Final Products for Hackathon

November 17, 2019

1 Visualization for Publibike Dataset

1.1 Interactive visualization of number of rides per station at different hours in one day

1. The average number of rides per station of all dates (1 graph)
2. The average number of rides per station of weekdays vs. weekends (2 graphs)
3. The average number of rides per station of each day during a week (7 graphs)

1.2 Publibike vs. Swisscom Commuters

Data description:

- Publibike: hourly data of number of rides;
- Swisscom Data:hourly data of number of commuters

1.2.1 Percentage of Publibike users among commuters

$$y_i I_{-W} = \beta_1 + \beta_2 x_i I_{-W} + \varepsilon_i, \tag{1}$$

where y_i represents the number of rides per hour, and x_i represents the number of Swisscom commuters per hour.

1.2.2 Substitution effect by buses

$$y_i I_{-W} I_D = \beta_1 + \beta_2 x_i I_{-W} I_D + \varepsilon_i \tag{2}$$
$$y_i I_{-W} I_E = \beta_1 + \beta_2 x_i I_{-W} I_E + \varepsilon_i \tag{3}$$
$$y_i I_{-W} I_N = \beta_1 + \beta_2 x_i I_{-W} I_N + \varepsilon_i \tag{4}$$

where $I_D = 1$ represents day time $[6, 19]$ when there are enough buses, $I_E = 1$ represents evening time $[20, 23]$ when there are only a few buses and $I_N = 1$ represents night time $[0, 5]$ when there is no bus.

Table 1: Model (1)

	Dependent variable:
	# rides per hour
# Swisscom commuters per hour	0.005***
	(0.001)
Constant	24.172***
	(1.070)
Observations	768
R^2	0.072
Adjusted R^2	0.071
Residual Std. Error	18.833 (df = 766)
F Statistic	59.793*** (df = 1; 766)
Note:	*p<0.1; **p<0.05; ***p<0.01

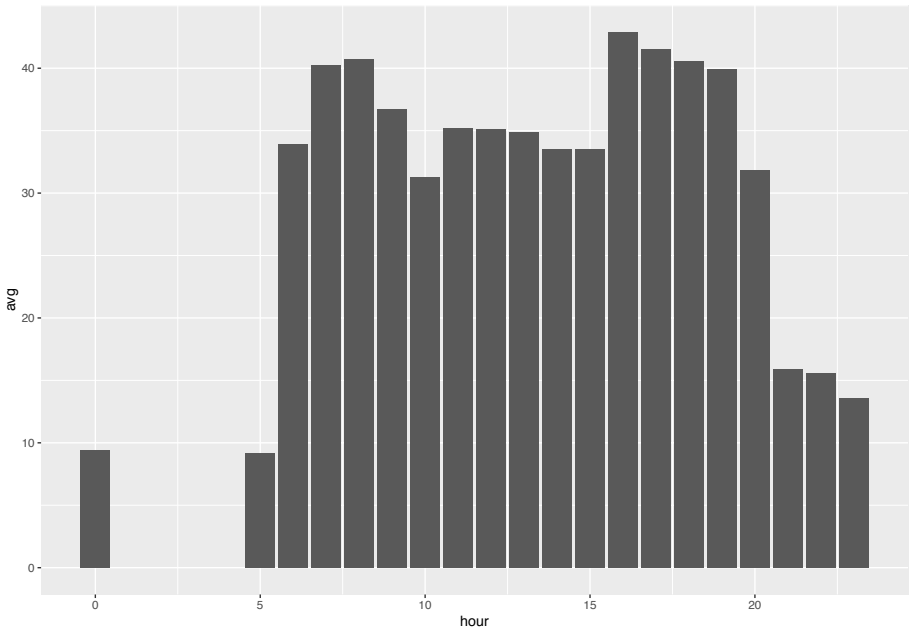

Figure 1: Histogram of number of Publibike rides by hour

Table 2: Model (2)-(4)

	Dependent variable: # rides per hour		
	6am - 8pm	9pm - 0am and 5am	1am - 4am
# Swisscom commuters	0.003***	0.088***	0.113**
	(0.001)	(0.013)	(0.043)
Constant	28.600***	7.525**	3.666*
	(1.657)	(3.729)	(2.150)
Observations	521	174	73
R^2	0.022	0.212	0.089
Adjusted R^2	0.020	0.207	0.076
Residual Std. Error	16.594 (df = 519)	21.521 (df = 172)	8.434 (df = 71)
F Statistic	11.759*** (df = 1; 519)	46.293*** (df = 1; 172)	6.902** (df = 1; 71)
Note:			*p<0.1; **p<0.05; ***p<0.01

1.2.3 Control for number of buses on the road

$$y_i I_{-W} = \beta_1 + \beta_2 x_i I_{-W} + \beta_3 z_i I_{-W} + \varepsilon_i, \tag{5}$$

where z_i is the number of buses that are on the road.

Table 3:

	Dependent variable:		
	rides		
	(1)	(2)	(3)
commutes	−0.0004	0.001	0.077***
	(0.001)	(0.001)	(0.020)
buses	0.401***	0.266**	0.489
	(0.068)	(0.114)	(0.742)
Constant	18.431***	20.116***	3.807
	(1.435)	(3.977)	(6.764)
Observations	768	521	174
R^2	0.112	0.032	0.214
Adjusted R^2	0.110	0.029	0.205
Residual Std. Error	18.437 (df = 765)	16.523 (df = 518)	21.556 (df = 171)
F Statistic	48.327*** (df = 2; 765)	8.680*** (df = 2; 518)	23.288*** (df = 2; 171)
Note:			*p<0.1; **p<0.05; ***p<0.01