CodeOp DA1

Tech challenge

(5-6 Sep 2020)

Hyojung Lee

Introduction

Traffic state information by sections of the city of Barcelona (Open Data BCN)

Field information:

1. **idTram** Section identification number
2. **data** Date & time of observation (format: YYYY-MM-DD)
3. **estatActual** Current traffic status
   (0 = sense dades / 1 = molt fluid / 2 = fluid / 3 = dens / 4 = molt dens / 5 = congestió / 6 = tallat)
4. **estatPrevist** Expected traffic status after 15min
   (0 = sense dades / 1 = molt fluid / 2 = fluid / 3 = dens / 4 = molt dens / 5 = congestió / 6 = tallat)

Factors to consider

PROXIMITY TO RELEVANT FEATURES

- primary / secondary roads
- bus stops
- speed cameras
- traffic signals

TIME

- of the day (rush hour)
- of the week (Mon - Sun)

Geospatial data

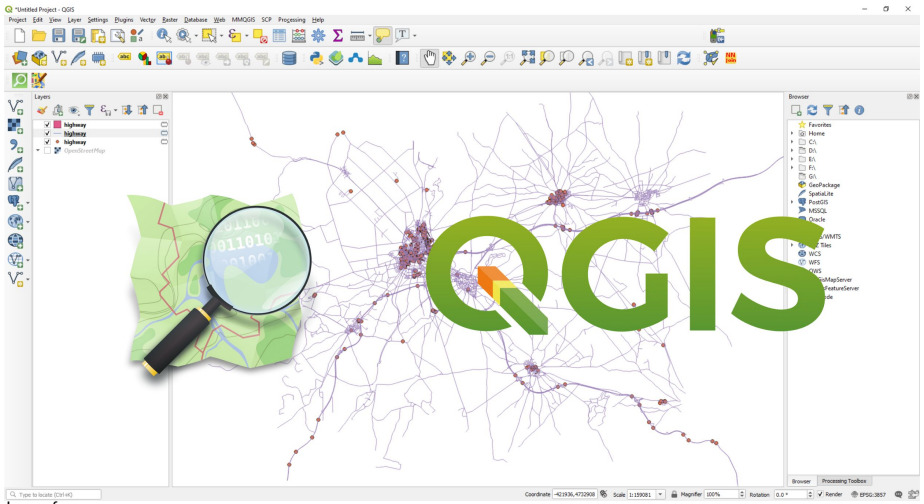


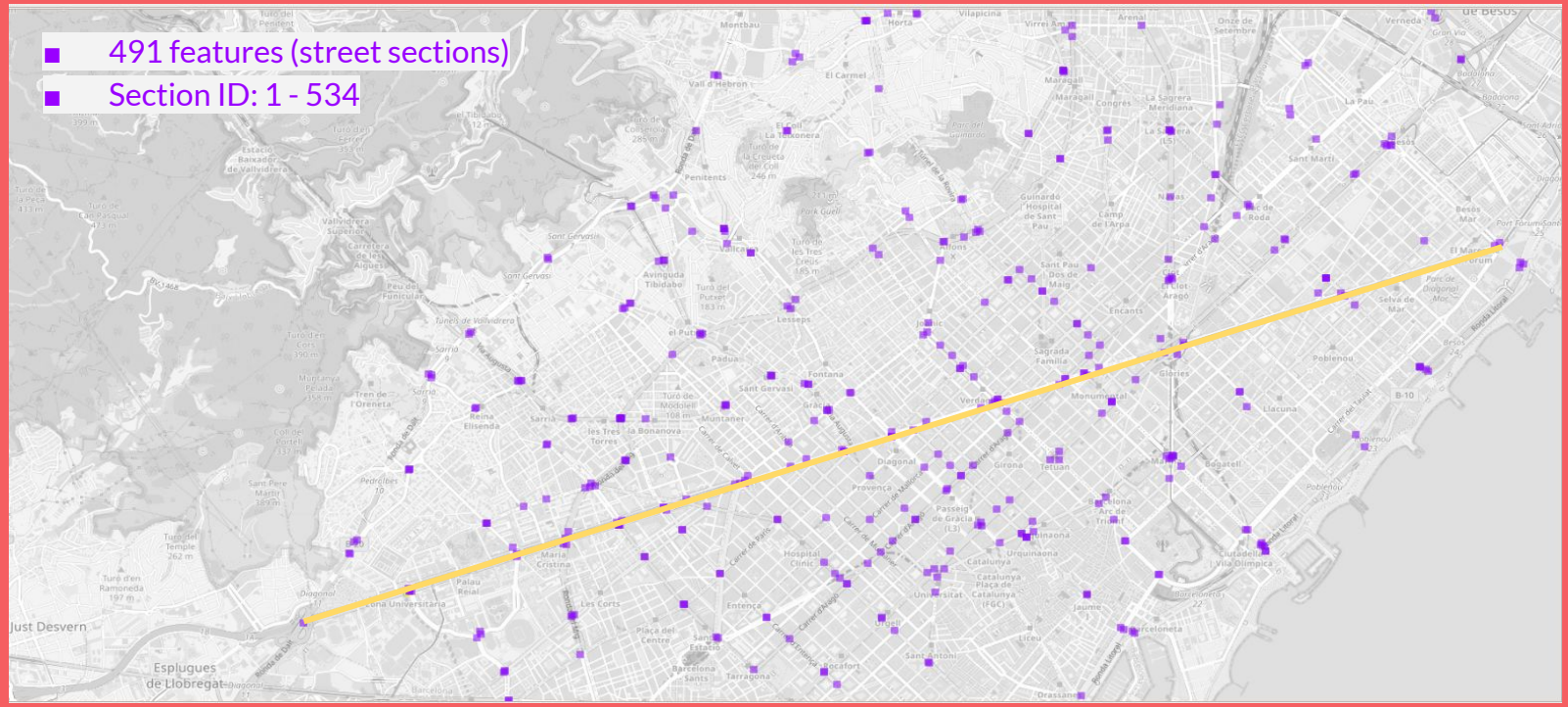
Image from:
https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.giscourse.com%2Fhow-to-download-osm-data-using-quickosm-plugin-in-qgis%2F&psig=AOvVaw2iM52GnYplbCoHEs_WMnU3&ust=1599468471250000&source=images&cd=vfe&ved=0CAIQjRxqFwoTCMj7rcKS1OsCFQAAAAAdAAAAABAP

A free and open-source platform of geographic information system (GIS) application that supports viewing, editing, and analysis of geospatial data.

Additional open-source data:

- OpenStreetMap (open-source web map)
- Coordinates of each section (OpenDataBCN)

Relación de tramos de la vía pública de la ciudad de Barcelona



Avinguda Diagonal's length:

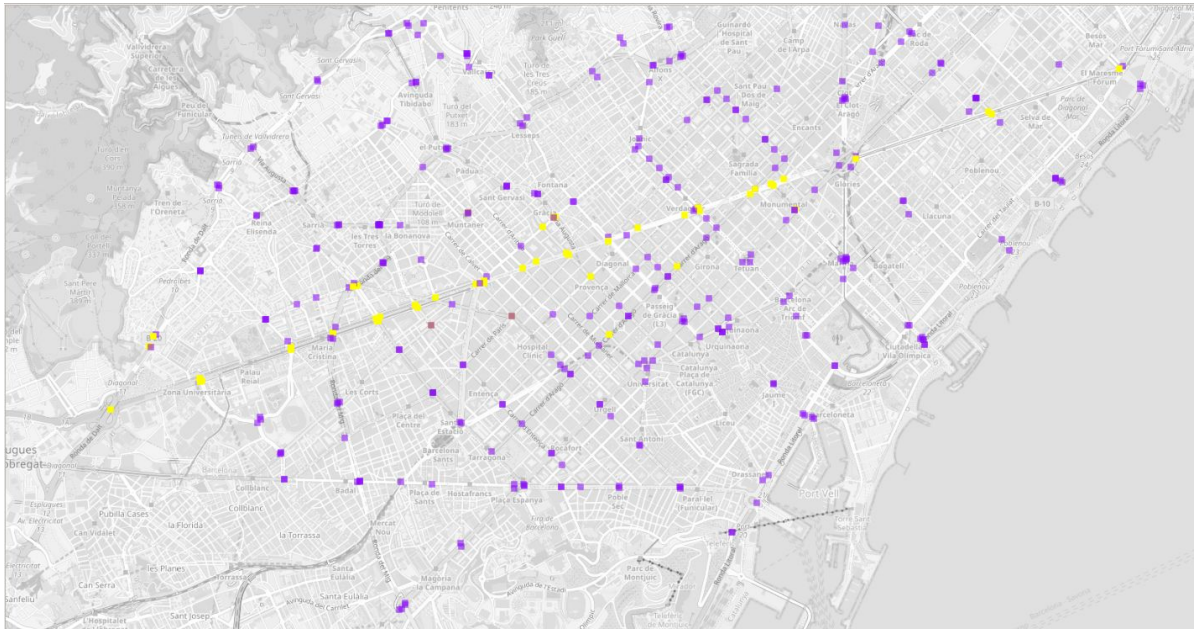
- Google: 11km
- My measurement: 14km

Let’s specify our focus:

I. Where

II. When

Select sections of Av. Diagonal - I: Query

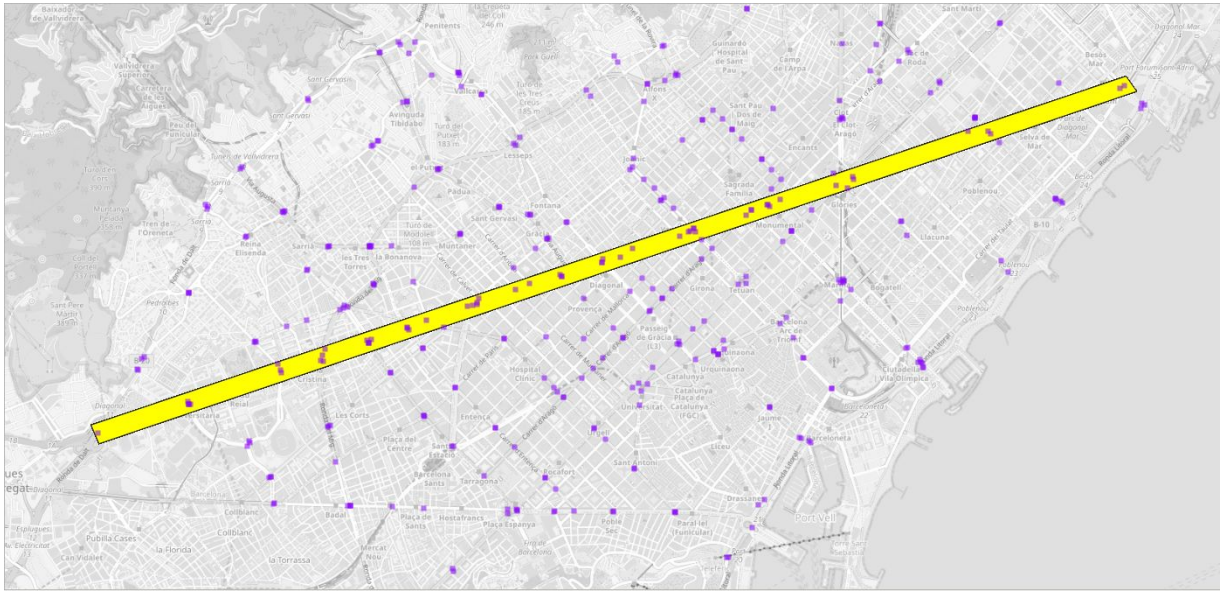


- Purple boxes: all street sections
- Yellow boxes: queried sections (‘Diagonal’)

- Section names containing ‘Diagonal’ : 56 features
- Not all sections laid on Diagonal avenue were selected
- Some sections far from Diagonal avenue were also included

-> Not what we want

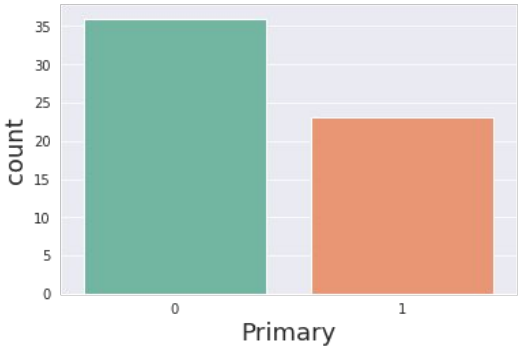
Select sections of Av. Diagonal - II: Join by location

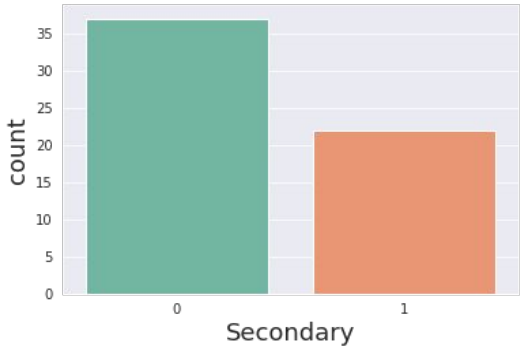


- Purple boxes: all street sections
- Yellow boxes: queried sections (‘Diagonal’)

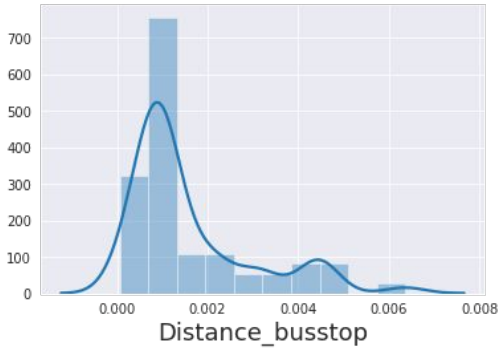
- Retrieve section ID which is laid within the Diagonal
- ‘Join-Attribute by location’ method
- Advantages:
  - Specific IDs to focus (59 IDs)
  - Intersection info

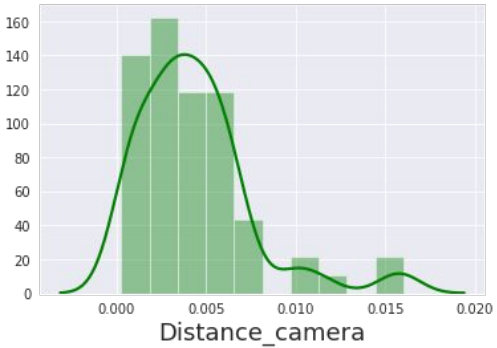
59 sections in diagonal:

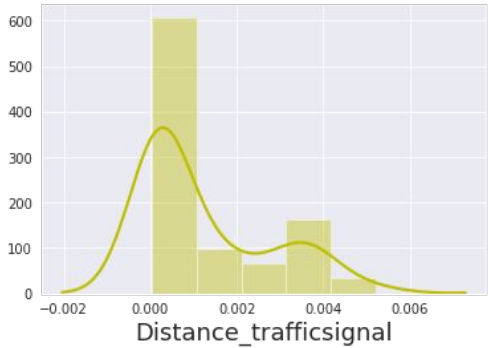




- Primary roads (e.g. Numancia, Arago)
- Secondary roads (e.g. Balmes)







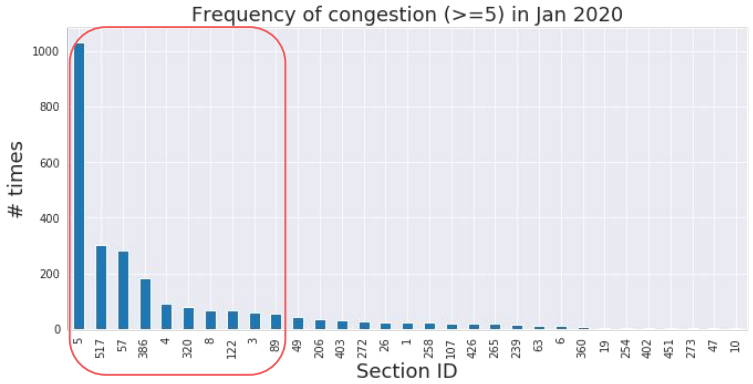
Let’s specify our focus:

I. Where

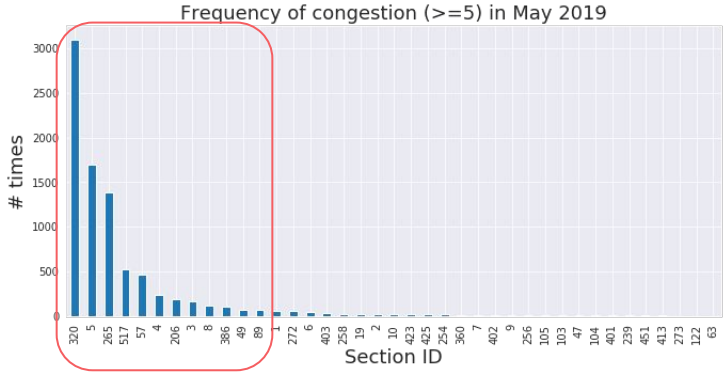
II. When

Most congested sections of Av. Diagonal

I. in Jan 2020



II. in May 2019

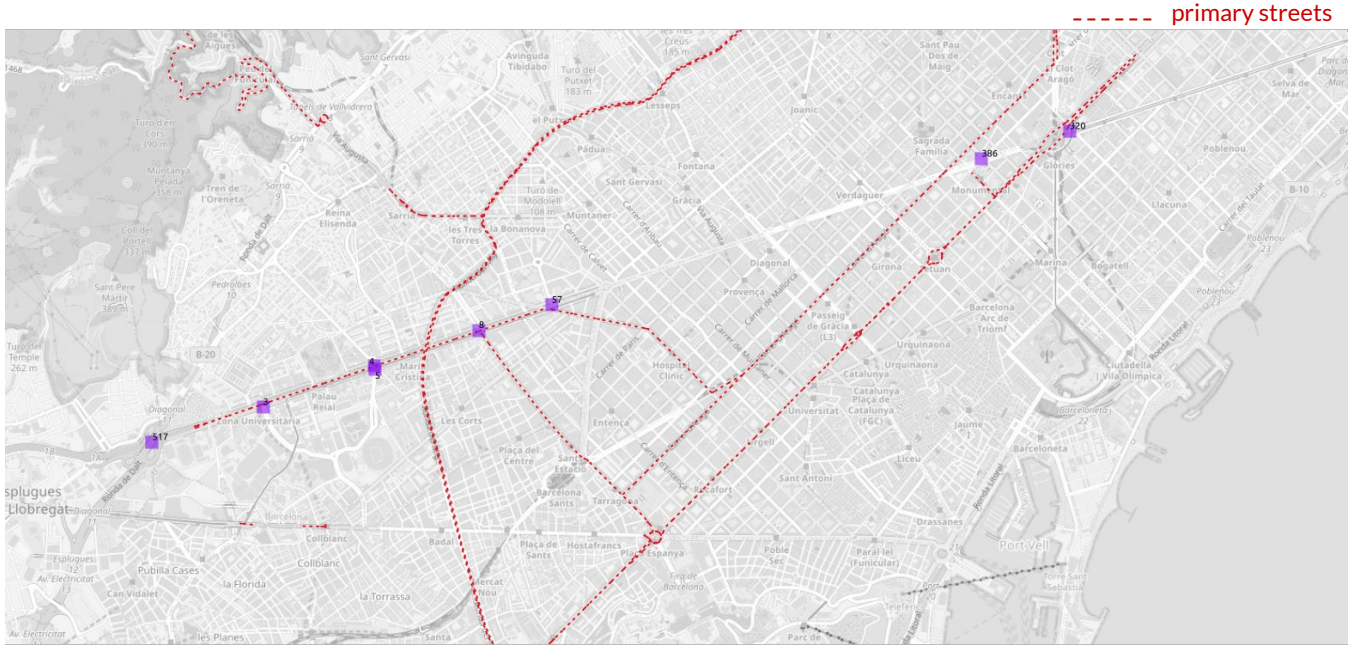


The busiest 10 sections : Where & When?

Most frequently congested sections of Av. Diagonal: 1. Where

8 overlapping IDs (of TOP 10):

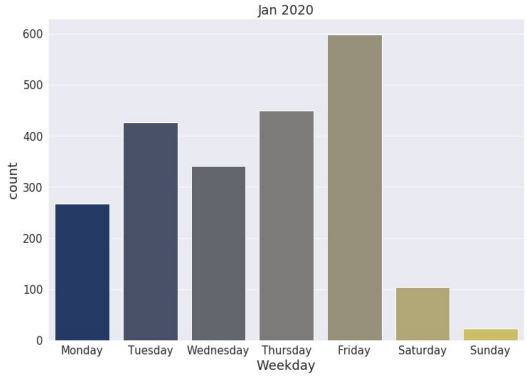
Tram	Description
3	Diagonal (Doctor Marañón a Pl. Pius XII)
4	Diagonal (Pl. Pius XII a Doctor Marañón)
5	Diagonal (Pl. Pius XII a Pl. Maria Cristina)
8	Diagonal (Numància a Pl. Maria Cristina)
57	Ganduxer (Diagonal a Pl. Gregori Tramaturg)
320	Meridiana (Pl. de les Glòries a Aragó)
386	Consell de Cent (Av. Diagonal a Av. Meridiana)
517	Ronda de Dalt (Diagonal a Carretera d'Esplugues)

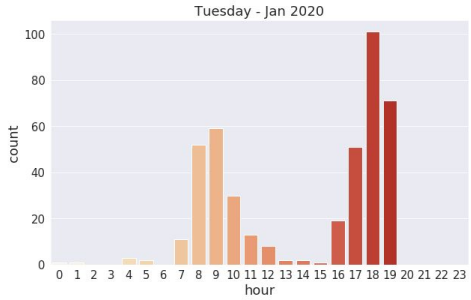


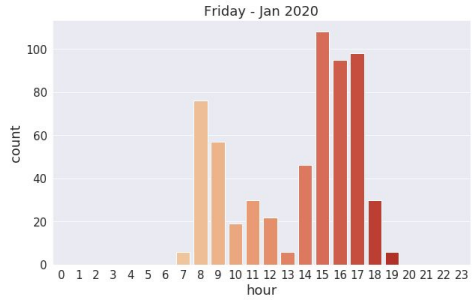
The most frequently congested section of the Diagonal is closely related to the layout of the primary roads of the city

Most frequently congested sections of Av. Diagonal: 2. When

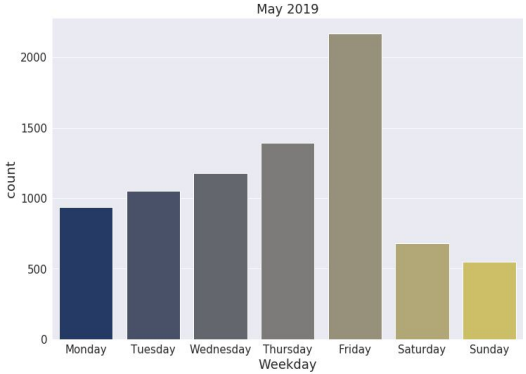
I. in Jan 2020

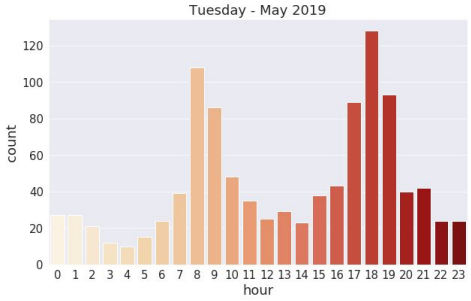


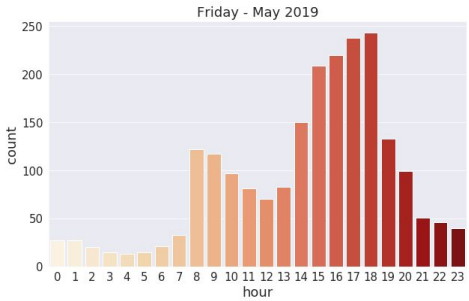




II. in May 2019







Data Modeling

Which model to use?

What to predict?

- traffic state (1 - 6) of 8 diagonal sections (Jan-July 2019)
- Features including:
  - time (month, weekday, hour)
  - how many intersections with other big roads
  - max speed of each section
  - how close to bus stops, traffic signals, cameras

Multiclass Classification

- Output: categorical values (1- 6)
- Time will be used as one of other features
- DecisionTreeClassifier
  &
  RandomForestClassifier

Model performance

Decision Tree Classifier

	precision	recall	f1-score
very fluid	0.89	0.82	0.85
fluid	0.88	0.65	0.75
dense	0.28	0.60	0.39
very dense	0.37	0.51	0.43
congestion	0.57	0.73	0.64
blocked	0.83	0.94	0.88

Random Forest Classifier

	precision	recall	f1-score
very fluid	0.89	0.84	0.87
fluid	0.89	0.70	0.78
dense	0.31	0.60	0.41
very dense	0.41	0.51	0.46
congestion	0.61	0.77	0.68
blocked	0.92	0.93	0.92

*class distribution: fluid (50%) > very fluid (25%) > congestion (9%) > dense (8%) > very dense (7%) > blocked (0.1%)

Conclusion

Traffic in Av. Diagonal	• 8 sections with consistent heavy traffic
Congestion: Where	• Just outside city center • Connection to suburb area of BCN
Congestion: When	• Peak on Friday • Rush hours (8am & 6pm | except Friday: 3pm)
Traffic State Prediction	• Easier to predict the extreme (fluid & congested) • More featured should be considered to predict traffic density 3-4