Vol 3 Issue 2 Nov 2013

Impact Factor : 1.6772 (UIF)

ISSN No : 2249-894X

Monthly Multidisciplinary Research Journal

Review Of Research Journal

Chief Editors

Welcome to Review Of Research

RNI MAHMUL/2011/38595 **ISSN No.2249-894X**

Review Of Research Journal is a multidisciplinary research journal, published monthly in English, Hindi & Marathi Language. All research papers submitted to the journal will be double - blind peer reviewed referred by members of the editorial Board readers will include investigator in universities, research institutes government and industry with research interest in the general subjects.

Advisory Board

Address:-Ashok Yakkaldevi 258/34, Raviwar Peth, Solapur - 413 005 Maharashtra, India
Cell : 9595 359 435, Ph No: 02172372010 Email: ayisrj@yahoo.in Website: www.isrj.net

Review Of Research
Vol. 3, Issue. 2, Nov. 2013
ISSN:-2249-894X
Available online at www.ror.isrj.net

ORIGINAL ARTICLE

ROAD TRAFFIC ANALYTICS FOR DYNAMIC TRAFFIC FLOW CONTROL USING MATLAB®& SAS®

OMKAR KULKARNI

Symbiosis Centre of Management: Human Resource Development,
Symbiosis international University, Pune.

Abstract:

Urban areas are developing so are the metropolitan cities. However due to the stringent laws and impromptu regional transport ministry, the ever increasing and expanding cities' roads have exhausted their traffic handling capacities. Hitherto there have been no major congestions and traffic imbroglio; though traffic jams have become common. But in no more time, this situation will be replaced with constant jams, congestions and superflous queuing at signals and long waiting lines. To avert this situation and mollify the traffic flow, the Government officials are planning to make new roads by utilizing natural resources like hills, grass plains, forests and grazing lands meant for livestock farming. Obviously it will ease the traffic flow temporarily; making a more complex situation in the near future, as it will just add a new congested traffic route disrupting the eco-system, residential life in that area. To address this issue, quantitative optimization with a combination of analytics can become an arbitrator to solve the problem. This paper attempts to derive an algorithm and prove that statistical analysis/analytics can be the best possible method to solve the issue. The algorithm is explained with help of a live case study for better understanding.

We present research on developing models that regulate traffic flow and congestion in the Kothrud suburbs. We review the modeling effort and describe experiments probing the predictive accuracy of the models. Finally, we present research on building models that can identify current and future surprises, via efforts on modeling and controlling unexpected situations.

KEYWORDS:

Traffic Analytics , Dynamic Traffic , metropolitan cities , eco-system.

INTRODUCTION

Machine learning and intelligence are being applied in multiple ways to addressing difficult challenges in multiple fields, including transportation, energy, and healthcare. Research scientists at Microsoft Research have been engaged in efforts in all of these areas.[1] They are focussing on multiyear efforts at Microsoft Research to infer and forecast the flows of traffic. The work leverages machine learning to build services that make use of both live streams of sensed information and large amounts of heterogeneous historical data. This has led to multiple prototypes and real-world services such as traffic-sensitive directions Maps. Focused work in this realm also stimulated new efforts in related areas, such as privacy and routing.

LITERATURE SURVEY:

Traffic has been growing in major cities around the world given the increase in densities of cars on

Title: ROAD TRAFFIC ANALYTICS FOR DYNAMIC TRAFFIC FLOW CONTROL USING MATLAB®& SAS®
Source: Review of Research [2249-894X] OMKAR KULKARNI yr:2013 vol:3 iss:2

roads and the slow development of road infrastructure. With research starting in 2002, research scientist and developer teams at Microsoft Research pioneered the use of machine learning methods to build predictive models for traffic[2]. The work led early on to prototypes that can infer and predict the flow of traffic at different times into the future based on the analysis of large amounts of data on traffic over months and years. The work was leveraged in revolutionary services, such as traffic maps that show users how traffic is evolving over time, as well as in services that provide traffic-sensitive directions by considering the inferred speeds on roads that are not sensed directly.

Research on machine learning for traffic spanned several projects and has focused on both on principles and applications. Multiple technical and empirical studies were performed as part of this work. On the fielding of applications, the research efforts sit behind the traffic-sensitive directions in Bing Directions within Bing Maps and the mobile directions service on the Windows Mango phones. A portion of Microsoft Research's methods, tools, and software on predictive analytics for traffic were licensed externally in 2004 to traffic startup Inrix shortly after the company was formed, helping to slingshot that company into the world as a leading international provider of traffic analyses and predictions.

As part of efforts on learning about traffic flows from data, researchers have explored methods that enhance the safety and privacy of people who wish to help with the “crowdsourcing” of real-time flows of road data from their mobile GPS data. Principles of community sensing have been developed. These principles center on working with people under a “privacy budget” based on the use of the computations of the value of information for understanding flows over time on the road network.

However, several methods, steps and techniques have been developed to monitor traffic, detect congestions and prevent jams. A combination of analytics and statistics to provide better traffic management is still in infant stage. We have attempted to combine both for ultimate traffic flow management and coined a new term 'TRAFFIC ANALYTICS'.

SCENARIO OF TRAFFIC CONGESTION:

The Law College Road and Karve Road are two such roads in Kothrud suburbs which are heavily conjested because of nasty human population. Unfortunately, every attempt to reduce the congestion has been failed and hence they are planning to widen it or build new roads. However the affected apartments have filed a PIL stating that the widening is against their interest. Similarly their second plan, to develop a ring road attaching Senapati Bapat Road and Paud road[3] is too kept on stay because of the public's demand. Hence with the help of ex-commissioner of Pune and Mayor, we have initiated the project with the help of transportation experts who will give a ground level idea and a specific clearance of how the traffic flows.

1. Maximum Volume of Traffic Directional maximum traffic volume at peak hours should not exceed 5465 vehicles/hour. Actually, this is exceeded many times.
2. Existing carrying capacity of Law College Road The minimum number of the vehicles (emitting carbon monoxide) i. e. 12739 is considered as carrying capacity of Law College Road. At times, about 70,000 vehicles/day ply on the road.
3. Cost Analysis NPV of the proposed road project is Rs. 5.7 crores for 5 years at the rate of 12% per annum. IRR is 0.25.
4. Carbon footprint As per the calculation based on NOx and VOC emissions, the total carbon footprint of Law College Road is 41,66.142.49 tons/year (in terms of Rs. 467 crores). Construction of proposed road will have its impact and it estimated at 2 – 3% only.

Hence we can conclude for the following main points which will clearly lead us to a better understanding of our problem statement:
1. The carrying capacity of the Law College Road is completely exhausted due to ever-increasing traffic volume which directs explicitly towards the need for alternatives, and new roads.
2. As per the weighted analysis of the impacts, the proposed road has negative impacts more even though the positive effects on Law College Road are compensated. If the proper mitigation measures are taken to reduce the negative impacts due to development activities then the project will become environmentally sustainable as given in the Impact Analysis Statement.
3. Environmentally planned with adequate funding and monitoring by City level committee (CLC) / project evaluation agency (PEA) of the Road Construction Project will lead to sustainable traffic management in the city.
4. After implementation of all the mitigation measures, based on the rating criteria, the project can be considered as Non-Polluting Project because absolute value of impacts is +13 due to use of alternate roads of Bhandarkar and Prabhat road with proper mitigation measures and positive impacts due to decongestion

on Law College Road absolute value becomes +37. 5. All these scientific and planning studies, suggest the need of alternatives for the Law College Road[3].

3.1 Vehicle Classification and Configuration:

Although there are various classification approaches available to various institutions/organisations or end users of this Guideline, the responsible organisation/institution may use whichever method is deemed appropriate and applicable for the purpose at hand, depending on the quality of data required[4].

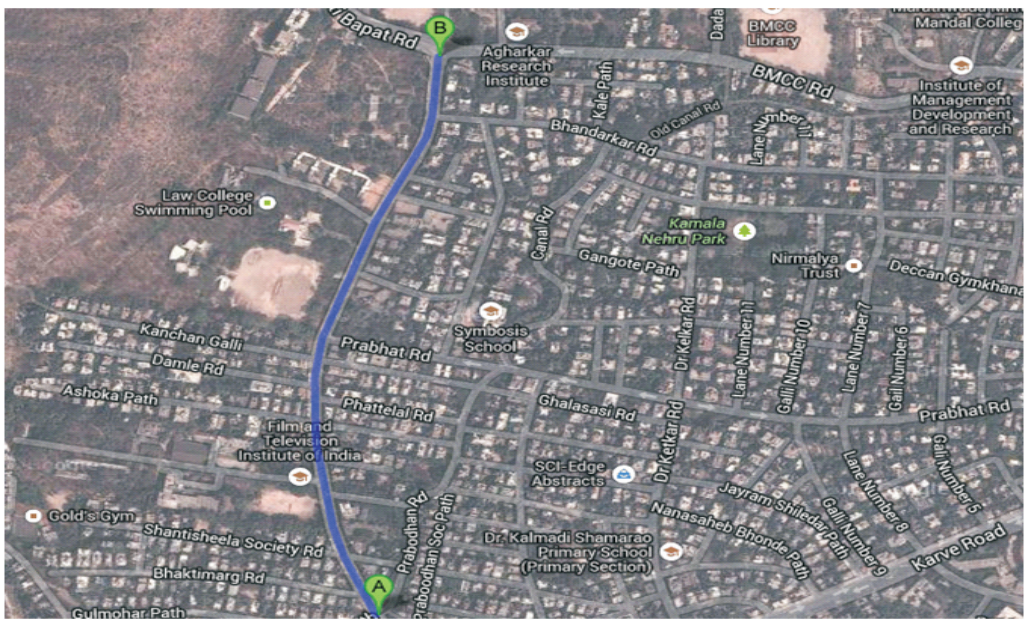
Manual Count classification Manual traffic fiow count is categorised by a visual assessment of the vehicle size and confi guration of axles. The current manual traffic fiow data collection system in Pune classifies vehicles into nine categories as follows:

Passenger Cars	(1M)
Pick up or Van less than 1.5 tons	(2M)
Trucks, more than 1.5 and less than 5 tons	(3M)
Light Lorries, more than 5 and less than 10 tons	(4M)
Medium Lorries, more than 10 and less than 20 tons	(5M)
Heavy Lorries, more than 20 tons.....	(6M)
Heavy Busses.....	(7M)
Mini Busses	(8M)

Trucks- two axles, 4 x 4 vehicles of 1.5 tonnes but less than 5 t onnes.
Light Lorries- commercial vehicles of two axles. These are vehicles with 5 tonnes but less than 10 tonnes.
Medium Lorries- vehicles with 3 or 4 axles and having weight of more than 10 tons but less than 20 tons.
Heavy Lorries - all lorries with 5 axles or more and weight more than 20 tons.
For our analysis, we will primarily use 7 major vehicle types viz., 2 wheeler, 3 wheeler, 4 wheeler, LCV's, trucks, buses and cycles.

PRELMINARY ANALYSIS:

Before we start with any analysis, we prepare a demographic map of the road; whose traffic analysis will be made. The following map shows the two ends of the roads; the names of the places are immaterial. So we will keep the conventions as place A and place B. The map has been downloaded from Google Maps, with the lates GPS mapping[5]. The blue highlighed line is the Law College Road, out primary experimental road.



Source: Google Maps, 30 September, 2013

The problem statement is to deduce a smooth traffic flowing mechanism without utilizing more resources. Hence we attack the problem statement in a conventional business analytics way, thus by breaking down the data by different day time segments. Initially we tabulate the data using PROC TABULATE to undertand the patterns and bahaviour of the traffic.

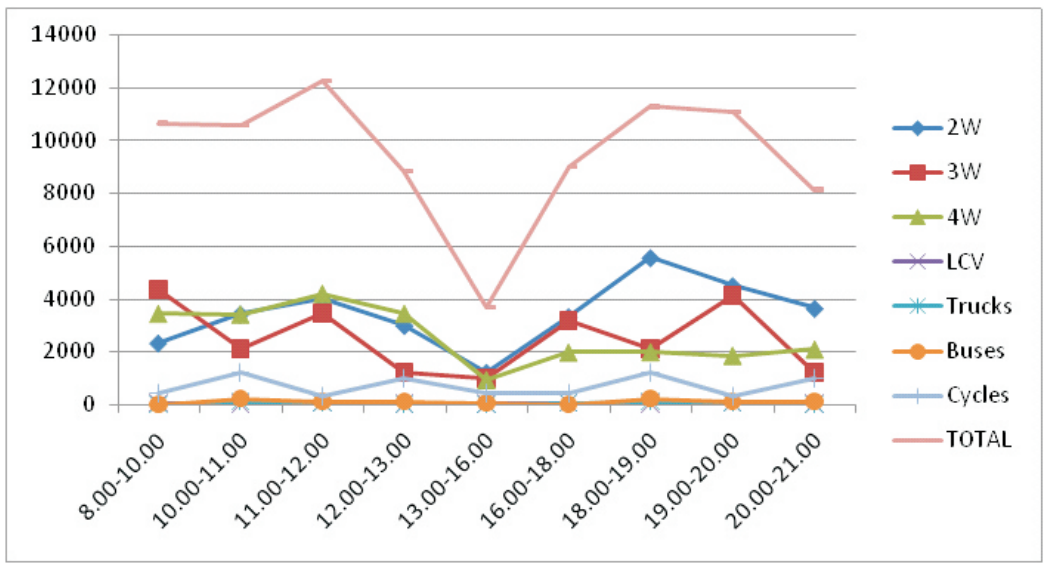
There are few tables below which show the data analysis in a very systematic and legitimate way. We have made the data representation as cogent as possible so that the analysis will be easy. There are two tables below; the first one explaining a day's analysis and second table shows and aggregate of 60 days time span across which the data was collected.

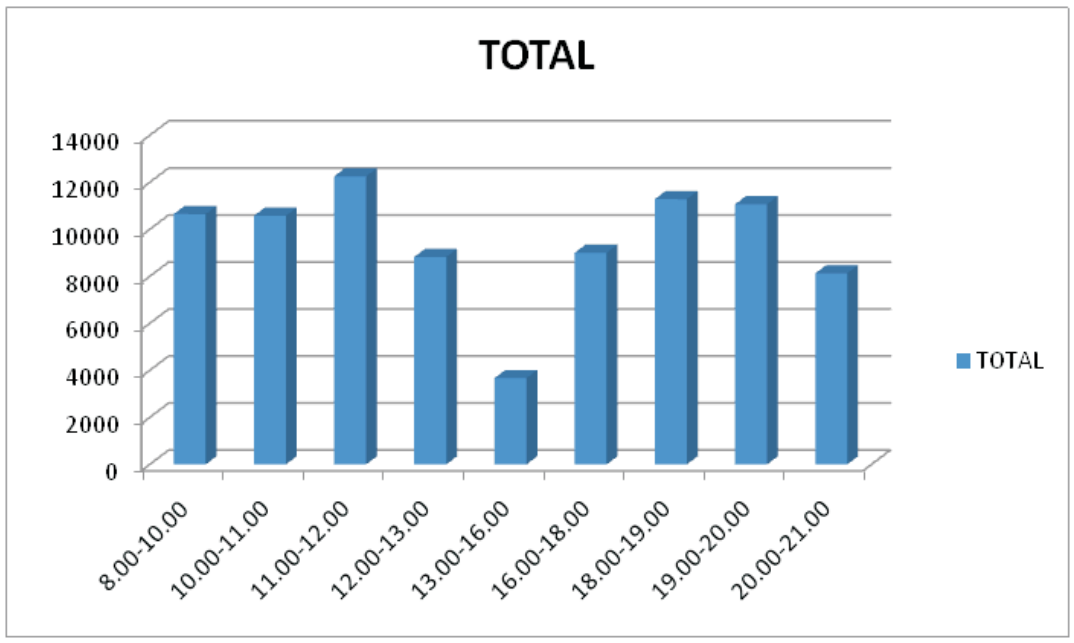
Following is a sample table collected for Monday 30th Spetember 2013. The table displays the number of different vehicles and their summation[6].

Table 4.1:

Time	2W	3W	4W	LCV	Trucks	Buses	Cycles	TOTAL
8.00-10.00	2324	4354	3453	56	32	12	423	10654
10.00-11.00	3422	2123	3432	23	123	223	1231	10577
11.00-12.00	5567	3455	4210	87	56	123	322	13820
12.00-13.00	2991	1233	3454	43	23	99	987	8830
13.00-16.00	1212	988	944	45	12	45	425	3671
16.00-18.00	3324	3154	2993	56	50	12	423	10012
18.00-19.00	3422	2123	2900	23	123	223	1231	10045
19.00-20.00	5001	4110	3241	93	56	123	322	12946
20.00-21.00	3641	1233	2104	43	23	99	987	8130

The line graphs and the bar graphs for the table 4.1 are as follows:





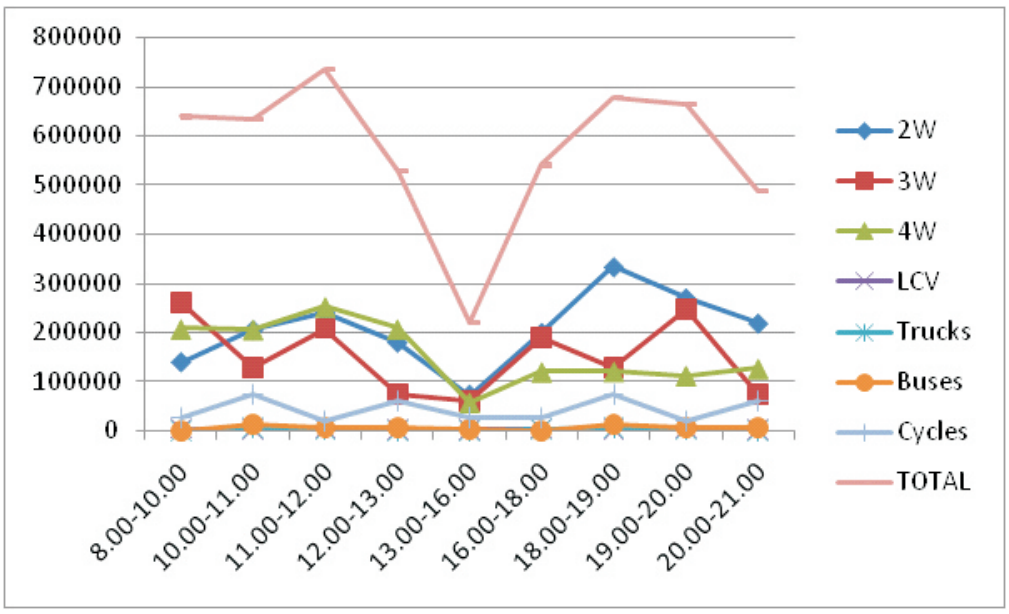
However, we need to assess that the same behaviour and pattern is being folloed daily, as such a poor random sampling will lead us to haphazard and anomolous results. It is imperative to analyze and speculate result on the basis of wide sampling otherwise the actual results may vary from the one obtained from a single sampling.

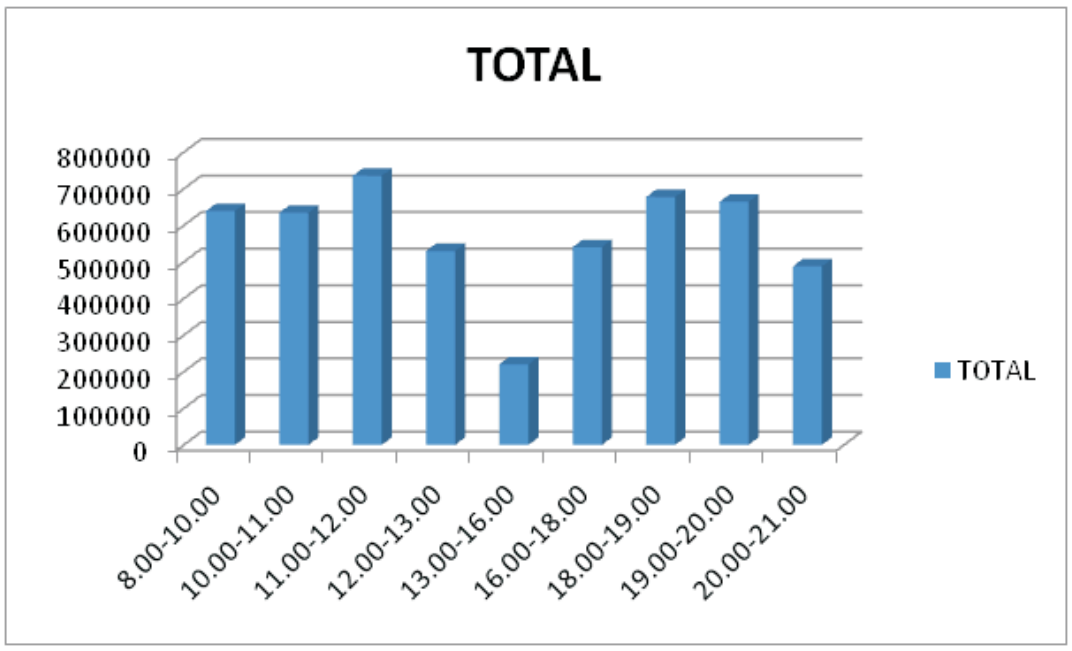
So, we have consolidated and aggregated the data for a span of 2 months ore precisely 60 days. Following are the results:

Table 4.2:

Time	2W	3W	4W	LCV	Trucks	Buses	Cycles	TOTAL
8.00-10.00	139440	261240	207180	3360	1920	720	25380	639240
10.00-11.00	205320	127380	205920	1380	7380	13380	73860	634620
11.00-12.00	240120	207300	252600	5220	3360	7380	19320	735300
12.00-13.00	179460	73980	207240	2580	1380	5940	59220	529800
13.00-16.00	72720	59280	56640	2700	720	2700	25500	220260
16.00-18.00	199440	189240	119400	3360	3000	720	25380	540540
18.00-19.00	334020	127380	119880	1380	7380	13380	73860	677280
19.00-20.00	270000	246600	111900	5580	3360	7380	19320	664140
20.00-21.00	218460	73980	126240	2580	1380	5940	59220	487800

Following are the graphs for table 4.2:



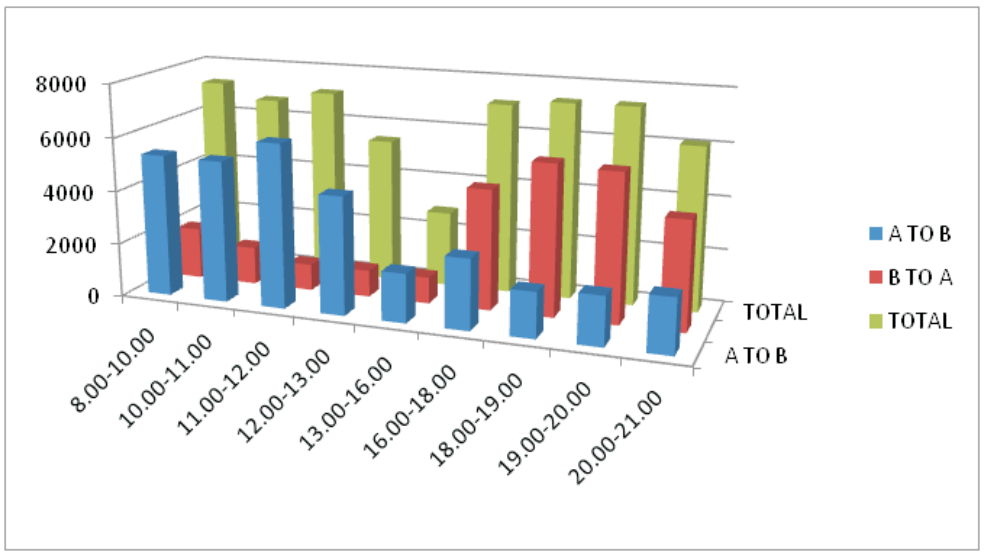


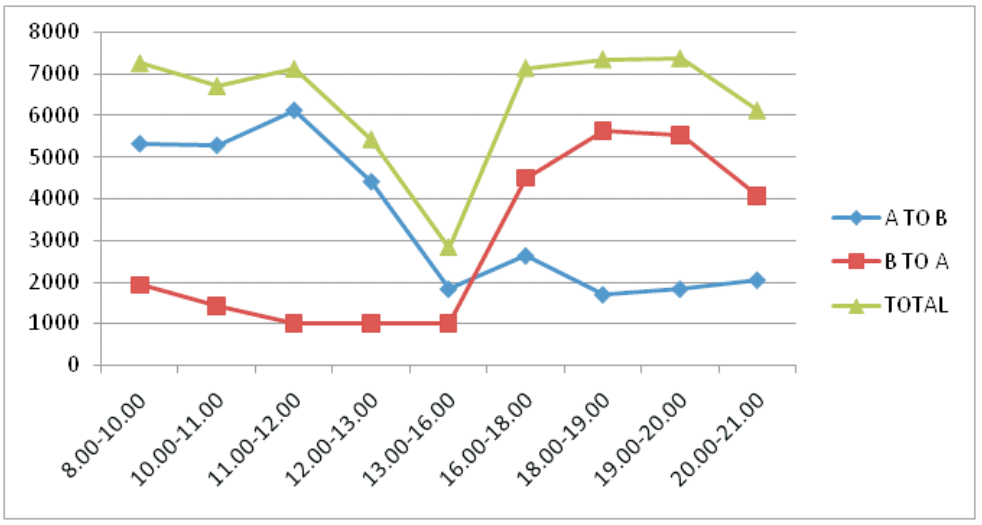
Some observations are lucidly stated from the above data and graphs for table 4.1 and 4.2 are:
1.Traffic is highly concentrated during the peak hours in the morning and evening; from 1000 HRS to 1300 HRS and from 1600 HRS to 2100 HRS.
2.There is a steep fall in the traffic afternoon.
3.The concentration of traffic in the specific hours makes us suspecious about the flow of the traffic; making it imperative to scrutinize the flow of the traffic specifically in both separate lanes.
4.The heavy concentration and slow moving traffic gives rise to more studies about:
1.Width of the road
2.Total intersecting lanes
3.Parked vehicles and pedestrians
4.Low lying unused footpaths
Hence we collected the indivisual data of the same day for two different lanes. The results were as follows:

Table 4.3:

Time	A TO B	B TO A	TOTAL
8.00-10.00	5327	1923	7250
10.00-11.00	5288.5	1411	6699.5
11.00-12.00	6127.5	989	7116.5
12.00-13.00	4415	1002	5417
13.00-16.00	1835.5	988	2823.5
16.00-18.00	2631	4504.5	7135.5
18.00-19.00	1700	5644	7344
19.00-20.00	1834	5534.5	7368.5
20.00-21.00	2049	4065	6114

Following are the line and bar graphs for table 4.3:





It becomes lucid to understand now that there is a certain changing pattern of traffic on the Law College Road,Kothrud.

USE OF STATISTICS:

Now that it has been clear that this problem has certain co-relations, we will apply co-relation to

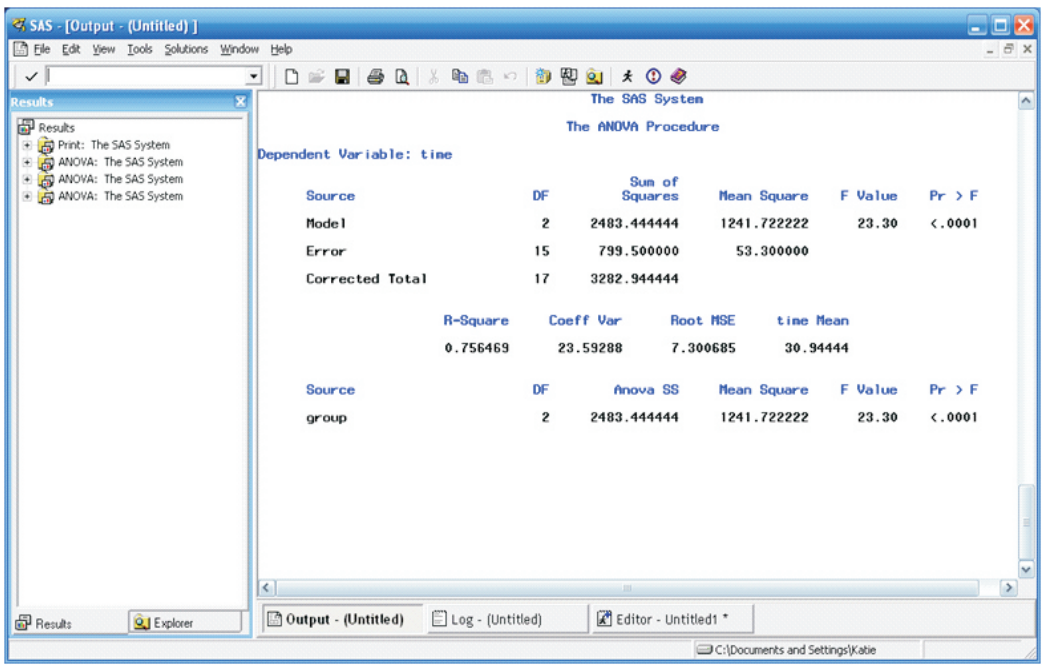
determine the co-relation co-efficient and the apply Proc UNVARIATE to determin several statistical terms.

Algorithm:
1.An ANOVA tests Ho: all group means are equal vs. Ha: at least one group's mean is different.
The ANOVA results do not tell you which group is different, only whether a difference exists.
2.For testing our hypothesis we will be using SAS code as follows;

```
PROC ANOVA DATA = relief;
        class group;
        model time = group;
RUN;
QUIT;
```

"class" tells SAS the classification variable. In general, this is going to be the effect that you are studying. In this case, the effect is "group."
"model" tells SAS the dependent variable. The general format is "model Y = X" where Y is the dependent variable, and X is the independent variable. In this case, time to relief is dependent on treatment group.
Often a "quit" statement is necessary, because SAS may continue to run a procedure until either another one has been run, or SAS has been told to quit.

Following is a sample output of the ANOVA procedure when applied on the said datasets of building materials:

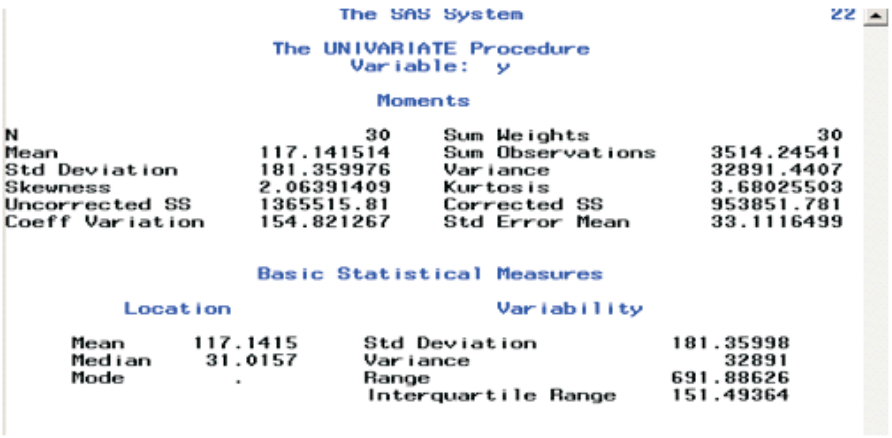


The "Between SS" is under "Model" and has a value of 2483.44.
The "Within SS" is under "Error" and has a value of 799.50.
The F*= MSB/MSW = 1241.72/53.30 = 23.30. The p-value of this F* is found under "Pr>F" and $p < 0.0001$.
Because the p-value for the test statistic (F*) is less than alpha (0.05), we reject the null hypothesis and conclude that at least two of the groups' means differ on the annual prices of material hence we will continue with the study of price hike.

1.2nd ANOVA procedure will be applied on structured intra-quarter vehicle data.

2.OR Apply Proc Unvariate for extensive descriptive statistics:

PROC UNIVARIATE DATA=one;
VAR x y;
CLASS Group;
OUTPUT OUT=Desc N=N_xN_y
MEAN=mean_xmean_y
STD=sd_xsd_y;
RUN;
Following is the procedure of the output of PROC UNVARIATE:



The SAS System 22

The UNIVARIATE Procedure
Variable: y

Moments

N	30	Sum Weights	30
Mean	117.141514	Sum Observations	3514.24541
Std Deviation	181.359976	Variance	32891.4407
Skewness	2.06391409	Kurtosis	3.68025503
Uncorrected SS	1365515.81	Corrected SS	953851.781
Coeff Variation	154.821267	Std Error Mean	33.1116499

Basic Statistical Measures

Location		Variability	
Mean	117.1415	Std Deviation	181.35998
Median	31.0157	Variance	32891
Mode	.	Range	691.88626
		Interquartile Range	151.49364

The same results as above with changes in numerical values occur which imply that there is definitely a co-relation between the number of vehicles and the time segments.
Correlation is a measure of the strength of relationship between random variables. The population correlation between two variables X and Y is defined as:

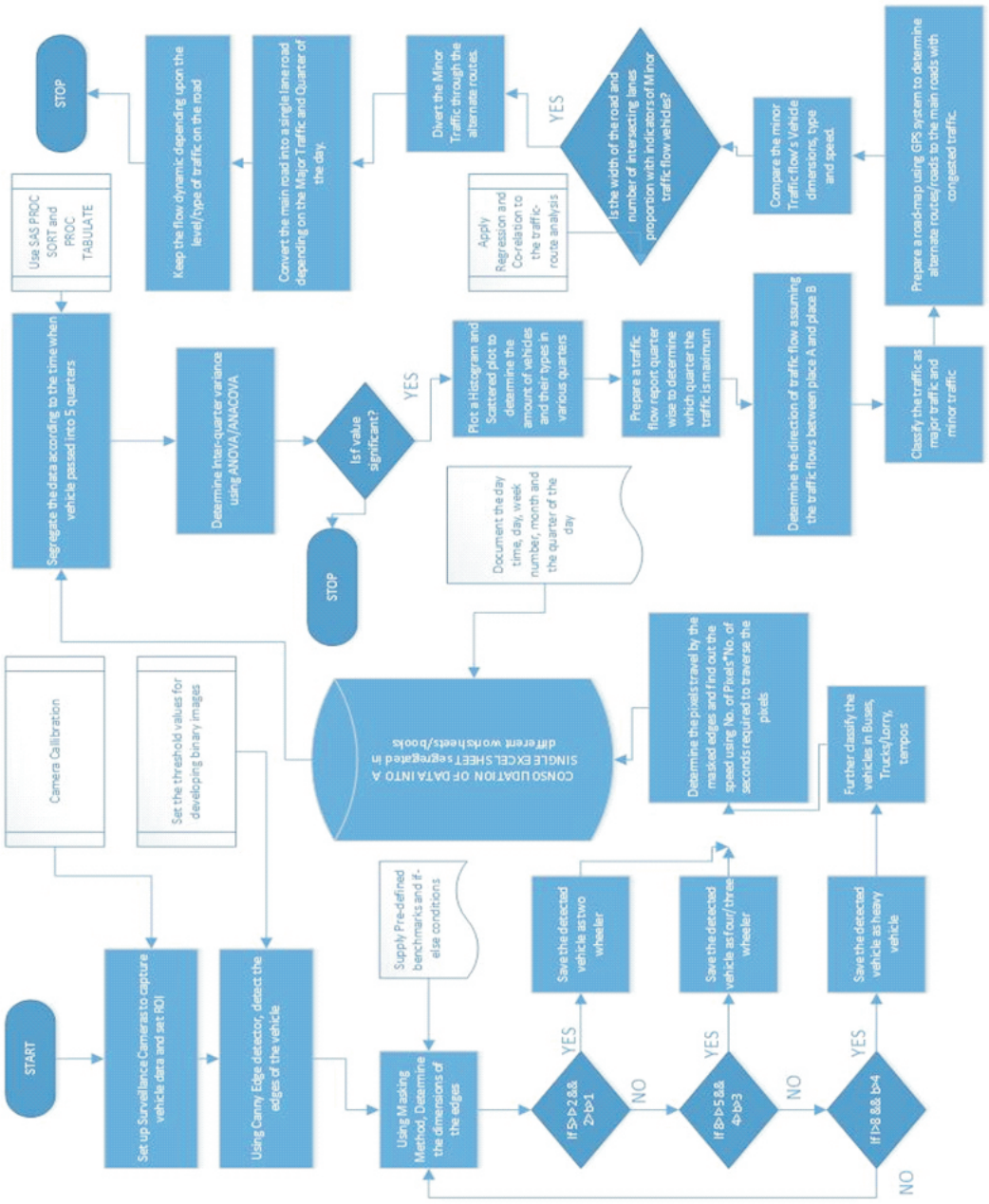
$$\rho(X, Y) = \text{Covariance}(X, Y) / \{\text{Variance}(X) * \text{Variance}(Y)\}^{1/2}$$

ρ is called the Product Moment Correlation Coefficient or simply the Correlation Coefficient. It is a number that summarizes the direction and closeness of linear relations between two variables. The sample value is called r, and the population value is called ρ (rho). The correlation coefficient can take values between -1 through 0 to +1. The sign (+ or -) of the correlation defines the direction of the relationship. When the correlation is positive (r > 0), it means that as the value of one variable increases, so does the other.
SAS codes (v8.2) for correlation coefficient, hypothesis test and confidence interval are:

```
proc corr data=a outp=corr;                    *output with Pearson correlation coefficient;
var x y;
run;
data corr_ci;
set corr (rename=(x=corr) drop=y _name_);
retain n;
if _type_='N' then n=corr;
if _type_='CORR' and corr ^= 1;
fishersz=0.5*(log(1+corr)-log(1-corr));        *Fisher Z transformation;
Sigmaz=1/sqrt(n-3);                            *variance;
l95=fishersz-1.96*sigmaz;                      *α=0.05, i.e. at 95% level;
u95=fishersz+1.96*sigmaz;
l95=(exp(2*l95)-1)/(exp(2*l95)+1);             *inverse of Fisher Z ;
u95=(exp(2*u95)-1)/(exp(2*u95)+1);             *transformation to get CI; run;
proc print data=corr_ci;
run;
```

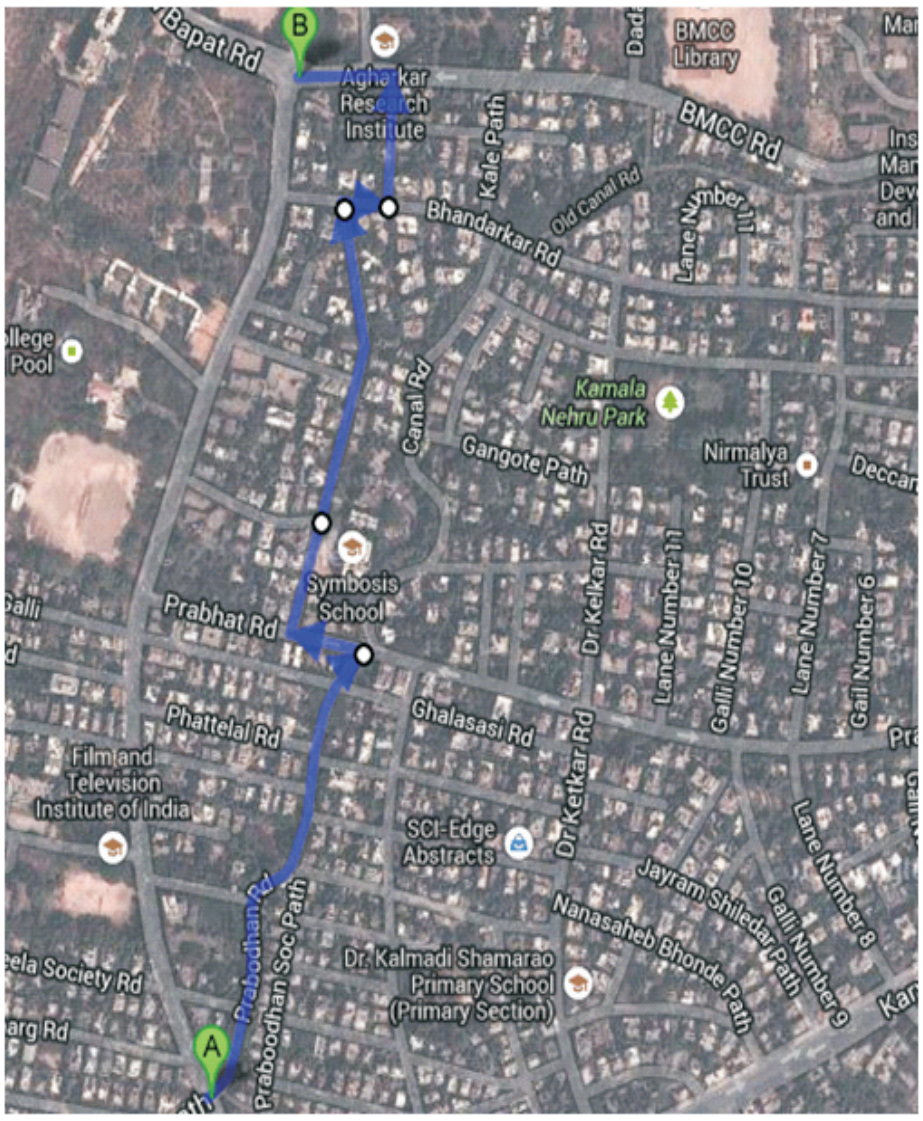
The correlation value (r^2)1. is less than 0.5; this implies that the direction of the major traffic flow and the number of vehicles in it are significantly related leading us to a conclusion of traffic being concentrated in a single direction.

2.Stop.
Pictorial Representation of the algorithm for better visualization of the flow of steps:



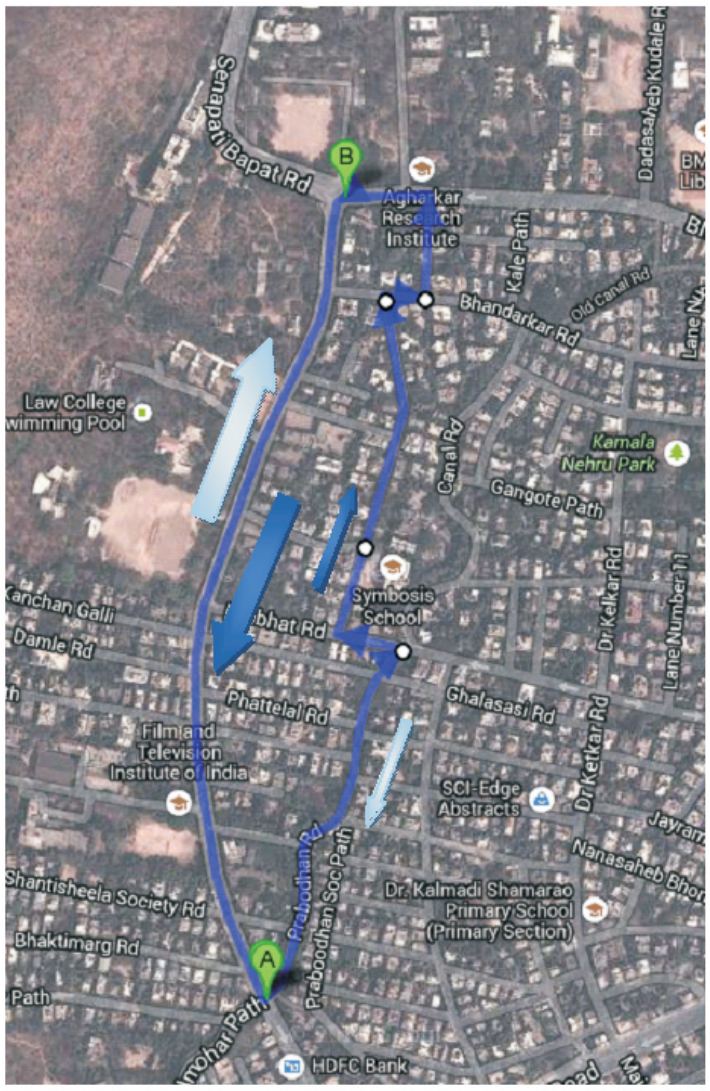
Interpretation of Results/CONCLUSION:

1.The f value being significant clearly states the variance between the vehicle movement and behavior.
2.Further, the co-relation which is greater than 0.5 states that the waiting time is directly proportional to the width of the road, clearly indicating that the road will have to be made one way to channelize the traffic smoothly, without disrupting the available resources.
3.Hence an alternate road, as shown in the figure below will be used as an alternate route for diverging the minor traffic and the main road will be used only for the major traffic flow.



Alternate route connecting place A and place B

1.At the peak hours the two traffics (major and minor) will be kept separate on main and alternate roads respectively. This will not only help the congestion to be totally nullified, but will also help the local transport authority to smoothly vigil the behavior and monitor the same.
2.However, as mentioned the drivers will be incurring an extra driving of 300 meters if, they are in the minor traffic; still increasing the speed if the driver has to go in opposite direction.
3.This will further increase fuel and time efficiency, making it easy for the commuters to manage their resources wisely.



MAJOR TRAFFIC GOING FROM A TO B

MINOR TRAFFIC GOING FROM B TO A

MAJOR TRAFFIC GOING FROM B TO A

MINOR TRAFFIC GOING FROM A TO B

ACKNOWLEDGEMENT:

We are extremely grateful towards the Regional Transport Office, Ex-Deputy Mayor and IAS Officer Dr. RamanathJha for his recognition of our efforts and continued support towards this project. We are thankful to all the faculty of Symbiosis for helping hand in theoretical proofing of the optimization concept.

REFERENCES:

[1]http://research.microsoft.com/en-us/projects/clearflow/Predictive Analytics for Traffic

[2] Paper on predicting traffic flows: E. Horvitz, J. Apacible, R. Sarin, and L. Liao (2005). Prediction, Expectation, and Surprise: Methods, Designs, and Study of a Deployed Traffic Forecasting Service, Twenty-First Conference on Uncertainty in Artificial Intelligence, UAI-2005, Edinburgh, Scotland, July

2005.
[3] Environmental Impact Statement & Traffic Survey Study For Proposed Paud Road – BalBharati Link Road, Pune (For Hon. High Court appointed Experts' Committee) August 2012
[4] Regional Transport Office, Pune, Norms for vehicle dimensions and regulation.
[5] Google Maps, Pune City (18 Degrees 31' N, 73 Degrees 51' E).
[6] 'Vehicle Detection Using Masking Method', Omkar Kulkarni, International Journal of Knowledge News, Agra.