



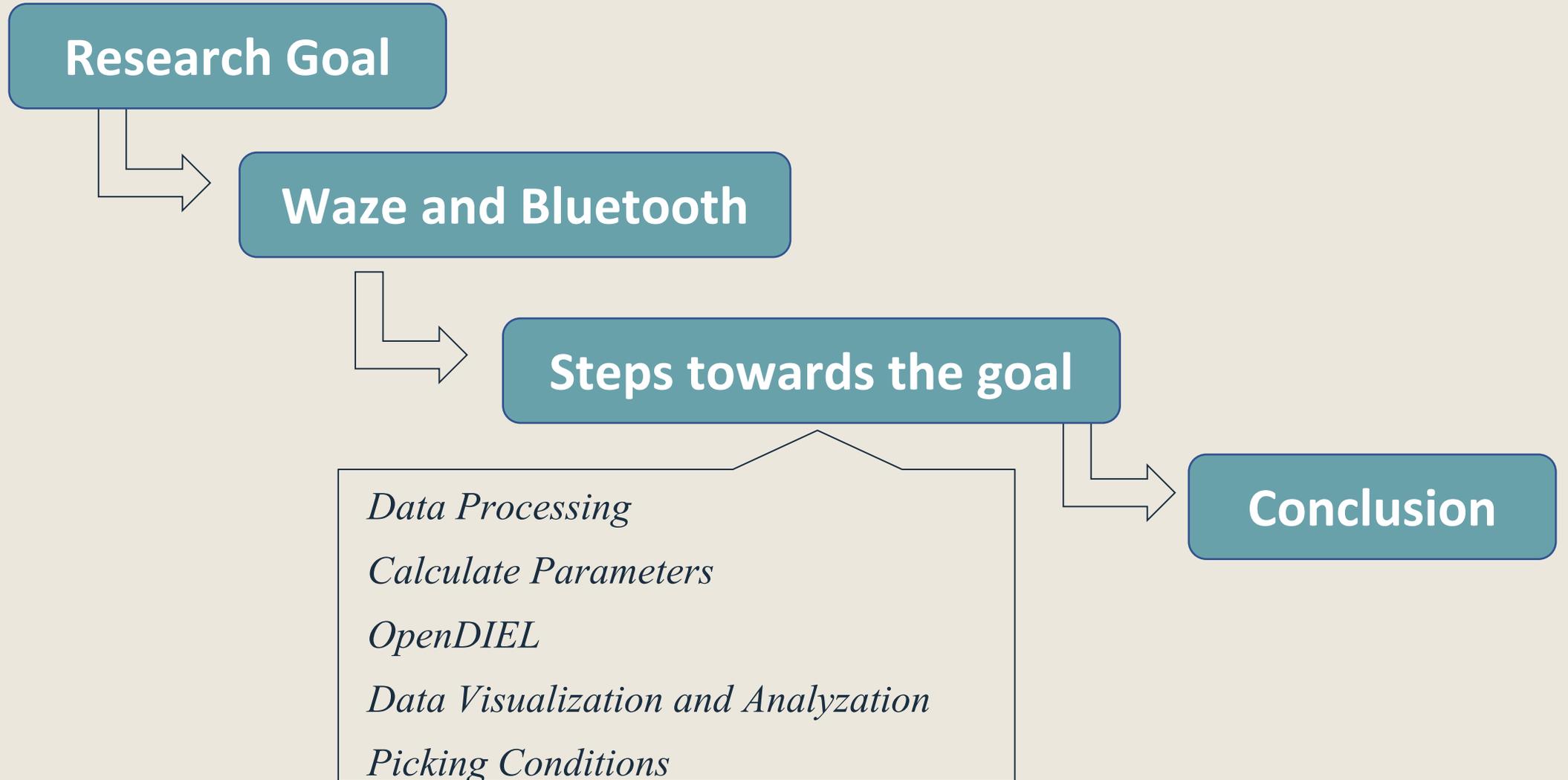
TRAFFIC DATA ANALYTICS

WAZE VS. BLUETOOTH

Students: Brooklynn Hauck (SRU); Tracy Liu (CSUST)

Mentors: Dr. Lee Han; Nima Hoseinzade; Yuandon Liu; Dr. Kwai Wong

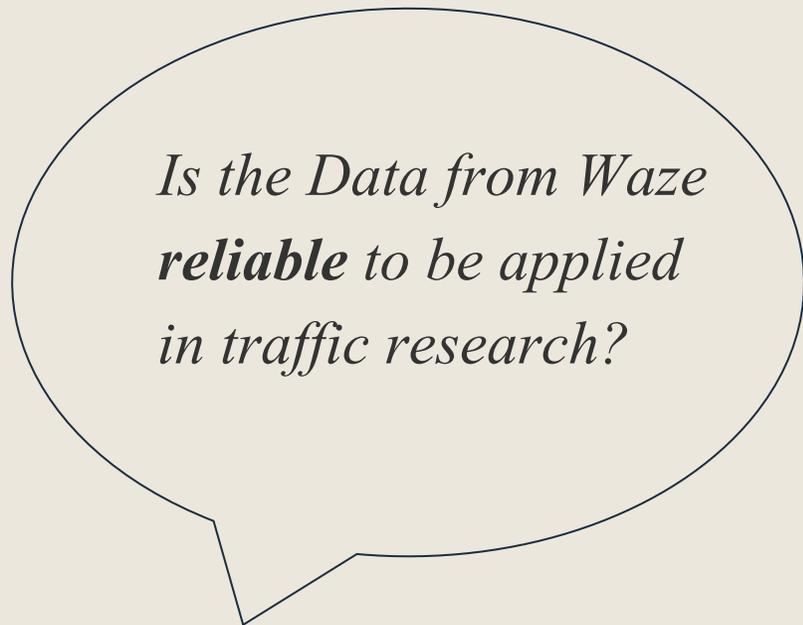
Introduction



Research Goal

Use **Bluetooth Data** as ground truth

Started out as a very open goal:



*1 What **day of the week** is the Waze Data most reliable?*

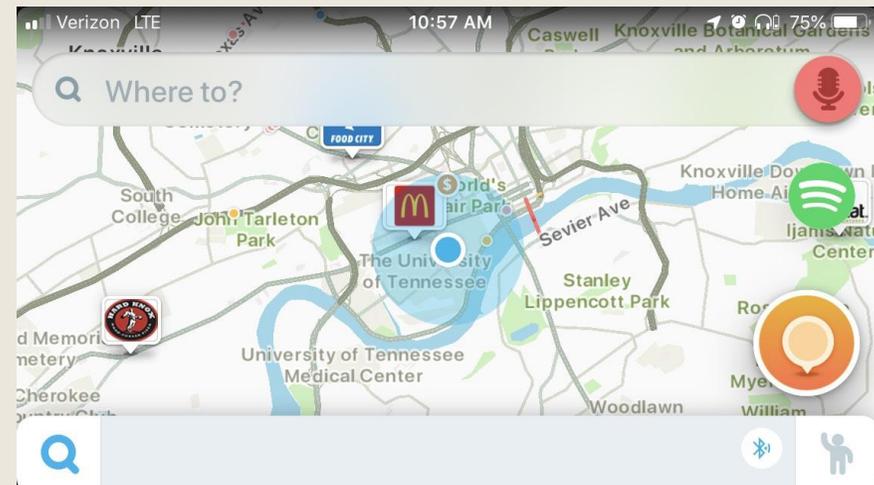
*2 What **time of day** is Waze Data most reliable?*

*3 Which **factors** affect the reliability of Waze?*

(Traffic Volume, Speed or Length of Segment)

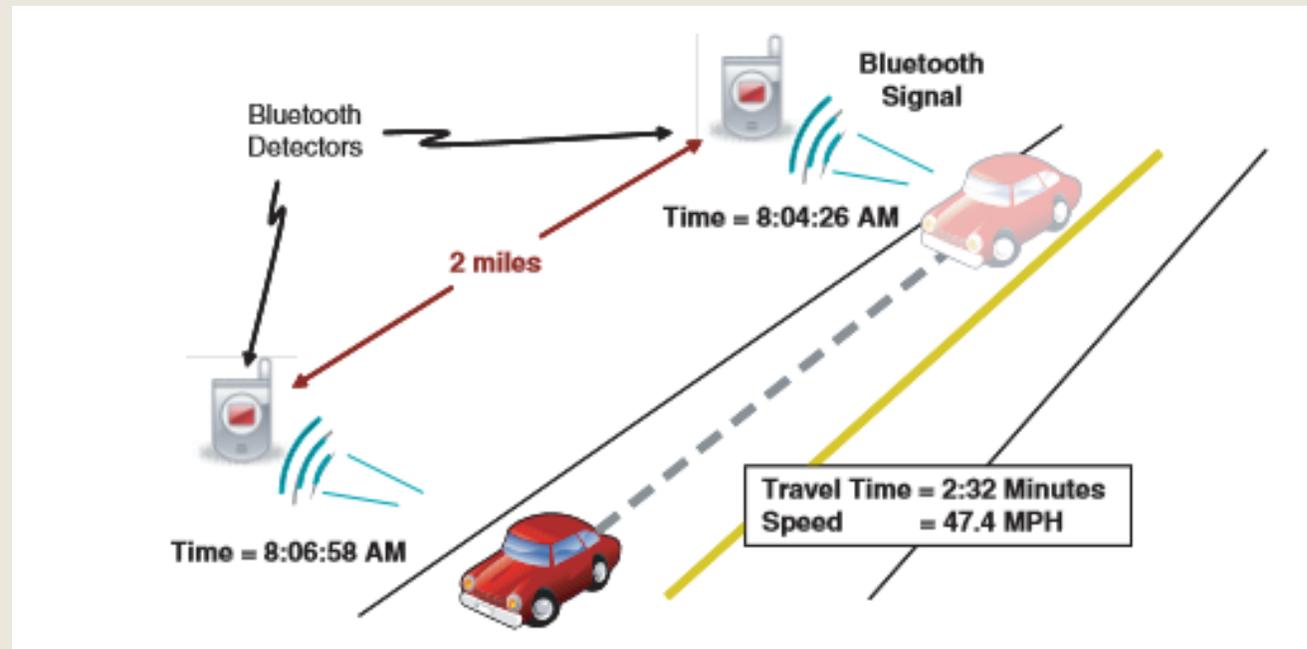
What is Waze?

- A downloadable GPS app for both IOS and Android, provided by Google, that provides turn-by-turn navigation information and user submitted Travel Time and Route Details
- Data is collected by Online data source provided by Waze
- Data has not been tested and is an unreliable source, though Universities have been using it for Research purposes.

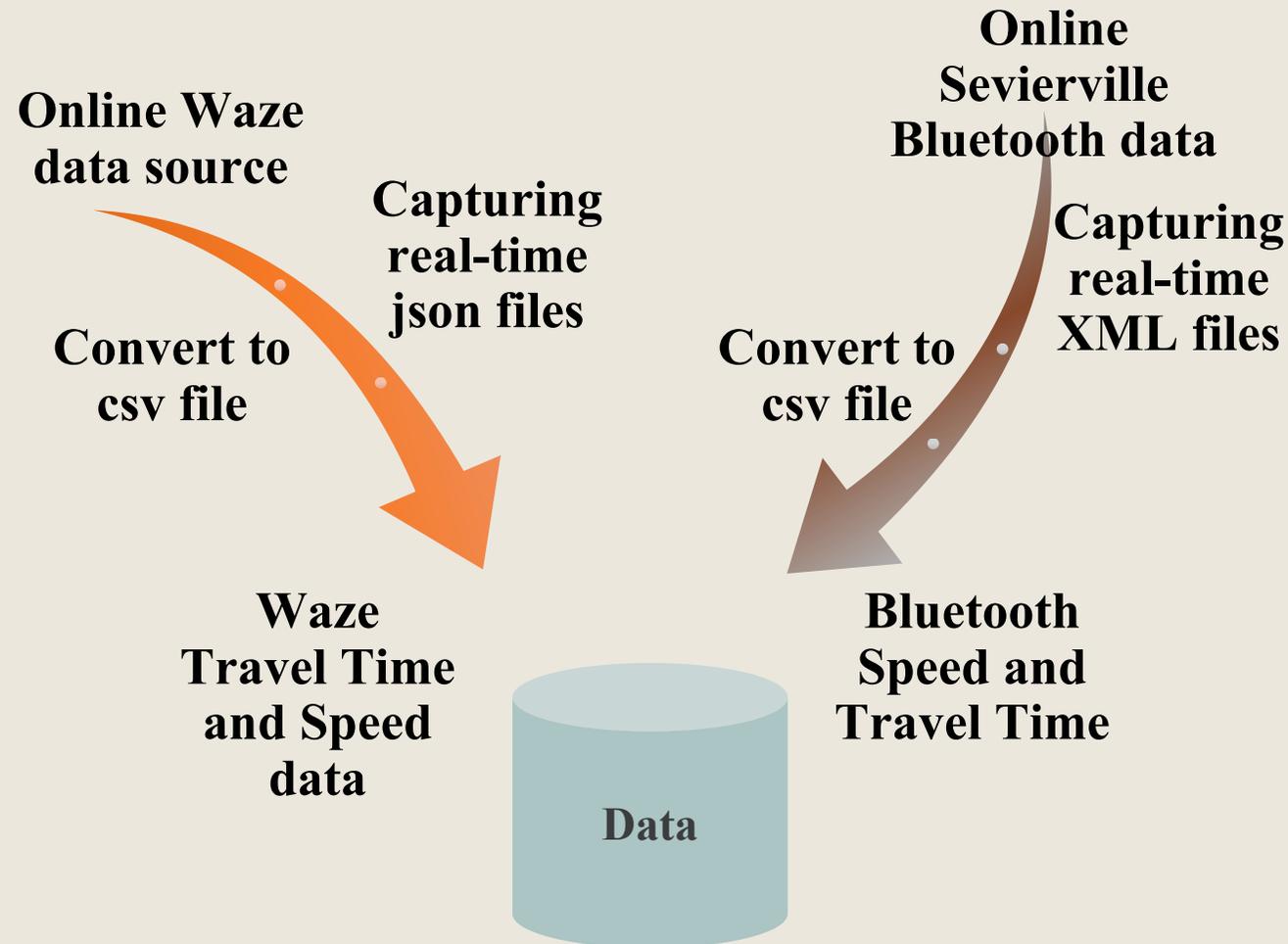


What is Bluetooth?

- A wireless technology standard for exchanging data between fixed and mobile devices over short distances
- Data is collected by two Bluetooth detectors with a set distance between the two.
- Bluetooth has been proven to be a reliable source of data.



About Data

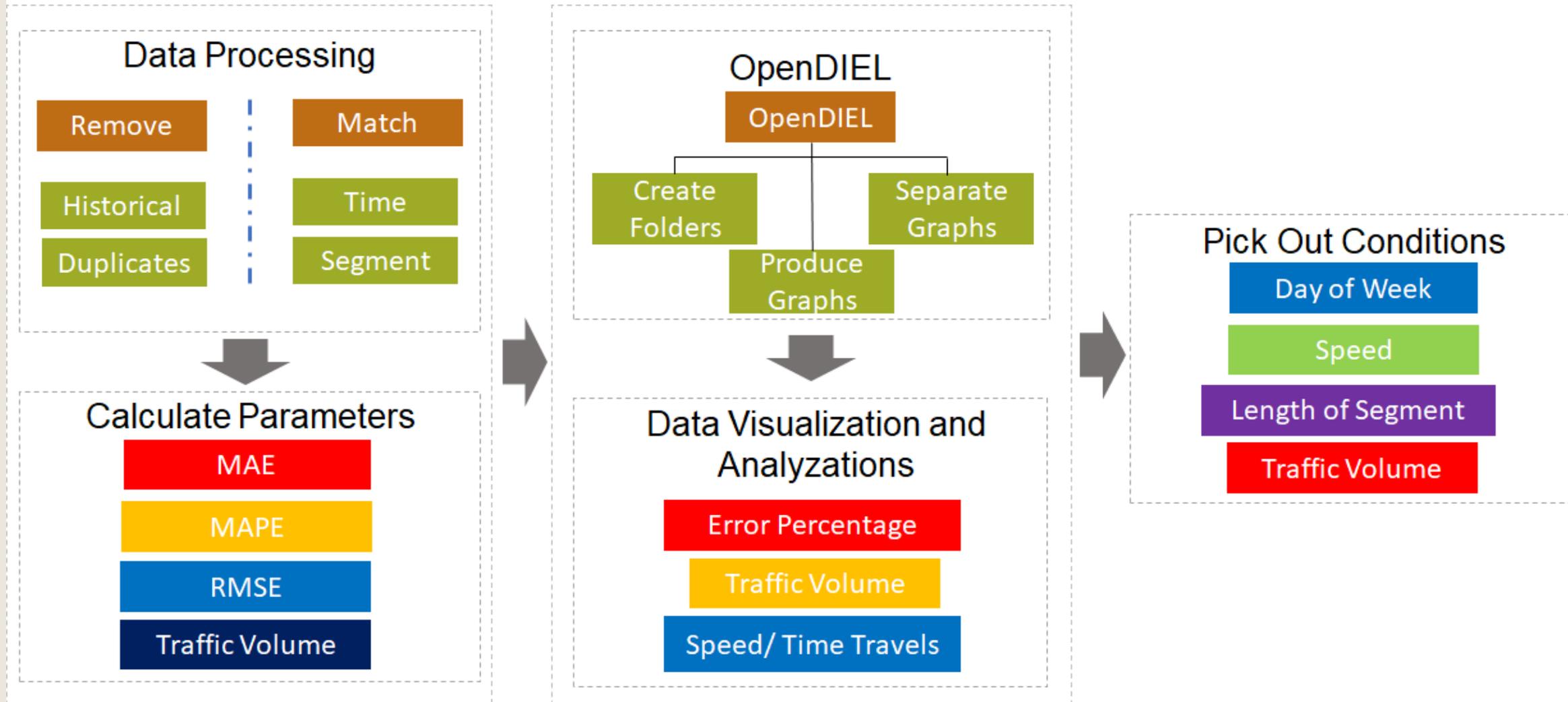


Speed & Travel Time

One month

Each second

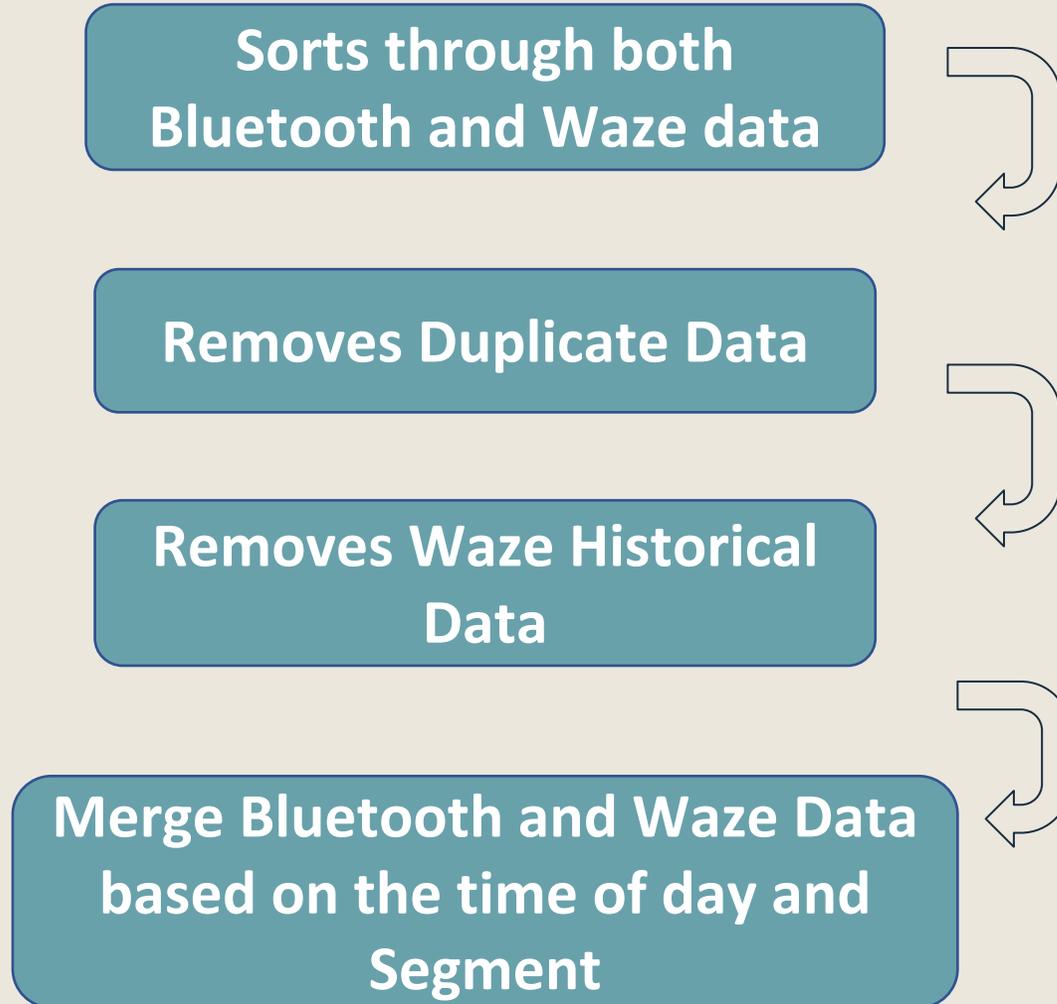
Steps Towards the Goal



Data Processing

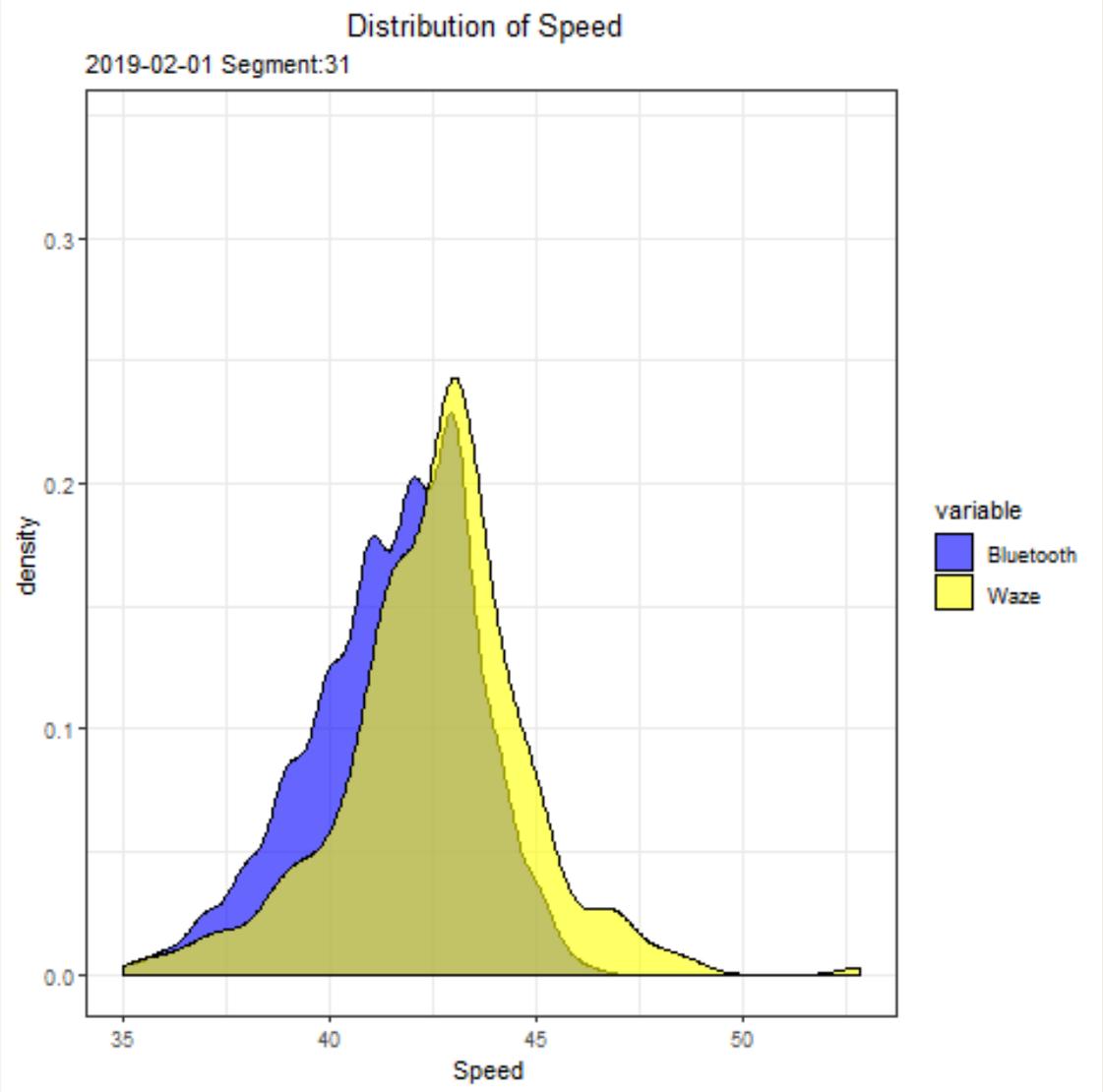
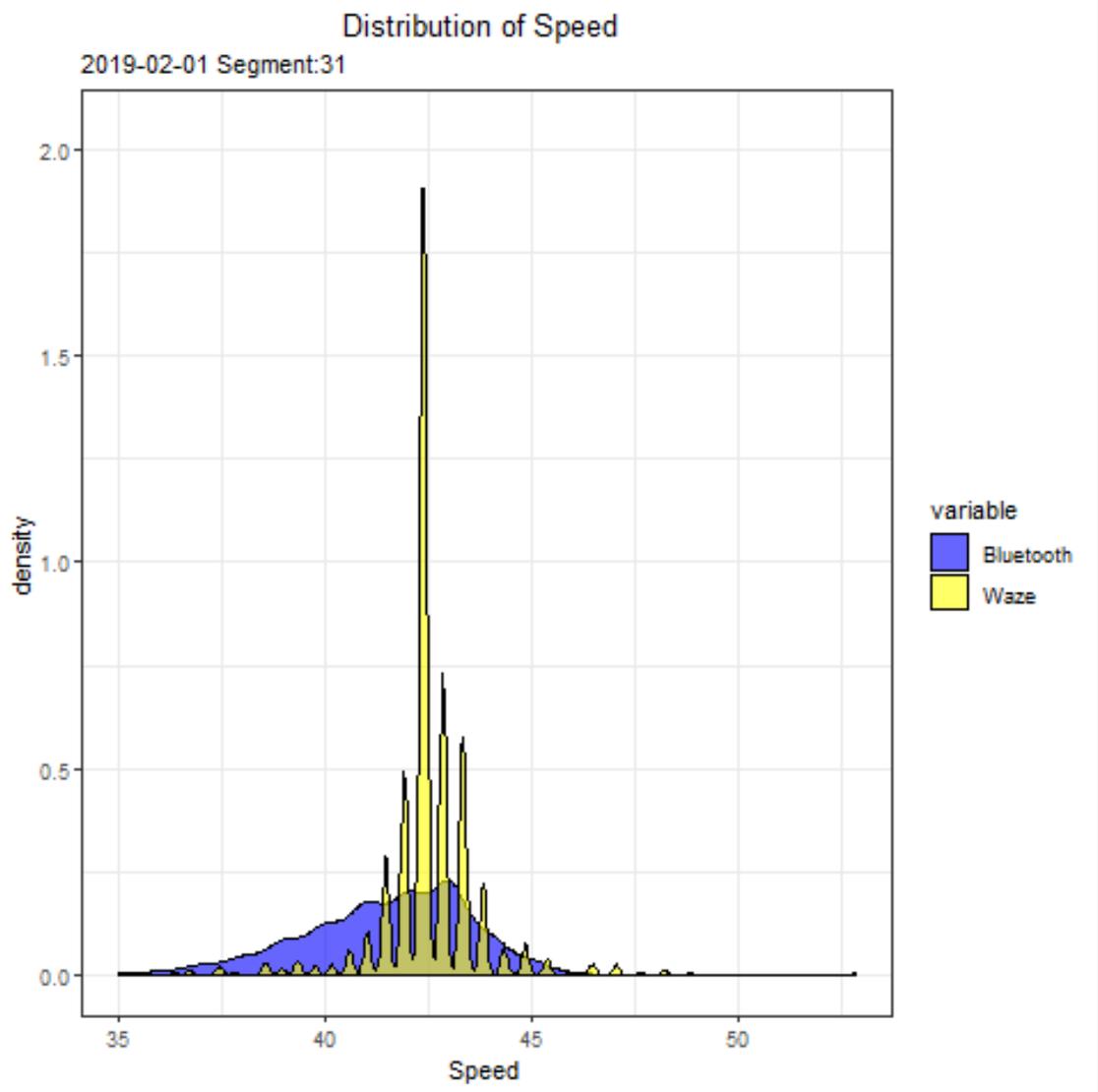


Code was written in R



Time	bITT	bIS	waTT	waS
00:00:00	NA	NA	NA	NA
00:00:01	NA	NA	NA	NA
00:00:02	NA	NA	NA	NA
00:00:03	NA	NA	NA	NA
00:00:04	NA	NA	NA	NA
00:00:05	NA	NA	NA	NA
00:00:06	NA	NA	NA	NA
00:00:07	NA	NA	NA	NA
00:00:08	NA	NA	NA	NA
00:00:09	NA	NA	NA	NA
00:00:10	NA	NA	NA	NA
00:00:11	NA	NA	NA	NA
00:00:12	NA	NA	NA	NA

Removing Historical Data



Calculate Parameters

- We choose these three error formulas as parameters:
 - *MAE* (Mean Absolute Error)
 - *RMSE* (Root Mean Standard Error)
 - *MAPE* (Mean Absolute Percentage Error)

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y - \bar{y}|$$
$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y - \bar{y})^2}$$
$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y - \bar{y}}{y} \right|$$

Time	bIS	waS
00:01:11	NA	NA
00:01:12	NA	34
00:01:13	45	NA



Time	bIS	waS
00:01	40	38
00:02	44	42
00:03	34	36

Aggregate in
1 minute

y = Bluetooth speed sample(after aggregate) in each hour

\bar{y} = Waze speed sample(after aggregate) in each hour

N = number of Bluetooth/Waze speed samples(after aggregate) in each hour

	MAE	RMSE	MAPE
Segment 1	2.1892	2.6288	0.0469
Segment 2	2.4704	2.9115	0.0515
Segment 3	2.2663	2.6334	0.0505
Segment 4	7.8430	8.7186	0.2052
Segment 5	6.0807	6.9048	0.1573
Segment 6	8.2280	8.9959	0.2684
Segment 7	4.2072	4.7272	0.1366
Segment 8	16.3164	16.5005	0.4805
Segment 9	6.0667	6.3625	0.1341
Segment 10	7.2523	8.1692	0.1924
Segment 11	7.6703	8.5503	0.2293
Segment 12	3.1369	3.6451	0.0907
Segment 13	3.0609	3.5602	0.0640
Segment 14	2.0104	2.3866	0.0429
Segment 15	1.9895	2.3519	0.0419

Segment 16	7.1739	8.0068	0.1721
Segment 17	2.8045	3.3861	0.0599
Segment 18	3.1959	3.6432	0.0766
Segment 19	4.4692	4.9151	0.1192
Segment 20	7.4335	8.3444	0.1743
Segment 21	7.3808	8.5761	0.3010
Segment 22	6.0856	6.7456	0.2412
Segment 23	5.7263	6.2203	0.1603
Segment 24	3.7685	4.2955	0.1057
Segment 25	9.6134	10.4328	0.3068
Segment 26	10.6924	11.7192	0.2537
Segment 27	2.7086	3.2314	0.0749
Segment 28	11.5111	12.2354	0.3003
Segment 29	3.1410	3.5524	0.0804
Segment 30	3.6801	3.9767	0.1214
Segment 31	1.6250	1.9510	0.0386

- The MAE and RMSE of most segments do not exceed 10, the MAPE does not exceed 0.1
- There is big difference in different segments, which may be caused by other factors

OpenDIEL

- R Code was ran on Comet in OpenDIEL to produce the distribution graphs
 - *Over 1800 graphs total*
- 3 different codes to produce all 1800 graphs
 - *Create Folders*
 - *Produce Graphs*
 - *Separate Graphs (Segments; Days of the week)*

OpenDIEL Codes

Create Folders

- Folder for one Segment and One Day (930 Folders)
- Copies data Specific for that day
- Creates a text file with the Day and Segment

Produce Graphs

- Sorts through Bluetooth and Waze Data
- Pulls out all the data for a Specific Day and Segment
- Prints out distribution graphs
 - *Speed*

Separate Graphs

- Puts graphs into folder based on the Segment and Day of the week

Data Visualization

- Printed graphs based on:

{
Speed
Error Percentage
Traffic Volume
}

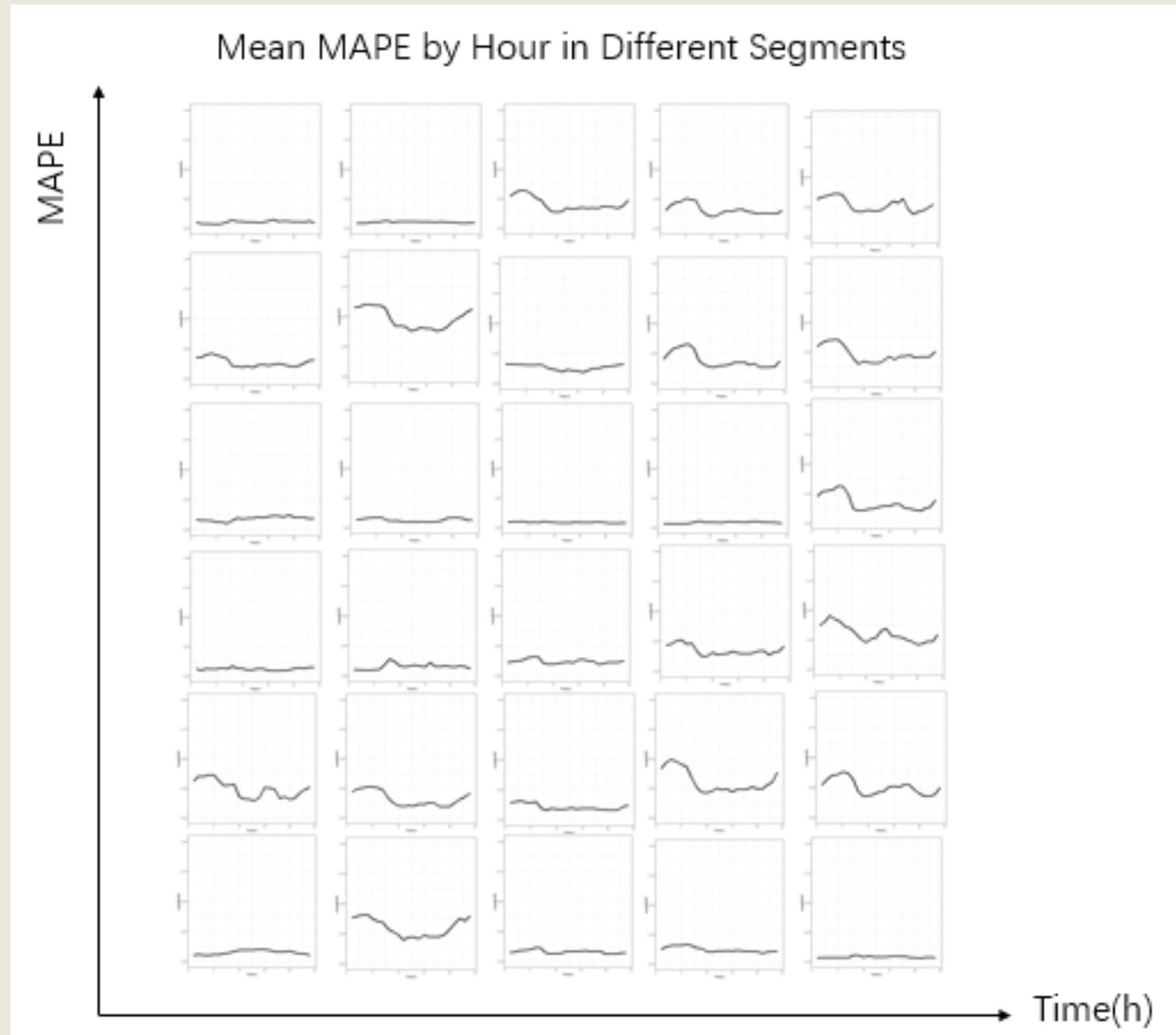
- Types of Graphs

{
Distribution Graph
Line Graph
Box Plot
Scatter Plot
}

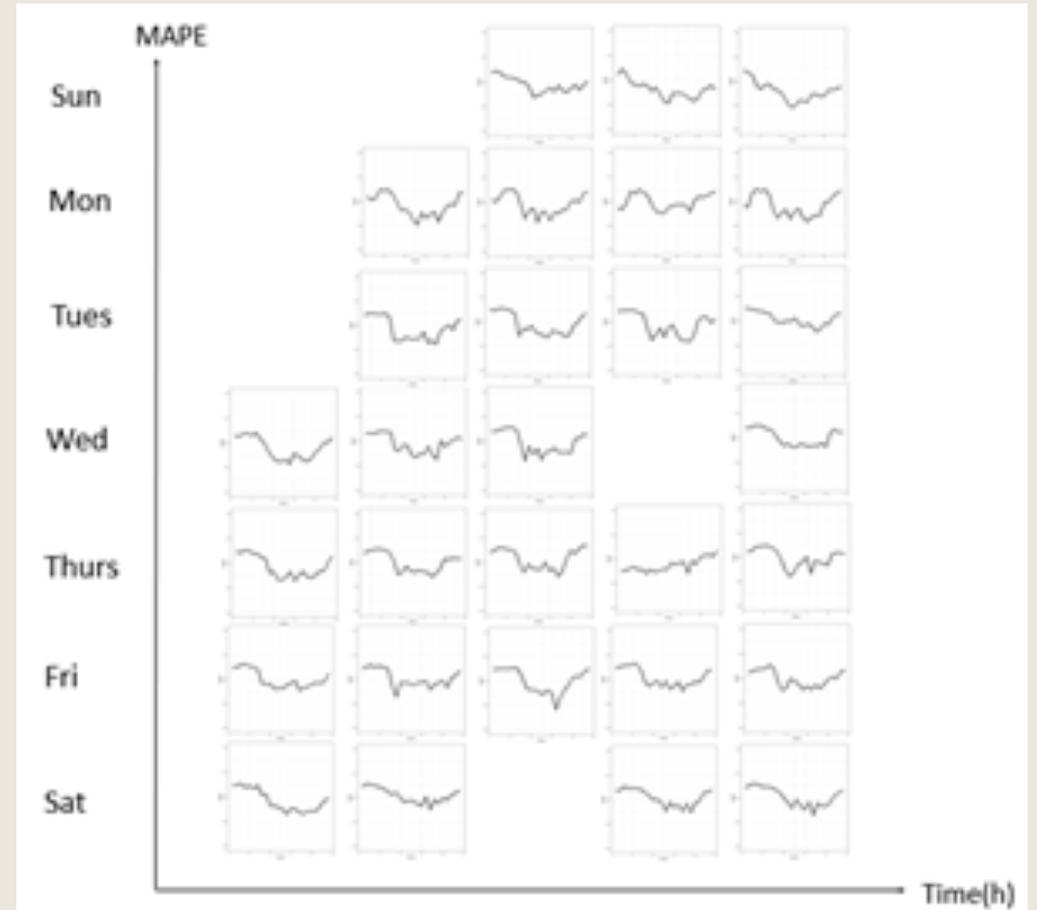
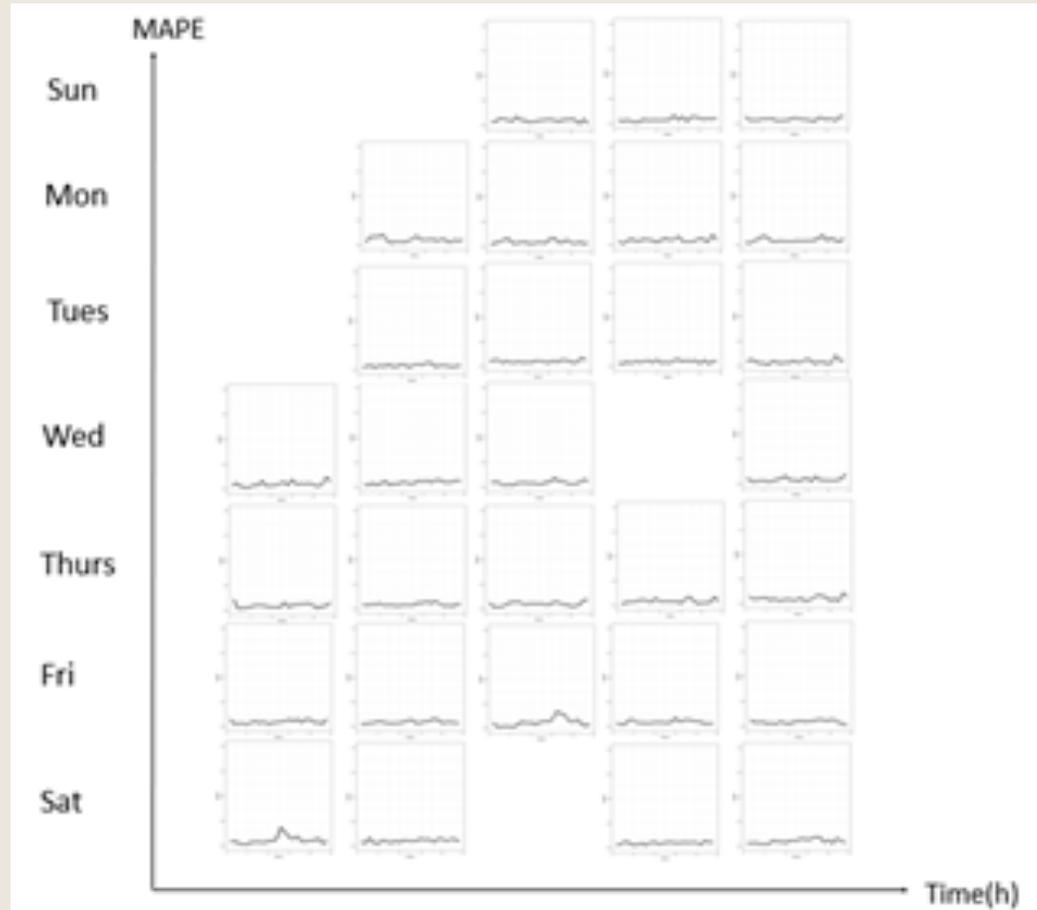
Data Analyzation

- What we did?
 - *Compare based on different segments*
 - *Compare based on the day of the week*
 - *Compare based on the Traffic Volume*
 - *Picked two segments with good MAPE and two segments with bad MAPE*
 - *Compare based on the Length of the Segment*
 - *Compare based on the Speed*

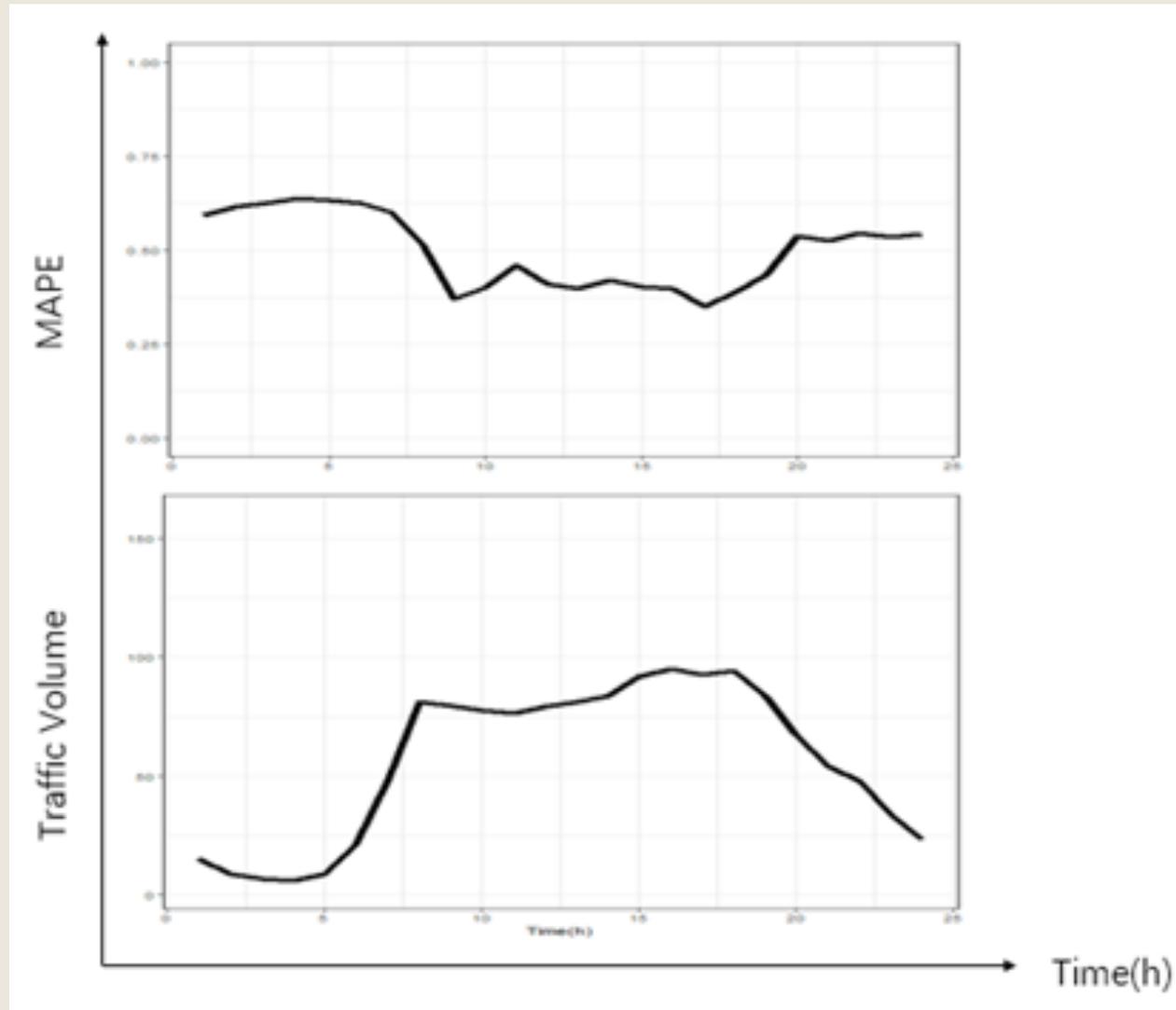
Mean MAPE in 31 Segments



Comparison based on the days of the week

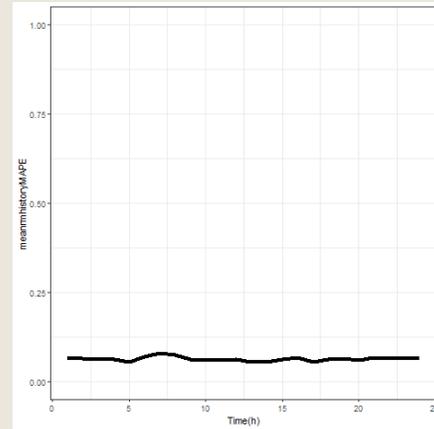


Comparison based on the Traffic Volume

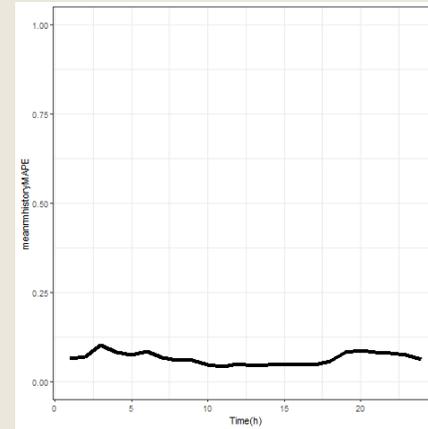


Segments Used for Comparison

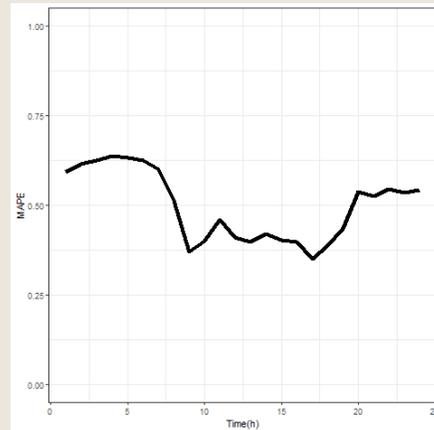
2



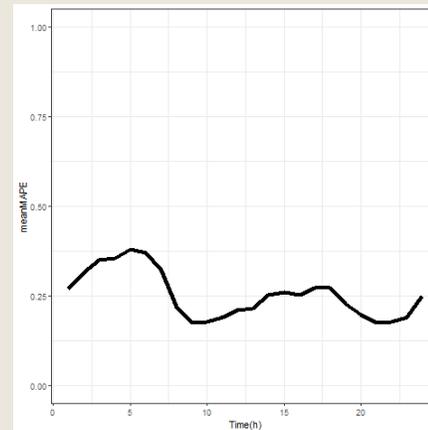
18



8

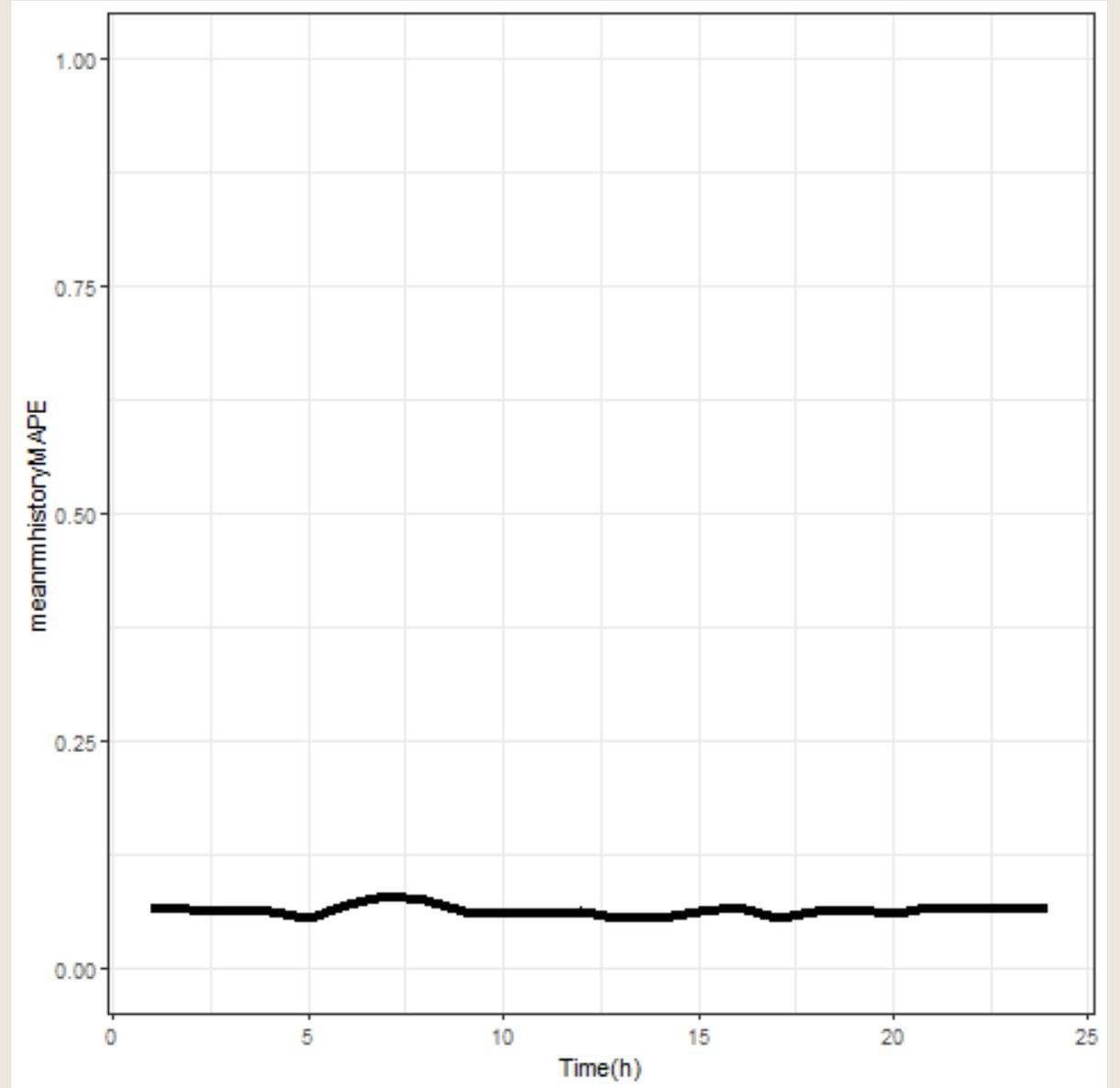
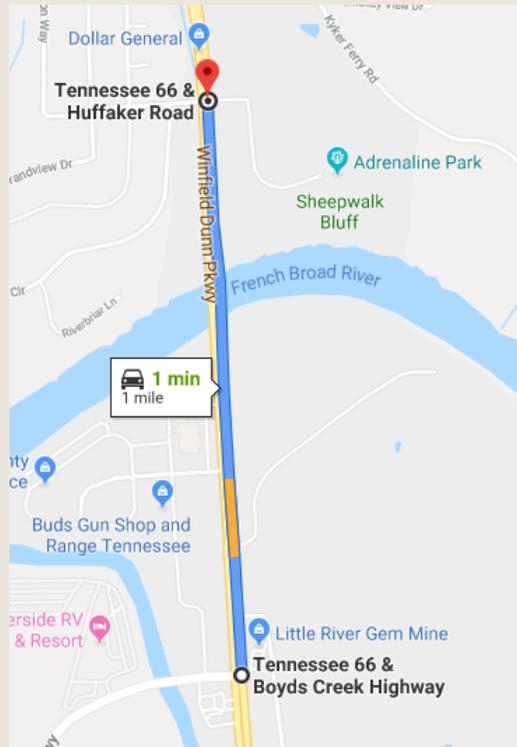


26



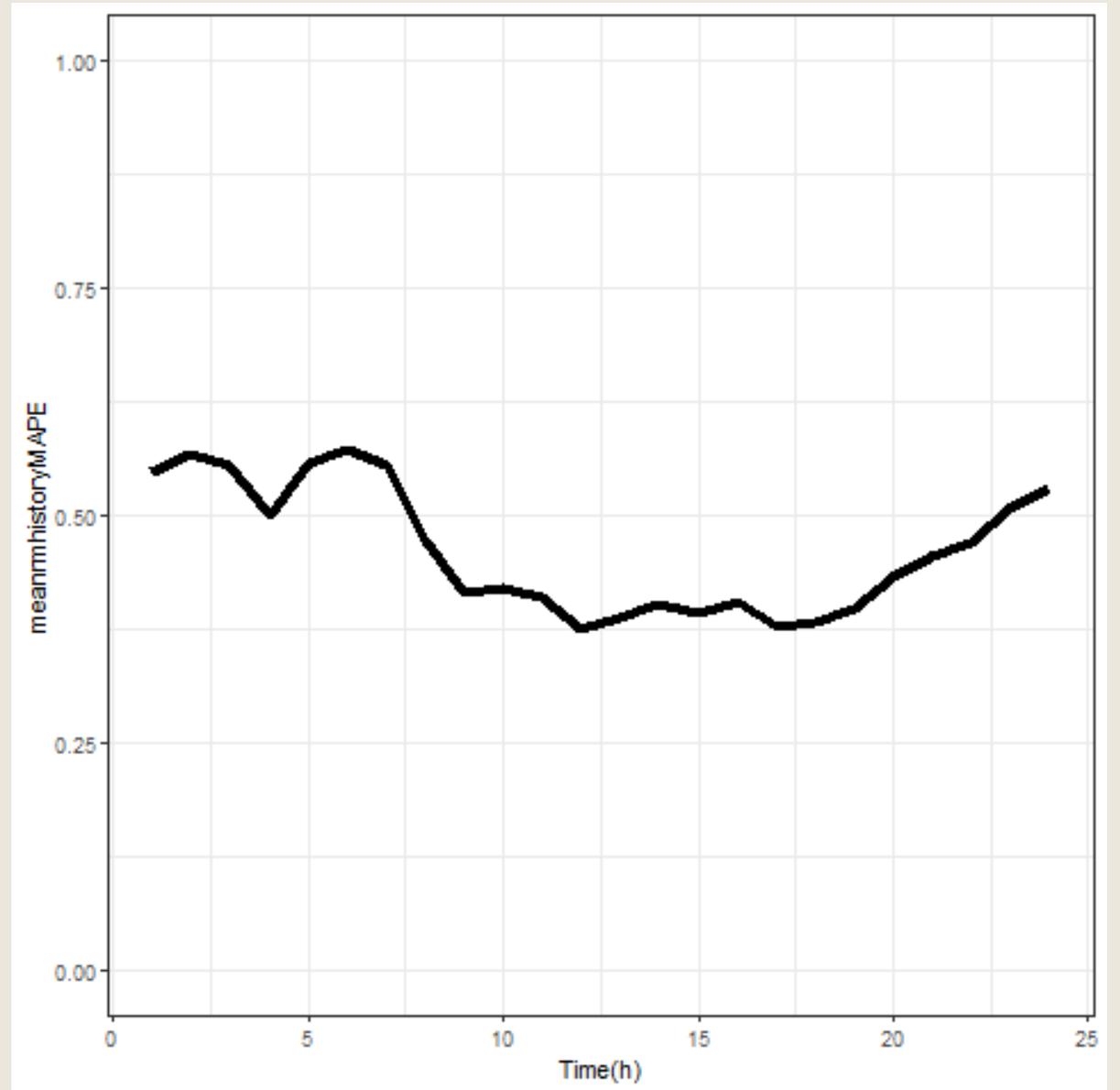
Segment 2

- TN-66 & Boyds Creek Rd to TN-66 & Huffaker Rd (Sevierville, TN)
- Speed Limit: 45 mph
- Length of Segment: 1 mile



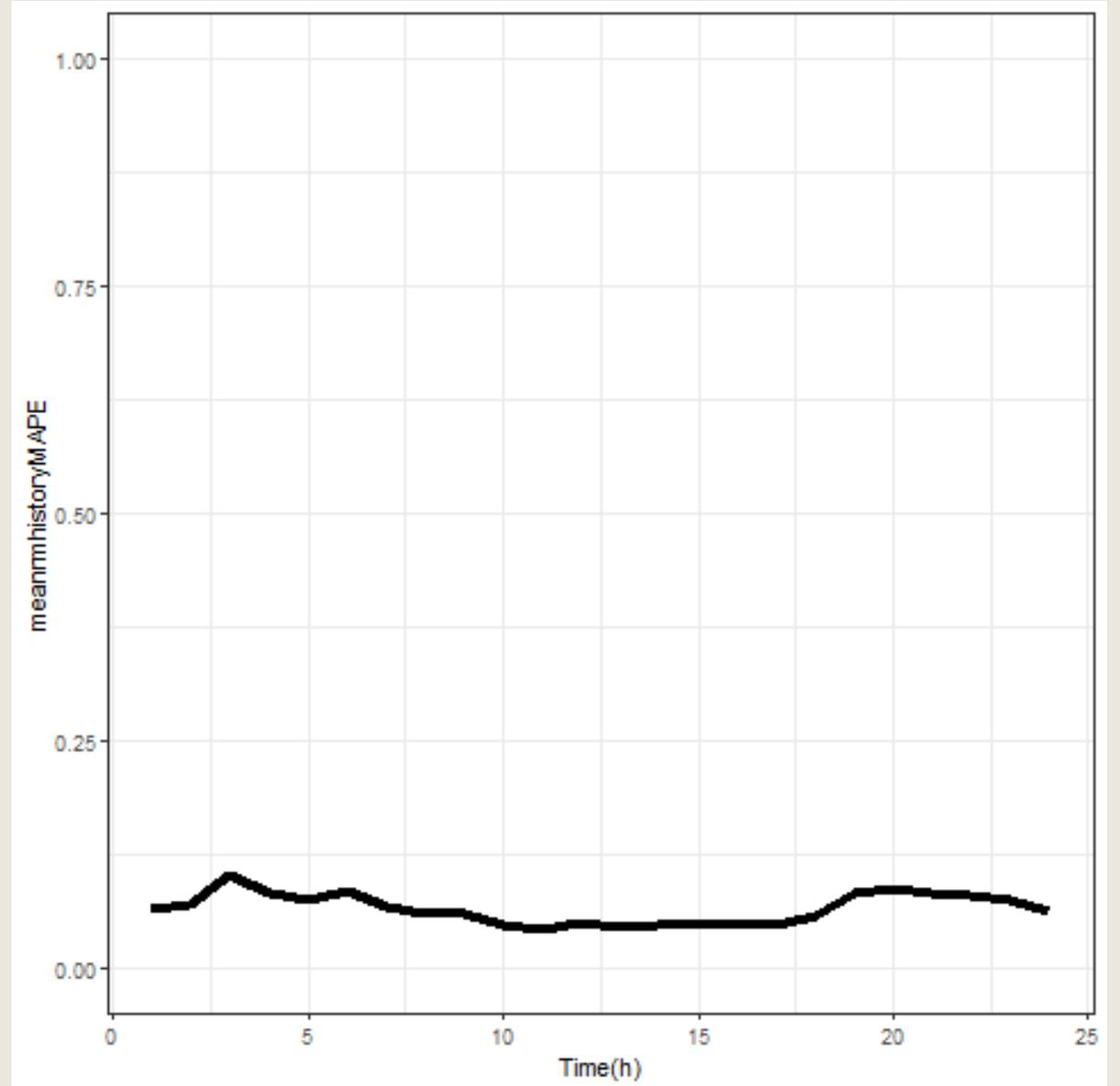
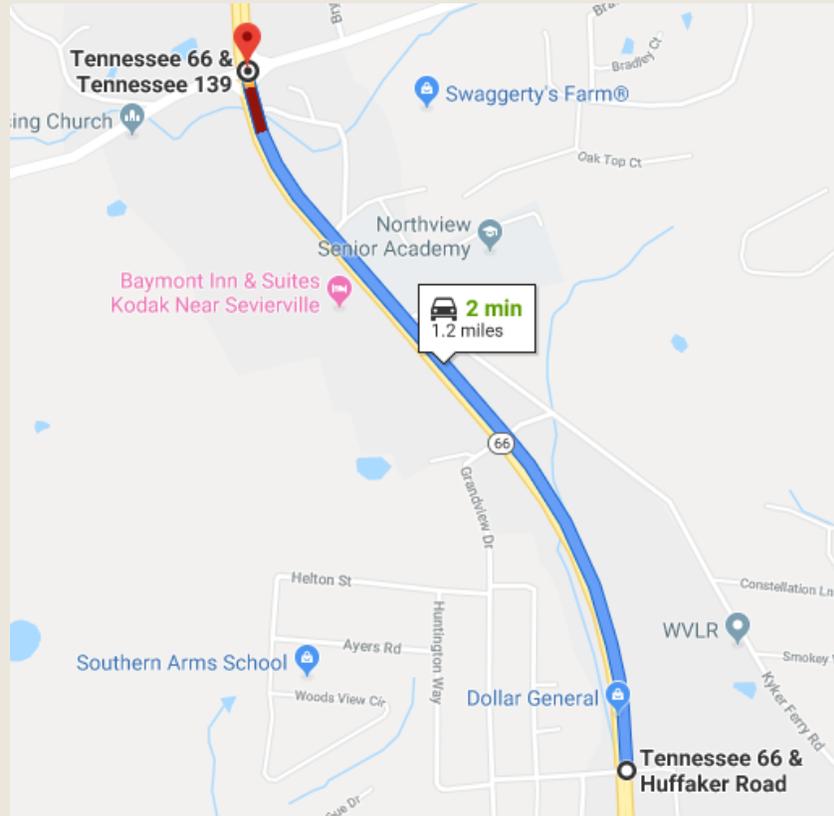
Segment 8

- Dolly Parton Pkwy & Veterans Blvd to Dolly Parton Pkwy & Food City (Sevierville, TN)
- Speed Limit: 35 mph
- Length of Segment: 0.4 mile
- Red Lights Passed: 1



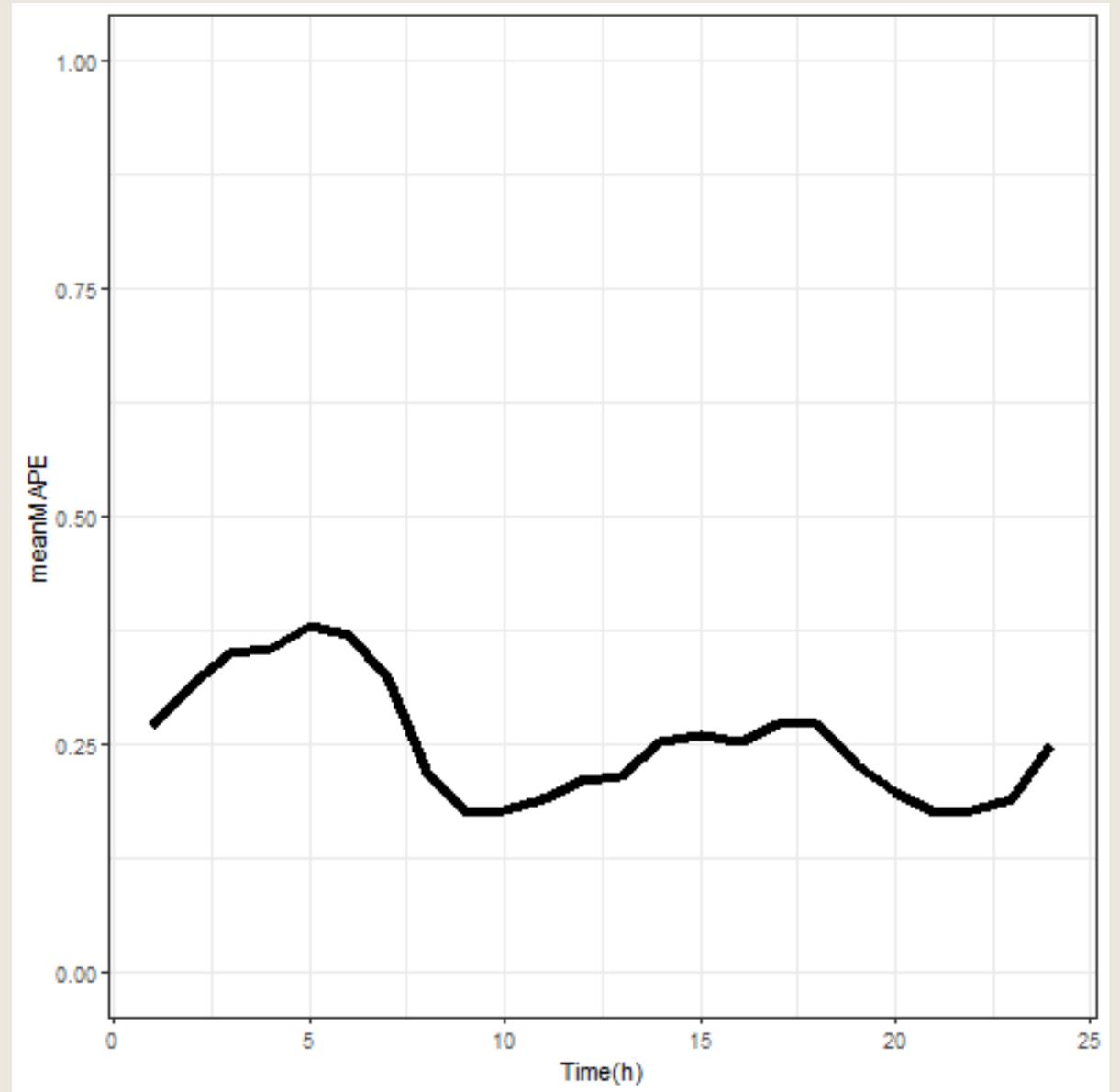
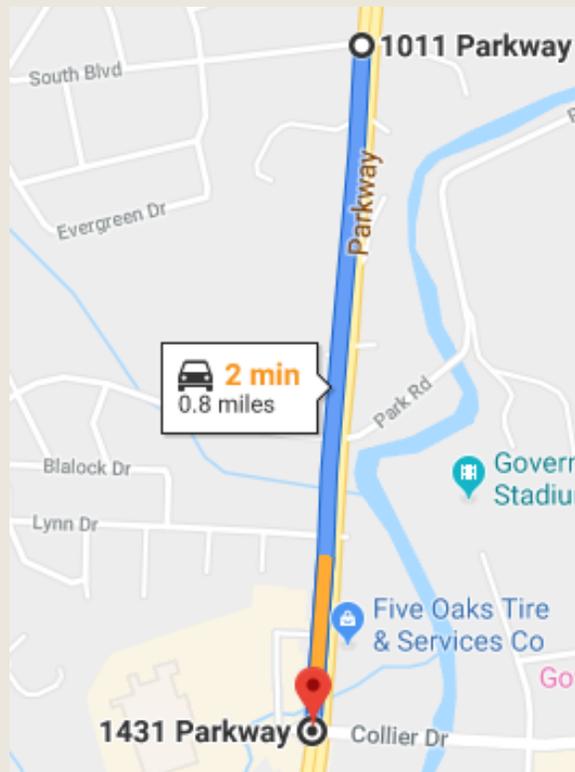
Segment 18

- Tn-66 & Huffaker Rd to TN-66 & TN-139 (Sevierville, TN)
- Speed Limit: 45 mph
- Length of Segment: 1.2 mile



Segment 26

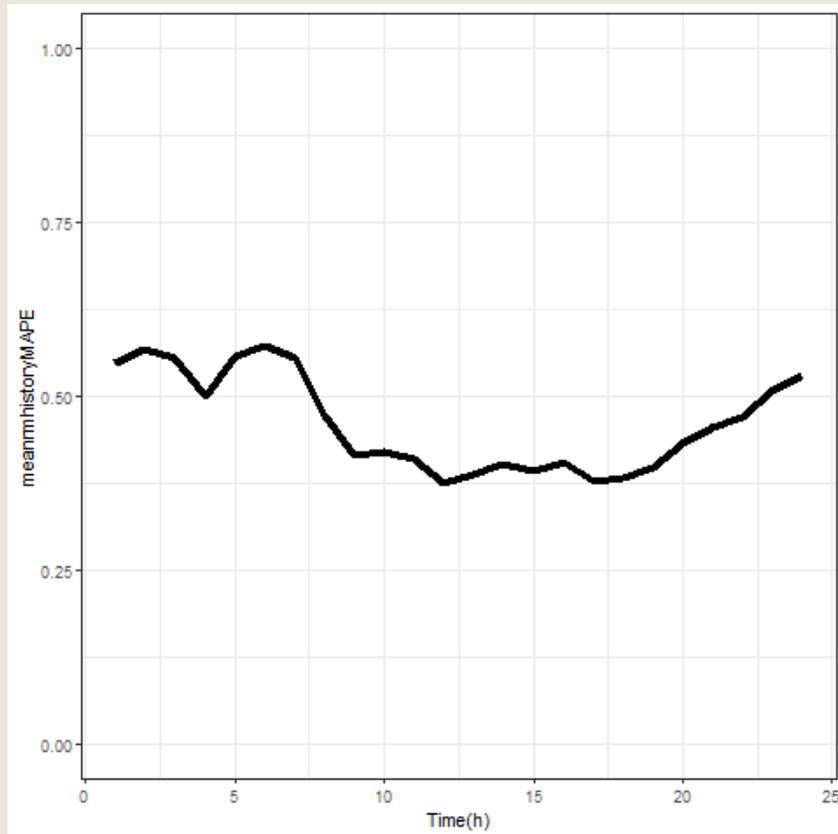
- South Blvd to Collier Dr. (Pigeon Forge, TN)
- Speed Limit: 35 mph
- Length of Segment: 0.8 mile
- Red Lights Passed: 1



Comparison based on the Length of a Segment

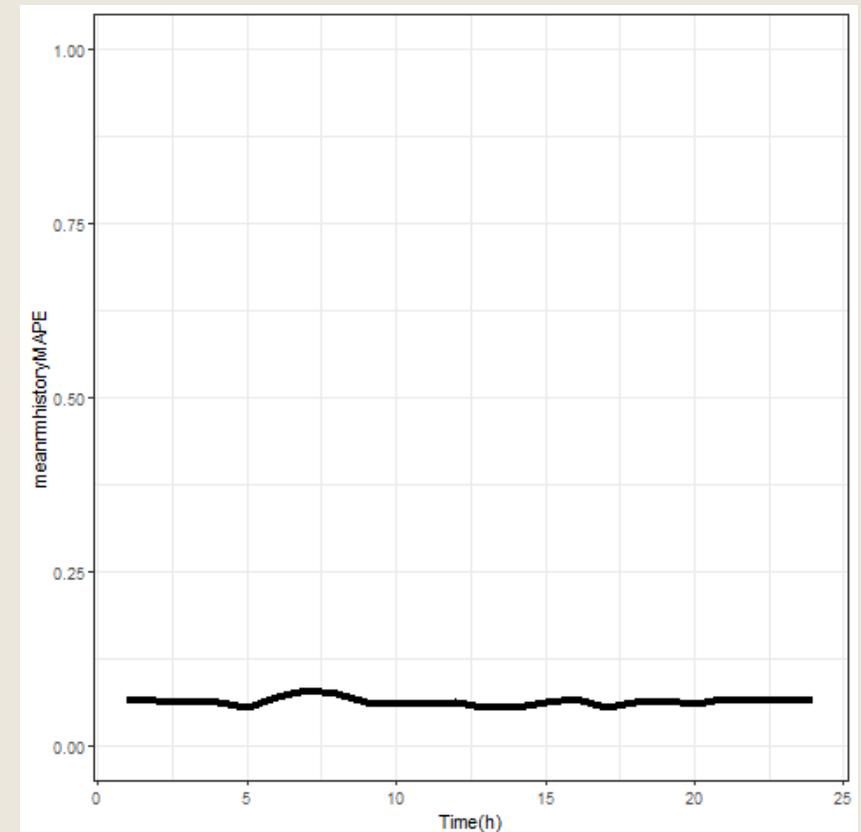
Segment 8

- Length of Segment: 0.4 mile
- Lights Passed: 1



Segment 18

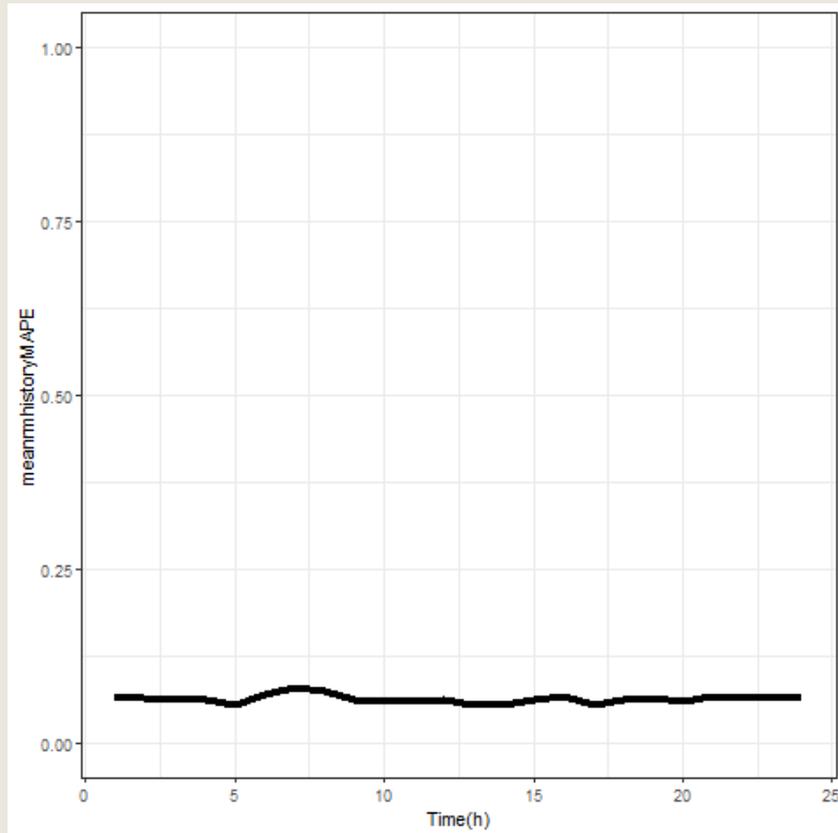
- Length of Segment: 1.2 mile
- Lights Passed: 1



Comparison based on the Length of a Segment (Number of Stop lights)

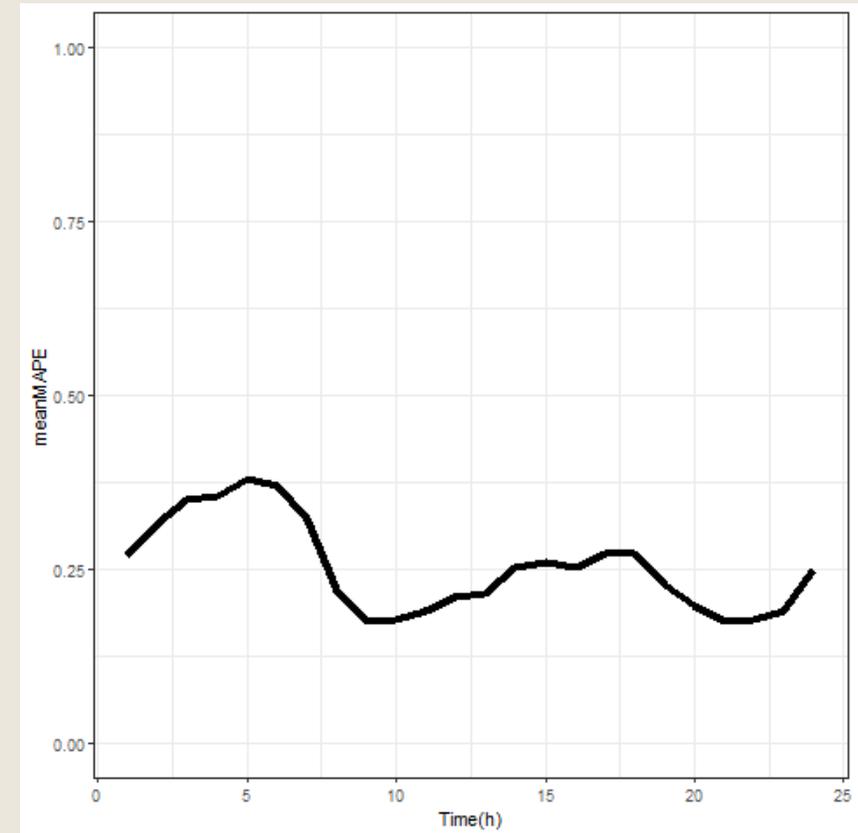
Segment 2

- Length of Segment: 1 mile
- Lights Passed: 0

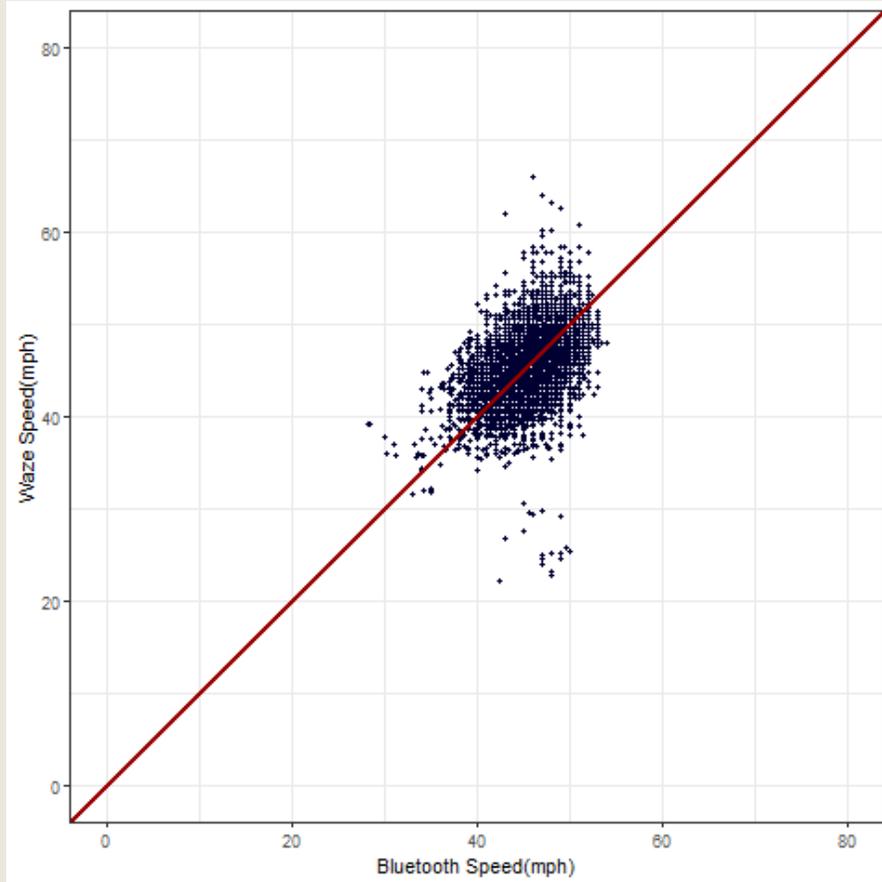


Segment 26

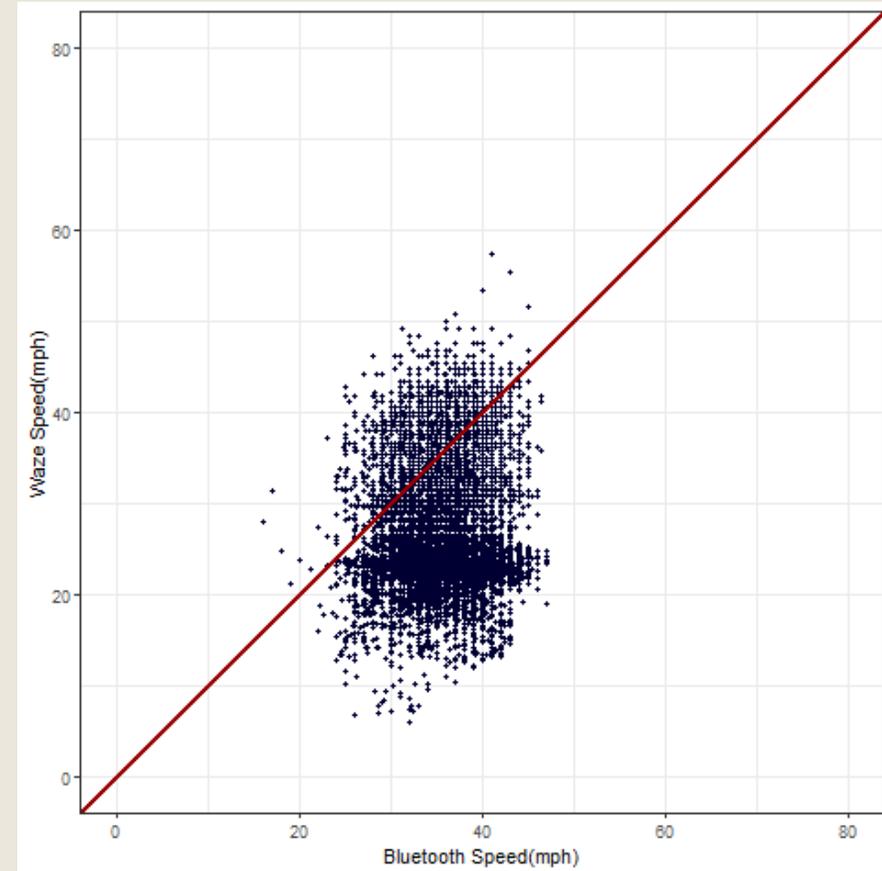
- Length of Segment: 0.8 mile
- Lights Passed: 1



Comparison based on the Speed

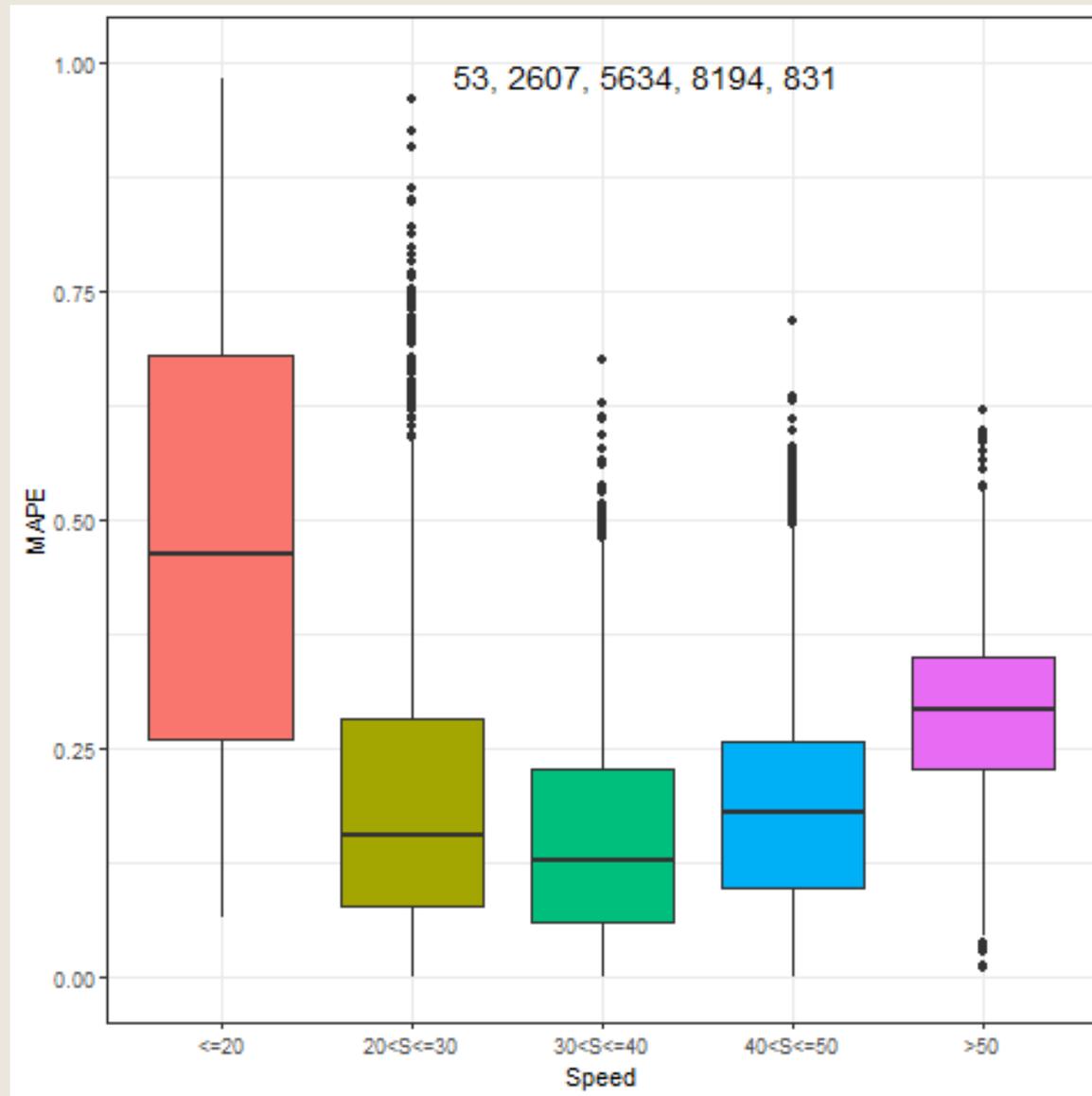


- 1) Speed points are close to the perfect line (Bluetooth speed = Waze speed)
- 2) Range 40~55 mph



- 1) Range 25~45 mph
- 2) Waze Speed < Bluetooth Speed

Comparison based on the Speed



Grouping in 5 speed levels:

$\text{speed} \leq 20$

$20 < \text{speed} \leq 30$

$30 < \text{speed} \leq 40$

$40 < \text{speed} \leq 50$

$50 < \text{speed}$

when speed **less than or equal to 20**, or **more than 50 mph**, MAPE is higher than other speed levels.

When speed is **from 30 to 40 mph**, MAPE is very low

Is Waze Reliable?

Yes, but only under these conditions:

- Day of the week?
 - *Any day of the week works!*
- Time of day?
 - *Day time; Roughly between 7am and 8pm*
- Traffic Volume?
 - *Traffic Volume need to be High*
 - *The higher the Traffic Volume the lower the MAPE*
- Speed?
 - *Speed needs to be between 30 mph and 45 mph*
- Length of the segment?
 - *The Length of the segment should long*
 - *Longer Segments have lower MAPE*
 - *Preferably with no stop lights to pass through*

Acknowledgement

This project has been done during the tenure of the RECSEM program of 2018. The program took place at The University of Tennessee. The program is funded by JICS, NSF, and UTK. Nothing could have been accomplished without the guidance of our mentors, Dr. Lee Han, Dr. Kwai Wong, as well as the graduate students, Nima Hoseinzadeh, Yuandong Liu.

References

Ding, F., Chen, X., He, S., Shou, G., Zhang, Z., & Zhou, Y. (2019). Evaluations of Wi-Fi Signal Based System for Freeway Traffic States Monitoring: An Exploratory Field Test.

Haghani, A., Hamed, H., Sadabadi, K. F., Yound, S., & Tarnoff, P. (n.d.). Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors.

Yang, S., Brakewood, C., Nicolas, V., & Sion, J. (2019). Bikeshare Users on Budget? Trip Chaining Analysis of Bikeshare User Groups in Chicago. doi:10.1177/0361198119838261

Any Questions?

