

GROUP 3

TIME SERIES DATA

# TRAFFIC CONGESTION PREDICTION AND ANOMALY DETECTION

USING DATA SCIENCE FOR SMARTER  
URBAN MOBILITY

ELECTIVE II

## INDUSTRY/SECTOR

---

What type of organization is using the data?



### GOVERNMENT AGENCIES

Urban mobility management & policy-making



### LOGISTICS COMPANIES

Route optimization & delivery efficiency



### URBAN PLANNERS

Road infrastructure & transport planning



### PUBLIC SAFETY ORGANIZATIONS

Traffic enforcement & emergency response

PROBLEM

---

METRO MANILA LOSES

₱3.5 BILLION

DAILY DUE TO TRAFFIC CONGESTION

SOURCE: DOF.GOV

## OTHER PROBLEMS

---

Why Does Traffic Congestion Matter?



### LONGER TRAVEL TIMES

Commuters lost approximately **127 hours** annually due to rush-hour traffic.

STATISTA.COM



### LOGISTICS COMPANIES

Average fuel economy of private vehicles decreases from 8.97 km/l at 39.79 km/h to 5.26 km/l at 16.25 km/h

ENGINEERINGRESEARCH.COM



### URBAN PLANNERS

In 2023, Metro Manila recorded over 92,500 road accidents

MANILASTANDARD.NET

# GOAL

Smarter, Safer, and More Efficient Roads



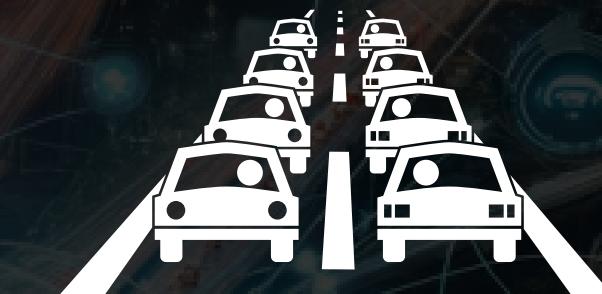
## FORECAST FUTURE TRAFFIC TRENDS

Predict congestion hotspots before they happen.



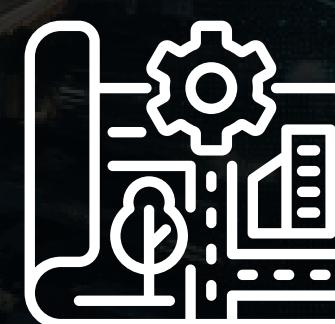
## IDENTIFY HIGH RISK ROADS

Pinpoint areas prone to accidents and delays.



## IMPROVE REAL-TIME TRAFFIC MANAGEMENT

Help agencies like MMDA optimize routes.



## ENHANCE URBAN PLANNING

Assist policymakers in designing better road networks.

TIME SERIES DATA

---

# DATA DESCRIPTION

Datasets Used for Traffic Prediction

## DATA SET

# KAGGLE MANILA TRAFFIC INCIDENT

Contains data on date and time of the incident, city and road location, lanes blocked due to the accident, type of accident, and involved vehicles

How it's collected: via a Python script parsing MMDA Twitter alerts

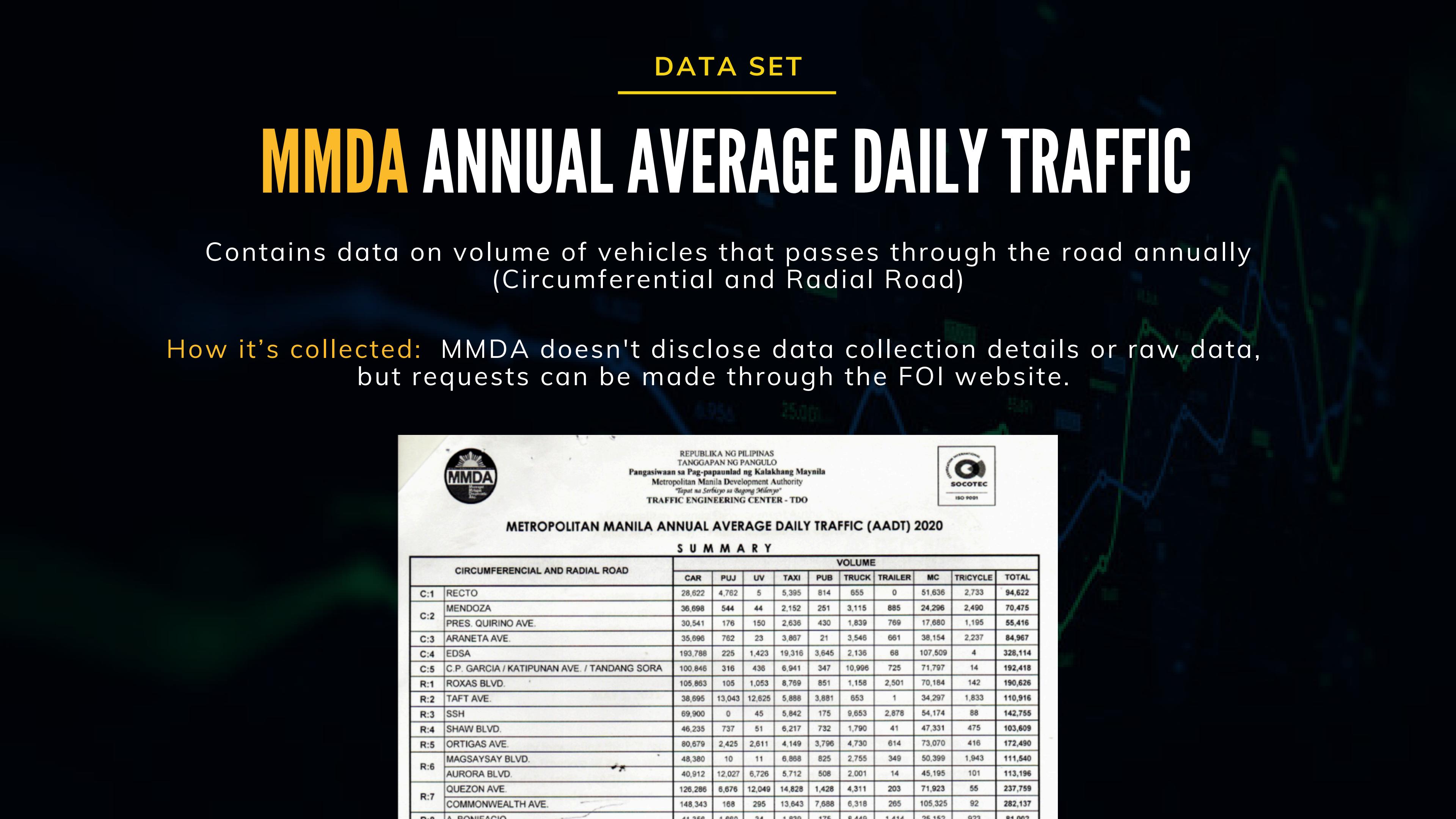
Date	Time	City	Location	High_Accuracy	Direction	Type	Lanes_Blocked	Involved
2018-08-20	7:55 AM	Pasig City	ORTIGAS EMERALD	1EB		VEHICULAR ACCIDENT		1 TAXI AND MC
2018-08-20	8:42 AM	Mandaluyong	EDSA GUADIX	1NB		STALLED L300 DUE TO MECHANICAL PROBLEM		1L300
2018-08-20	9:13 AM	Makati City	EDSA ROCKWELL	1SB		VEHICULAR ACCIDENT		1 SUV AND L300
2018-08-20	8:42 AM	Mandaluyong	EDSA GUADIX	1NB		STALLED L300 DUE TO MECHANICAL PROBLEM		1L300
2018-08-20	10:27 AM	San Juan	ORTIGAS CLUB FILIPINO	1EB		VEHICULAR ACCIDENT		12 CARS
2018-08-20	11:18 AM	Makati City	C5 KALAYAAN	1SB		VEHICULAR ON FIRE		1 BUS
2018-08-20	12:33 PM	Quezon City	EDSA ORTIGAS ROBINSONS	1NB		VEHICULAR ACCIDENT		1 CAR AND BUS
2018-08-20	12:39 PM	Mandaluyong	EDSA LIGHT MALL	1NB		MULTIPLE COLLISION		23 CARS
2018-08-20	12:39 PM	Mandaluyong	EDSA LIGHT MALL	1NB		MULTIPLE COLLISION		23 CARS
2018-08-20	2:51 PM	Quezon City	EDSA FARMERS	1NB		VEHICULAR ACCIDENT		1 BUS AND CAR
2018-08-20	4:14 PM	Pasig City	C5 LANUZA	1SB		STALLED TRUCK DUE TO MECHANICAL PROBLEM		1 TRUCK
2018-08-20	5:17 PM	Quezon City	C5 ATENEO KATIPUNAN	1NB		MULTIPLE COLLISION CSNEO KAT		1 CLOSED VAN, TRUCK, CAR AND MC
2018-08-20	6:21 PM	Quezon City	EDSA BONI	1NB		STALLED BUS DUE TO MECHANICAL DEFECT		1 BUS
2018-08-20	7:19 PM	Markina	MARCOS HIGHWAY LRT SANTOLAN	1EB		VEHICULAR ACCIDENT		1 PUJ AND AUV
2018-08-20	7:44 PM	Pasay City	EDSA HERITAGE	1NB		VEHICULAR ACCIDENT		1 TAXI AND TRUCK
2018-08-20	8:58 PM	Pasig City	C5 QRTIGAS FLYOVER	1NB		STALLED CLOSED VAN DUE TO MECHANICAL DEFECT		1 CLOSED VAN
2018-08-20	8:58 PM	Pasig City	C5 QRTIGAS FLYOVER	1NB		STALLED CLOSED VAN DUE TO MECHANICAL DEFECT		1 CLOSED VAN
2018-08-20	7:05 AM	Quezon City	EDSA ERMIN GARCIA	1SB		STALLED BUS DUE TO MECHANICAL PROBLEM		1 BUS
2018-08-20	7:05 AM	Quezon City	EDSA ERMIN GARCIA	1SB		STALLED BUS DUE TO MECHANICAL PROBLEM		1 BUS
2018-08-21	9:50 AM	Makati City	EDSA GUADALUPE PET PLANS	1SB		ONGOING DPWH ROAD PATCHING	1	
2018-08-21	10:00 AM	Quezon City	COMMONWEALTH DILIMAN	1EB		VEHICULAR ACCIDENT		12 MC
2018-08-21	10:30 AM	Parañaque	ROXAS BACLARAN	1SB		ONGOING DPWH FOOTBRIDGE PAINTING	1	
2018-08-21	10:47 AM	Quezon City	C5 EASTWOOD	1SB		VEHICULAR ACCIDENT		1 MC AND SUV
2018-08-21	11:05 PM	Pasay City	EDSA CABRERA	1SB		ONGOING DPWH ROAD PATCHING	1	
2018-08-21	11:28 PM	Manila	LACSON F/O	0SB		ONGOING DPWH ASPHALT LAYING	1	
2018-08-21	3:03 PM	Quezon City	NORTH AVE. AGHAM	0EB		VEHICULAR ACCIDENT		2 VAN AND PUJ
2018-08-21	6:20 AM	Pasig City	MARCOS HIGHWAY LIGAYA	1WB		VEHICULAR ACCIDENT		1 VAN AND SUV
2018-08-21	6:30 AM	Quezon City	QUEZON AVENUE	0WB		VEHICULAR ACCIDENT		1 BUS AND CAR
2018-08-21	6:30 AM	Quezon City	QUEZON AVENUE	0WB		VEHICULAR ACCIDENT		1 BUS AND CAR
2018-08-21	6:39 AM	Makati City	EDSA AYALA TUNNEL	1NB		STALLED L300 DUE TO MECHANICAL PROBLEM		1L300
2018-08-21	6:41 AM	Quezon City	EDSA NEPA Q-MART	1NB		VEHICULAR ACCIDENT		1 BUS AND CAR
2018-08-21	6:42 AM	Quezon City	EDSA P. TUAZON	0SB		VEHICULAR ACCIDENT		12 SUV'S
2018-08-21	6:48 AM	Quezon City	C5 GREENMEADOWS	1SB		VEHICULAR ACCIDENT		1 CAR AND MC
2018-08-21	7:16 AM	Quezon City	COMMONWEALTH LUZON	1WB		VEHICULAR ACCIDENT		1 SUV AND PUJ
2018-08-21	7:20 AM	Quezon City	1170N F/O	0SR		VEHICULAR ACCIDENT		1 CAR AND SUV

## DATA SET

# MMDA ANNUAL AVERAGE DAILY TRAFFIC

Contains data on volume of vehicles that passes through the road annually  
(Circumferential and Radial Road)

How it's collected: MMDA doesn't disclose data collection details or raw data, but requests can be made through the FOI website.



SUMMARY											
CIRCUMFERENTIAL AND RADIAL ROAD		VOLUME									
		CAR	PUJ	UV	TAXI	PUB	TRUCK	TRAILER	MC	TRICYCLE	TOTAL
C:1	RECTO	28,622	4,762	5	5,395	814	655	0	51,636	2,733	94,622
C:2	MENDOZA	36,698	544	44	2,152	251	3,115	885	24,296	2,490	70,475
	PRES. QUIRINO AVE.	30,541	176	150	2,636	430	1,839	769	17,680	1,195	55,416
C:3	ARANETA AVE.	35,696	762	23	3,867	21	3,546	661	38,154	2,237	84,967
C:4	EDSA	193,788	225	1,423	19,316	3,645	2,136	68	107,509	4	328,114
C:5	C.P. GARCIA / KATIPUNAN AVE. / TANDANG SORA	100,846	316	436	6,941	347	10,996	725	71,797	14	192,418
R:1	ROXAS BLVD.	105,863	105	1,053	8,769	851	1,158	2,501	70,184	142	190,626
R:2	TAFT AVE.	38,695	13,043	12,625	5,888	3,881	653	1	34,297	1,833	110,916
R:3	SSH	69,900	0	45	5,842	175	9,653	2,878	54,174	88	142,755
R:4	SHAW BLVD.	46,235	737	51	6,217	732	1,790	41	47,331	475	103,609
R:5	ORTIGAS AVE.	80,679	2,425	2,611	4,149	3,796	4,730	614	73,070	416	172,490
R:6	MAGSAYSAY BLVD.	48,380	10	11	6,868	825	2,755	349	50,399	1,943	111,540
	AURORA BLVD.	40,912	12,027	6,726	5,712	508	2,001	14	45,195	101	113,196
R:7	QUEZON AVE.	126,286	6,676	12,049	14,828	1,428	4,311	203	71,923	55	237,759
	COMMONWEALTH AVE.	148,343	168	295	13,643	7,688	6,318	265	105,325	92	282,137
R:8	A. RONIEACIO	41,366	1,660	24	1,829	175	8,449	1,414	25,152	923	81,002

TIME SERIES DATA

---

# DATA MINING SOLUTION

Extracting Insights from data to Enhance Traffic Forecasting and Decision-Making

## DATA MINING TECHNIQUES

---

01



### TIME SERIES ANALYSIS

Analyzing a sequence of data points collected over an interval of time like peak congestion hours and daily or weekly trends.

02



### CLASSIFICATION

Grouping road segments with similar traffic characteristics.

03



### ANOMALY DETECTION

Identifying unusual traffic patterns, like slowdowns or unexpected congestion.

## DATA MINING TECHNIQUES

# TIME SERIES ANALYSIS

Use past vehicle volume trends and incident frequency to predict the future traffic congestion per road

City	Location	High_Accuracy	Direction	Type
Pasig City	ORTIGAS EMERALD	1EB		VEHICULAR ACCIDENT
Mandaluyong	EDSA GUADIX	1NB		STALLED L300 DUE TO MECHANICAL
Makati City	EDSA ROCKWELL	1SB		VEHICULAR ACCIDENT
Mandaluyong	EDSA GUADIX	1NB		STALLED L300 DUE TO MECHANICAL
San Juan	ORTIGAS CLUB FILIPINO	1EB		VEHICULAR ACCIDENT
Makati City	C5 KALAYAAN	1SB		VEHICULAR ON FIRE
Quezon City	EDSA ORTIGAS ROBINSONS	1NB		VEHICULAR ACCIDENT
Mandaluyong	EDSA LIGHT MALL	1NB		MULTIPLE COLLISION
Mandaluyong	EDSA LIGHT MALL	1NB		MULTIPLE COLLISION
Quezon City	EDSA FARMERS	1NB		VEHICULAR ACCIDENT
Pasig City	C5 LANUZA	1SB		STALLED TRUCK DUE TO MECHANICA
Quezon City	C5 ATENEO KATIPUNAN	1NB		MULTIPLE COLLISION CSENEO KAT
Quezon City	EDSA BONI	1NB		STALLED BUS DUE TO MECHANICAL
Manila	MARCOS HIGHWAY LRT SANTOLAN	1EB		VEHICULAR ACCIDENT
Pasay City	EDSA HERITAGE	1NB		VEHICULAR ACCIDENT
Pasay City	C5 ORTIGAS FLYOVER	1NB		STALLED CLOSED VAN DUE TO MEC
Pasay City	C5 ORTIGAS FLYOVER	1NB		STALLED CLOSED VAN DUE TO MEC
Quezon City	EDSA ERMIN GARCIA	1SB		STALLED BUS DUE TO MECHANICAL
Quezon City	EDSA ERMIN GARCIA	1SB		STALLED BUS DUE TO MECHANICAL
Makati City	EDSA GUADALUPE PET PLANS	1SB		ONGOING DPWH ROAD PATCHING
Quezon City	COMMONWEALTH DILUMAN	1EB		VEHICULAR ACCIDENT
Panatagaque	ROXAS BACLARAN	1SB		ONGOING DPWH FOOTBRIDGE PAINT
Quezon City	C5 EASTWOOD	1SB		VEHICULAR ACCIDENT
Pasay City	EDSA CABRERA	1SB		ONGOING DPWH ROAD PATCHING
Manila	LACSON FIO	0SB		ONGOING DPWH ASPHALT LAYING
Quezon City	NORTH AVE. AGHAM	0EB		VEHICULAR ACCIDENT
Pasay City	MARCOS HIGHWAY LIGAYA	1WB		VEHICULAR ACCIDENT
Quezon City	QUEZON AVENUE	0WB		VEHICULAR ACCIDENT
Quezon City	QUEZON AVENUE	0WB		VEHICULAR ACCIDENT
Makati City	EDSA AYALA TUNNEL	1NB		STALLED L300 DUE TO MECHANICAL
Quezon City	EDSA NEPA Q-MART	1NB		VEHICULAR ACCIDENT
Quezon City	EDSA P. TUAZON	0SB		VEHICULAR ACCIDENT
Quezon City	C5 GREENMEADOWS	1SB		VEHICULAR ACCIDENT
Quezon City	COMMONWEALTH LUZON	1WB		VEHICULAR ACCIDENT
Quezon City	EDSA FIO	0SR		VEHICULAR ACCIDENT

## KAGGLE INCIDENT REPORT

This dataset provides time-based accident data

METROPOLITAN MANILA ANNUAL AVERAGE DAILY TRAFFIC (AADT) 2020									
SUMMARY									
CIRCUMFERENCIAL AND RADIAL ROAD									
	CAR	PUJ	UV	TAXI	PUB	TRUCK	TRAILER	MC	TRICYCLE
C:1 RECTO	26,622	4,762	5	5,395	814	655	0	51,636	2,733
C:2 MENDOZA	36,699	544	44	2,152	251	3,115	885	24,296	2,490
C:2 PRES. QUIRINO AVE.	30,541	176	150	2,636	430	1,839	769	17,680	1,195
C:3 ARANETA AVE.	35,696	762	23	3,867	21	3,546	661	38,154	2,237
C:4 EDSA	193,788	225	1,423	19,316	3,645	2,136	68	107,509	4
C:5 C.P. GARCIA / KATIPUNAN AVE. / TANDANG SORA	100,846	316	438	6,941	347	10,996	725	71,797	14
R:1 ROXAS BLVD.	105,863	105	1,053	8,759	851	1,158	2,501	70,184	142
R:2 TAFT AVE.	38,695	13,043	12,625	5,898	3,881	653	1	34,297	1,833
R:3 SSH	69,900	0	45	5,842	175	9,653	2,878	54,174	88
R:4 SHAW BLVD.	46,238	737	51	6,217	732	1,790	41	47,331	475
R:5 ORTIGAS AVE	80,879	2,425	2,811	4,149	3,796	4,730	614	73,070	416
R:6 MAGSAYSAY BLVD.	48,380	10	11	6,868	825	2,755	349	50,399	1,943
R:7 AURORA BLVD.	40,912	12,027	6,726	5,712	508	2,001	14	45,199	101
R:7 QUEZON AVE.	126,286	8,876	12,049	14,828	1,428	4,311	203	71,929	56
R:7 COMMONWEALTH AVE.	148,343	168	295	13,643	7,688	6,318	295	105,325	92
R:8 A. BONIFACIO	41,356	1,060	34	1,839	175	8,449	1,414	25,152	923
R:9 RIZAL AVE.	30,086	9,988	101	2,204	449	2,938	109	40,721	956
R:10 DEL PAN	45,911	668	28	2,304	254	6,281	7,342	42,621	2,166
MARCOS HIGHWAY	102,714	12,696	3,313	7,762	454	5,125	225	90,421	453
MICARTHUR HIGHWAY	33,129	6,300	41	2,720	2,244	3,022	183	59,340	190
<b>TOTAL</b>	<b>1,364,674</b>	<b>75,276</b>	<b>41,064</b>	<b>130,052</b>	<b>28,968</b>	<b>82,372</b>	<b>19,247</b>	<b>1,121,225</b>	<b>18,506</b>

## MMDA ANNUAL AVERAGE DAILY TRAFFIC (AADT)

This dataset gives the total vehicle volume per road for the year

# DATA PROCESSING STEPS

## MERGING TRAFFIC & INCIDENT DATA

- Combine the AADT dataset with MMDA Incident Reports for better insights.
- Correlate vehicle volume with accident occurrences.



## ESTIMATING DAILY TRAFFIC VOLUME

- Distribute AADT per day ( $\text{AADT}/365$ ) to analyze daily trends.
- Obtain a raw dataset for accurate volume distribution.

## TRAINING THE PREDICTION MODEL

- Use historical data to train the model for traffic forecasting.
- Learn from past trends to improve future congestion predictions.

# KEY VARIABLES

01

## DATE & TIME

Traffic behavior varies by hour & season

02

## ROAD NAME

Location-specific congestion patterns

03

## PAST TRAFFIC VOLUME

Historical congestion trends

04

## INCIDENT FREQUENCY

How often accidents cause congestion

05

## LANES BLOCKED

Road capacity affected by accidents

06

## TIME OF DAY

Rush hour vs. off-peak traffic patterns

# USING ARIMA FOR TRAFFIC FORECASTING

A time-series forecasting method that captures long-term trends in traffic volumes and identifies patterns like rush hour congestion

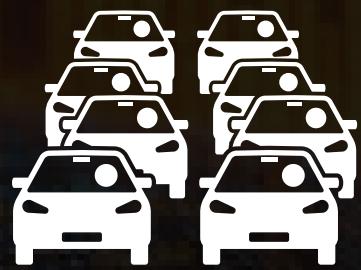
Ideal for time-dependent traffic pattern and useful for daily and seasonal forecasting

Example Prediction: EDSA will experience high congestion at 6PM



TIME SERIES ANALYSIS

# EXPECTED OUTCOME



## TRAFFIC CONGESTION PREDICTION

Identify which roads will experience heavy traffic before it happens.



## TRAFFIC VOLUME & INCIDENT FORECASTING

Predict traffic volume trends for any given road and incidents based on historical patterns.

# CLASSIFICATION TECHNIQUE

Use traffic volume + accident data for congestion levels



## HEAVY TRAFFIC

High vehicle count & frequent accidents



## MODERATE TRAFFIC

Medium volume and occasional accidents



## LIGHT TRAFFIC

Low vehicle count and no accidents

**Expected Output:** Real time classification of roads into heavy, moderate, or light congestion levels

## CLASSIFICATION TECHNIQUE

# FEATURES USED FOR CLASSIFICATION

01

### VEHICLE COUNT

AADT Adjuster per day/hour,  
higher volume increases  
congestion

02

### INCIDENT FREQUENCY

Frequent accidents means  
higher congestion risk

03

### LANES BLOCKED

If a major accident blocks  
lanes means traffic worsens

04

### TIME OF DAY

Since rush hours naturally  
produce more congestions

05

### TYPE OF ACCIDENT

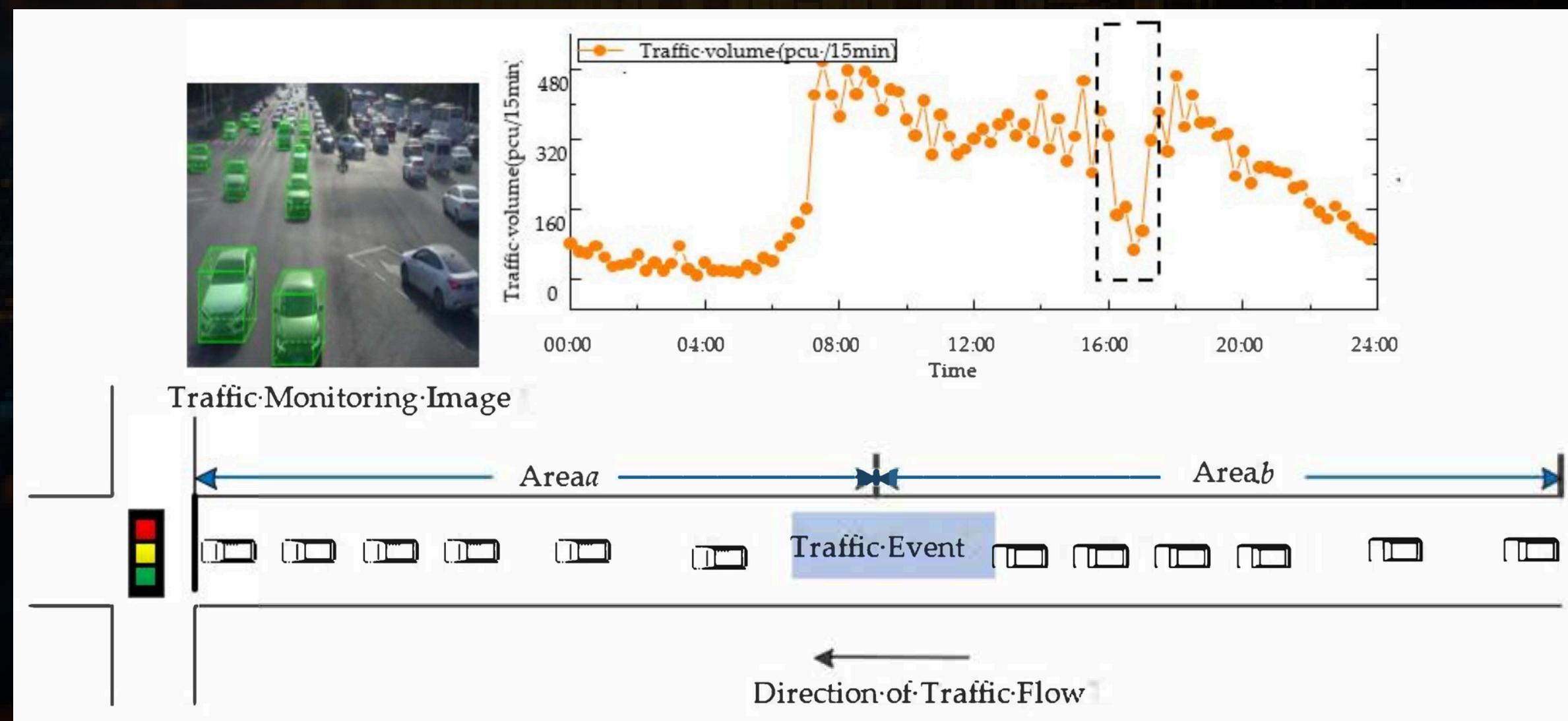
Collisions cause more delays  
than minor breakdowns

## CLUSTERING TECHNIQUE

# USING RANDOM FOREST

A machine learning algorithm that uses multiple decision trees to make predictions

Classify traffic conditions (Works well for real time traffic classification)



# ANOMALY DETECTION TECHNIQUE

Identify unexpected congestion spikes based on incidents and volume



High traffic occurs during  
rush hours

If traffic suddenly spikes  
outside normal hours  
means there is an accident

Detect unusual congestion  
spikes using anomaly  
detection models

## ANOMALY DETECTION

# FEATURES USED FOR ANOMALY DETECTION

01

### CURRENT TRAFFIC VOLUME VS. HISTORICAL VOLUME

Detects unusual spikes or drops in traffic flow compared to past trends

02

### REAL-TIME INCIDENT MONITORING

Identifies accidents, roadworks, or unexpected disruptions affecting traffic.

03

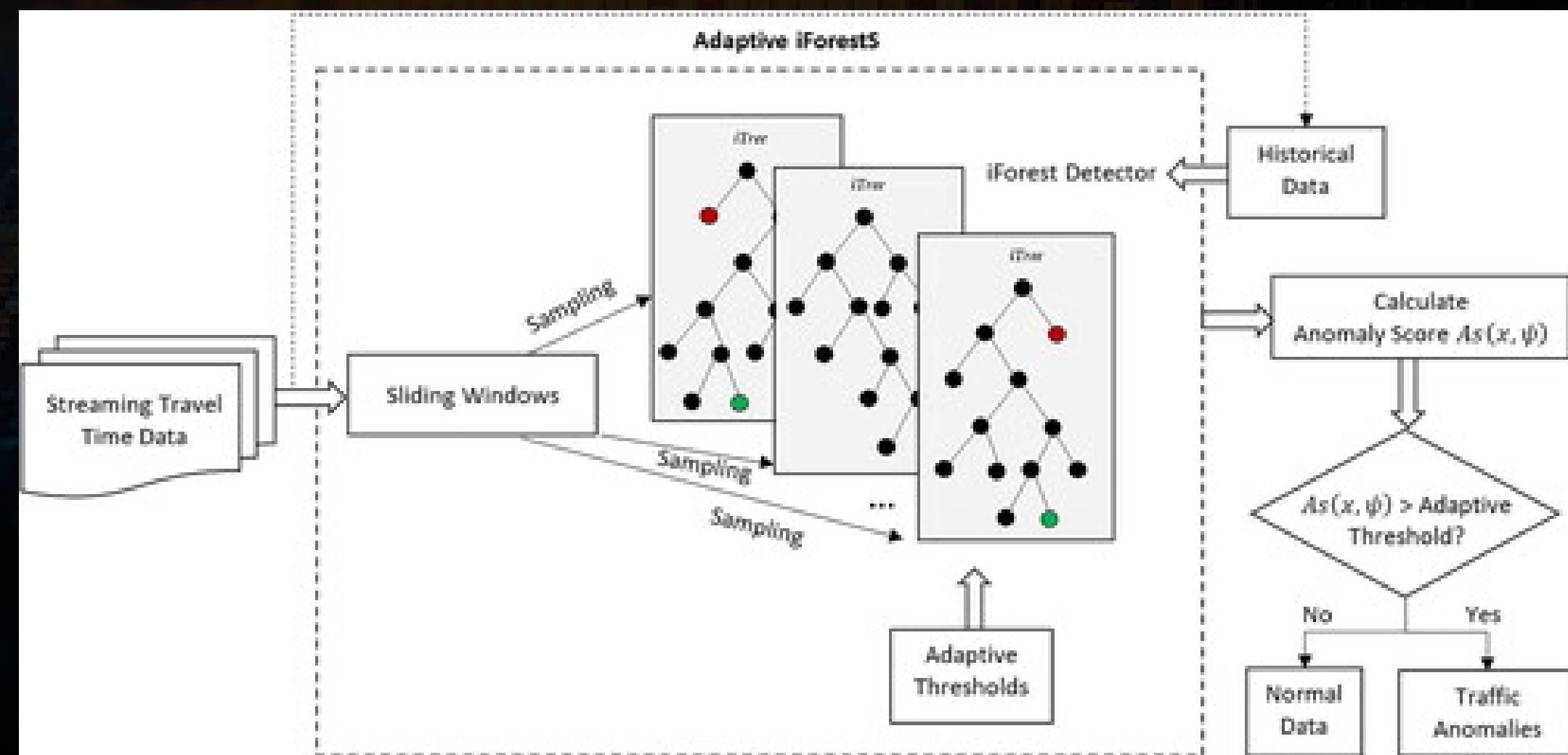
### LANES BLOCKED

Measures the effect of lane closures on congestion levels and rerouting needs.

# USING ISOLATION FOREST

Isolate data points that don't follow the normal pattern

If at midnight there is a spike traffic at a road that does not receive much traffic at that time it will be flagged as anomaly. Lightweight and can easily filter false alarms and only highlight unusual congestion events



## ANOMALY DETECTION

Problem: This can't explain why an anomaly occurred

# USING SHAP FOR IDENTIFYING ANOMALY CAUSE

Example: Rule-Based Insights



TRAFFIC SPIKE +  
NO INCIDENTS

Possibly a special event



TRAFFIC SPIKE +  
INCIDENT REPORTED

Likely due to an accident



TRAFFIC DROP +  
ROAD CLOSED

Likely due to a blockage

## ANOMALY DETECTION

# EXPECTED OUTCOME



### TRAFFIC CONGESTION PREDICTION

Predict future traffic congestion levels for each road based on historical patterns.



### ANOMALY DETECTION & CAUSE IDENTIFICATION

Detect unusual traffic spikes and determine their root causes (e.g., accidents, events, roadblocks).



### REAL-TIME TRAFFIC MANAGEMENT

Improve route planning and rerouting strategies for better traffic flow.



### HIGH-RISK ROAD IDENTIFICATION

Identify roads with frequent accidents and congestion to enhance road safety.

## BENEFITS

---



- 01
- 02
- 03
- 04
- 05

### Improved Traffic Forecasting

Helps MMDA optimize traffic signals and road management.

### Enhanced Road Safety

Identifies high-risk areas for better enforcement.

### Better Emergency Response

Predicts high-risk areas for faster response times.

### Informed Urban Planning

Supports infrastructure and road network improvements.

### Economic Efficiency

Reduces fuel costs, delays, and vehicle maintenance costs.

## CHALLENGES

---

01



### Data Quality Issues

Missing/inaccurate data may affect predictions.

02



### Ethical Considerations

Bias in data collection can lead to unfair traffic policies.

03



### Scalability

Handling large real-time traffic data requires robust infrastructure.

04



### Weather Conditions Not Included

Lack of weather data may impact prediction accuracy.



GROUP 3

# THANK YOU

TRAFFIC CONGESTION PREDICTION AND  
ANOMALY DETECTION

ELECTIVE II