Development of Stay Time Prediction Model for Leisure Activity: Case Study of Tokyo Metropolitan

05. Hiroshima University
Yu Fujiwara
Taisei Yoshioka
ZAFIRAH BINTI ABDUL GANI
Keishi Fujiwara
Natsuki Nagasaka
Tetsuro Sakai
Seiya Tsukamoto

Background

With COVID-19

- **Lockdown**

1,Stay at home
2,Remote work
3,Online shopping

After COVID-19

- **Recovering**

1,Going shopping
2,Combination of both work style

Disappear former economic activity in central district
because people have no necessity to gather even though pandemic is end

Q. How can we increase stay time to restore people's gathering ?

Data Cleaning

Data PT data : 1,474,144

↓

Trip for leisure activity : 95,262

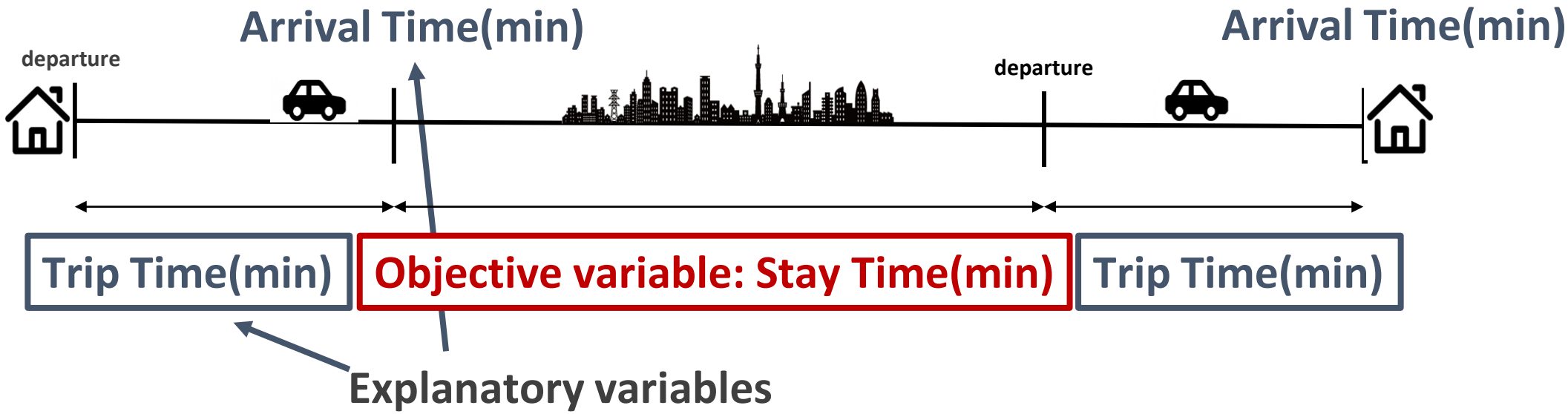
↓

Without commuters : 74,024

↓

Consider trip inside same zone : 59,818

Data Cleaning



Other

1. Man dummy
2. Under 40 dummy
3. Standardized number of retails in zone
4. Standardized number of labors in zone
5. Car or public transport use dummy

From commercial statistics H26

Basic Analysis 1

✔ Gender - Stay

Gender	No. of Person	Average Stay Time (min)
Male	25035	118.5
Female	34783	124.7

✔ Age - Stay

Age	No. of person	Average Stay Time (min)
under 10	2692	106.4
10th	3069	117.7
20th	3477	129.9
30th	7575	120.2
40th	5784	123.2
50th	6726	126.1
60th	16245	122.5
70th	11675	119.6
over 80	2575	134.9

Basic Analysis 2

✔ Time zone - Stay

Time zone of arrival	No. of person	Average Stay Time(min)
	4178	166.9
Afternoon (9am~4pm)	44142	124.0
Night (4pm~mid)	11498	98.5

✔ No. of Retail stores - Stay

No. of retail (per zone)	No. of person	Average Stay Time(min)
~1000	20672	119.0
~2000	24563	123.4
~3000	9521	123.0
~4000	3675	122.9
~5000	0	0
~6000	1387	138.6

Basic Analysis 3

✔ Travel mode- Stay

Mode	No. of person	Average Stay Time(min)
Train	12876	154.95
Bus	2222	131.24
Car	18043	131.36
Bike	616	136.26
Bicycle	8813	116.39
Walk	16819	88.60
Other	78	83.99
NA	351	119.01

※NA: Not selected

Basic Analysis 4

Multi-regression

R-squared: 0.1072
Adjusted R-squared: 0.1070

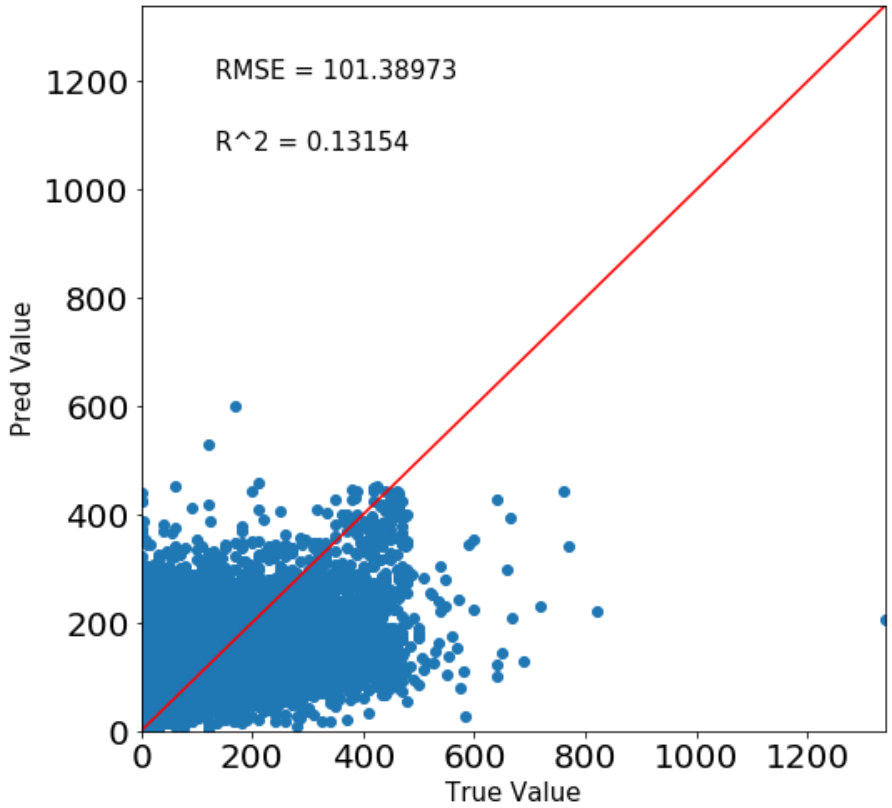
	Estimated	St.d	t	P
Intercept	222.3	2.204	101.2	<2e-16***
Trip Time	-0.084	0.011	-7.714	1.2e-14***
Arrival Time	-0.148	0.003	-58.02	<2e-16***
No.Retails	0.820	7.602	0.108	0.914
No.Labors	2.672	7.702	0.347	0.729
Under 40	16.94	1.194	14.19	<2e-16***
Man Dummy	-3.747	0.955	-3.926	8.7e-5***
Car Dummy	33.18	1.137	29.17	<2e-16***
PT Dummy	60.00	1.280	46.87	<2e-16***

training : test = 8 : 2

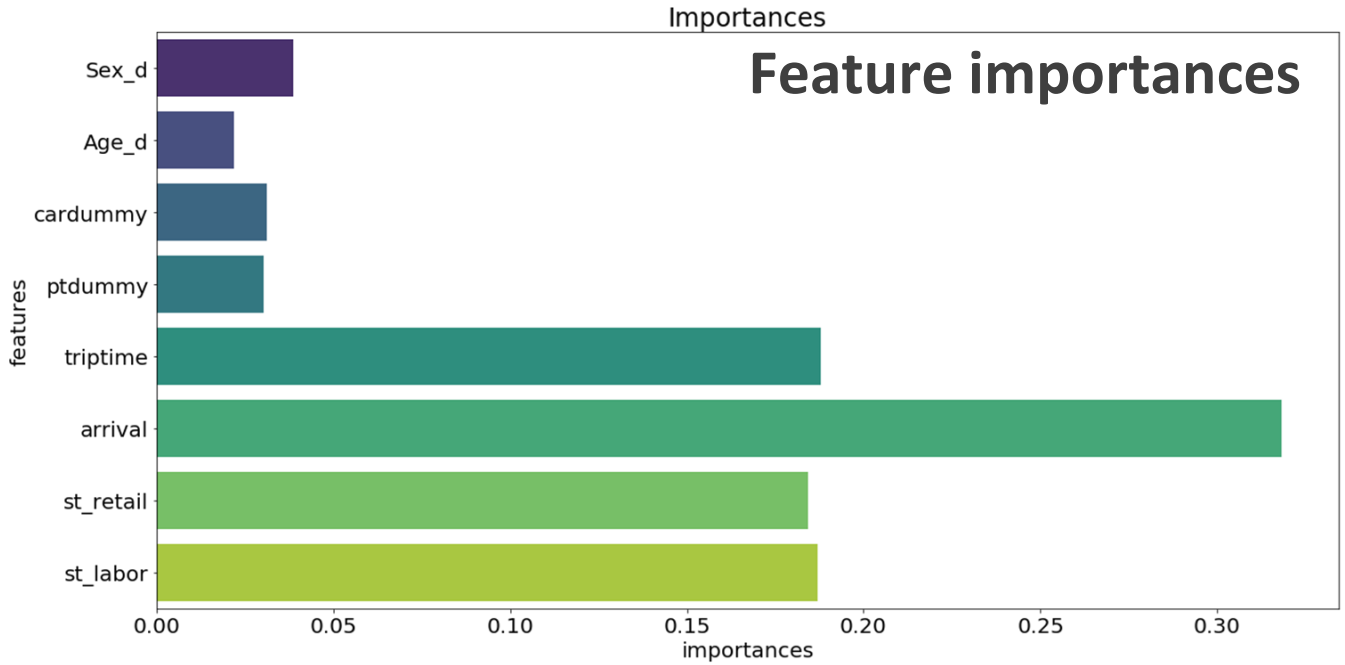
Results of Analysis: Random Forest

training : test = 8 : 2

Compare predicted value with true value



→Moderate good model



→Different from multi-regression result

Simulation

Idea for longer stay time
Short-term Policy Simulation

1. Change of Transport mode
Shift from cars to public transportation

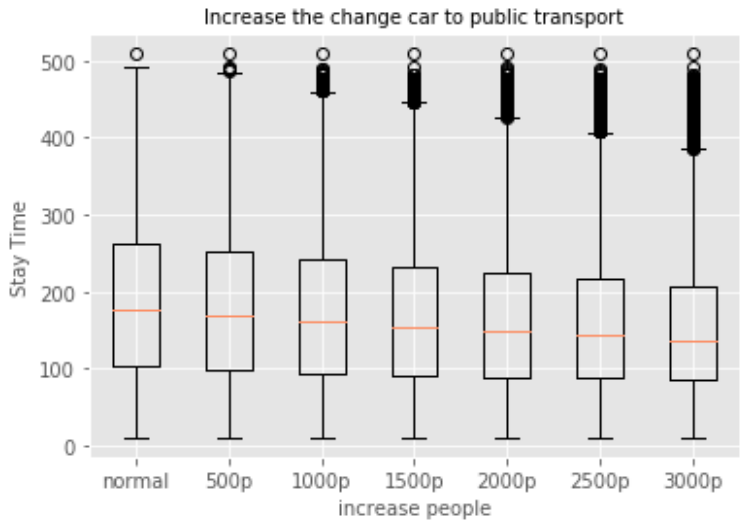
2. Change of Number of retail shops
Increasing shops affect to stay time?

Simulation

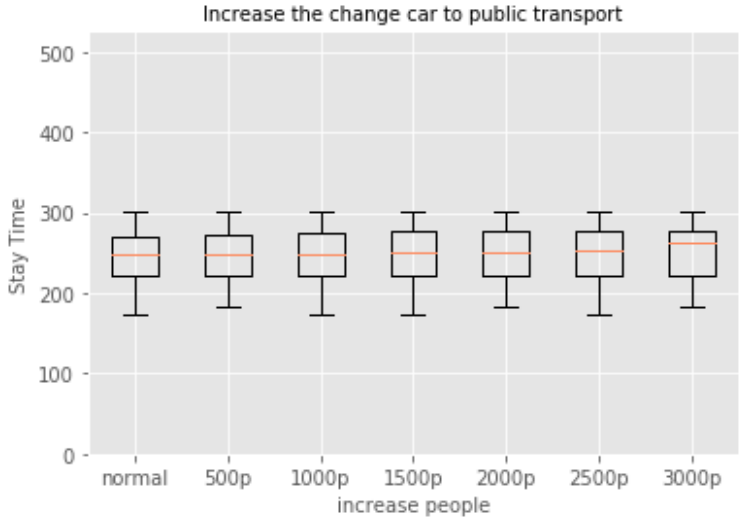
1. Change of Transport mode

Shift from cars to public transportation

Random Forest



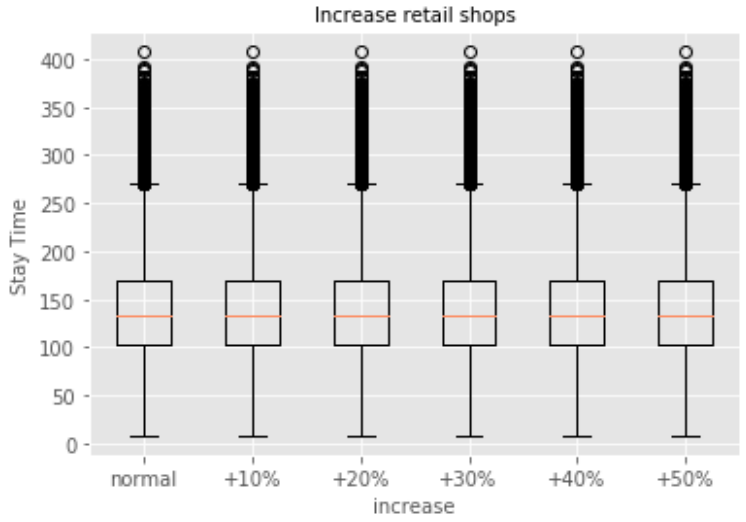
Linear Regression



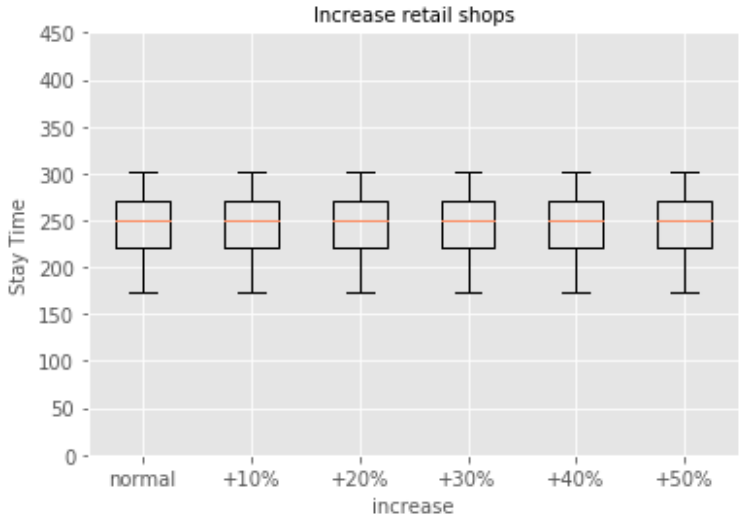
Simulation

2. Change of Number of retail shops
Increasing shops affect to stay time?

Random Forest



Linear Regression



Conclusion

Empirical Findings

- **Even if transport mode shift from cars to public transport, there is no change in stay time for both models**

- **Even if the number of retail shops increase, there is no change in stay time for both**

Methodological Findings

- **The possible range of predicted values for Random Forest are wide**
  →RF don't rely on constant terms, express large part of prediction by explanatory variables

- **The median of stay time for RF (130 minutes) is closer to that for cross tabulation (125 minutes) than that for Linear Reg. (250 minutes), hence RF is better for predicting stay time.**