

DATA SCIENCE

PREDICTION ON ROAD TRAFFIC ANALYSIS

CASE STUDY BOGOR

IMAMA LAVI

Agenda

1

Business Understanding

2

Data Understanding

3

Data Cleansing and Preprocessing

4

Modeling

5

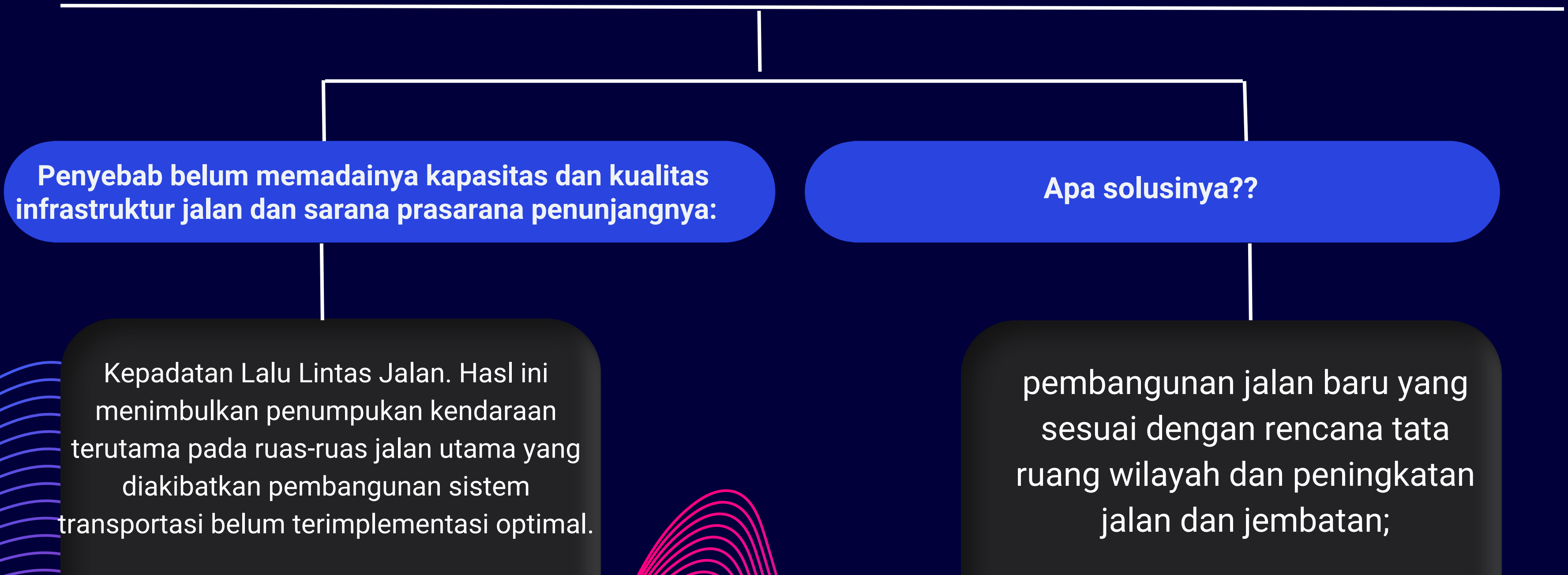
Referensi

Business Understanding



PERMASALAHAN DINAS PEKERJAAN UMUM DAN PENATAAN RUANG KOTA BOGOR

1. Kapasitas dan Kualitas infrastruktur jalan dan sarana prasarana penunjangnya belum memadai.



Research Question, Goals, Solution

Research Question

Bagaimana prediksi kondisi kemacetan jalan akibat dari padatnya lalu lintas ?

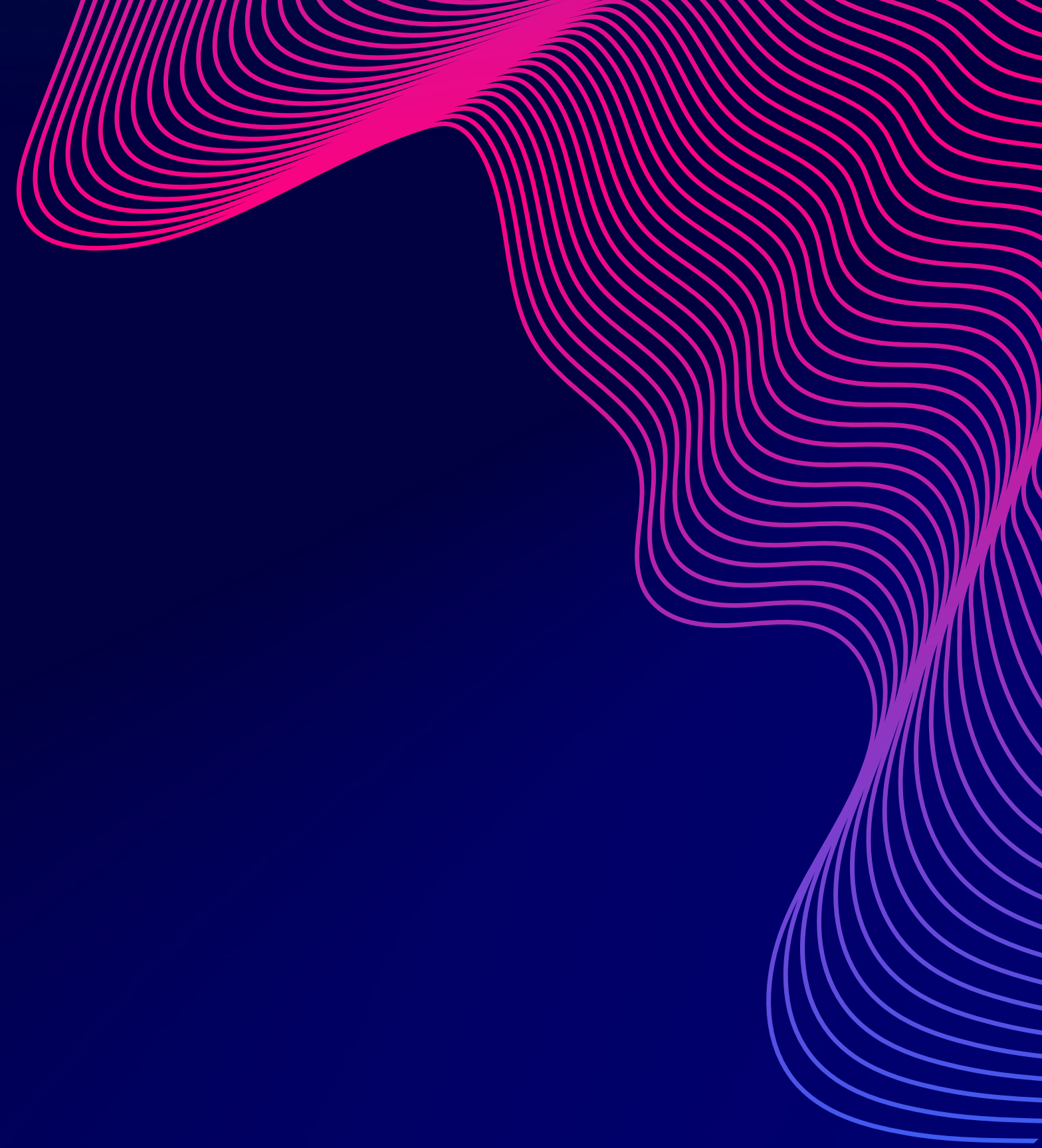
Goals

Mengetahui prediksi kondisi kemacetan di suatu jalan agar dijadikan salah satu pertimbangan untuk peningkatan jalan misalnya melalui pelebaran jalan, penyesuaian rekayasa jalan menjadi satu arah, dsb.

Proposed Solution

Prediksi level kemacetan jalan sebagai akibat dari kepadatan lalu lintas

Data Understanding



Data Understanding

Data

Data yang digunakan yaitu data
aggregate_median_irregularities_Kota Bogor.csv

Why

Hal ini disesuaikan dengan research question atau kebutuhan. Kolom yang digunakan yaitu street dan jam level

Data Cleansing and Preprocessing



Data Cleansing

Pada kolom yang dipilih yaitu street tidak terlihat data null jadi proses dilanjutkan dengan melakukan EDA dengan visualisasi

Quality

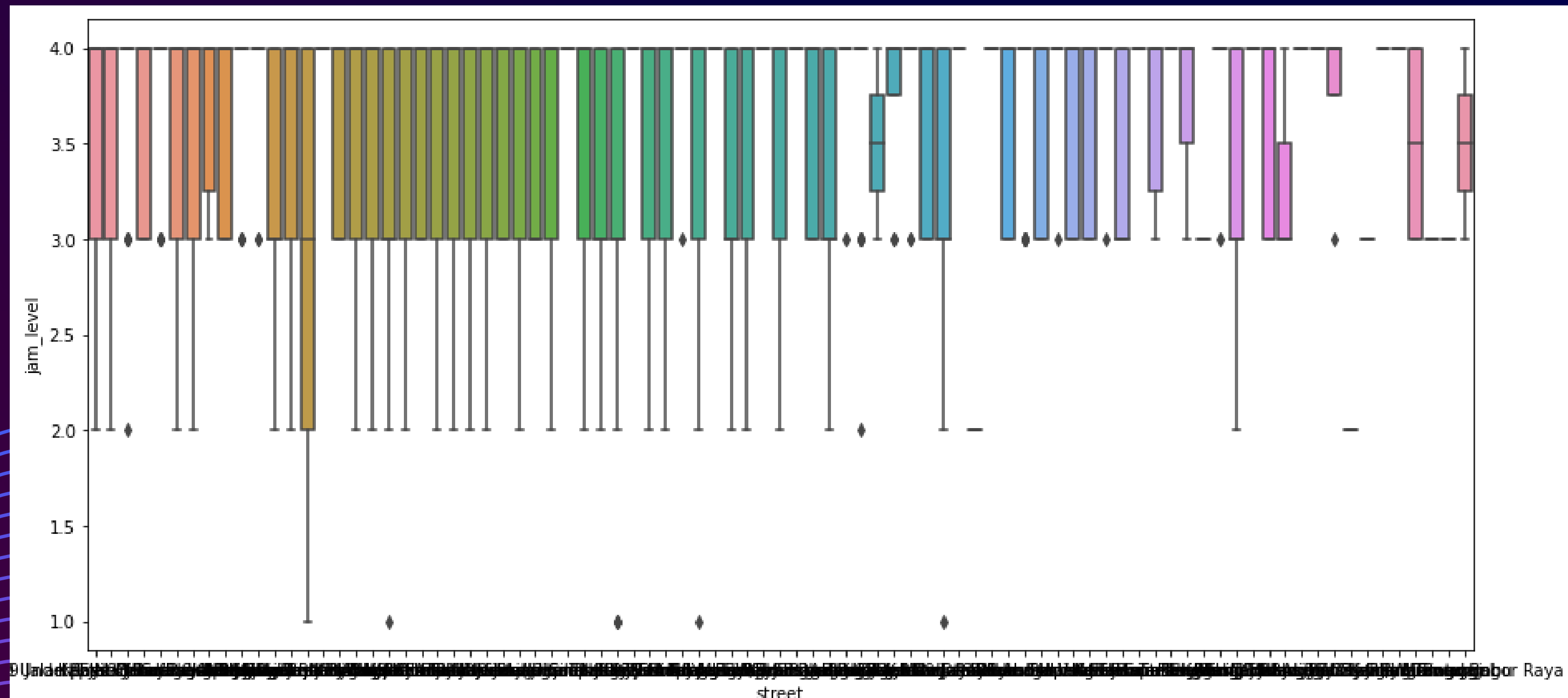
Delete NaN rows in table

```
[ ] df2[df2['street']!='NaN'].isnull().sum()

kemendagri_kabupaten_kode      0
kemendagri_kabupaten_nama      0
street                          0
jam_level                       0
median_length                   0
median_delay_seconds            0
median_regular_speed            0
total_records                   0
cause_type                      4051
median_seconds                  0
median_speed                     0
date                            0
median_jam_level                0
id                               0
geometry                        0
dtype: int64
```

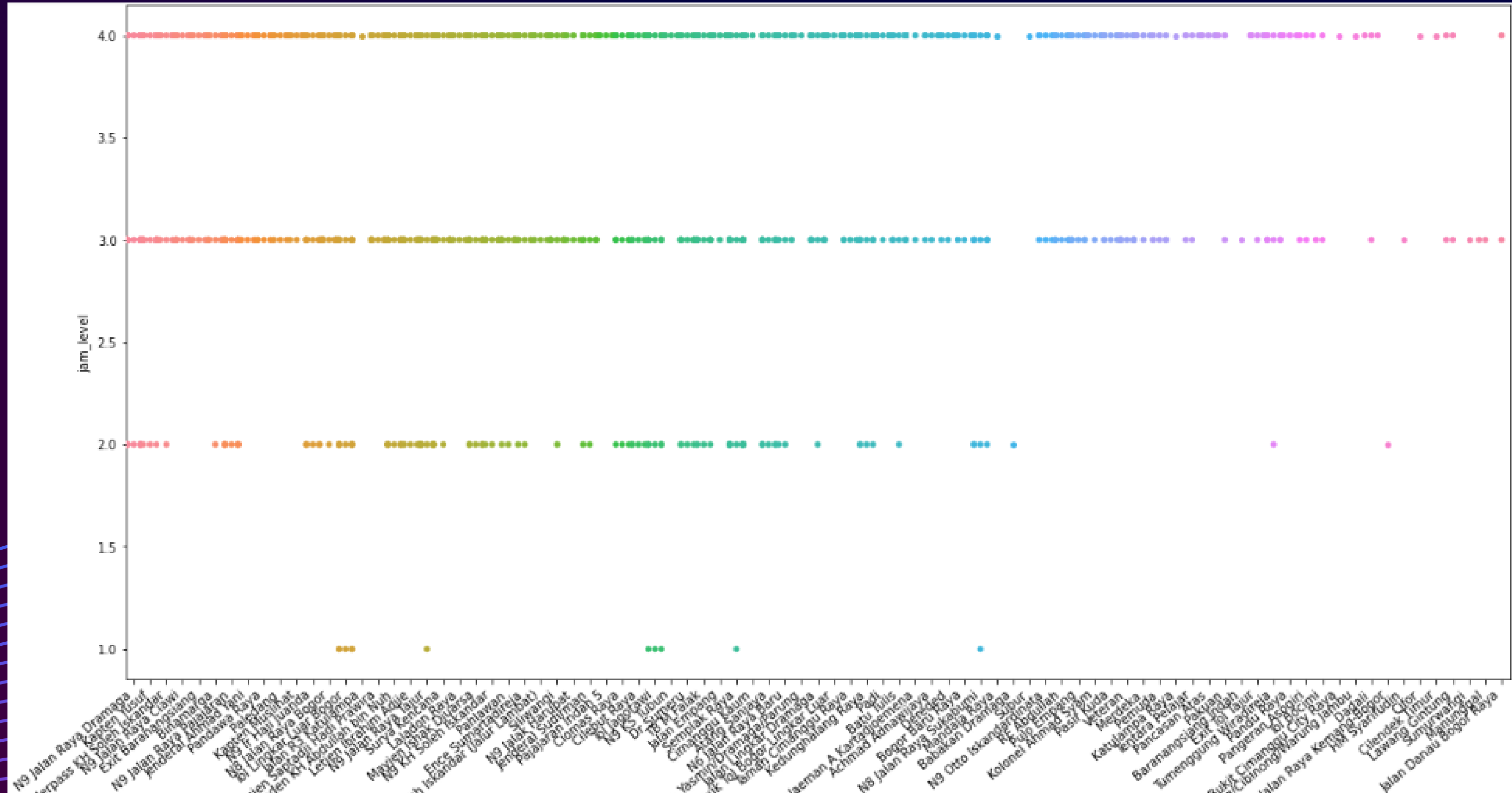
Rata-rata tingkat kemacetan ada di kisaran 3-4

Tahap selanjutnya adalah melihat korelasi dari street dan jam level. Visualiasi menggunakan boxplot berikut ini menampilkan ada beberapa data yang termasuk outlier untuk level kemacetannya yaitu di angka 1.



Rata-rata tingkat kemacetan ada di kisaran 3-4

Agar lebih jelas maka ada perubahan proses visualisasi menjadi swarmplot.



Modeling



Linear Regression

1. Splitting data menjadi 4 bagian

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.2)
```

2. Proses Pipeline dan training data

```
lr = LinearRegression()

pipe = make_pipeline(column_trans,lr) # Made pipeline and saved it in variable named pipe

pipe.fit(x_train, y_train) # Trained the pipe

Pipeline(steps=[('columntransformer',
                  ColumnTransformer(remainder='passthrough',
                                     transformers=[('onehotencoder',
                                                    OneHotEncoder(categories=[array(['Achmad Adnawijaya', 'Arah Masuk Tol Bogor Lingkar Luar',
'Atang Sanjaya', 'Babakan Dramaga', 'Baranangsiang Indah',
'Batu Tulis', 'Binamarga', 'Bogor Baru Raya',
'Brigjen Saptadji Hadi Prawira', 'Bukit Cimanggu City Raya',
'Cifon', 'Ci...',
'Pancasan Atas', 'Pandawa Raya', 'Pandu Raya', 'Pangeran Asogiri',
'Pasir Kuda', 'Pemuda', 'Pulo Empang', 'RE Abdullah',
'RH Soelaeman A Kartadjoemena', 'Semplak Raya', 'Siliwangi',
'Subur', 'Sumurwangi', 'Surya Kencana', 'TB M Falak',
'Taman Cimanggu Raya', 'Tentara Pelajar', 'Tol Bocimi',
'Tol Jagorawi', 'Tol Lingkar Luar Bogor', 'Tumenggung Wiradireja',
'Veteran'], dtype=object)]),
                  ['street']]]),
          ('linearregression', LinearRegression()))
```

Linear Regression

Hasil dari prediksi. Caranya adalah dengan input nama jalan lalu akan keluar prediksi nilai level kemacetannya.

```
street = input("Input street name: ")

PredictedStreet = pipe.predict(pd.DataFrame(columns=x_test.columns,data=np.array([street]).reshape(1,1)))

print("Predicted street is ", round(PredictedStreet[0],2))
```

```
Input street name: N9 Jalan Raya Dramaga
Predicted street is  3.54
```



Summary

Prediksi kemacetan bertujuan untuk mengetahui kondisi suatu jalan agar dijadikan salah satu pertimbangan untuk peningkatan kualitas jalan misalnya melalui pelebaran jalan, penyesuaian rekayasa jalan menjadi satu arah, dsb.

Rata-rata nilai kemacetan di Bogor ada di angka 3-4.

Referensi



- The relationship between road accident severity and Time-dependent resilience analysis of a road network in an extreme environment
- Road Network Resilience: How to Identify Critical Links Subject to Day-to-Day Disruptions
- Measures to evaluate post-disaster trip resilience on road networks
- A methodology for road traffic resilience analysis and review of related concepts