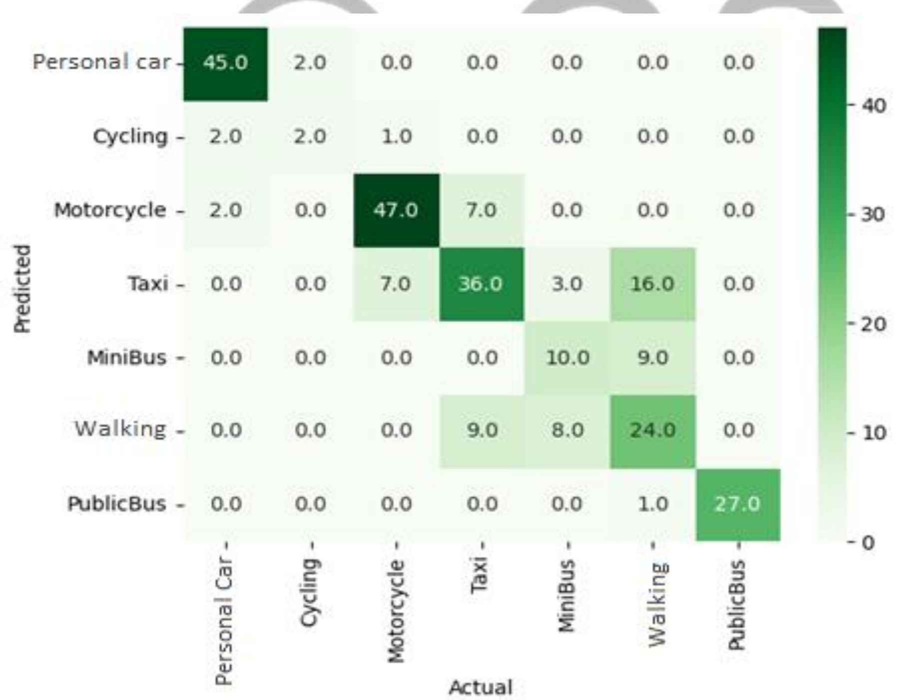

Figure 11: XGBoost Confusion matrix

b) <u>XGBoost Regression Model</u>

The XGBoost regression model exhibits a mixed performance across different evaluation metrics. The Mean Square Error (MSE) value of 859.81 indicates the average squared difference between predicted and actual values, reflecting a moderate level of error in the model's predictions. A lower MSE would suggest better accuracy in predicting target variables. The Mean Absolute Error (MAE) of 682.45 represents the average absolute difference between predicted and actual values. This metric provides insight into the magnitude of errors in the model's predictions. A lower MAE indicates better accuracy, so the relatively high MAE in this case suggests some level of variability or inconsistency in the model's predictions. XGBoost model explains approximately 65% of the variance in the data, indicating a moderate level of predictive power.

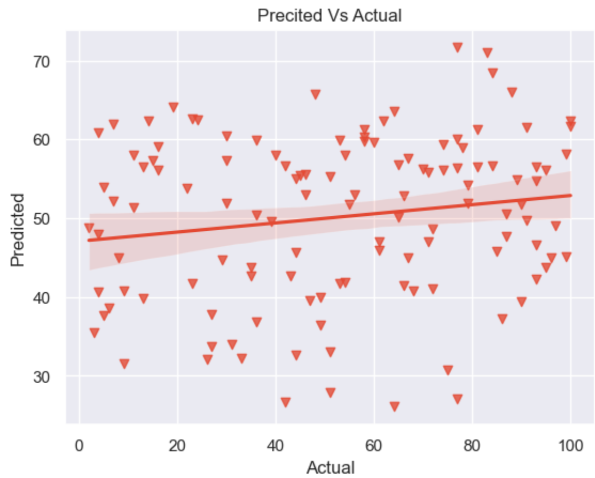

Figure 12: Model prediction of thr XGBoost

Table 6: XGboost Model Evaluation Metrics

Metrics	**XGboost Regression**
Mean Square error	859.81
Mean absolute error	682.45
Root mean square	29.32
R^2	0.65

Fitting the model with the XGboost model estimates, *eqn 4.1* can be written as;

$$\begin{aligned} mode\ rate = & -0.092(gender) - 0.116(Owncar) + 0.034(areawithmotorableroad) - 0.022(occupation) \\ & - 0.022(factorinfluencingmodechoice) - 0.035(workplace) + 0.087(schooloruni) + 0.200(market) \\ & + 0.008(hospitalorhealthcenter) - 0.04(shoppingcenterormall) - 0.046(religiousplace) \\ & - 0.072(parkorrecreationarea) - 0.123(governmentoffices) + 0.088(age) - 0.0025(monthlyincome) \end{aligned}$$

The model results shows that variables such as area with motorable road, school or university, market, hospital, and age positively influence the mobility or rate of road usage by the residents. Conversely, variables such as gender, owned car, occupation, factors influencing mode, workplace, mall, religious place, park, government office, and monthly income influence the rate of road usage by the residents.

4.0 Conclusion

- The most substantial demographic group among respondents, comprising 37.5% and consisting of 165 individuals, is within the age bracket of 18-30 years. Regarding occupation, self-employment is the dominant category at 32.0%, followed closely by employment at 31.5% among respondents.

- Males constitute 54.4% of respondents, while females account for 45.6%, and 27.7% of respondents earn monthly incomes ranging from 50,000 to 100,000 Naira.

- Taxis are the most preferred mode of transportation (30.6%), followed by motorcycles (23.5%) and personal cars (28.2%)

- Akure's transportation landscape significantly varies by zone density. High-density zones (4,144 weekly trips) rely heavily on taxis (31.2%) and motorcycles (23.5%) for navigating congested areas. In contrast, personal cars (20.75%) are the preferred mode in low-density zones (800 weekly trips) due to sparser populations and potentially longer distances. The medium-density zone (916 weekly trips) exhibits a more balanced mix with taxis (21.8%), personal cars (19.8%), motorcycles (15.1%), and tricycles (9.5%). This highlights the crucial role of density in planning sustainable transportation solutions for Akure.

- . While the Random Forest model attains a classification accuracy of 78% for transportation modes, the XGBoost model shows slightly lower accuracy at 74% for predicting transportation modes

- Overall, the Random Forest model exhibits better performance than the XGBoost model in terms of error metrics, featuring lower mean square error (805.76), mean absolute error (603.25), and root mean square (28.39). Moreover, the Random Forest model attains a higher R-squared value of 0.70, making it a preferred choice for transportation planning in Akure.

5.0 References

1) A. Agyemang, "Mode choice for long distance trip; Evidence from the greater Accra metropolitan area of Ghana," *Journal of Transportation Geography*, vol. *64*, no.1, pp.150–157, 2017
2) Q. Al-Salh and D. Esztergar-Kiss, "Linking mode choice with travel behaviour using logit model based on utility functions" *Sustainability,* vol. 13*, no.1*, 2021
3) M. Drliciak, J. Celko, M. Cingel and D. Jandacka, "Traffic volumes as a modal split parameter," *Sustainability*, vol. 12, no. 1, pp. 10252, 2020
4) J. O. Oyedepo, "Performance Analysis of Off-Street Parking around the Central Business District of Akure Southwest Nigeria" *Consilience: The Journal of Sustainable Development*, vol. 16, no. 1, pp. 91–105, 2016.
5) Y. Zhao, X. Han, and X. Hu, "Traffic Flow Prediction Model Based on the Combination of Improved Gated Recurrent Unit and Graph Convolutional Network" *Frontiers in Bioengineering and Biotechnology*, vol. 10, no. 2, pp. 1–17, 2022

6) J. E. Etu and O. J. Oyedepo, "Forecasting Trip Generation For High Density Residential Zones of Akure, Nigeria: Comparability of Artificial Neural Network And Regression Models." *Journal of Civil Engineering, Science and Technology*, vol. 9, no. 2, pp. 2-10, 2018
7) O. O. Samuel and O. A. Adetayo, "Assessment of Passengers' Satisfaction of Public Transport System in Akure-Owo Axis, Nigeria" *American International Journal of Multidisciplinary Scientific Research*, vol. 4, no. 1, pp. 1-16, 2018
8) A. Owolabi and D. Akinwunmi, "Operational Characteristics of Para-transit Modes in Akure, Nigeria" *FUTY Journal of the Environment*, vol. 6, no. 2, pp. 8-10, 2012
9) D. A. Hensher and K. J. Button, "Handbook of Transport modelling" Emerald Group Publishing Limited, 2007.