

# Toronto Bike-Sharing Demand Prediction

Group 7

Inez Suyono  
Ivan Coelho  
Shohesh Patel  
Xiaodan Du



UNIVERSITY OF TORONTO  
SCHOOL OF CONTINUING STUDIES

# Overview of The Project



**Correlation(s) between the different types of weather conditions and the demand on bike sharing**

# Project Scope

## Location



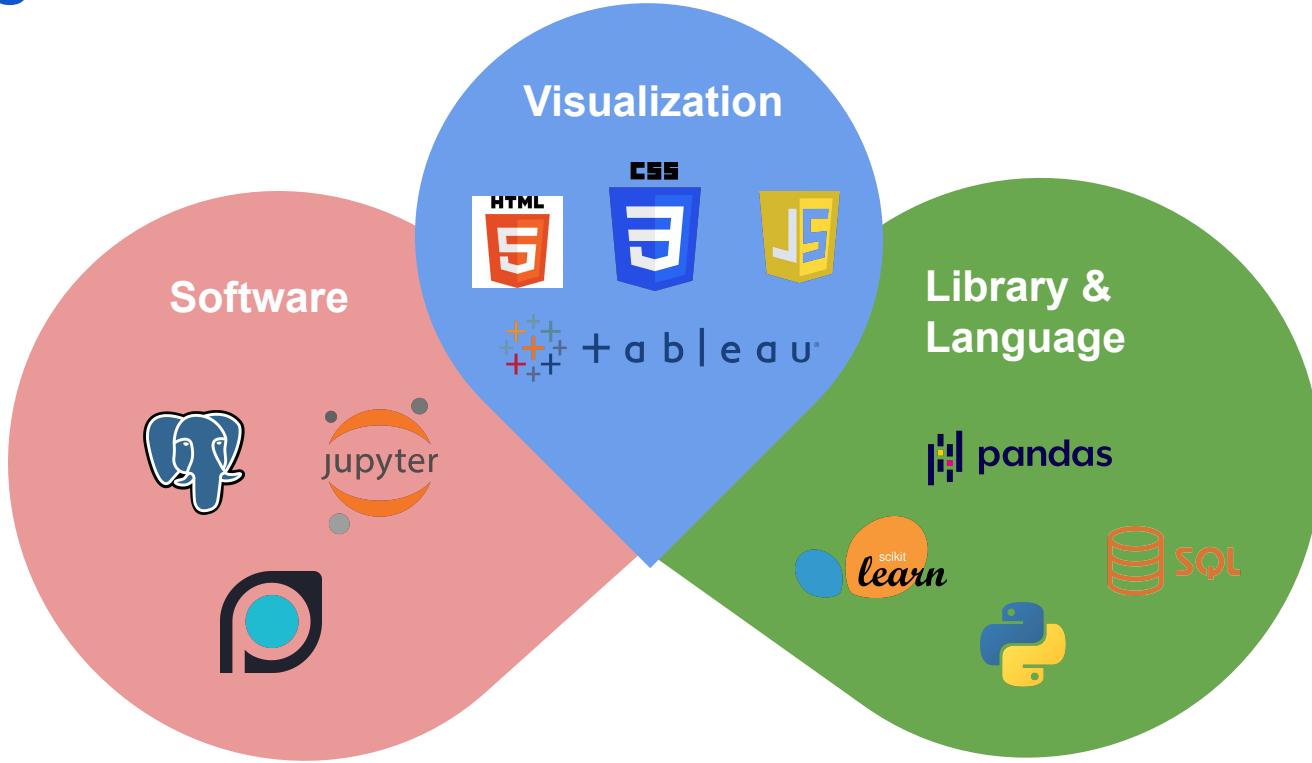
Toronto - Ontario- Canada

## Time



3 Years  
1 Jan 2019 to 31 Dec 2021

# Tools





# Data & Database

# Data Extraction- Bike Data

DOWNLOAD

https://ckan0.cf.opendata.inter.prod-toronto.ca/tr/dataset/bike-share-toronto-ridership-data

Data

Followers 0

Organization

City of Toronto

The January – December 2019, January – December 2020 & January 2021 Bike Share Toronto ridership dataset contains anonymized trip data. It includes:

Trip ID Trip Duration Trip Start Station ID Trip Start time Trip Start Station Location Trip End Station ID Trip End Time Trip End Station Location Bike ID User type

Bike Share Toronto Ridership Data

Data and Resources

- bikeshare-ridership-readme
- bikeshare-ridership-2014-q4-2015-q3
- bikeshare-ridership-2016-q3
- bikeshare-ridership-2016-q4
- bikeshare-ridership-2017

Bike Share Ridership 2018



A	B	C	D	E	F	G	H	I	J	K
1	Trip Id	Trip Dura	Start Statio	Start Time	End Statio	End Time	End Statio	Bike Id	User Type	
2	7452736	625	7162 02/01/202	Hayter St ,	7190 02/01/202	St. George	653	Annual Member		
3	7452737	426	7170 02/01/202	Spadina A	7239 02/01/202	Bloor St W	4682	Annual Member		
4	7452738	817	7475 02/01/202	Jarvis St /	7075 02/01/202	Queens Q	3852	Annual Member		
5	7452739	653	7099 02/01/202	Cherry St ,	7292 02/01/202	Granby St	4546	Annual Member		
6	7452740	189	7176 02/01/202	Bathurst S	7402 02/01/202	Wellington	4867	Annual Member		
7	7452741	942	7398 02/01/202	York St / F	7426 02/01/202	Fleet St / C	2597	Annual Member		
8	7452742	257	7014 02/01/202	Sherbourr	7113 02/01/202	Parliamen	3493	Annual Member		
9	7452743	237	7254 02/01/202	Borden St	7078 02/01/202	College St	111	Annual Member		
10	7452744	622	7002 02/01/202	St. George	7155 02/01/202	Bathurst S	2727	Annual Member		
11	7452745	527	7118 02/01/202	King St W ,	7322 02/01/202	King St W ,	481	Annual Member		
12	7452746	255	7031 02/01/202	Jarvis St /	7042 02/01/202	Sherbourr	3745	Annual Member		
13	7452747	1257	7385 02/01/202	Charles	7350 02/01/202	Broadview	1522	Annual Member		
14	7452748	911	7271 02/01/202	Yonge St /	7299 02/01/202	Mill St / Pe	993	Annual Member		
15	7452749	434	7005 02/01/202	King St W ,	7261 02/01/202	Queens Q	1339	Annual Member		
16	7452750	855	7023 02/01/202	College St	7026 02/01/202	Bay St / St	3560	Annual Member		
17	7452751	311	7003 02/01/202	Madison A	7239 02/01/202	Bloor St W	86	Annual Member		
18	7452752	1207	7012 02/01/202	Elizabeth S	7111 02/01/202	King St W ,	1283	Annual Member		
19	7452753	269	7199 02/01/202	College St	7184 02/01/202	Ossington	1437	Annual Member		
20	7452754	740	7239 02/01/202	Bloor St W	7072 02/01/202	Fleet St / E	4682	Annual Member		
21	7452755	407	7494 02/01/202	Davenpor	7271 02/01/202	Yonge St /	4610	Annual Member		

# Data Cleaning- Bike Data

```
In [2]: In [20]: #change column order  
df_final2=df_final2[['Date','counts_trips','counts_Bike','trip_duration','counts_member_annual','counts_member_casual']]  
df_final2
```

```
Out[20]:
```

	Date	counts_trips	counts_Bike	trip_duration	counts_member_annual	counts_member_casual
0	2019-01-01	1064	674	1121463.0	873	392.0
1	2019-01-02	2290	1135	2838818.0	2206	1029.0
2	2019-01-03	2279	955	1650338.0	2228	1081.0
3	2019-01-04	3290	1381	2269606.0	3130	160.0
4	2019-01-05	2160	1053	1468872.0	1537	223.0
...	...	...	...	...	...	...
1091	2021-12-27	1265	756	955753.0	873	392.0
1092	2021-12-28	3054	1466	2356616.0	2025	1029.0
1093	2021-12-29	3521	1668	2896184.0	2440	1081.0
1094	2021-12-30	4065	1785	2938443.0	2836	1229.0
1095	2021-12-31	4832	1973	4279555.0	3186	1646.0

1096 rows × 6 columns

- Combine All CSV
- Count bike
- Sum Trip Duration
- Count Member Annual
- Count Member Casual

# Data Extraction- Weather Data

SCRAPE  
FROM WEB

Time	Temperature (°F)			Dew Point (°F)			Humidity (%)			Wind Speed (mph)			Pressure (hPa)			Precipitat		
Jan	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Total		
1	41	37.5	30	41	34.1	25	10	80.0	75	23	13.0	5	29.5	29.5	29.4	0.00		
2	30	23.5	16	27	16.5	7	86	74.6	50	18	12.3	6	29.8	29.6	29.5	0.00		
3	27	17.2	10	18	7.3	-2	74	65.0	49	16	8.3	2	30.1	30.0	29.9	0.00		
4	36	30.2	25	28	22.0	16	76	72.1	64	20	13.4	2	30.0	29.9	29.6	0.00		
5	41	36.9	28	36	30.5	18	93	77.1	64	31	18.9	5	29.8	29.3	29.2	0.00		
6	25	22.4	21	16	14.0	12	80	70.1	59	26	21.6	16	29.6	29.5	29.4	0.00		
7	22	16.4	13	7	5.5	4	73	62.6	50	23	15.7	8	30.1	29.9	29.7	0.00		
8	36	22.0	10	28	13.6	-1	93	69.5	51	22	11.2	5	30.3	30.1	29.8	0.00		

A	B	C	D	E	F	G	H	I	J	K	L	M
1	W_Year	W_Month	W_Day	W_Max_EW	Avg_dW	Min_dW	Max_wW	Avg_wW	Min_wW	Max_hW	Avg_hW	Min_humid
2	1.527083	Jan	1	43	30.9	12	23	13.3	2	100	85.5	64
3	1.527083	Jan	2	32	22.9	12	13	7.4	5	100	84	74
4	1.527083	Jan	3	32	28.6	23	23	11.3	0	100	86	69
5	1.527083	Jan	4	36	32.2	30	18	11.7	5	87	80.7	70
6	1.527083	Jan	5	34	31.8	30	20	10.1	2	93	83.7	75
7	1.527083	Jan	6	37	24.6	7	18	12.2	7	93	68.6	46
8	1.527083	Jan	7	36	21.7	9	29	16.5	5	93	73.3	58
9	1.527083	Jan	8	45	38	36	20	7.9	0	100	90.2	76
10	1.527083	Jan	9	36	23.1	16	30	18.3	10	87	68.1	55
11	1.527083	Jan	10	18	11.9	7	26	16.5	7	79	66.8	54
12	1.527083	Jan	11	7	0.5	4	20	9.8	2	73	53.6	36
13	1.527083	Jan	12	12	9.5	5	14	8.1	0	68	58.8	50
14	1.527083	Jan	13	14	9.7	3	17	13.7	9	79	68.5	46
15	1.527083	Jan	14	21	18.2	14	15	7.8	2	86	74.3	64
16	1.527083	Jan	15	27	25.6	21	29	16.9	8	86	81	74
17	1.527083	Jan	16	30	23.3	3	26	20.3	5	93	79.1	43
18	1.527083	Jan	17	21	11.8	3	12	7.8	3	86	69.5	58
19	1.527083	Jan	18	27	22.1	12	22	9.4	0	93	75	51
20	1.527083	Jan	19	12	2.2	0	17	10.9	5	79	72.9	49
21	1.527083	Jan	19	12	2.2	0	17	10.9	5	79	72.9	49

A	B	C	D	E	F	G	H	I	J	K	
1	date	tavg	tmin	tmax	precip	snow	wdir	wspd	wpgrt	pres	tsun
2	2019-01-01	-1.4	-8	5.2	0.6	323.6	24.5	58	1015.1		
3	2019-01-02	-4.9	-9.3	-0.5	3.6	0	68.1	10.1			1023
4	2019-01-03	-1	-2.8	0.8	0.2	30	249	20.2	46		1013.3
5	2019-01-04	3.6	-0.8	8	0	30	244.4	22.9	54		1007.6
6	2019-01-05	1.2	-1.7	4	0	0	302	13.4	33		1006.1
7	2019-01-06	-1.7	-7.3	4	0	0	343.1	23.4	52		1021.3
8	2019-01-07	-1.4	-8.4	5.7	2.8	0	100.3	21.3	63		1023.2
9	2019-01-08	5.3	2.4	8.2	3.8	0	246.8	12.9	50		1004.6
10	2019-01-09	-1.3	-4.9	2.3	0	0	299.2	32.2	71		1007.9
11	2019-01-10	-8	-11.4	-4.6	0	0	317.1	32.4	61		1018.8
12	2019-01-11	-10.7	-13.8	-7.5	0	0	332.3	17.1	41		1020.3
13	2019-01-12	-6.7	9	-4.3	0	0	27.1	10.6			1034.3
14	2019-01-13	-8.7	-11.9	-5.4	0	0	53.3	13.1			1032.4
15	2019-01-14	-4.3	-7.8	-0.8	0	0	284.8	8.7			1026
16	2019-01-15	-2	-4.1	0.1	0	0	238	19.5	54		1019.1
17	2019-01-16	-6.3	-12.2	-3.3	0.4	0	288.2	27.1	63		1019.2
18	2019-01-17	-8.4	-13.2	-3.5	0.6	0	86.6	7.8			1025
19	2019-01-18	-5.4	-10.4	-0.4	1.6	20	297.7	15.8	46		1018.4
20	2019-01-19	-13.6	-16.8	-10.3	5.4	0	18.7	21.6	42		1023.3
21	2019-01-20	-18	-21.8	-14.2	1.6	80	346.7	34.5	67		1015.3



# Data Cleaning- Weather Data

In [256...]

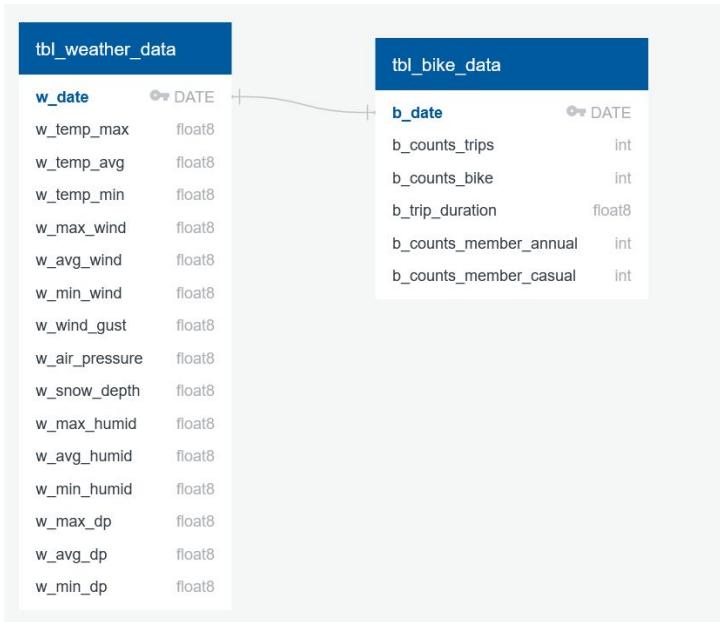
```
import pandas as pd
import glob
import os
# read file from wunderground.com
path = r'D:\bootcamp\week_20_Final_Project\BIke_sharing\weather\Wunderground'
all_files = glob.glob(os.path.join(path, "*.csv"))
```

- Combine CSV for result from scraping
- Change Date format

Out[18]:

	W_date	W_Temp_Max	W_Temp_Avg	W_Temp_Min	W_Max_wind	W_Avg_wind	W_Min_wind	W_Wind_Gust	W_Air_Pressure	W_Snow_Depth	W_Max_humid	W_Avg_humid	\
1	2019-01-01	5.2	-1.4	-8.0	37.0	21.4	3.2	58.0	1015.1	NaN	100.0	85.5	
2	2019-01-02	-0.5	-4.9	-9.3	20.9	11.9	8.0	NaN	1023.0	0.0	100.0	84.0	
3	2019-01-03	0.8	-1.0	-2.8	37.0	18.2	0.0	46.0	1013.3	30.0	100.0	86.0	
4	2019-01-04	8.0	3.6	-0.8	29.0	18.8	8.0	54.0	1007.6	30.0	87.0	80.7	
5	2019-01-05	4.0	1.2	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
6	2019-01-06	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
7	2019-01-07	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
8	2019-01-08	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
9	2019-01-09	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
10	2019-01-10	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
11	2019-01-11	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
12	2019-01-12	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
13	2019-01-13	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
14	2019-01-14	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
15	2019-01-15	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
16	2019-01-16	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
17	2019-01-17	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
18	2019-01-18	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
19	2019-01-19	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
20	2019-01-20	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
21	2019-01-21	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
22	2019-01-22	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
23	2019-01-23	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
24	2019-01-24	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
25	2019-01-25	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
26	2019-01-26	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
27	2019-01-27	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
28	2019-01-28	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
29	2019-01-29	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
30	2019-01-30	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
31	2019-01-31	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
32	2019-02-01	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
33	2019-02-02	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
34	2019-02-03	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
35	2019-02-04	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
36	2019-02-05	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
37	2019-02-06	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
38	2019-02-07	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
39	2019-02-08	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
40	2019-02-09	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
41	2019-02-10	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
42	2019-02-11	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
43	2019-02-12	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
44	2019-02-13	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
45	2019-02-14	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
46	2019-02-15	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
47	2019-02-16	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
48	2019-02-17	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
49	2019-02-18	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
50	2019-02-19	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
51	2019-02-20	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
52	2019-02-21	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
53	2019-02-22	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
54	2019-02-23	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
55	2019-02-24	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
56	2019-02-25	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
57	2019-02-26	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
58	2019-02-27	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
59	2019-02-28	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
60	2019-02-29	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
61	2019-03-01	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
62	2019-03-02	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
63	2019-03-03	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
64	2019-03-04	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
65	2019-03-05	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
66	2019-03-06	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
67	2019-03-07	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
68	2019-03-08	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
69	2019-03-09	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
70	2019-03-10	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
71	2019-03-11	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
72	2019-03-12	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
73	2019-03-13	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
74	2019-03-14	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
75	2019-03-15	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
76	2019-03-16	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
77	2019-03-17	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
78	2019-03-18	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
79	2019-03-19	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
80	2019-03-20	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
81	2019-03-21	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
82	2019-03-22	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
83	2019-03-23	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
84	2019-03-24	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
85	2019-03-25	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
86	2019-03-26	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
87	2019-03-27	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
88	2019-03-28	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
89	2019-03-29	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
90	2019-03-30	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
91	2019-03-31	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
92	2019-04-01	1.2	-0.5	-1.7	32.2	16.3	3.2	33.0	1006.1	0.0	93.0	83.7	
93	2019-04-02	1.2	-0.5	-1.7	32.2	16.3	3						

# Entity Relationship Diagram



# Connection to Machine Learning

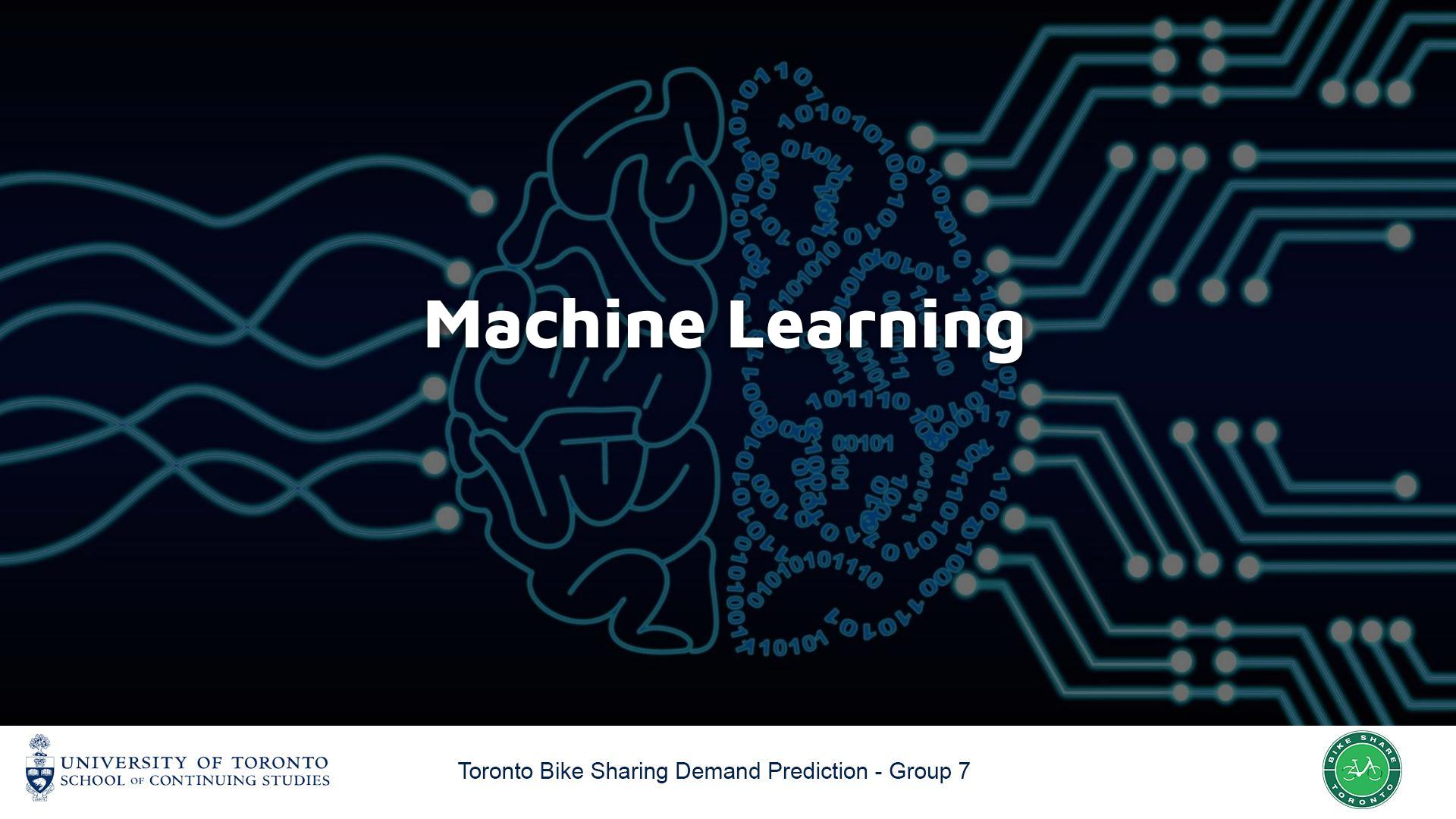
```
In [1]: # import pandas and create database engine
import pandas as pd
from sqlalchemy import create_engine
import psycopg2
from config import db_password
```

```
db_string = f"postgresql://postgres:{db_password}@localhost/Bike_sharing"
```

```
In [2]: # Create the database engine
engine = create_engine(db_string)
```

```
In [7]: # Join table tbl_bike_data and tbl_weather_data
Merged_df = pd.read_sql_query('select * from tbl_bike_data a join tbl_weather_data b on a.b_date=b.w_date', con=engine)
Merged_df
```

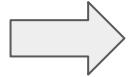
	b_date	b_counts_trips	b_counts_bike	b_trip_duration	b_counts_member_annual	b_counts_member_casual	w_date	w_temp_max	w_temp_avg	
0	2019-01-01	1064	674	1121463.0		897	167	2019-01-01	5.2	-1.4
1	2019-01-02	2290	1135	2838818.0		2207	83	2019-01-02	-0.5	-4.9
2	2019-01-03	2279	955	1650338.0		2228	51	2019-01-03	0.8	-1.0
3	2019-01-04	3290	1381	2269606.0		3130	160	2019-01-04	8.0	3.6



# Machine Learning

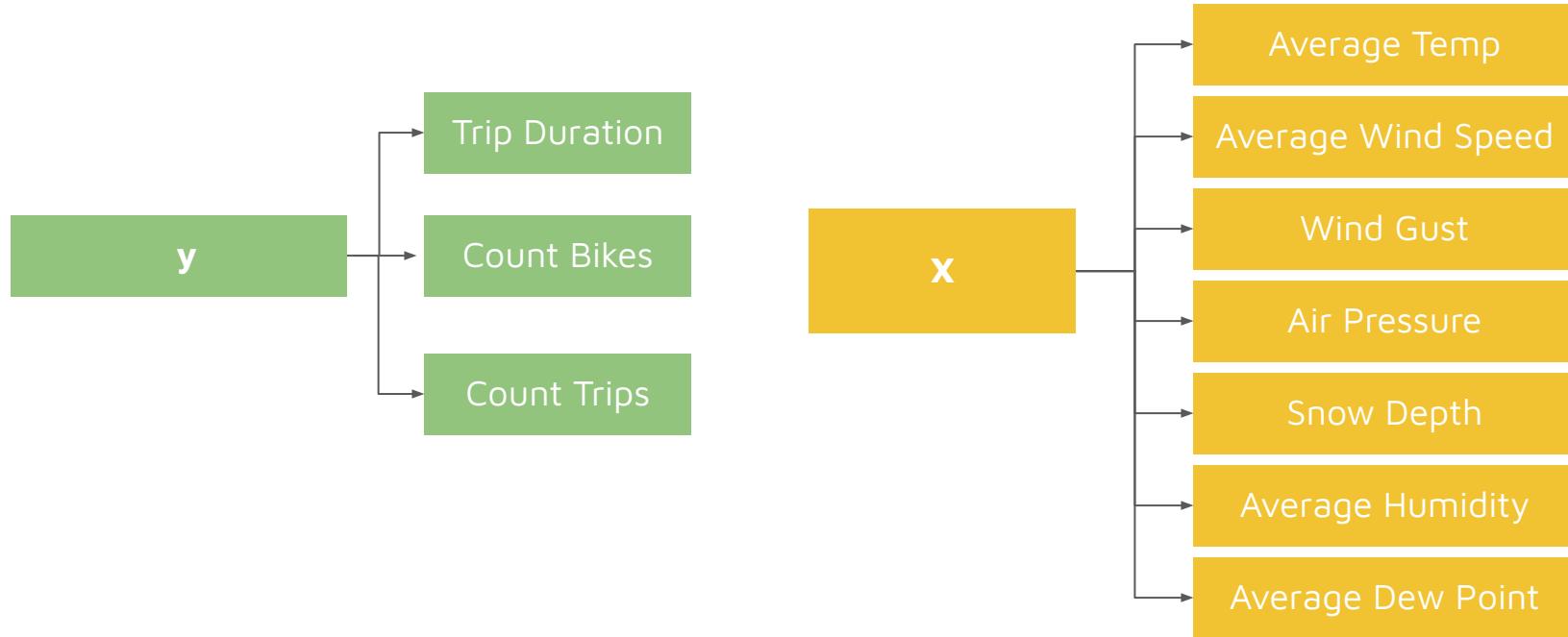
# Machine Learning pipeline

Data extraction &  
preprocessing

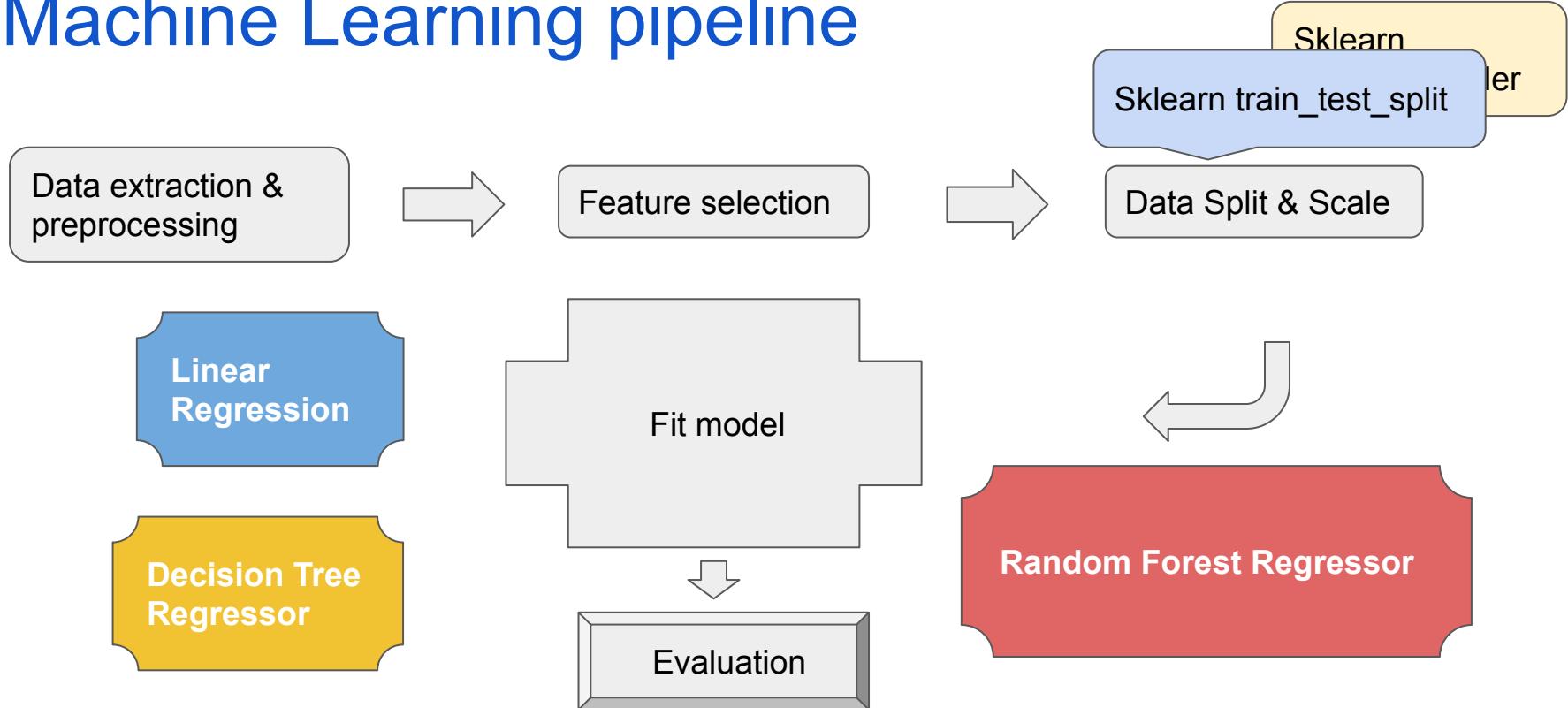


Feature selection

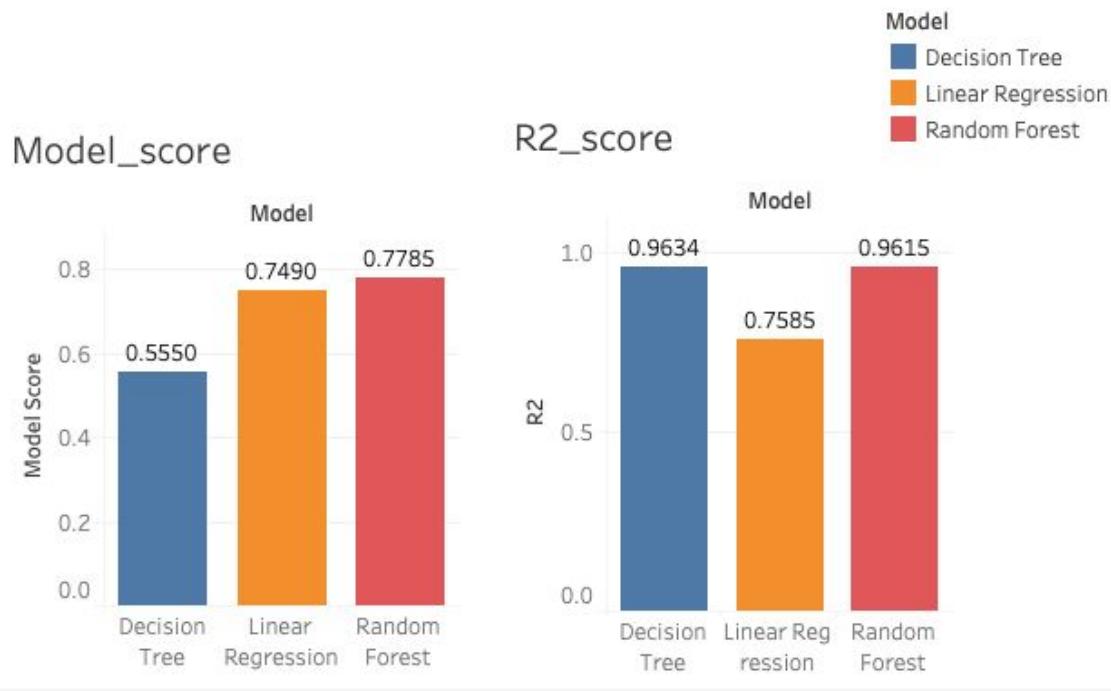
# Features selection



# Machine Learning pipeline



# Machine Learning - Evaluation Results



# Machine Learning - Evaluation Results

**model=74.90%**

MAE: 433.93  
MSE: 284,447  
RMSE: 533.33  
R2: 75.85%

Linear Regression

**model=55.50%**

MAE: 160.38  
MSE: 42,654  
RMSE: 206.53  
R2: 96.34%

Decision Tree Regressor

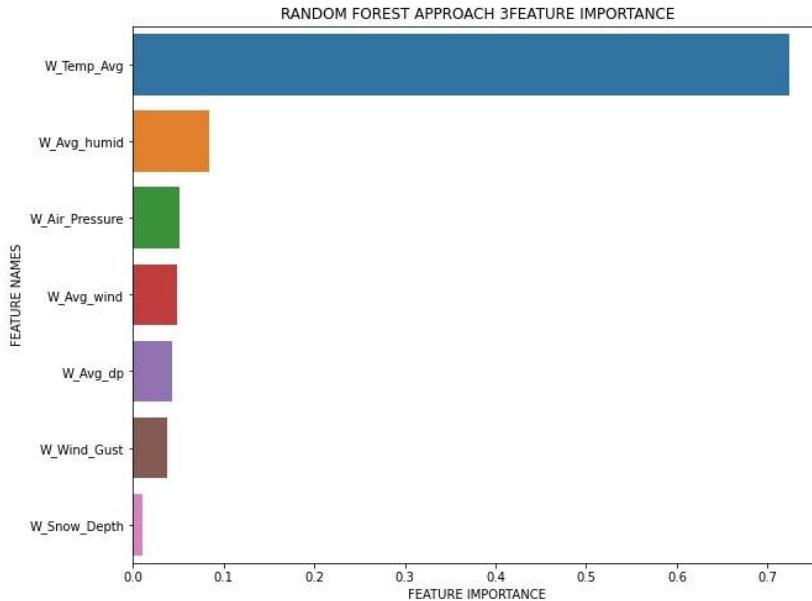
**model=77.85%**

MAE: 160.38  
MSE: 42,654  
RMSE: 206.53  
R2: 96.15%

1

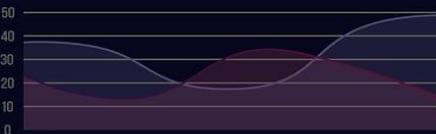
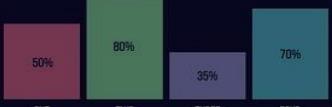
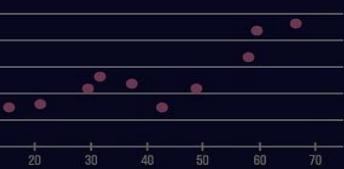
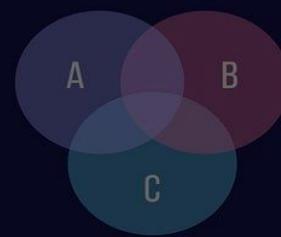
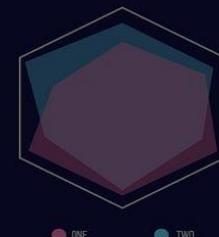
Random Forest Regressor

# Machine Learning - Summary



# Future analysis

- Improve accuracy :
  - BaggingRegressor
  - ExtraTreesRegressor
  - AdaBoostRegressor
- Adding more feature variables
- Prediction Equation



UNIVERSITY OF TORONTO  
SCHOOL OF CONTINUING STUDIES

Toronto Bike Sharing Demand Prediction - Group 7



# Website



## Toronto Bike Sharing Demand Prediction

Overview   Data & Database   Machine Learning   About Us

[CLICK HERE](#)



# THANK YOU!

