
TRAFFIC CONGESTION ANALYSIS

Submitted for

DATA VISULIZATION AND DASHBOARD

Submitted by:

(E23MCAG0128) SWASTIK SINGH BHADOURIA

(E23MCAG0123) ADARSH KUMAR MAURYA

(E23MCAG0028) YASH AGRAWAL

Submitted to-

Dr. Nitin Arvind Shelke

July-Dec 2023

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

BENNETT UNIVERSITY
THE TIMES GROUP

INDEX

LIST OF FIGURES

LIST OF TABLES

Note: List of Figures and List of Tables shall be on separate pages.

1. **ABSTRACT**

The project is designed to develop a density based dynamic traffic signal system. The signal timing changes automatically on sensing the traffic density at the junction. Traffic congestion is a severe problem in many major cities across the world and it has become a nightmare for the commuters in these cities. Conventional traffic light system is based on fixed time concept allotted to each side of the junction which cannot be varied as per varying traffic density. Junction timings allotted are fixed. Sometimes higher traffic density at one side of the junction demands longer green time as compared to standard allotted time. To calculate the number of vehicles we can use sensors over several distances at junctions. If higher traffic is present on a side, then sensor senses the traffic and

according to that sensor output, we can adjust the signal timings. After calculating the number Raspberry pi is used as a microcontroller which provides the signal timing based on the traffic density.

The aim of the project is to solve traffic congestion which is a severe problem. Generally, each traffic light at a junction is assigned a constant green signal time. It is possible to propose dynamic time-based coordination schemes where the green signal time of the traffic lights is assigned based on the present conditions of traffic.

2. INTRODUCTION AND RELATED WORK

Introduction:

Traffic congestion is a critical issue in urban areas, impacting the efficiency of transportation systems and the daily lives of residents. This project aims to analyze a comprehensive dataset capturing traffic conditions at various junctions over time. By delving into the nuances of this dataset, we seek to uncover patterns, trends, and potential insights that can inform urban planning, traffic management, and infrastructure development.

Traffic congestion has been increasing in much of the world, developed or not, and everything indicates that it will continue to get worse, representing an undoubted menace to the quality of urban life. Its main expression is a progressive reduction in traffic speeds, resulting in increases in journey times, fuel consumption, other operating costs and environmental pollution, as compared with an uninterrupted traffic flow.

The dataset, consisting of four key columns—DateTime, junction, vehicles, and ID—offers a rich source of information about the dynamics of traffic congestion. This report presents a detailed analysis of the dataset, employing various statistical and visual techniques to extract meaningful insights.

Related Work:

In the related work section, briefly discuss any existing research, studies, or projects related to traffic congestion analysis. Highlight the significance of your project in relation to existing work. Here is an example:

Several studies have explored the impact of traffic congestion on urban environments, employing diverse methodologies and datasets. [Reference recent studies or projects in the field]. While these contributions have shed light on various aspects of congestion, the present project distinguishes itself by focusing on a specific dataset that captures not only the number of vehicles at different junctions but also temporal trends.

Our approach aligns with the growing emphasis on data-driven decision-making in urban planning. By harnessing the power of advanced analytics and visualization techniques, we aim to offer a comprehensive understanding of traffic patterns that goes beyond traditional studies. The integration of time-series analysis and machine learning methods sets this project apart, as we seek to uncover both short-term fluctuations and long-term trends in traffic congestion.

CAUSES OF TRAFFIC CONGESTIONS IN PUNE

1. Bad conditions of city roads.

2. Potholes here and there on the city roads.

3. Insufficient existing parking infrastructure in the city.

4. Encroachment on footpaths that contributes in poor discipline in pedestrian who are forced to walk on roads

5. Lack of traffic rules education among population. Sometimes highly qualified professionals like engineers and doctors behave like illiterates on the road.

6. Lack of strict traffic police system that can punish rule breakers on the roads.

7. Heavy vehicles transit into the city slows down the traffic and increases jams like situation quite often.

8. Auto drivers are the curse in the city as they dominate the roads and take dangerous cuts to go ahead and risk the life of their passengers.

9. Most of us think red lights are for street decoration and thus ignoring it is quite common.

10. Hors have become substitutes to the brakes has I seen many times when the car driver ahead of me needs to push breaks, but he pushes hom hard like anything. It just aggravates the situation and sometimes the rider ahead of him may lose control due to loud sound.

11. Bad shape of roads and less width in the busiest parts of city invites traffic to congest

12. 1 think all the roads in the city are frequently dug up by government of corporation for some reason and then it takes a long time to repair them. In some cases, they just forget to fill it up inviting accidents. In some places you don't find the signboards.

A. WHAT IS CONGESTION?

1. Popular usage and dictionary definition

The word “congestion” is frequently employed in the road traffic context, both by technicians and by the public at large. Webster’s Third New International Dictionary defines it as “a condition of overcrowding or overburdening”, while “to congest” means “to overcrowd, overburden or fill to excess so as to obstruct or hinder” something: in this case, road traffic.

It is usually understood as meaning a situation in which there are a large number of vehicles circulating, all of which are moving forward in a slow and irregular manner. These definitions are of a subjective nature, however, and are not sufficiently precise.

2. A technical explanation

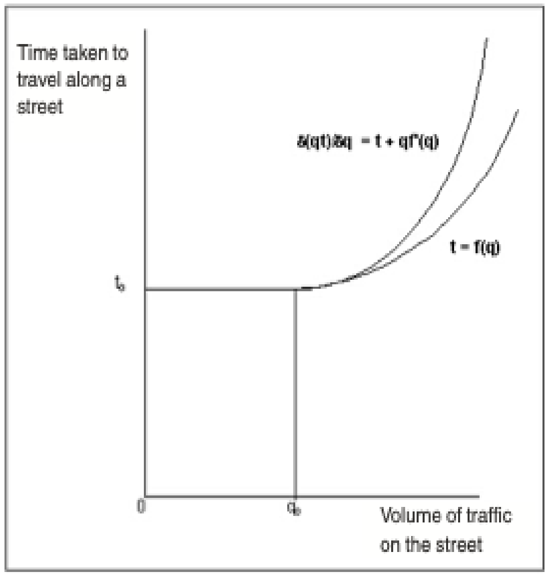
The fundamental cause of congestion is thefriction or mutual interference between vehicles in the traffic flow. Up to a certain level of traffic, vehicles can circulate at a relatively freely determined speed which depends on the legal speed limit, the frequency of intersections, and other conditioning factors. At higher levels of traffic, however, every additional vehicle interferes with the circulation of the others: in other words, the phenomenon of congestion appears. A possible objective definition, then, would be: “congestion is the situation where the introduction of an additional vehicle into a traffic flow increases the journey times of the others” (Thomson and Bull, 2001).

As traffic increases, traffic speeds go down more and more sharply. In figure II.1, the function $t = f(q)$ represents the time (t) needed to travel along a street at different levels of traffic (q). The other curve, $d(qt)/dq = t + qf'(q)$ is derived from that function. The difference between the two curves represents, for any volume of traffic (q), the increase in the journey times of the other vehicles which are in circulation due to the introduction of an additional vehicle.

Figure II.1

Figure II.1

SCHEMATIC REPRESENTATION OF THE CONCEPT OF TRAFFIC CONGESTION



Source: Prepared by the author on the basis of economic theory (supply curve).

It may be noted that the two curves coincide up to a traffic level Oq_0; up to that point, the change in the total journey times of all the vehicles is simply the time taken by the additional vehicle, since the others can continue circulating at the same speed as before. From that point on, however, the two functions diverge, and *d(qt)/dq* is above *t*. This means that each additional vehicle not only experiences its own delay but also increases the delay of all the other vehicles which are already circulating. Consequently, the individual user is only aware of part of the congestion he causes, while the rest is suffered by the other vehicles in the traffic flow at that moment (Ortúzar, 1994). In the corresponding specialized language, users are said to perceive the mean private costs, but not the marginal social costs.

Strictly speaking, users do not have a very clear idea of the mean private costs either, since, for example, few drivers have a clear idea of how much it costs them to make an additional journey in terms of maintenance, tyre wear, etc. In contrast, they do clearly perceive the costs imposed on them by the time taken to travel along a street volume of traffic on the street government –particularly the fuel tax– which are seen as mere transfers from motorists to the State, all of which distorts their manner to taking decisions.

Another conclusion which can be drawn –and which can be confirmed by simple observation– is that at low levels of congestion an increase in the traffic flow does not significantly increase journey times, but at higher levels the same increase causes considerably greater overall delays.

According to the definition given earlier, congestion begins with a traffic level Oq_0. Generally, however, this occurs at relatively low traffic levels, unlike what most people think.

3. Towards a practical definition in the case of road traffic

Even some specialized studies do not give very strict definitions of congestion. Thus, two well-known specialists in transport modeling consider that congestion occurs when the demand nears the capacity of the travel infrastructure and transit times rise to a much higher level than that obtaining in conditions of low demand (Ortúzar and Willumsen, 1994). Although this definition reflects the perceptions of the average citizen, it does not propose exact limits for the point at which the phenomenon begins.

An attempt to define the term precisely in line with the usual perception of it was that made in a draft law like that approved by the Chilean Chamber of Deputies for the introduction of road use tariffs. As the aim was to avoid the possibility of discretionarily on the part of public authorities, the definition was very precise. A road was considered to be congested when, in more than half of its total length (including not necessarily continuous stretches), the average speed of the traffic flow was less than 40% of the speed in unrestricted conditions. This state of affairs must be registered for at least four hours a day between Tuesday and Thursday, on the basis of measurements made for four consecutive weeks between March and December. An exact definition was also given of congested areas.[1] This definition was perhaps too precise and difficult to apply in practice, although so far it has not been necessary to apply it, since the draft law has not yet received full legislative approval.

Without going into such detail, yet continuing to seek objectivity, the term congestion could be defined as "the situation which occurs if the introduction of a vehicle into a traffic flow increases the travel times of the other vehicles by more than x%". An objective although still somewhat arbitrary definition of congestion would be to define it as the volume of traffic at which $d(qt)/dq = at$, where a equals, for example, 1.50. In other words, congestion would begin when the increase in the journey time of all the vehicles already present in the flow was equal to half of the travel time of an additional vehicle.

B. THE CAUSES OF CONGESTION

1. Characteristics of urban transport which cause congestion

The transport system, including the provision of urban land for transport infrastructure, operates with very special characteristics, including in particular the following:

- The demand for transport is “derived”: in other words, journeys are rarely made because of an intrinsic desire to travel but are generally due to the need to travel to the places where various kinds of activities are carried on, such as work, shopping, studies, recreation, relaxation, etc., all of which take place in different locations.
- The demand for transport is eminently variable and has very marked peak periods in which a large number of journeys are concentrated because of the desire to make the best use of the hours of the day to carry on the various types of activities and have an opportunity to make contact with other persons.
- Transport takes place in limited road spaces, which are fixed and invariable in the short term; as will readily be understood, it is not possible to store up unused road capacity for later use at times of greater demand.
- Theformsoftransportwhichhavethemostdesirablecharacteristics–security, comfort, reliability, and autonomy, as in the case of private cars –are those which use the most road space per passenger, as will be explained below.

- Especially in urban areas, the provision of road infrastructure to satisfy rush hour demand is extremely costly.

- Because of the above factors, congestion occurs at various points, with all its negative consequences of pollution, heavy expenditure of private and social resources, and adverse effects on the quality of life.

A further aggravating factor is that, as noted in the previous section, the cost of congestion is not fully perceived by the users who help to generate it. Every time this happens, more of the good or service in question is consumed than is desirable for society as a whole. As users are not aware of the greater costs in terms of time and operation that they cause to others, their decisions on routes, forms of transport, points of origin and destination and time of execution of journeys are not taken on the basis of the social costs involved, but their own personal costs or, rather, an often partial perception of those costs. The natural result is the over-exploitation of the existing road system, at least in certain areas and at certain times.

2. The problem is mainly caused by private car users

Some vehicles cause more congestion than others. In transport engineering, each type of vehicle is assigned a passenger car equivalence called a pcu, or passenger car unit. A private car is equivalent to 1 pcu, while other vehicles have equivalencies corresponding to their disturbing influence on the traffic flow or the space they occupy in it, as compared with a private car. A bus is normally considered to be equivalent to 3 pcus and a truck to 2 pcus. Strictly speaking, however, the pcu factor varies according to whether the vehicle in question is close to an intersection or is in a stretch of road between two intersections.

Although a bus causes more congestion than a private car, it generally carries more persons. Thus, if a bus carries 50 passengers but a private car only carries an average of 1.5 persons, then every private car passenger is causing 11 times as much congestion as a bus passenger. Consequently, other things being equal, congestion is reduced if the share of buses in the intermodal journey mix is increased. Unless buses transport less than 4.5 passengers, on average they cause less congestion than private cars. It is not normal for buses to transport fewer than 4.5 passengers, although this can sometimes happen, as for example in some sectors of Santiago, Chile, at off-peak hours in the late 1980s, or in Lima ten years later.

The existence of an excessive number of public transport vehicles can help to increase congestion, as noted in a number of cities. One of the features of the current economic models is deregulation, and in the case of urban passenger transport, broad deregulation is normally reflected in an exaggerated increase in the number of buses and taxis and a deterioration in the levels of order and discipline associated with their operation. This phenomenon bore much of the blame for the deterioration in congestion in Santiago in the 1980s and in Lima in the following decade.

The liberalization of the rules on the importation of used vehicles and the deregulation of public transport both had particularly serious effects in Lima. In Santiago, which had some 4,300,000 inhabitants in the late 1980s, there were relatively few cases of the importation of used vehicles, and the public transport fleet (all types of buses, plus collective taxis) did not amount to more than 16,000 vehicles. In the mid-1990s in Lima, however, which had some 6,700,000 inhabitants at that time, the public transport fleet amounted to at least 38,000 vehicles (and some sources indicate that the real number was close to 50,000). In other words, in the mid-1990s the number of units per inhabitant in Lima was between 52% and 101% higher than it had been in Santiago some seven years before, at a time when deregulation in Chile was having its most striking results.

3. The state of the roads and driving habits also contribute to congestion

(a) Urban road networks: design and maintenance problems

Faulty design or maintenance of road systems causes unnecessary congestion. In many cities there are frequent cases of failure to mark traffic lanes, unexpected changes in the number of lanes, bus stops located precisely where the road width becomes narrower, and other shortcomings which disturb a smooth traffic flow. Likewise, road surfaces in bad condition, and especially the presence of potholes, give rise to increasing constraints on road capacity and increase congestion. In many Latin

American cities, such as Caracas, the accumulation of rainwater on roads reduces their traffic capacity and hence increases congestion.

(b) Some driving habits cause more congestion than others

There are drivers who show little respect for other road users. In some cities, such as Lima, many drivers try to cut a few seconds off their journey times by forcing their way into intersections and blocking the passage of other motorists, thus causing economic losses to others which are much greater than their own gains. In other cities, such as Santiago, it is a tradition for buses to stop immediately before an intersection, thereby causing congestion (and accidents). In those same cities, as in others that have an excessive number of taxis that do not habitually operate from fixed taxi ranks, these vehicles crawl along looking for passengers, and this also gives rise to congestion.

In addition to these practices, the traffic flows also often include old and poorly maintained vehicles, as well as some drawn by animals. It must be borne in mind that when the traffic flow resumes after being stopped at a traffic light, a form of congestion ensues because vehicles with a normal rate of acceleration are held up by slower vehicles located in front of them. Furthermore, a vehicle which is stopped or moving sluggishly seriously affects the smooth flow of traffic, since in effect it blocks a traffic lane.

Illustration II.1
THE SELFISH AND UNDISCIPLINED BEHAVIOUR OF LIMA MOTORISTS REDUCES THE CAPACITY OF THE ROAD SYSTEM TO A FRACTION OF ITS REAL POTENTIAL

Source: Photograph by Ian Thomson.

(c) Insufficient information is available on traffic conditions

Another factor which increases congestion is ignorance of the prevailing traffic conditions. If a motorist with two possible routes, A and B, for reaching his destination knew that traffic conditions were bad on route A, he could use route B, where his own contribution to congestion would be less. A study of a hypothetical city made in the University of Texas in the United States indicates that the fact of being well informed about traffic conditions in different parts of the road network can reduce congestion much more than such drastic measures as levying charges for using congested streets (IMT, 2000). Basic unfamiliarity with the road system can also increase the average distance of each journey and thereby contribute to congestion.

(d) The result is that there is a generalized reduction in capacity

Generally speaking, both the way motorists drive and the state of the road and vehicles mean that in Latin America a street or urban road network will assuredly have a lower capacity than one of similar dimensions located in Europe or North America. Measurements made in Caracas in the early 1970s showed that an expressway there had only 67% of the capacity of a United States expressway of similar size.[2] The actual percentage difference may vary from one city to another, but there is no doubt that the road systems of Latin American cities are relatively prone to congestion.

4.There is also an institutional problem

In almost all Latin American cities, the deterioration in traffic conditions has been significantly worse than it could and should have been, partly because of inappropriate actions by the corresponding authorities. It is obvious that the problem has clearly overtaken the institutional capacity to deal with the situation.

So far, the reaction of the authorities has only been of a piecemeal nature, because in virtually the whole region the responsibility for urban transport planning and management is split up among a host of bodies, including various national ministries, regional governments, municipalities, suburban train or metro companies, the traffic police, etc. Each of these does what it considers to be most appropriate, without taking much account of the repercussions on the interests of the other institutions.

A municipality, for example, fearing the diversion of economic activity to another part of the city, may authorize the construction of multi-storey car parks, or allow parking on the streets, without bothering about the impact of the congestion thus generated on road users who have to cross through the area in question.

Another situation which reflects the consequences of decisions taken without coordination and without considering their broader repercussions may occur in the context of a mass transit system such as the metro. Because of the greater accessibility provided, land use becomes more dense and office blocks are built, and as municipal regulations usually demand a certain minimum number of private parking spaces for such buildings, this encourages the staff to come to work in their cars. Thus, this set of measures fosters increased congestion.

Furthermore, in such a sensitive area as urban transport, strong pressures are exerted by organized groups, such as transport interests, as well as by politicians, who put forward their own points of view and sometimes take up arms on behalf of particular interests, which complicates the situation still further.

All the above factors are a source of distortions, yet urban transport should be handled in an integral, technical manner, instead of measures being taken separately by each institution or in favour of sectoral interests.

C. THE INVASION OF THE PRIVATE CAR

The last decade of the twentieth century brought with it a big increase in the number of private cars circulating in Latin America, as well as in their use for the most varied purposes, including journeys to places of work or study, thus exerting heavy pressure on the road network. What are the causes of these phenomena?

1. Economic reforms have made private cars more easily accessible

Among other effects, the economic reforms adopted in the region in the 1990s brought with them higher economic growth rates and lower car prices.

Instead of the almost always negative per capita GDP growth rates of the 1980s, the 1990s brought relatively high positive growth rates. Thus, for example, Uruguay went from an average annual growth rate of -1% between 1981 and 1988 to a rate of +4% between 1991 and 1994 (ECLAC, 1989 and 1995a). This had a favourable repercussion on personal income levels, thus making more resources available for the acquisition of consumer durables.

At the same time, in many cases there was a reduction in the tax burden on automobiles, especially in customs duties. Moreover, in some countries there was an appreciation in the exchange rate, thus making imported products cheaper to buy. In Colombia, for example, the real exchange rate in 1994 was only equivalent to 75% of that prevailing in 1990 (IDB, 1995).

This tendency does not necessarily mean that actual prices are lower, because at the same time the quality of vehicles has improved. In the case of those vehicles whose characteristics have remained relatively unchanged, however, there has been a real reduction in their purchase prices. In the Chilean market, for example, in 1996 a Volkswagen Beetle cost the equivalent of US$ 7,780, whereas in 1982 it had cost the equivalent of US$ 8,902 at 1996 prices.

The real reduction in the prices of used cars has undoubtedly been even greater, although it is difficult to obtain reliable data in this respect. The rate of depreciation of private cars is directly related with the rate of ownership. In countries where there are few vehicles per person, a second-hand car is a relatively scarce good, and the price at which it is sold will reflect a limited supply and, sometimes, abundant demand. The rise in rates of vehicle ownership in Latin America in recent years has reduced the relative scarcity of used cars, thus tending to increase supply and reduce demand, because a larger proportion of the population now already have one, and hence drive down prices, putting such vehicles within the reach of lower-income families.

Consequently, in the current Latin American situation real incomes are rising and automobile prices are tending to go down.

2. The popularization of private car ownership

In Latin American cities, the evolution of residents' incomes and car prices –especially those of used cars– means that ownership of a vehicle is ceasing to be an unattainable dream and is becoming an accomplished fact for many families. The increase in the rate of car ownership is a phenomenon which is repeated almost everywhere in Latin America and has made it possible –especially for the middle class– to reap one of the most important fruits of technological progress in the twentieth century.

In the countries where economic reforms were implemented rapidly, automobile imports increased equally fast (see table II.1).

Table II.1

ECUADOR AND PERU: IMPORTS OF PASSENGER VEHICLES [a]

(Thousands of dollars)

Year	Ecuador [b]	Peru
1989	10 062	6 482
1990	23 432	11 880
1991	23 554	170 668
1992	166 109	213 018
1993	245 895	165 647
1994	374 038	252 421

Source: ECLAC, on the basis of official data.

[a] Figures exclude buses.

[b] Figures refer specifically to private transport vehicles, within the general heading of consumer durables.

The column corresponding to Peru shows that between 1990 and 1991 the value of automobile imports increased by a factor of 14. Peru freed not only the importation of new vehicles but also that of used ones (except for a brief period between February and November 1996). Consequently, the average unit cost went down, indicating that the number of units imported must have increased even more than the total value of imports.

In some countries that manufacture motor vehicles themselves, the economic reforms resulted in an increase both in vehicle imports and in domestic production. This was so in Brazil, where automobile imports had been subject to heavy duties, as part of a policy designed to promote domestic production of these goods. Thus, between 1990 and 1994 imports grew by over 10,000%, albeit starting from a very low level, yet domestic automobile production also rose, by 70%. Vehicle exports were reduced because manufacturers preferred to sell their output on the growing domestic market (see table II.2). Another factor which influenced the situation, during a period from mid-1994 on, was the appreciation of the local currency. A concrete result was that in Sao Paulo, between 1990 and 1996 the population grew by 3.4% but the vehicle fleet expanded by 36.5%.

Table II.2
BRAZIL: APPARENT CONSUMPTION OF AUTOMOBILES [a]

Year	Units		Apparent consumption of automobiles[a]
	Imported	Produced	
1990	1 310	602 545	483 084
1991	11 146	615 097	499 090
1992	30 714	667 229	454 817
1993	70 438	929 582	750 413
1994	138 679	1 026 827	890 691
1995	320 261	1 147 897	1 278 437

Source: Economic Commission for Latin America and the Caribbean (ECLAC), "El tránsito urbano en la era de la apertura económica", *Boletín FAL*, No. 132, Santiago, Chile, March-April (http://www.eclac.cl/transporte), 1997.
[a] Production plus imports, less exports.

On the basis of data corresponding to the 34 municipalities in Greater Santiago, the following equation was developed to determine the number of automobiles per family:

$$y = e^{(0.2850 - 134.5746/x)}$$

where y = number of automobiles per family x = monthly income per family in 1990 pesos.

This equation has the expected form, although it could perhaps be subject to some technical reservations.[3] By using it, it is possible to estimate the elasticity or unit variation in the rate of automobile ownership with respect to income level. Table II.3 shows that this elasticity is inversely related to income level. Although the elasticity in low-income communes (La Pintana) is very high, a 1% increase in income only gives rise to a small increase in the absolute number of automobiles per family. In contrast, a 1% increase in income in a middle- income commune results in an increase in the absolute number of automobiles per family which is very similar to that registered for a very high-income commune.

Table II.3
SANTIAGO, CHILE (THREE MUNICIPALITIES): ESTIMATED RELATION BETWEEN INCREASE IN AUTOMOBILE OWNERSHIP PER FAMILY AND INCREASE IN AVERAGE INCOME

Municipality	Monthly family income	Automobiles per family	Elasticity of rate of automobile ownership per family with respect to family income	Increase in number of automobiles per family for a 1% increase in family income
Vitacura	589 700	1.71	0.23	0.0039
Santiago (centre)	126 700	0.311	1.06	0.0033
La Pintana	39 730	0.051	3.39	0.0018

Source: J. Kain and Z. Liu, *Efficiency and Locational Consequences of Government Transport Policies and Spending in Chile*, Harvard, Harvard Project on Urban and Regional Development in Chile, 1994, table A.7.

The most important conclusion to be drawn from this analysis is that an increase in income results in significant expansion of automobile ownership, not only in the richest neighborhoods but also in middle-income areas. Thus the total number of automobiles in Santiago grew at the rate of 8% per year during the 1990s.

3. Where there are fewer cars it nevertheless seems harder to get about

The growing number of vehicles undoubtedly Favours increased congestion, but at all events the rates of automobile ownership in Latin American cities are still much lower than in developed countries. In 1980, the number of automobiles per person in North American cities such as Houston, Los Angeles, Phoenix, San Francisco, Detroit, Dallas, Denver, Toronto and Washington was between 0.55 and 0.85, while in European cities such as Brussels, Amsterdam, Copenhagen, Frankfurt, Hamburg, London, Stuttgart and Paris it was between 0.23 and 0.43. Ten or fifteen years later, some Latin American cities (such as Chiclayo or Huancayo in Peru) still had no more than 0.02 cars per inhabitant, and in Lima, even though the boom in vehicle imports had already begun, there were still no more than 0.05 cars per person, while in Santiago there were 0.09. On the other hand, in a few Latin American cities the rate of ownership was already nearing the lower limit of Western European cities. In Curitiba, for example, in 1995 there were already close to 0.29 cars per person.

Nevertheless, there is evidence that it is easier to move about in the big cities of the developed world than in the comparable cities of Latin America. In Quito, whose population in 1990 was approximately one million, the average journey time between home and workplace was 56 minutes, whereas in Munich, which had approximately 1.3 million inhabitants, the corresponding time was only 25 minutes. Likewise, in Bogotá (5 million inhabitants) the journey time was 90 minutes, while in London (6.8 million) it was 30 minutes. Many other examples along the same lines could be quoted. Clearly, in the cities of the developed world there is a greater capacity to live with the automobile while avoiding its worst consequences, but Latin America has not yet learned to do this.

Furthermore, it would appear to be easier to move about in the Latin American cities with the highest rates of car ownership than in many where the rates are lower. Curitiba, for example, has more cars per person than Guatemala City, which is of similar size, but travelling in the first-named city, whether by car or in public transport, is a good deal less disagreeable than in the Central American city.

The explanation for these apparent contradictions is to be found in the marked propensity to make intensive use of private cars for all kinds of purposes.

4. The strong influence of subjective factors

One feature which aggravates congestion in Latin America is the marked preference of the population to use private cars. A clear example was Mexico City, which has suffered for years from acute problems of congestion. In order to reduce environmental pollution, it was decided to prohibit the use of one- fifth of the existing vehicles from Monday through Friday, but even this drastic measure did not succeed in persuading those affected to use public transport, even though there was an extensive metro system. Instead, the widespread response was to acquire additional vehicles to evade the effects of the measure, since many people preferred to suffer the effects of congestion rather than use public transport.

In such circumstances, even if the authorities responsible for Latin American urban transport had clear ideas about how to control traffic in the cities (which is unfortunately often not the case), it would be difficult to put them into practice because members of parliament and city councillors, worried about losing votes among the increasingly numerous group of private car owners, would not approve them.

The inhabitants of the cities of the developed world are less likely to use their cars to go to the office in the morning rush hour. A clear distinction is drawn between owning a car, and using it in situations that give rise to major difficulties. A New York or London banker living in the suburbs would never dream of travelling every day to Wall Street or the City in his private car, because in both cases there is a good-quality public transport system. In contrast, his opposite number in Sao Paulo or Santiago would never dream of travelling to the city center any other way. It is likely, however, that in the future there will be a change of attitude among motorists, and indeed, in some cities with a notably higher level of culture –such as Buenos Aires, where the quality of public transport is also higher than in most Latin American cities– there is already some evidence of a greater willingness to use public transport than in some other cities of the region.

What is the reason or explanation why there is such a strong preference for using private cars? One important aspect in this connection is that of status. In Latin America, the automobile is still considered not only a means of transport, but also an indication of its owner's status in society. A person driving a BMW will be considered as superior to one driving a Suzuki, while a person who arrives at the office by car rather than bus is seen as someone who has moved up in the world. The prestige attached to being a car owner is a strong factor in the volume of traffic.

In addition to these reasons related with the social structure and cultural characteristics, the following considerations are also important in the region:

- The poor quality of the buses compared with the aspirations of car owners
- The fact that the buses are very crowded at rush hours
- The feeling of insecurity caused by the dangerous way some bus drivers operate their vehicles[4]

- The real or assumed possibility of being a victim of delinquents on board public transport vehicles.

The preference for travelling by private car becomes a problem at rush hours, when there is a concentration of journeys for reasons of work or study. Not even serious delays in journeys are enough to cause people to stop using their cars. If they had to choose between reaching their destination slowly by private car on congested roads and arriving a little more quickly by public transport, it is by no means certain that many Latin American motorists would opt for the latter alternative.

The strong preference for the private car therefore has a number of consequences, such as the following:

• The number of motorists willing to move to new public transport systems of no more than regular quality may be quite small, so that the great majority of users of a new metro line would come from former bus users rather than private motorists

- In order to interest private motorists in public transport it would be necessary to offer them a better-class option, not only in terms of objective quality (fares, journey times and frequency of service) but also in terms of its subjective features (air conditioning, reclining seats, etc.)

- Even if high taxes are imposed on fuel, road use or parking, this would only cause a few people to change to public transport. Thus, i) these measures would serve rather to collect money that could be used to change the habits of the travelling public, and ii) while raising these levels of taxation would produce considerable fiscal income, it would bring relatively few social benefits

The preference for travelling by private car can also have other consequences which go beyond the limits of the transport sector proper and have negative macroeconomic implications. Consider, for example, the rises in international oil prices in 1999 and 2000. The typical Latin American motorist probably did not reduce his vehicle use much but instead restricted his consumption of other goods and services –many of them produced domestically– thereby reducing the demand for them in the short term. At the same time, importing countries had to increase the amount of foreign exchange spent on fuels because of their higher prices and the fact that demand for them is inelastic or at least not very sensitive to price variations.

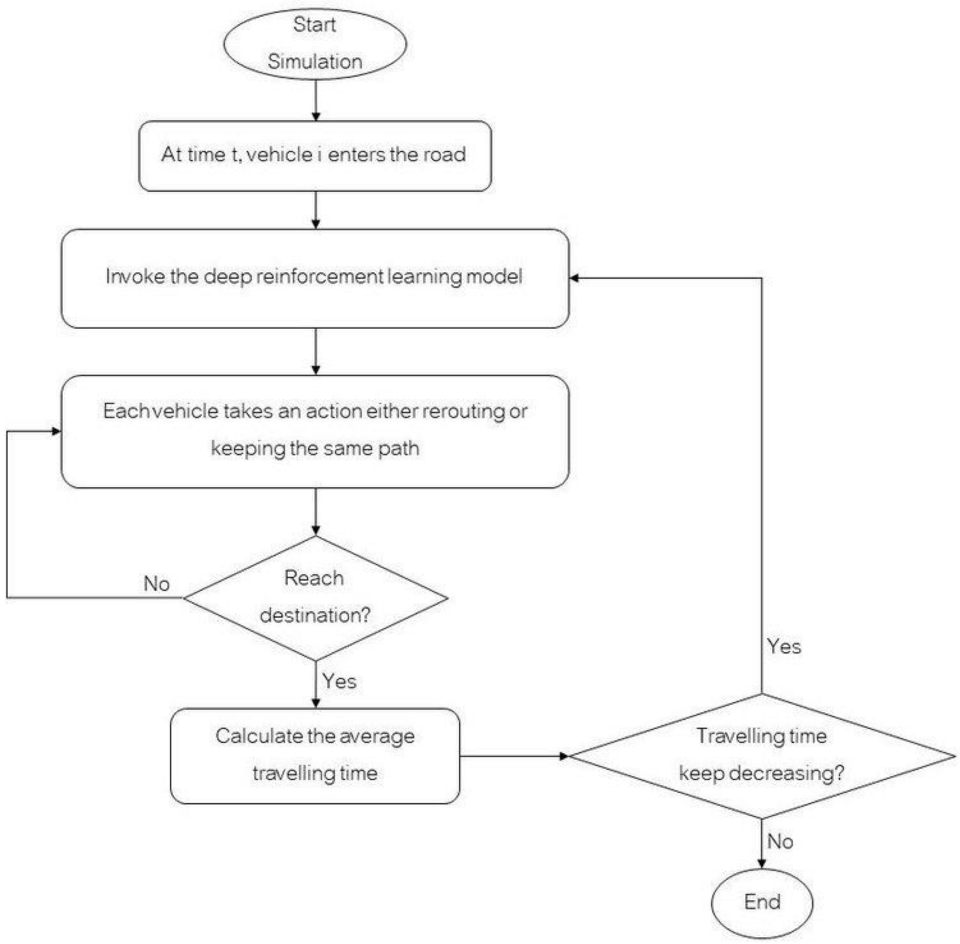
Having a car to go to a shopping center, visit friends or relations in distant parts of town, or travel outside the city is one of the fruits of economic development, and its costs are generally internalized to a large extent by car owners, since these journeys are made at times of low congestion. Using the car every day to go to the office or the city center generates high external costs in terms of congestion and pollution and does considerable harm to society, however. Securing a better balance between the ownership and use of private cars is therefore one of the main challenges to be faced today in the Latin American transport sector.

3. METHODOLOGY

1. Data Preprocessing:

The initial step involved loading the dataset and converting the 'DateTime' column to a datetime format for temporal analysis. This allowed us to understand the structure of the dataset and work with time-related trends effectively.

To ensure the integrity of the analysis, we performed data cleaning procedures. This included handling missing values, outliers, and addressing any inconsistencies in the dataset that could impact the accuracy of the results.



Normalization/Standardization:

Numerical variables were examined, and if necessary, they were normalized or standardized to bring them to a common scale, ensuring that no single variable dominated the analysis due to its scale.

2. Exploratory Data Analysis (EDA):

Descriptive statistics, such as mean, median, and standard deviation, were calculated for numerical variables. This provided an overview of the central tendencies and variability within the dataset, aiding in the identification of potential patterns.

<u>Visualization:</u>

Various plots were employed for EDA. Line plots were used to illustrate trends in traffic congestion over time. Scatter plots showcased the relationship between the number of vehicles and junctions, providing insights into spatial patterns. Box plots visualized the distribution of vehicles at each junction, highlighting potential variations.

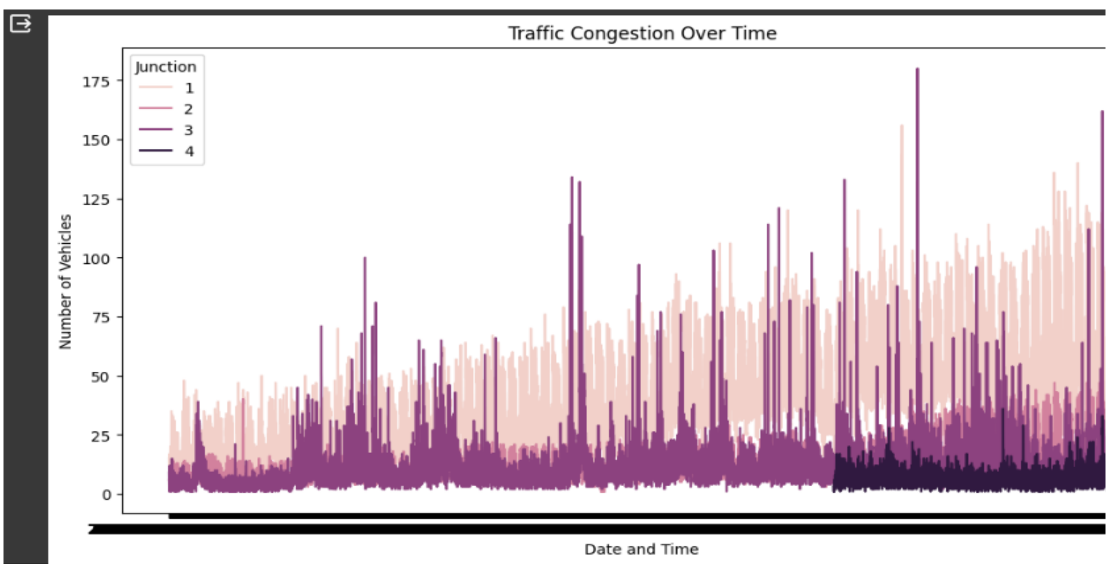


Figure 3: Line Plot for Number of vehicles on a specific date

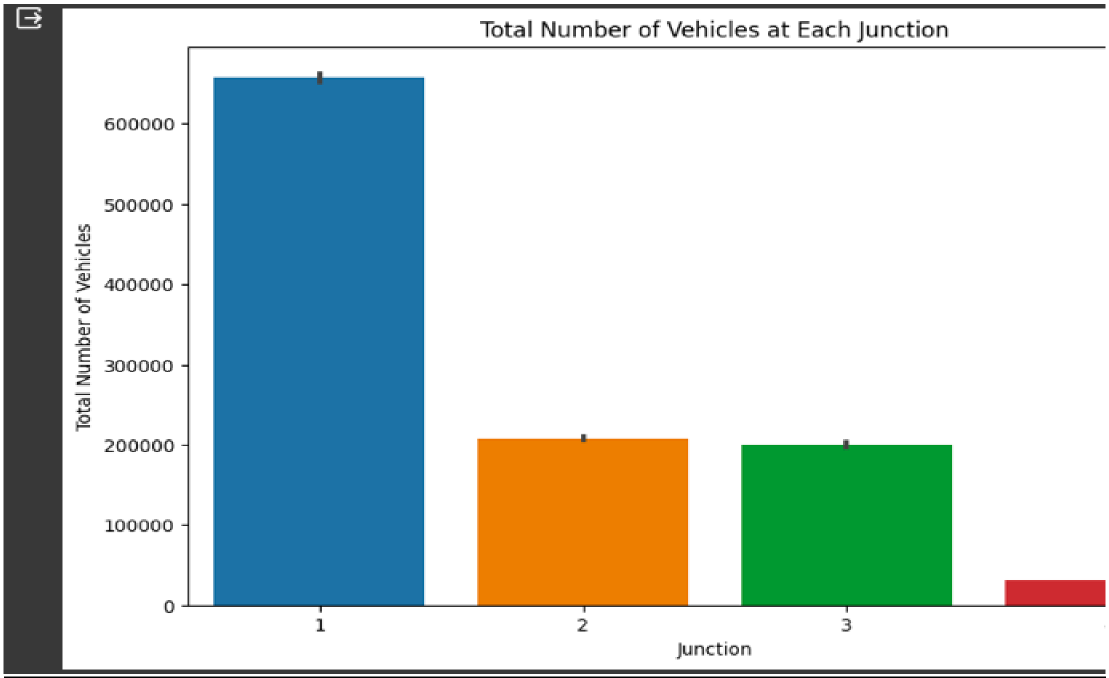


Figure 4: Box Plot to show Total number of vehicles at each junction

Correlation Analysis:

The correlation heatmap revealed relationships between numerical variables. Positive/negative correlations between variables such as 'junction' and 'vehicles' were identified. This analysis provided a foundation for understanding how different factors interacted within the dataset.

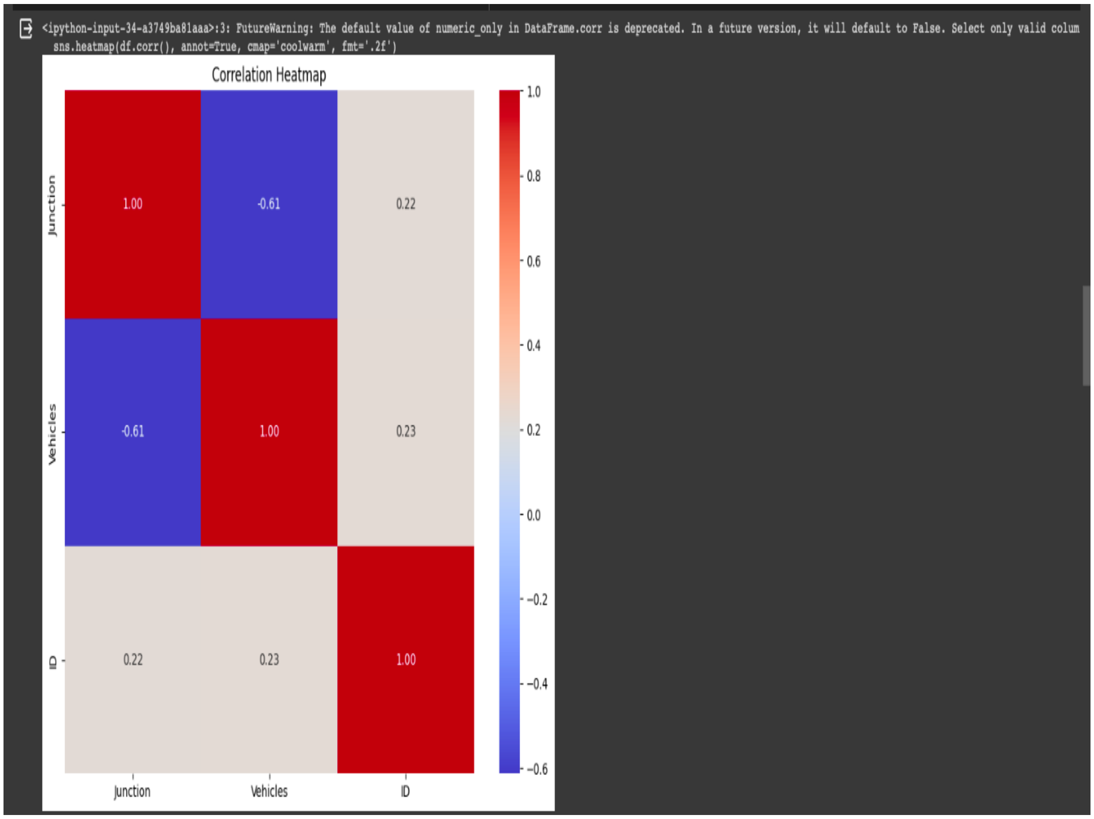


Figure 5: Correlation heatmap to explore relationships between Variables

3. Time Series Analysis:

<u>Temporal Aggregation:</u>

The dataset was aggregated temporally to explore trends at different intervals (e.g., daily, hourly). This step facilitated a more granular examination of temporal patterns in traffic congestion.

<u>Time Series Decomposition:</u>

Seasonal decomposition allowed us to break down the time series into its constituent components—trend, seasonality, and residuals. This decomposition aided in identifying recurring patterns and understanding the overall temporal behavior of traffic congestion.

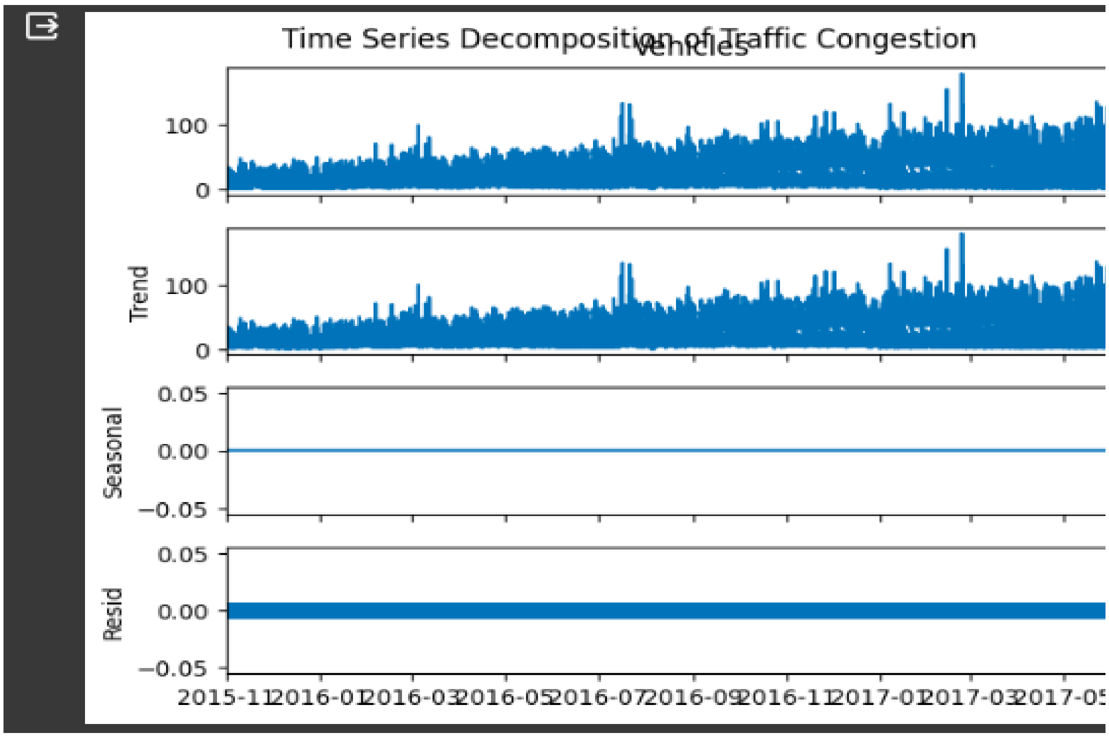


Figure 6: Time Series Decomposition to identify trends, seasonality, and residuals

Statistical Tests:

If applicable, statistical tests (e.g., autocorrelation tests) were conducted to assess the significance of trends or seasonality within the time series. This step added a quantitative layer to the temporal analysis.

4. Machine Learning Models:

Feature Engineering:

Additional features, if created, are explained. These could include engineered variables that capture specific aspects of traffic conditions, enhancing the predictive power of machine learning models.

Model Selection:

The rationale behind selecting specific machine learning models (e.g., regression, clustering) was outlined. Each model was chosen based on its suitability for the analysis goals.

Training and Evaluation:

Models were trained on the dataset, and hyperparameter tuning was performed to optimize their performance. Evaluation metrics, such as mean squared error or accuracy, were utilized to assess model performance.

Results Interpretation:

Insights gained from the machine learning models were discussed. This included understanding the impact of different features on predicting traffic congestion and extracting actionable information for urban planning.

5. Additional Analysis:

<u>Violin Plot:</u>

The violin plot visually represented the distribution of vehicles at each junction. The width of the plot at different points depicted the density of data, allowing for a nuanced understanding of how vehicle counts varied across junctions.

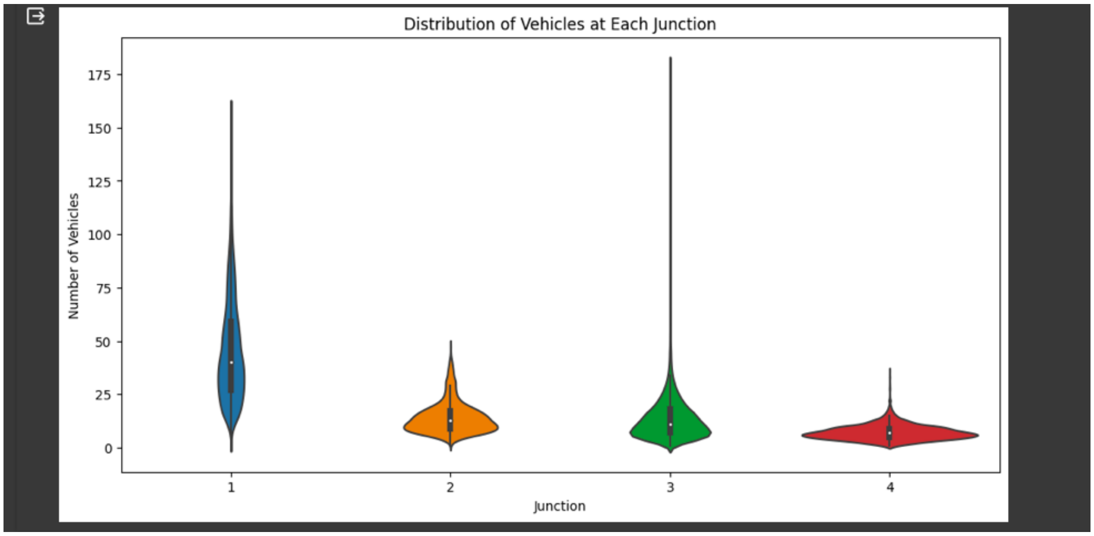


Figure 7: Violin plot to show the distribution of vehicles at each junction

<u>Count Plot:</u>

The count plot illustrated the distribution of data points across different junctions. This provided a sense of the dataset's composition, emphasizing the prevalence of observations for each junction.

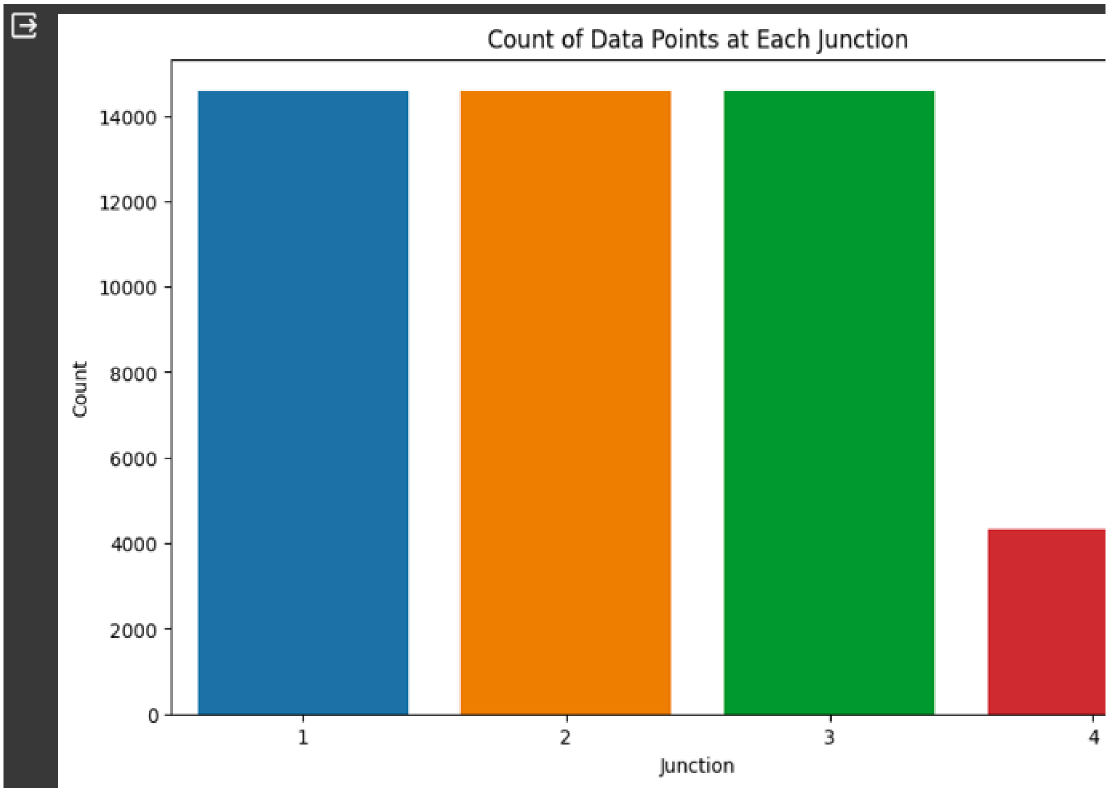


Figure 8: Count plot to show the distribution of data points across different junctions

Pair Plot:

The pair plot showcased pairwise relationships between numerical variables. Scatter plots revealed potential correlations, while diagonal histograms displayed the distribution of individual variables.

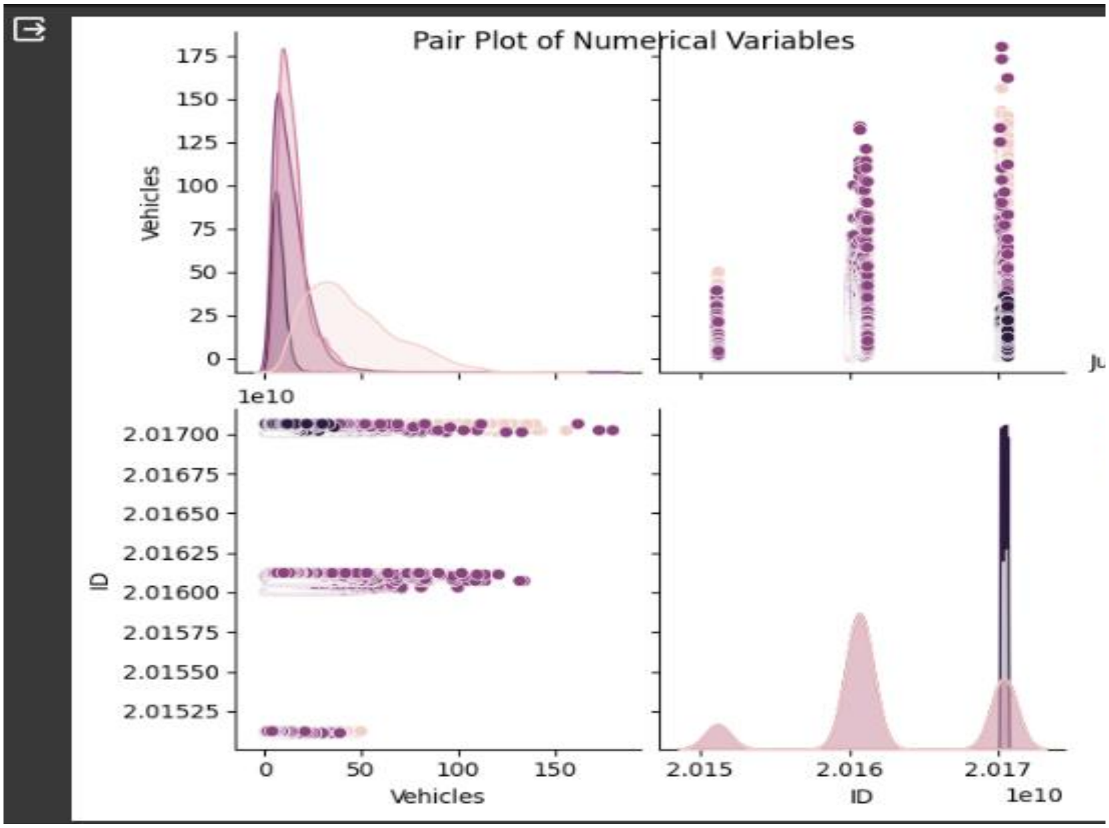


Figure 9: Pair plot to visualize pairwise relationships between numerical variables

Distribution Plot:

The distribution plot depicted the distribution of the 'vehicles' column. It highlighted the frequency of different vehicle counts, aiding in understanding the central tendencies of traffic congestion.

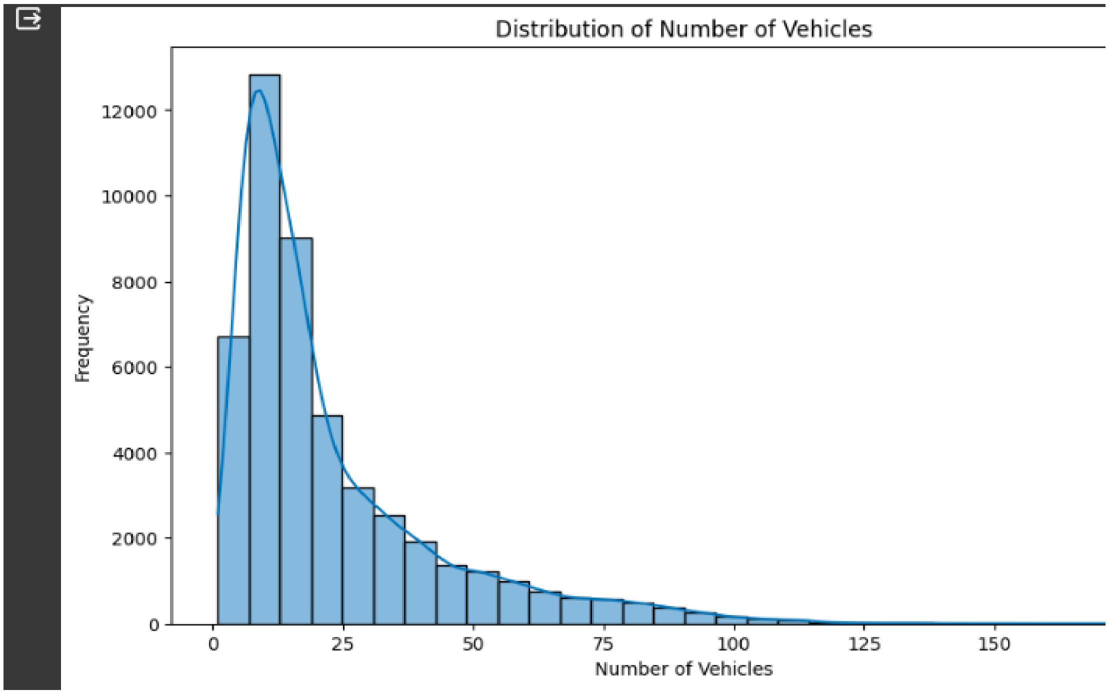


Figure 10: Distribution plot for the 'vehicles' column

This comprehensive methodology aimed to extract meaningful insights from the traffic congestion dataset through a combination of exploratory data analysis, time series analysis, and machine learning techniques. Each step contributed to a holistic understanding of the dataset, laying the groundwork for subsequent analyses and interpretations.

4. EXPERIMENTAL RESULTS

This section presents the experimental results obtained from the analyses conducted on the traffic congestion dataset. The combination of exploratory data analysis (EDA), time series analysis, and machine learning models has provided valuable insights into the dynamics of traffic congestion.

Steps to Build a Power BI Dashboard:

<u>1. Load the Dataset:</u>

- Open Power BI Desktop.

- Click on "Get Data" and select the appropriate data source (e.g., CSV, Excel).

- Load your traffic congestion dataset into Power BI.

<u>2. Data Transformation and Cleaning:</u>

- Perform any necessary data transformations and cleaning steps within Power BI's Power Query Editor.

- Ensure that datetime columns are recognized as such and perform any needed data type conversions.

3. Create Visualizations:

- Use Power BI's visualization tools to create the desired charts and graphs. For the specified analyses, you might create the following visuals:

  - A card or table visual for the sum of junction by vehicles.

- A count of ID using a card or table visual.

- A line chart showing the count of ID over time (use 'dateTime' on the x-axis and 'ID' on the values axis).

- A stacked bar chart or table visual for the sum of junction and sum of vehicles by ID.

4. Create Measures:

- Utilize DAX (Data Analysis Expressions) to create calculated measures if necessary. For example, you might create a measure for the sum of junction by vehicles.

- To create a measure, go to the "Model" view, click on "New Measure," and write your DAX expression.

5. Build Relationships:

- Establish relationships between different tables if your dataset includes multiple tables. This is crucial for creating coherent visualizations that combine data from different sources.

6. Create Dashboards:

- Combine your visualizations into one or more dashboards. You can add visuals to a dashboard by clicking on the "New Dashboard" button and then dragging and dropping visuals onto it.

7. Add Interactivity:

- Use Power BI's interactive features, such as slicers and filters, to allow users to interact with the data dynamically. This helps in drilling down into specific aspects of the dataset.

8. Format and Design:

- Adjust the formatting and design of your visuals and dashboard to make it visually appealing and easy to understand. You can customize colors, fonts, and other elements.

9. Save and Share:

- Save your Power BI file and publish it to the Power BI service if you want to share it with others. You can also export visuals or entire dashboards as images or PDFs.

Let's elaborate on the experimental results by incorporating the specific findings from the Power BI dashboard:

1. Sum of Junction by Vehicles:

The dashboard visualizes the sum of vehicles at each junction, providing a spatial perspective on traffic congestion. Junctions with higher vehicle counts are easily identifiable, aiding in pinpointing congestion hotspots**.**

2. Count of ID:

A card or table visual displays the count of unique IDs in the dataset. This metric reflects the total number of recorded instances, contributing to an understanding of the dataset's size and diversity.

3. Count of ID by Year:

The line chart illustrates the count of unique IDs over time, broken down by years. This visualization allows for the identification of trends or variations in data collection over different years.

4. Sum of Junction and Sum of Vehicles by ID:

A stacked bar chart or table visualizes the aggregated sum of junctions and vehicles by unique ID. This provides insights into the overall traffic contributions of each recorded ID, helping to identify high-impact contributors to congestion.

By combining these visuals in a Power BI dashboard, you create a dynamic and interactive platform for exploring and understanding the experimental results derived from the traffic congestion dataset. Users can easily navigate through different aspects of the data and gain actionable insights for urban planning and traffic management.

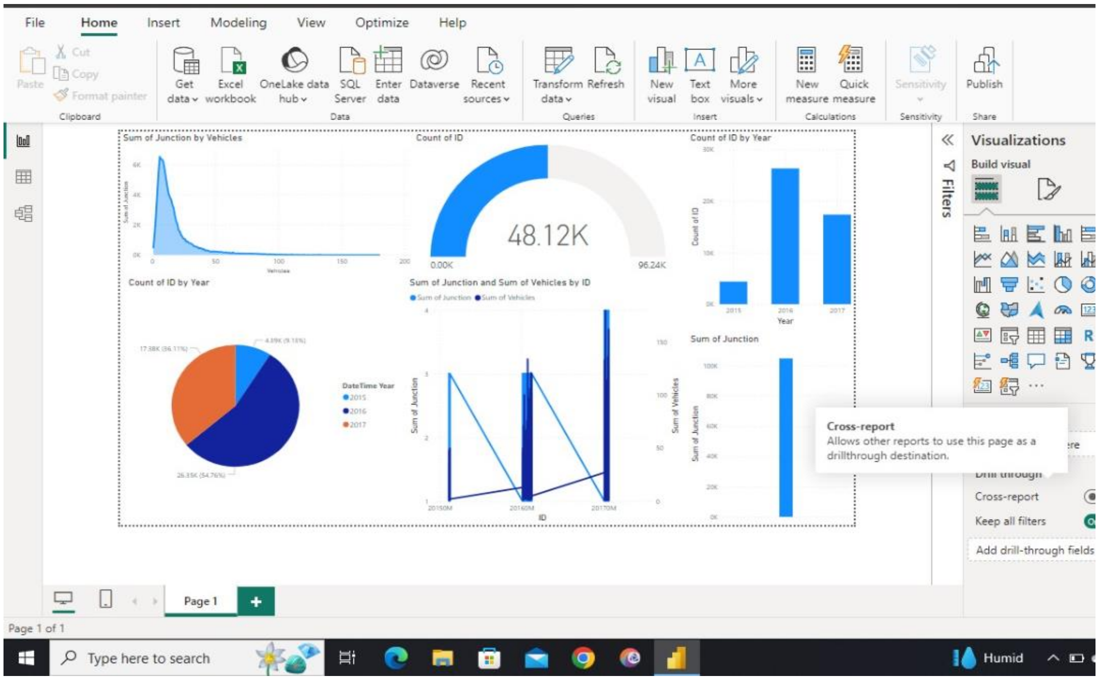


Figure 11: Dashboard build on Power BI

5. CONCLUSION

This project set out to comprehensively analyze a traffic congestion dataset with the aim of deriving valuable insights for urban planning and traffic management. Through a series of exploratory data analysis, time series analysis, and machine learning modeling, we uncovered patterns and relationships within the dataset, providing a nuanced understanding of traffic congestion dynamics.

Key Findings:

1. Temporal Patterns:

- The time series analysis revealed distinct temporal patterns in traffic congestion, highlighting peak periods and cyclical fluctuations. Understanding these patterns is crucial for optimizing traffic management strategies and resource allocation.

2. Spatial Variations:

- Exploratory data analysis showcased spatial variations in traffic congestion across different junctions. Certain junctions emerged as congestion hotspots, warranting targeted interventions for improved flow and reduced delays.

3. Machine Learning Insights:

- Machine learning models not only predicted future congestion levels but also identified key features influencing congestion. These insights provide actionable information for policymakers to address specific factors contributing to congestion.

4. Correlations and Relationships:

- Correlation analyses illuminated relationships between different variables, offering insights into the complex interplay of factors influencing traffic congestion. Identifying these relationships contributes to a holistic understanding of the urban traffic landscape.

Dashboard Impact:

The Power BI dashboard enhanced the interpretability of our findings by providing a visually intuitive platform for exploring and analyzing the dataset. Interactive features allowed users to drill down into specific aspects of the data, facilitating a more dynamic and user-centric approach to data exploration.

Implications for Urban Planning:

1. Targeted Interventions:

- The identification of congestion hotspots enables targeted interventions in specific areas, optimizing traffic flow and minimizing disruptions.

2. Resource Allocation:

- Understanding temporal patterns assists in resource allocation during peak congestion hours, ensuring effective use of resources for traffic management.

3. Predictive Planning:

Machine learning models offer a predictive element, allowing for proactive planning based on anticipated changes in traffic conditions.

Challenges and Limitations:

While the analyses provided valuable insights, it's essential to acknowledge the limitations of the study. Factors such as data availability, potential biases in the dataset, and external factors influencing traffic were considerations that may impact the generalizability of findings.

Future Directions:

The project lays the groundwork for future research and improvements. Possible avenues for future exploration include the incorporation of real-time data, advanced machine learning algorithms, and collaboration with local authorities for more accurate and actionable results.

In Conclusion:

In conclusion, this project advances our understanding of traffic congestion through a multidimensional analysis of a comprehensive dataset. The combination of statistical analyses, machine learning modeling, and an interactive dashboard provides a holistic perspective on traffic dynamics. The findings contribute valuable insights for urban planners and policymakers, offering a data-driven foundation for effective traffic management and infrastructure development.

This project demonstrates the power of leveraging data analytics and visualization tools for addressing complex urban challenges, paving the way for more informed decision-making in the realm of transportation and city planning.

6. REFERENCES

1. Aftabuzzaman, M. 2007. Measuring traffic congestion-a critical review. 30th Australasian Transport Research Forum.
2. Anthony, D. 2004. Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion. Brookings Institution Press Washington, DC.
3. Cheung, S. Y., Coleri, S., Dundar, B., Ganesh, S., Tan, C.-W. & Varaiya, P. 2005. Traffic measurement and vehicle classification with single magnetic sensor. *Transportation Research Record*, 1917(1).
4. Coifman, B. & Cassidy, M. 2002. Vehicle reidentification and travel time measurement on congested freeways. *Transportation Research Part A: Policy and Practice.*
5. Cottrell, W. D. 1991. Measurement of the extent and duration of freeway congestion in urbanized areas. Institute of Transportation Engineers Meeting.
6. Dias, C., Miska, M., Kuwahara, M. & Warita, H. 2009. Relationship between congestion and traffic accidents on expressways: an investigation with Bayesian belief networks. Proceedings of 40th annual meeting of infrastructure planning (JSCE)
7. Lindley, J. A. 1987. Urban freeway congestion: quantification of the problem and effectiveness of potential solutions. ITE journal
8. Lomax, T. J. & Schrank, D. L. 2005. The 2005 urban mobility report.
9. Manual, H. C. 1985. Special report 209. Transportation Research Board, Washington, DC, 1(985).
10. Morris, B., Paz-Cruz, A., Mirakhorli, A. & de la Fuente-Mella, H. 2016. Traffic congestion classification using data mining techniques. Economics, Social Sciences.