A Statistical Approach for Investigation into the Parking Lot Problems in VIT, Vellore

[1]Digjoy Samanta, [2]Dr. D. Anuradha,
[1]Undergraduate student, [2]Assistant Professor (Senior)
[1]School of Mechanical Engineering, VIT, Vellore
[2]Department of Mathematics, School of Advanced Sciences, VIT, Vellore

***Abstract*— Across the globe, a lot of studies are being on parking systems for making them more robust and user-friendly. The increasing number of cars every day is a major cause of imbalance in the demand and supply of parking lots in VIT, Vellore. This paper aims at verifying the causes by using statistical techniques. The study also highlights the other contributing factors in parking lot availability (PLA) problems so that a mathematical model can be obtained for finding the solution.**

***Index Terms*— Parking Lot Problems; Chi-square testing; Hypothesis testing**

I. INTRODUCTION

In an innovative and proactive approach to determine the causes of the parking lot availability problems in VIT Vellore, we present a statistical evidence which is consistent with our theoretical findings, as obtained from the survey. Testing of hypothesis plays a very important role in analytical studies for a given problem. While using statistical inference tests, statisticians prefer to conclude results on the basis of the data obtained from population. Here, it must be noted that the entire population cannot be tested individually. This can be due to the large volume of data present or the time required to analyze each data separately. Therefore, we need to extract a subset of the population to make our analysis easy. This brings us to the idea of sampling of our data. Pearson [1] proposed a methodology to check goodness of fit using chi-square distribution and to assess the homogeneity and independence of the given categorical data. Fatai et al. [2] used statistical testing of hypothesis to examine the relationship between crude oil rents and fiscal balances, while controlling other parameters. Meier and Hicklin [3] used testing of hypothesis to analyze employee turnover and organizational performance. Compton et al. [4] used statistical testing to analyze whether people adaptively sharpen attentional focus by learning from performance mistakes, using current theories of cognitive control. Jog et al. [5] using hypothesis testing brought out the implications of smart parking technologies in the general public. Okubo et al. [6] used testing of hypothesis to find trends in tourist behavior for determining a possibility of a collaboration between local government, tourism industry and visitors after conducting a survey amongst tourists.

In this paper, we conducted a survey to bring out the causes and solutions of the parking lot availability problems in VIT Vellore. Through the survey, we determined the causes of the parking lot availability problems which are increasing number of vehicles, lack of regulations in parking vehicles, utilization of parking lots for commercial activities and vehicles not parked in a proper fashion. Now, using hypothesis testing we will verify the causes for the parking lot problems. The paper has been subdivided into different sub-sections as Section 2 discusses the statistical technique used for analyzing the data, Section 3 discusses the possible causes and their impact analysis using the statistical hypothesis test mentioned above, followed by the conclusion.

II. METHODOLOGY

We will verify the causes of the PLA problems in VIT, Vellore by using Null Hypothesis Significance Testing (NHST). The data obtained from the survey is nominal in nature. Thereby, chi-square testing for independence of attributes is used. This has been done by incorporating the randomly sampled data from the survey.

Steps involved in Hypothesis Testing

The steps involved in NHST for this case can be stated as follows:

1. State the hypothesis.
2. Devise a method for analysing data.
3. Analyse the data obtained from survey.
4. State the conclusions and draw inferences from the same.

Stating the Hypothesis

We have two nominal variables X and Y having `q' and `t' levels respectively. The null hypothesis states that variable X has no effect on variable Y that is they are independent variables. Whereas, the alternate hypothesis states that variable X has an effect on variable Y, hence they are not independent. This can be mathematically stated as:

H_0: The variables are independent.

H_1: The variables are not independent.

Devising a Method for Analysis of Data

After the hypothesis has been stated, a suitable data analysis technique needs to be selected on the basis of which the hypotheses can be accepted or rejected. The plan must specify the following:

1. Significance Level at which Data will be analysed: The investigators must choose any value between 0 and 1 depending upon the accuracy level required. The most preferred values are 0.01, 0.10, and 0.05.
2. Test Method: Since the variables in this case are categorical in nature, chi-square testing is used to analyse the data. The conditions of simple random sampling and categorical data has been met successfully.

Analysis of Data

Based on the method chosen, we can now analyse the data using the mathematical calculations involved in it. Chi-square testing involves determination of degrees of freedom and test statistic value to assess the independence of attributes. Using the data from the survey, we can calculate the degrees of freedom, predictable frequencies, test value and the p-value associated with the test.

1. Degrees of Freedom(*df*) can be determined by using the following formula:

$$df = (r-1)\times(c-1) \quad (1)$$

Where '*r*' is the number of levels in X and '*c*' is the number of variables in Y.

2. Test Statistic for chi-square testing can be formulated as follows:

$$\chi^2 = \frac{N\times((A\times D)-(C\times B))^2}{(A+C)\times(B+D)\times(A+B)\times(C+D)} \quad (2)$$

Where *A* - Number of times feature '*q*' and class label '*t*' co-occurs
B - Number of times '*q*' appears without '*t*',
C - Number of times '*q*' appears without '*t*',
D - Number of times neither '*q*' nor '*t*' appears,
N -Total number of records

Deriving Conclusions and Inferences

Based on the p-value obtained, we can accept or reject the Null Hypothesis and accordingly accept or reject the Alternate Hypothesis. Thereby, completing the NHST.

III. RESULTS AND DISCUSSION

We used chi-square test for analysing whether means of transport plays a role in problems faced by user in parking their vehicles during the day. The results can be summarised as follows:

1. Means of transport is not a cause for the parking lot problems.
2. The category of user for parking lot and time slot do have an impact on parking lot availability.
3. There is not much evidence to support that mode of trans-port used is one of the causes of PLA problems.

Means of transport is not a cause for the parking lot problems

In order to verify the above statement, we perform hypothesis testing on the data obtained from the survey. Considering the following hypothesis:

H_0: The means of transport used for travelling and time slot do not have an impact on parking lot availability.
H_1: The means of transport used for travelling and time slot have an impact on parking lot availability.

Now, the following table is considered for hypothesis test:

Table 1: Time Slot vs. Mode of Transport

Time Slot	Bus	Car	Motorcycle/ Scooty	Bicycle
9:30 AM – 11:00 AM	2	5	18	2
11:00 AM – 12:30 PM	0	1	4	0
2:00 PM – 3:30 PM	2	0	1	2
3:00 PM – 4:30 PM	0	0	1	0
4:00 PM – 5:30 PM	0	0	1	1

The category of user for parking lot and time slot do have an impact on parking lot availability

In order to validate the above statement, we perform hypothesis testing on the data obtained from the survey. Considering the following hypothesis:

H_0: The category of user for parking lot has an impact on the parking lot availability with respect to time.
H_1: The category of user for parking lot does not have an impact on the parking lot availability with respect to time.

Now, the following table is considered for hypothesis test:

Table 2: Category of Users vs. Time Slot

Time Slot	Student	Faculty	Staff
9:30 AM – 11:00 AM	6	18	3
11:00 AM – 12:30 PM	2	3	0
2:00 PM – 3:30 PM	5	0	0
3:00 PM – 4:30 PM	0	1	0
4:00 PM – 5:30 PM	2	0	0

There is not much evidence to support that mode of transport used is one of the causes of PLA problems

We perform hypothesis testing on the data obtained from the survey to verify the above statement. Considering the following hypothesis:

H_0: The PLA problems faced by user depends on the mode of transport used for communicating.
H_1: The PLA problems faced by user does not depends on the mode of transport used for communicating.

Now, we perform hypothesis test on the following table:

Table 3: Problems Faced vs. Mode of transport

Mode of Transport Used	Yes	No
Bus	0	4
Car	3	3
Motorcycle/Scooty	14	11
Bicycle	3	2

Outcome percentages of the given pro les for the above table is not significant at 0.05 level of confidence. Therefore, we can conclude that there is not much evidence to support that mode of transport used is one of the causes of PLA problems.

Causes of the PLA Problems Based on Survey

Based on the above results, we can conclude that the causes of the PLA problems are due to the improper utilisation of the existing resources. Thereby, based on the survey, we can enlist the causes as follows:

1. Increasing number of vehicles inside the campus.
2. No proper parking regulations.
3. Vehicles are not being parked effectively.

Therefore we have now successfully established the causes of the PLA problems in VIT, Vellore. Now, to reduce the PLA problems we can frame a mathematical model of the current system and the simplified system by keeping in mind all the factors and constraints determined. This can be evaluated and modified according to real time situations, thereby successfully solving the given problem.

IV. Conclusions

After the statistical investigation, we now have successfully ascertained the causes of the parking lot availability problems inside the campus. The main solution lies in developing a parking methodology to ensure that everyone gets a parking space at the desired time slots in such a way that the time required and the distance travelled between each building is minimized. We need to optimize the cur-rent demand and supply of the parking lot availability in such a way that mini-mum additional cost is incurred and a new parking management system is developed. Using the feedback obtained from users, now we can successfully create and evaluate the success of mathematical models in solving such real time problems.

V. Acknowledgment

The authors of this paper are indebted to every student, staff and faculty member who took out their time in voicing out their concerns regarding the parking lot availability problems in Vellore Institute of Institute of Technology. Thereby, fulfilling the primary objective of this paper, which will help in making sound data driven decision making approach for solving this problem

REFERENCES

[1] K. Pearson, On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophy Magazine*, 50(5) (1900).
[2] K. Fatai, L. Oxley and F. G. Scrimgeour, Modeling and forecasting the demand for electricity in New Zealand: A comparison of alternative approaches, *Energy Journal*, **24**(2003).
[3] Meier, K. J., Hicklin, A., Employee turnover and organizational performance: Testing a hypothesis from classical public administration. *Journal of Public Administration Research and Theory*, **18(**2007).
[4] R. J. Compton, J. Hofheimer, and R. Kazinka, Stress regulation and cognitive control: evidence relating cortisol reactivity and neural responses to errors, *Cogn Affect Behav Neurosci*. **r; 13**(1) (2013).
[5] Y. Jog, A. Sajeev, S. Vidwans and C. Mallick, Understanding Smart and Automated Parking Technology, International Journal of u- and e- Service, *Science and Technology*, **Vol.8, No.2**(2015),.
[6] A. Okubo, K. Takeyasu, An Analysis of the Questionnaire Investigation on Tourists Behavior, Busi-ness and Management Research Vol 6, No 4 (2017).