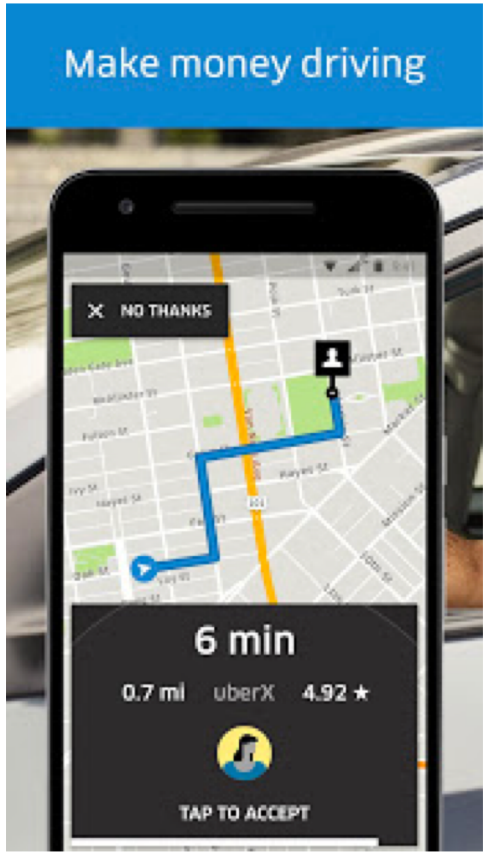
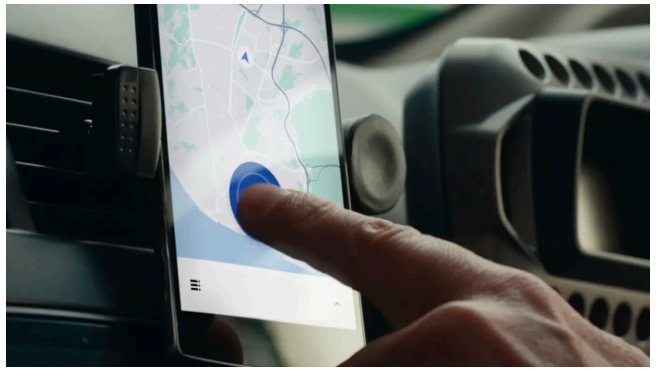
Analysis of Uber rides in different weather conditions

Team: **Ignacio Sanz, Jimca Muñoz, Micah Wu and Raul Hennings**
Data Analytics Bootcamp
UC Berkeley Extension
November 28, 2018.

The Uber app

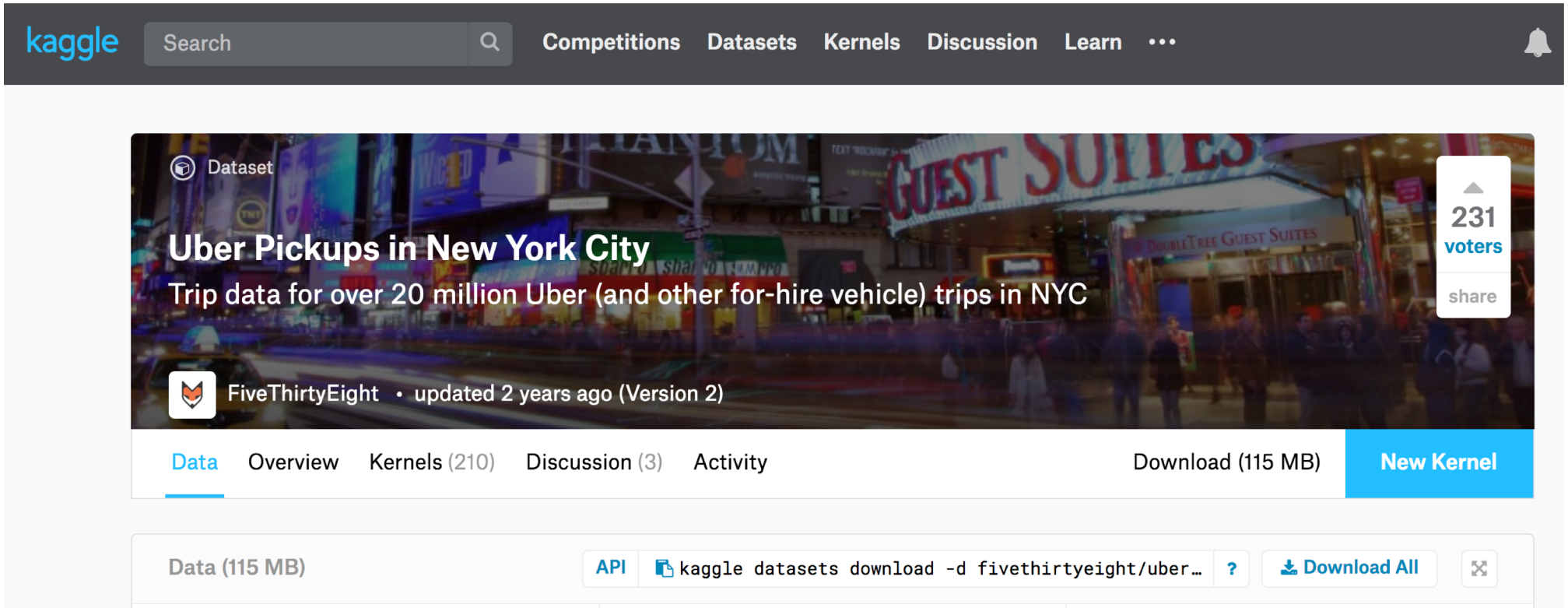
- Uber is a location based app that provides taxi service. It uses on-demand hiring of private drivers.
- 15 million Uber trips are completed each day
- More than 5 billion trips have been completed worldwide
- We looked at Uber rides data under different weather conditions





Our dataset: Uber Pickups in NY City

- On the Kaggle site: Uber Pickups in New York City.
- Raw data on pickups from a non-Uber FHV company.
- Couldn’t use Weather API because historical data is not free.



Performed 3 different analysis

1. Rain vs. No Rain
2. Rides on Cloudy (Rainy) days
3. Statistical analysis with a test hypothesis

1. Rain and no-rain

To start off…

We found the data of NYC Uber Rides for every hour of everyday starting from 2014 July 1st to 2014 October 1st.

Then we found data of the weather in NYC of every hour of every day from 2014 July 1st to 2014 October 1st

Together, we merged all of this information into one dataframe.

Then adding up all the hours = 2209 Hours

And total # of rides = 91,712

Average # of rides per hour = 41.5174287

	Temperature F	Number of Rides	Weather Description
2014-07-01 00:00:00	82.238	20	sky is clear
2014-07-01 01:00:00	80.222	33	broken clouds
2014-07-01 02:00:00	77.432	21	few clouds
2014-07-01 03:00:00	76.298	15	few clouds
2014-07-01 04:00:00	75.596	19	few clouds
2014-07-01 05:00:00	75.182	21	few clouds
2014-07-01 06:00:00	74.948	31	few clouds
2014-07-01 07:00:00	74.192	28	few clouds
2014-07-01 08:00:00	73.076	62	scattered clouds
2014-07-01 09:00:00	72.788	62	scattered clouds
2014-07-01 10:00:00	71.420	47	sky is clear
2014-07-01 11:00:00	72.068	27	sky is clear
2014-07-01 12:00:00	73.022	45	sky is clear
2014-07-01 13:00:00	74.516	33	broken clouds

Uber Rides when is Raining (Water Involved)

	Number of Hours Rained	Total Number of Rides
drizzle	1	135
fog	56	2394
haze	41	1923
heavy intensity drizzle	1	47
heavy intensity rain	12	412
light intensity drizzle	10	509
light rain	160	6862
mist	413	16847
moderate rain	36	1454
proximity thunderstorm with rain	1	14
thunderstorm with heavy rain	2	98
thunderstorm with light rain	3	144
thunderstorm with rain	2	59

Total # of rides = 30890

/

Total # of hours it Rained = 738

Average # of rides per hour
(During Rain)
= 41.86720867208672

Uber Rides when there is no Water/Rain at all

	Number of Hours with no Rain	Total Number of Rides
broken clouds	530	21964
dust	11	496
few clouds	172	6524
overcast clouds	163	7000
proximity thunderstorm	15	827
scattered clouds	439	17764
sky is clear	130	5818
smoke	3	134
thunderstorm	8	287

Total # of rides = 60814

/

Total # of hours with no rain = 1471

Average # of rides per hour (No Rain)
= 41.34194425560843

Rain vs No Rain

- Since the average remained the same per hour, We can conclude that whether its rains or not, it does not really have an impact on the number of Uber rides in NYC during 2014 July to October.

2. Rides on cloudy conditions

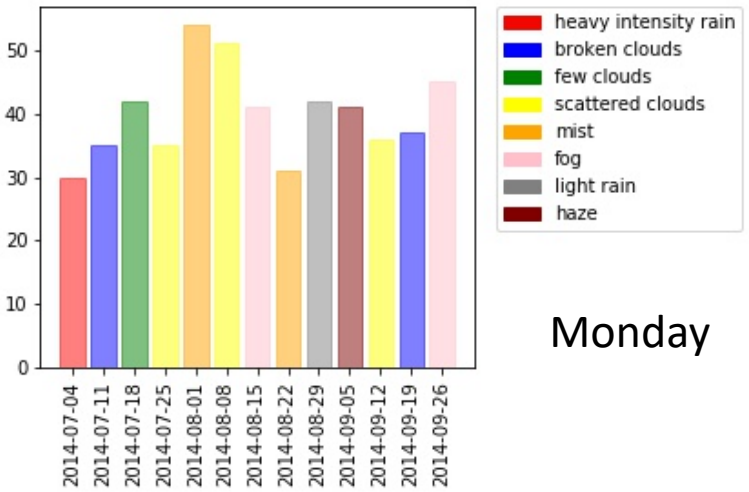
- 2 scenarios
  - Commute time
  - Night club closing time

- 2 questions
  - Did cloudy conditions and rain have any impact on the number of Uber rides taken on specific day?
  - Can we recommend Uber to increase the number of drivers in NYC during peak times based on weather conditions?

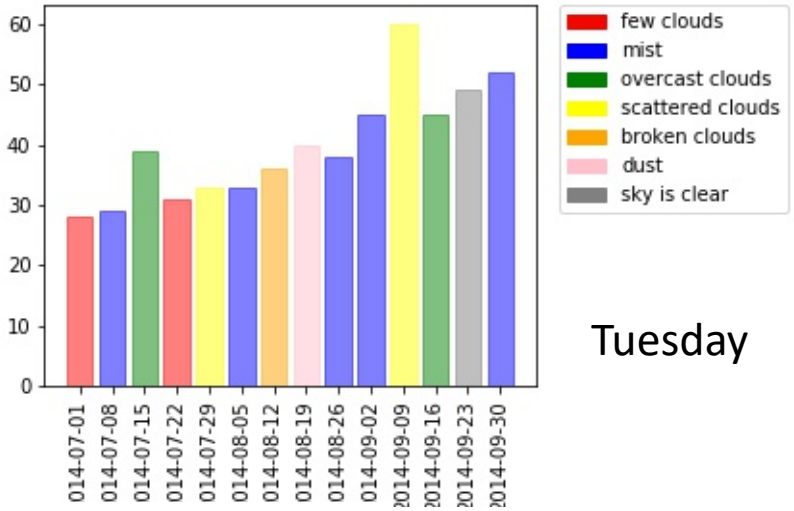
1st scenario – peak hours

- Morning commute
  - 7am – 9am Mon-Fri
- Lunch break
  - 12pm Mon – Fri
- Afternoon commute
  - 4pm – 6pm Mon-Fri

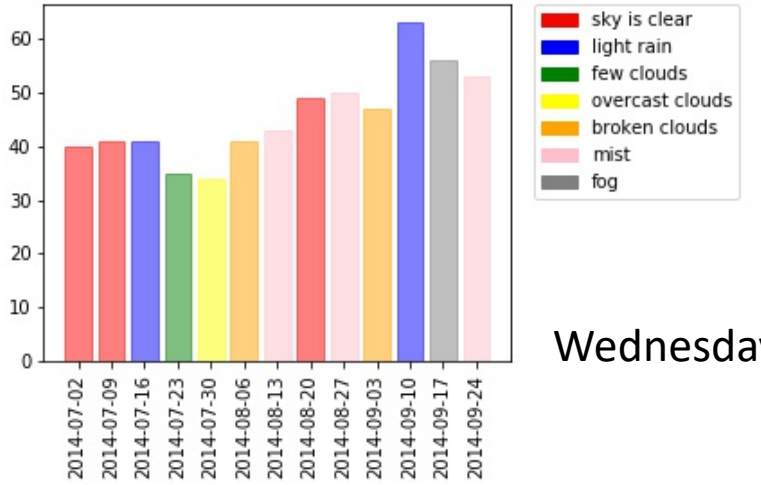
07:00:00



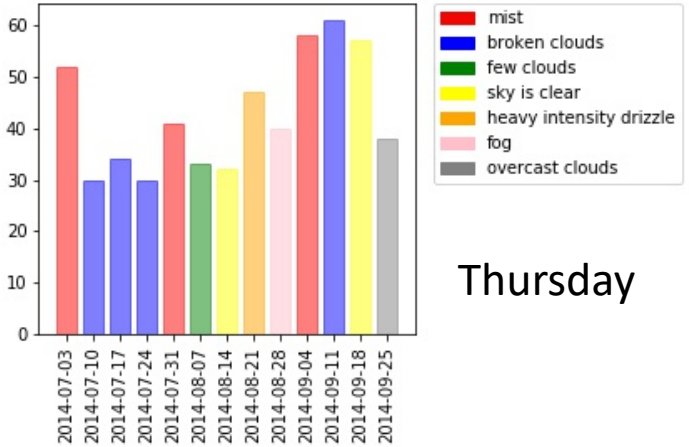
Monday



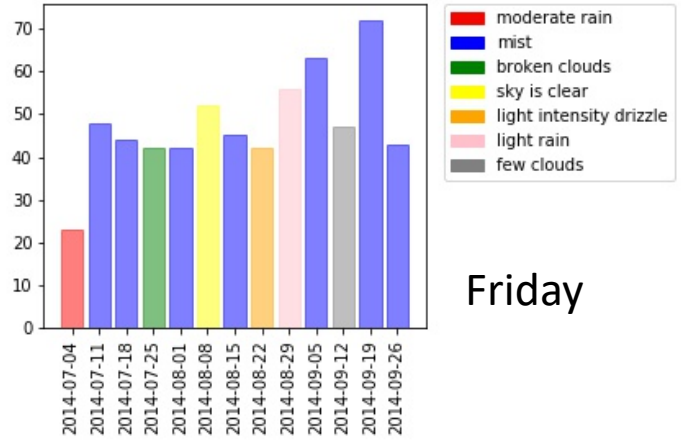
Tuesday



Wednesday

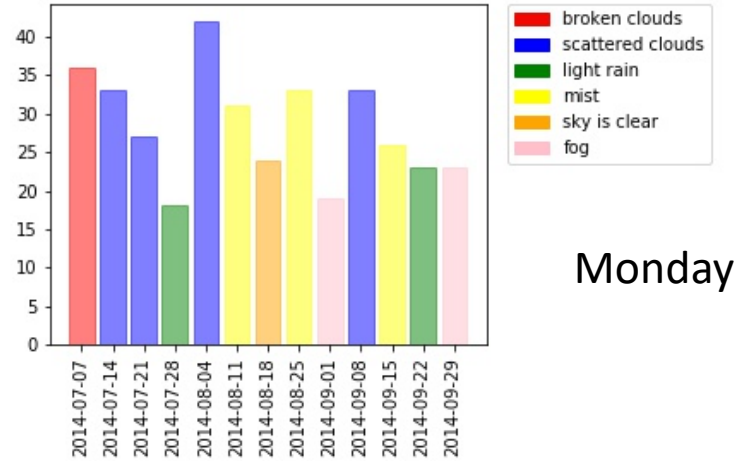


Thursday

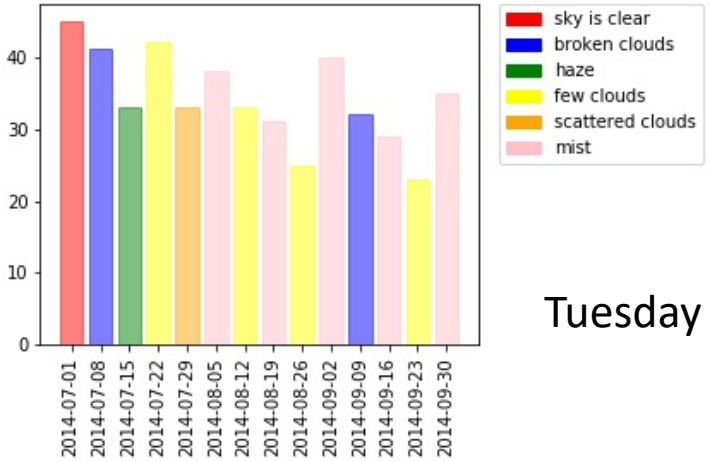


Friday

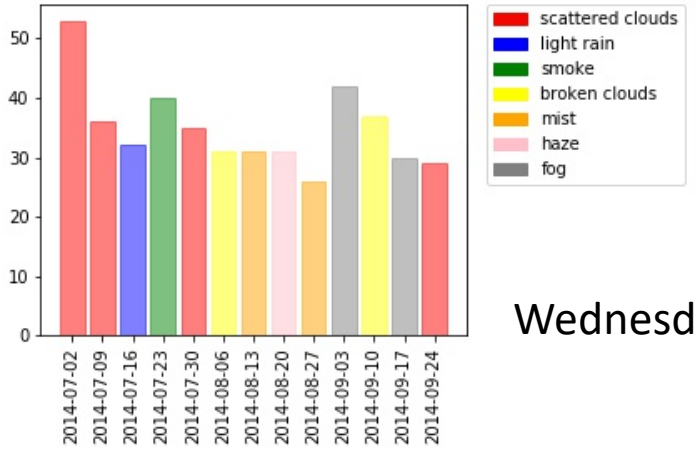
12:00:00



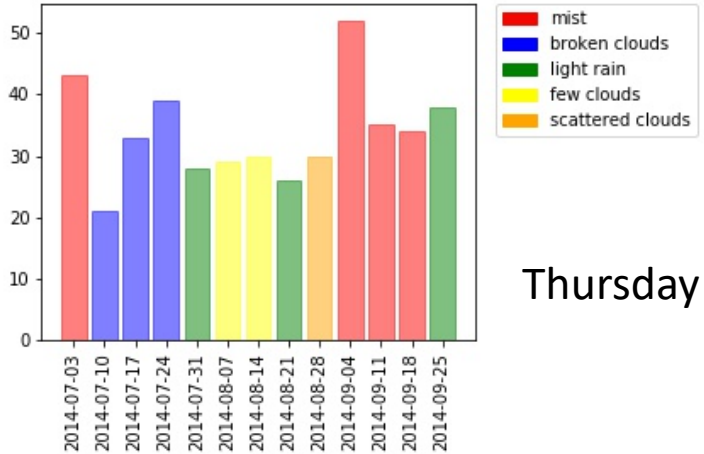
Monday



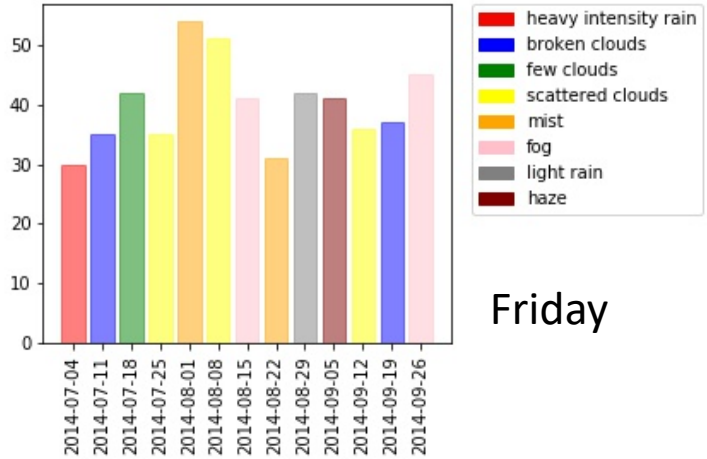
Tuesday



Wednesday

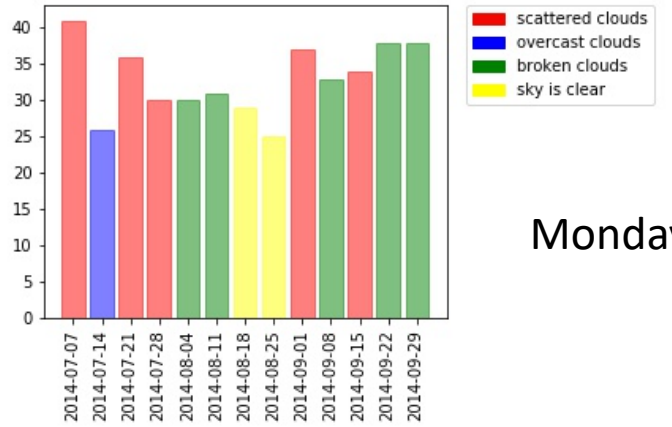


Thursday

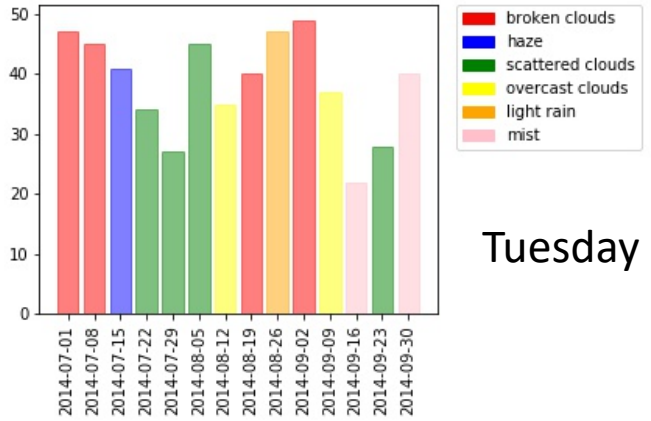


Friday

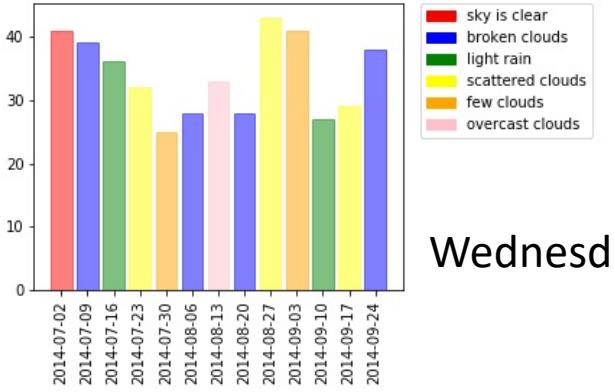
16:00:00



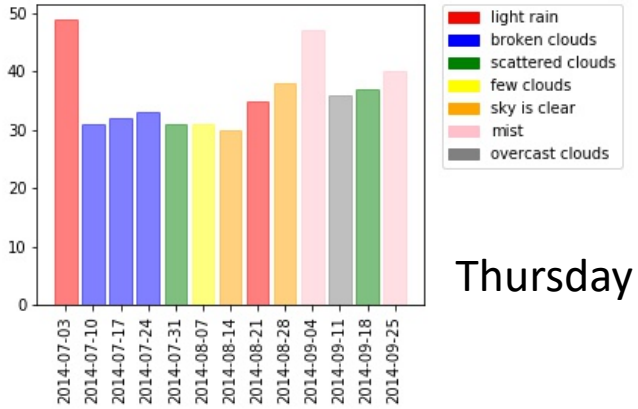
Monday



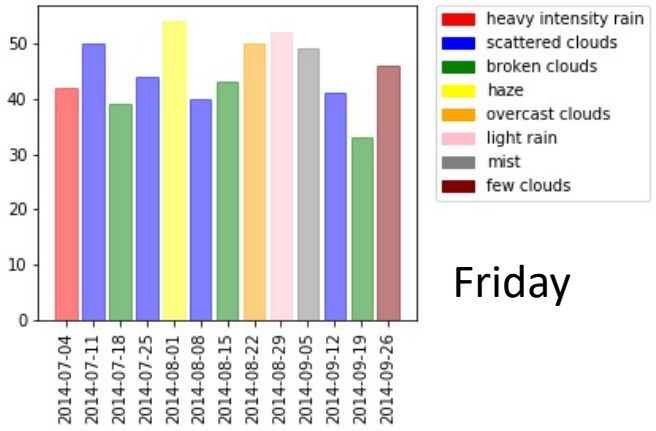
Tuesday



Wednesday



Thursday



Friday

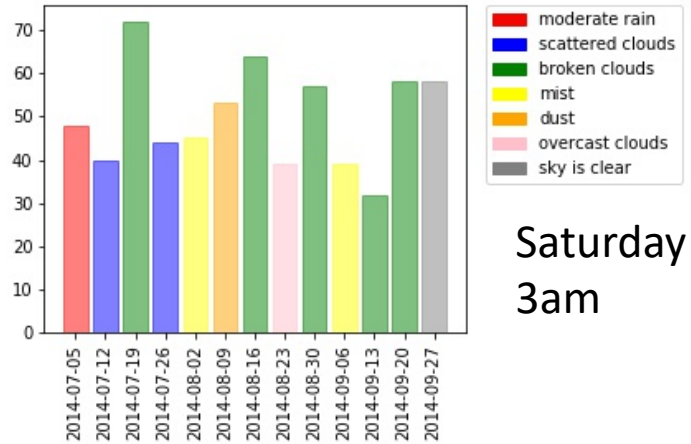
Weekday	Date	Time	Number of Rides	Weather Description	Temperature F
Monday	9/1/14	7:00:00	23	mist	73.544
Monday	8/18/14	7:00:00	28	sky is clear	59.7164
Monday	8/4/14	7:00:00	32	mist	68.612
Monday	7/21/14	7:00:00	33	sky is clear	63.014
Monday	8/11/14	7:00:00	34	mist	66.02
Monday	7/28/14	7:00:00	35	thunderstorm with rain	72.086
Monday	8/25/14	7:00:00	38	mist	60.224
Monday	7/7/14	7:00:00	41	broken clouds	73.292
Monday	7/14/14	7:00:00	47	broken clouds	72.932
Monday	9/22/14	7:00:00	57	mist	64.4
Monday	9/29/14	7:00:00	63	mist	62.816
Monday	9/8/14	7:00:00	67	broken clouds	60.188
Monday	9/15/14	7:00:00	67	sky is clear	50.54
Weekday	**Date**	**Time**	**Number of Rides**	**Weather Description**	**Temperature F**
Monday	9/1/14	8:00:00	20	fog	73.544
Monday	8/25/14	8:00:00	46	mist	59.288
Monday	8/18/14	8:00:00	56	sky is clear	60.19572545
Monday	7/21/14	8:00:00	63	mist	62.78
Monday	8/4/14	8:00:00	65	mist	68.846
Monday	7/14/14	8:00:00	68	light rain	73.67
Monday	9/22/14	8:00:00	68	mist	63.158
Monday	7/7/14	8:00:00	71	overcast clouds	72.932
Monday	7/28/14	8:00:00	72	moderate rain	70.34
Monday	8/11/14	8:00:00	74	mist	65.03
Monday	9/8/14	8:00:00	74	mist	59.666
Monday	9/29/14	8:00:00	75	mist	62.006
Monday	9/15/14	8:00:00	81	mist	49.964
Weekday	**Date**	**Time**	**Number of Rides**	**Weather Description**	**Temperature F**
Monday	9/1/14	9:00:00	18	mist	73.238
Monday	8/25/14	9:00:00	45	mist	58.766
Monday	7/7/14	9:00:00	51	overcast clouds	72.212
Monday	7/14/14	9:00:00	52	broken clouds	72.572
Monday	9/22/14	9:00:00	52	scattered clouds	62.186
Monday	7/28/14	9:00:00	58	light rain	69.314
Monday	7/21/14	9:00:00	59	mist	62.204
Monday	9/29/14	9:00:00	59	overcast clouds	61.718
Monday	8/11/14	9:00:00	60	moderate rain	64.652
Monday	9/15/14	9:00:00	63	fog	49.514
Monday	8/4/14	9:00:00	66	mist	68.504
Monday	9/8/14	9:00:00	69	broken clouds	58.694
Monday	8/18/14	9:00:00	80	mist	61.178

Monday morning at glance

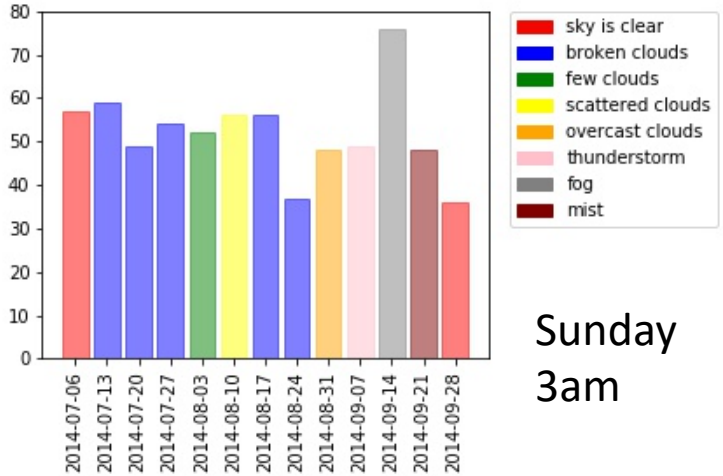
2nd scenario – night club closing

- Night club closing
  - 1am – 4am Sat-Sun

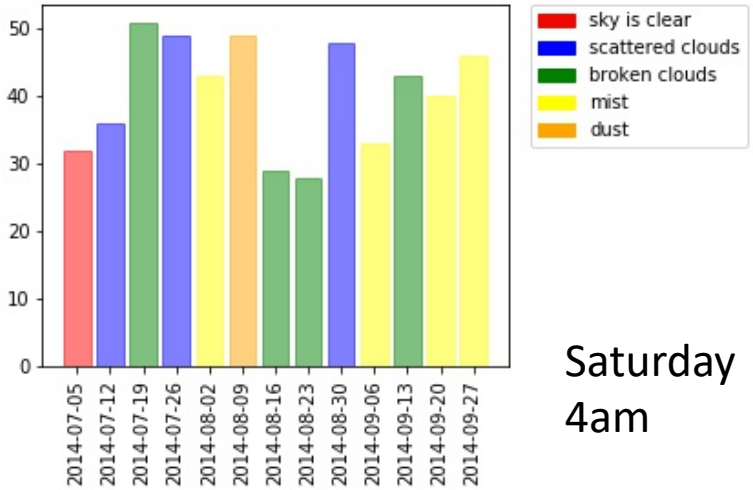
Night club closing



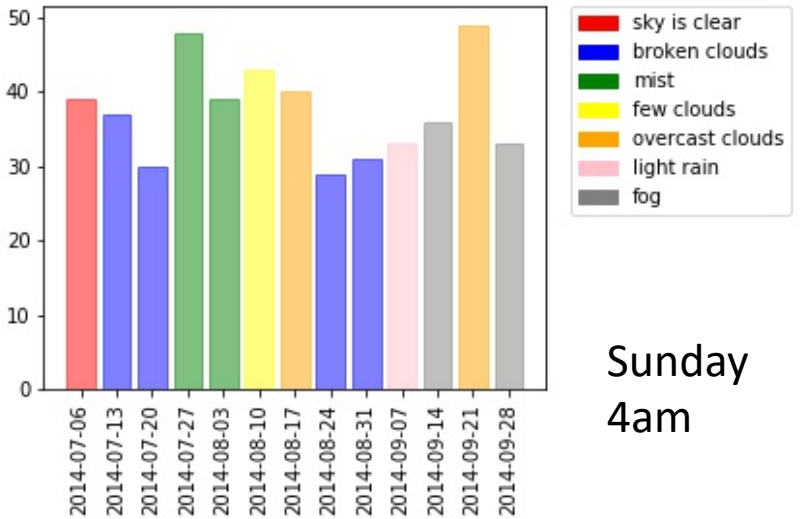
Saturday
3am



Sunday
3am



Saturday
4am



Sunday
4am

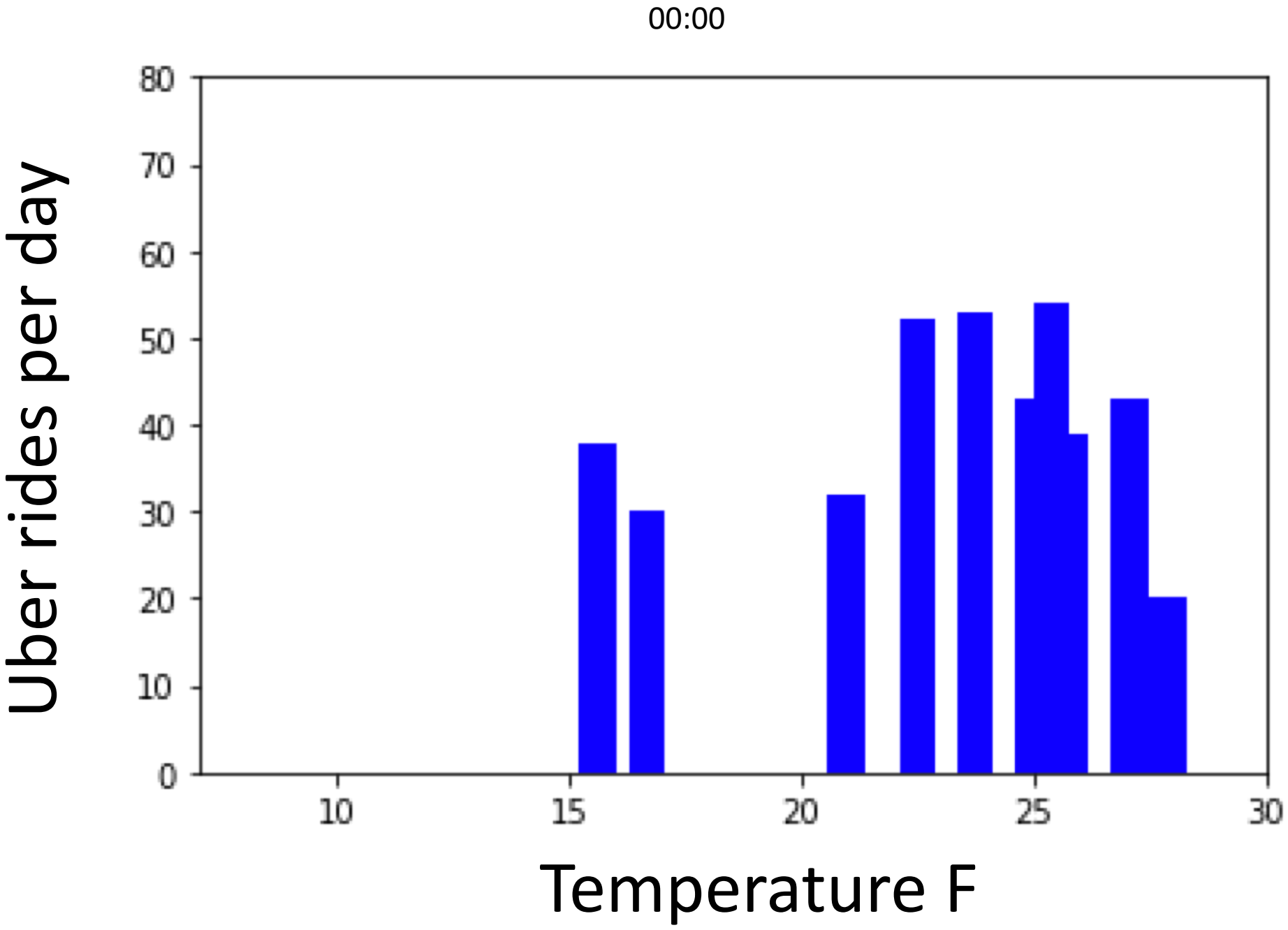
- 2 answers
  - Did cloudy conditions and rain have any impact on the number of Uber rides taken on specific day? - NO
  - Can we recommend Uber to increase the number of drivers in NYC during peak times based on weather conditions? – NO

3. Statistical analysis with a test hypothesis

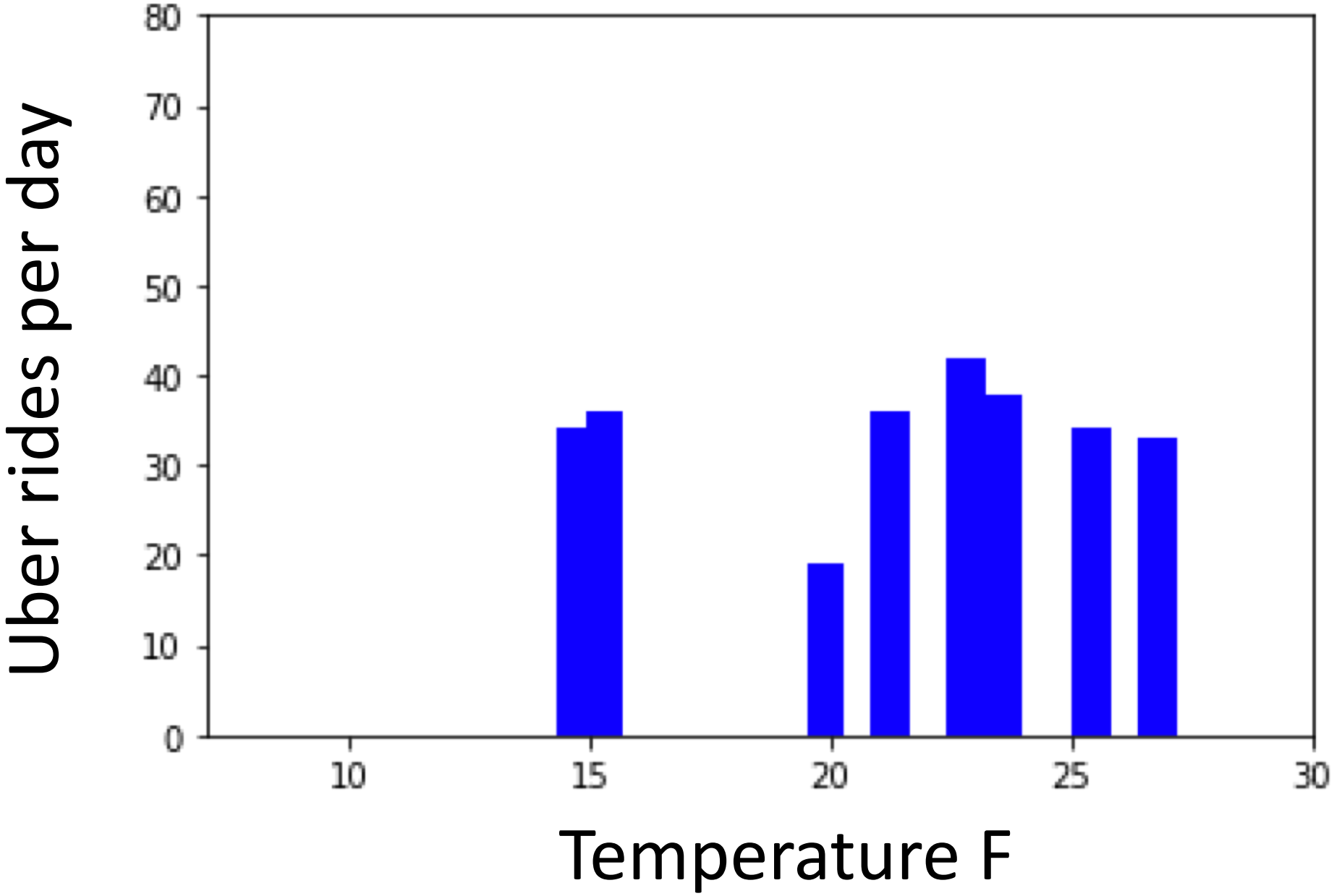
Does the weather temperature affect the number of Uber rides in the city of New York?

- **Hypothesis** - If the weather is related to the number of Uber rides, then a consistent and numeric correlation will show how the number of rides increases or decreases depending on the weather temperate over a given period of time (July to Sep) in a given location (NYC).
- **Null Hypothesis** - If the weather is not related to the number of Uber rides, then there will be no correlation and the increase or decrease (if any) of Uber rides will be independent to the weather temperature.

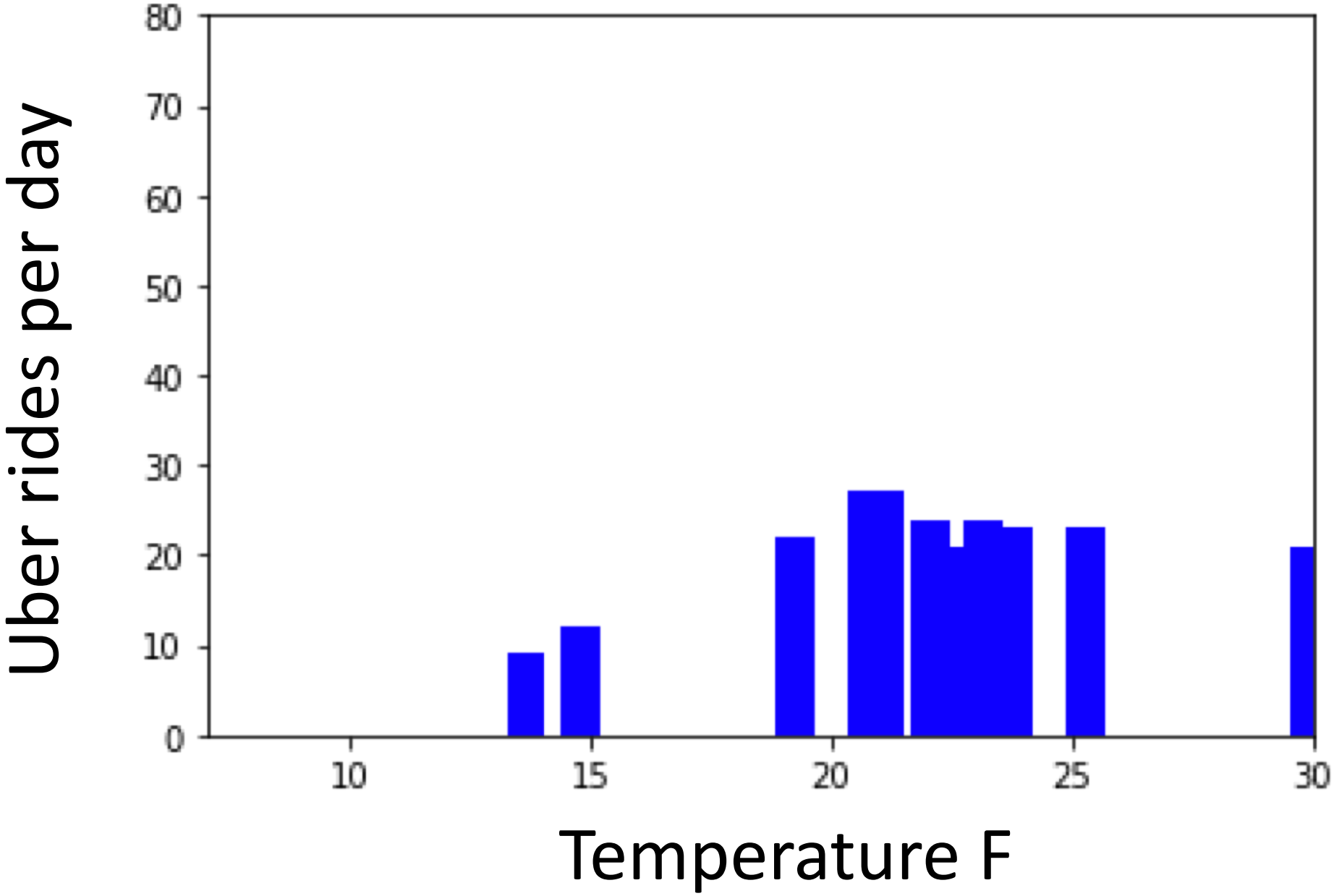
TUESDAY

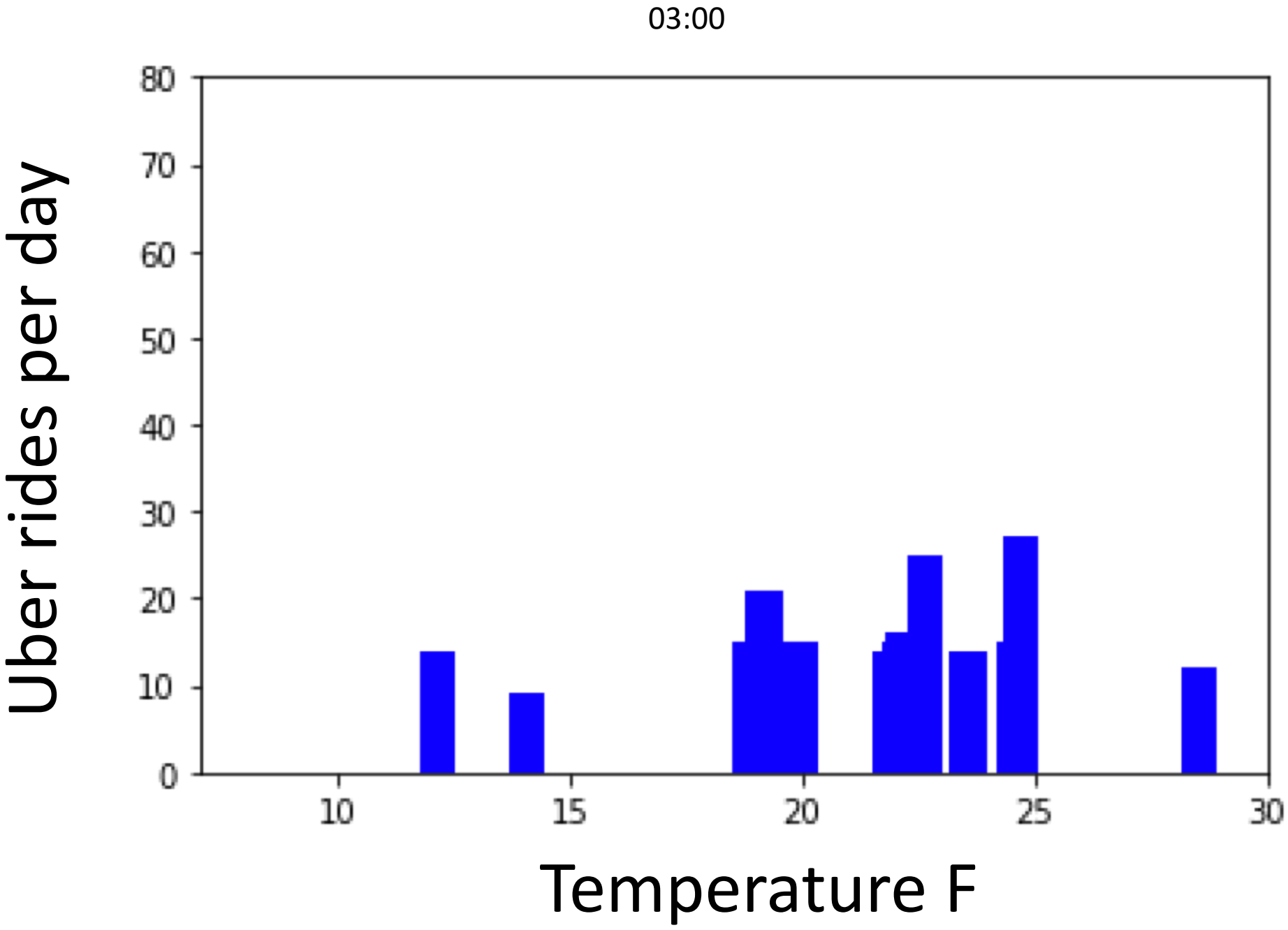

00:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

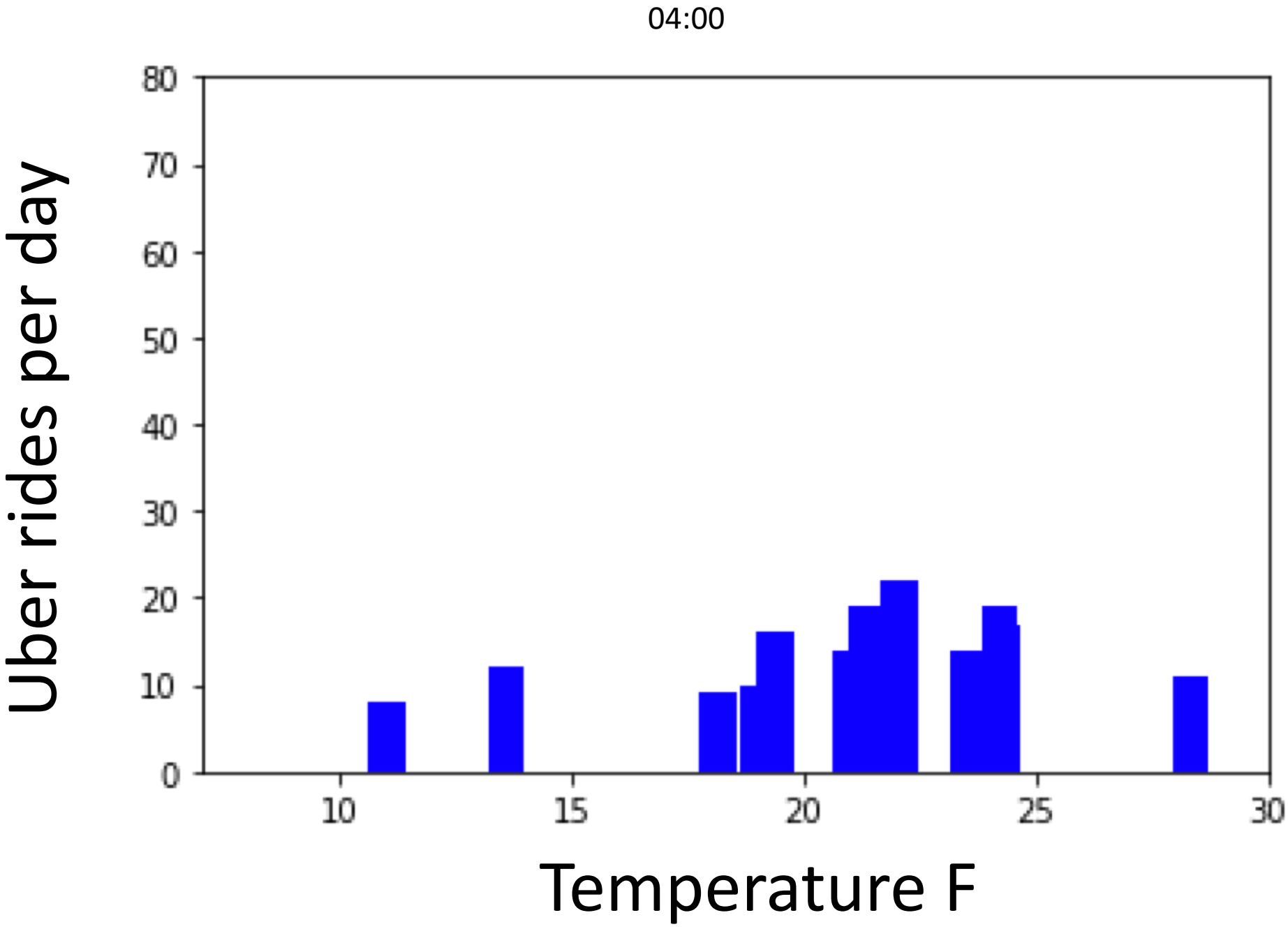
01:00

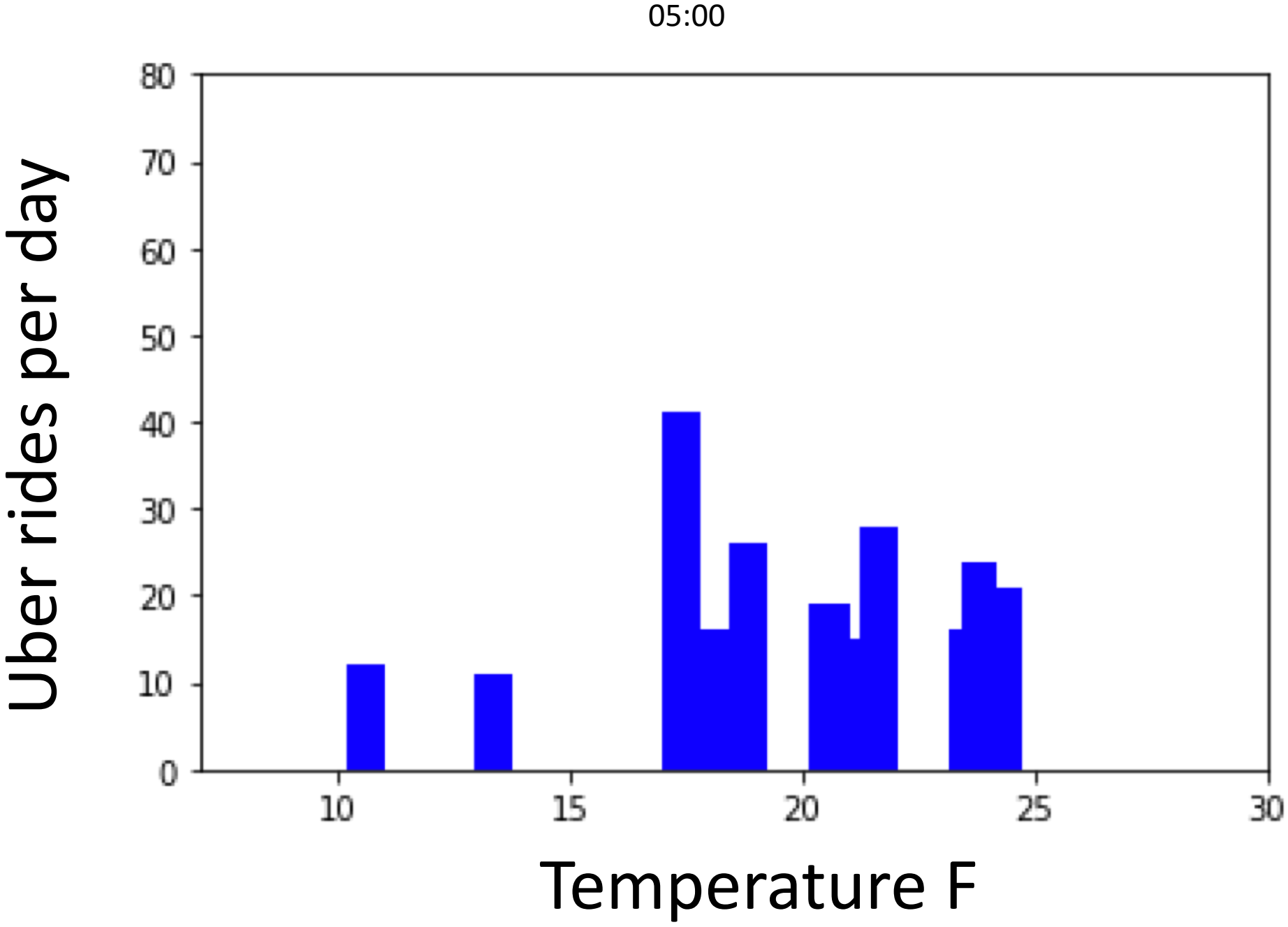

Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

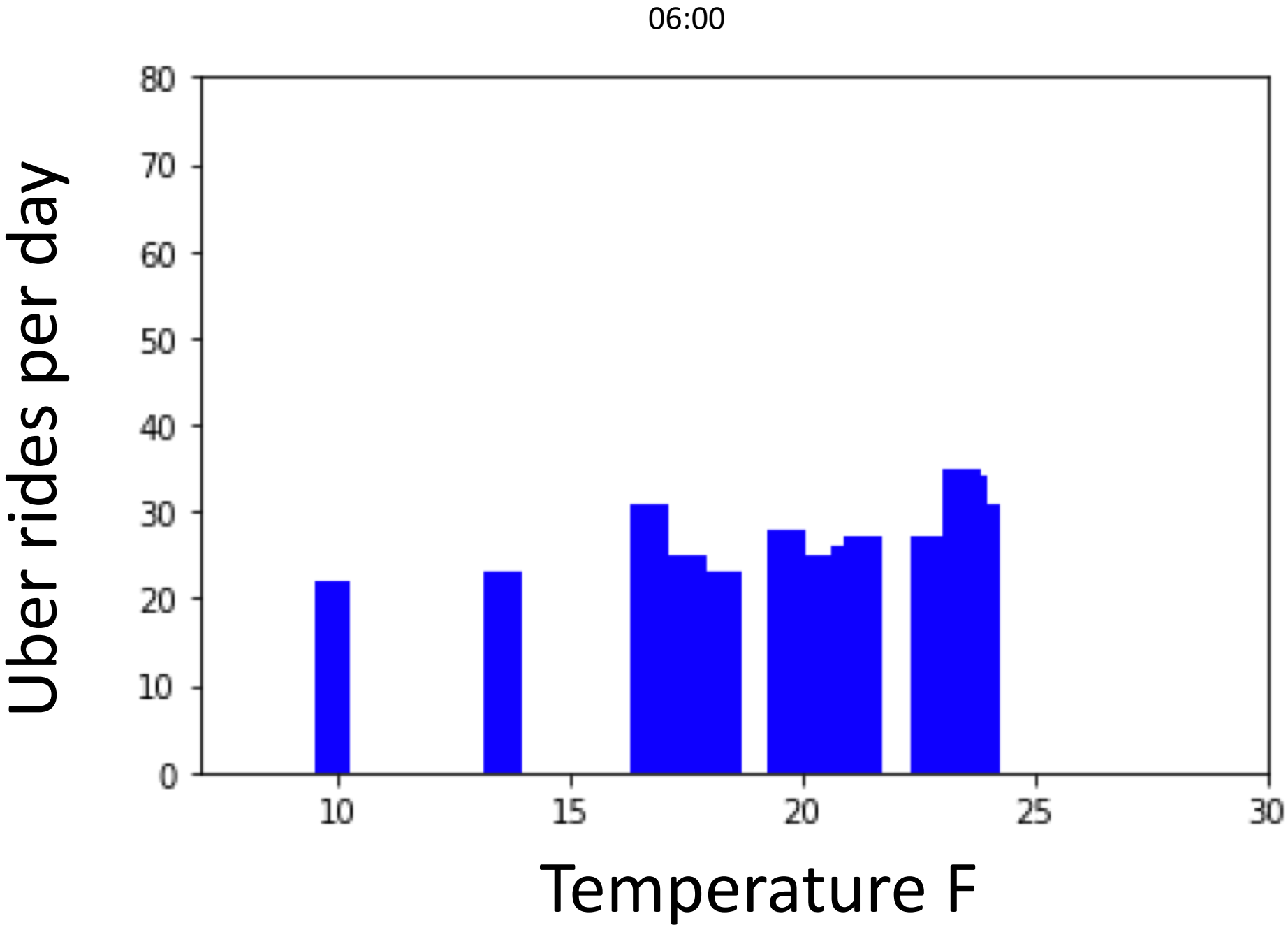
02:00

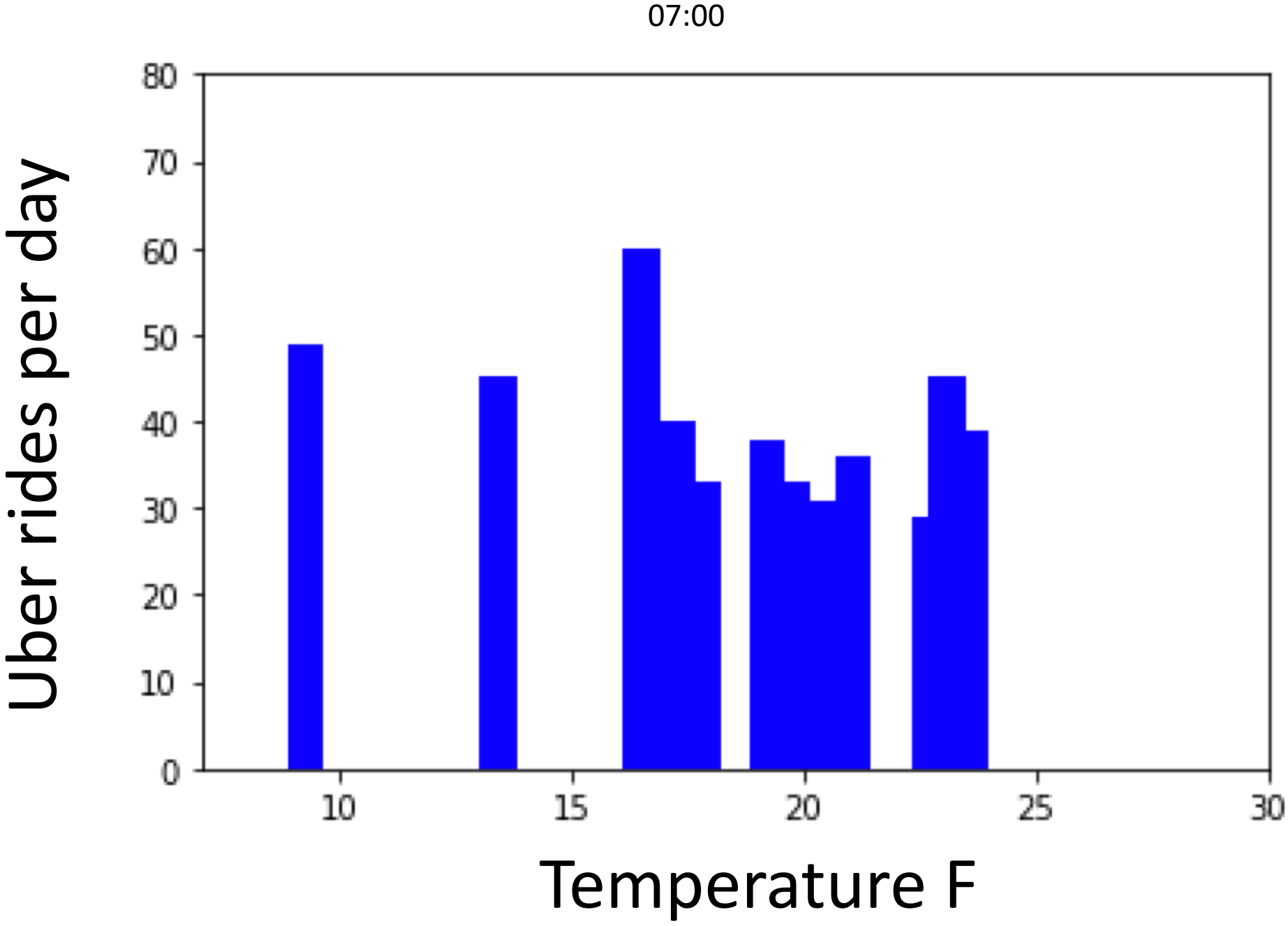

Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

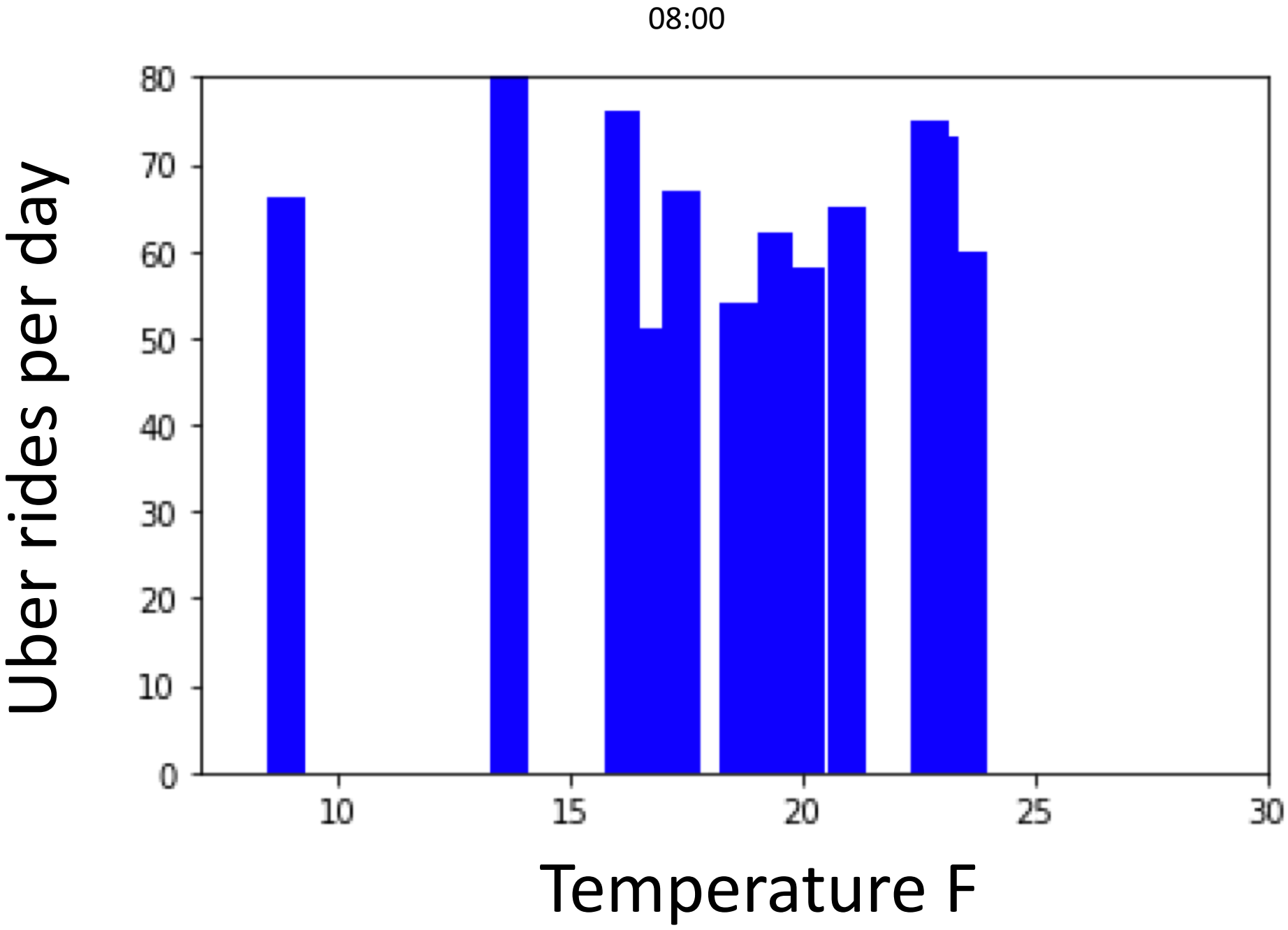

03:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

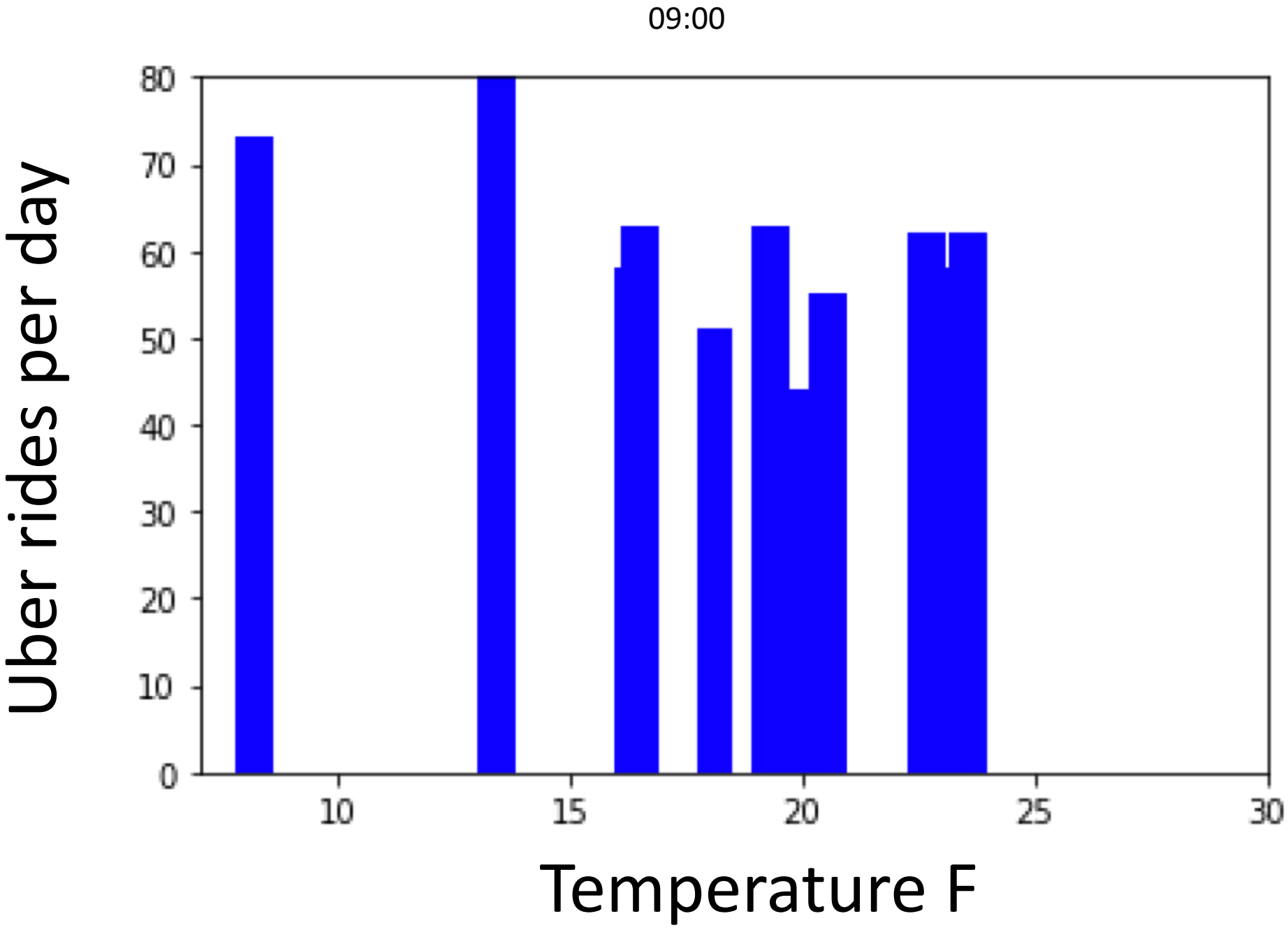

04:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F


05:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

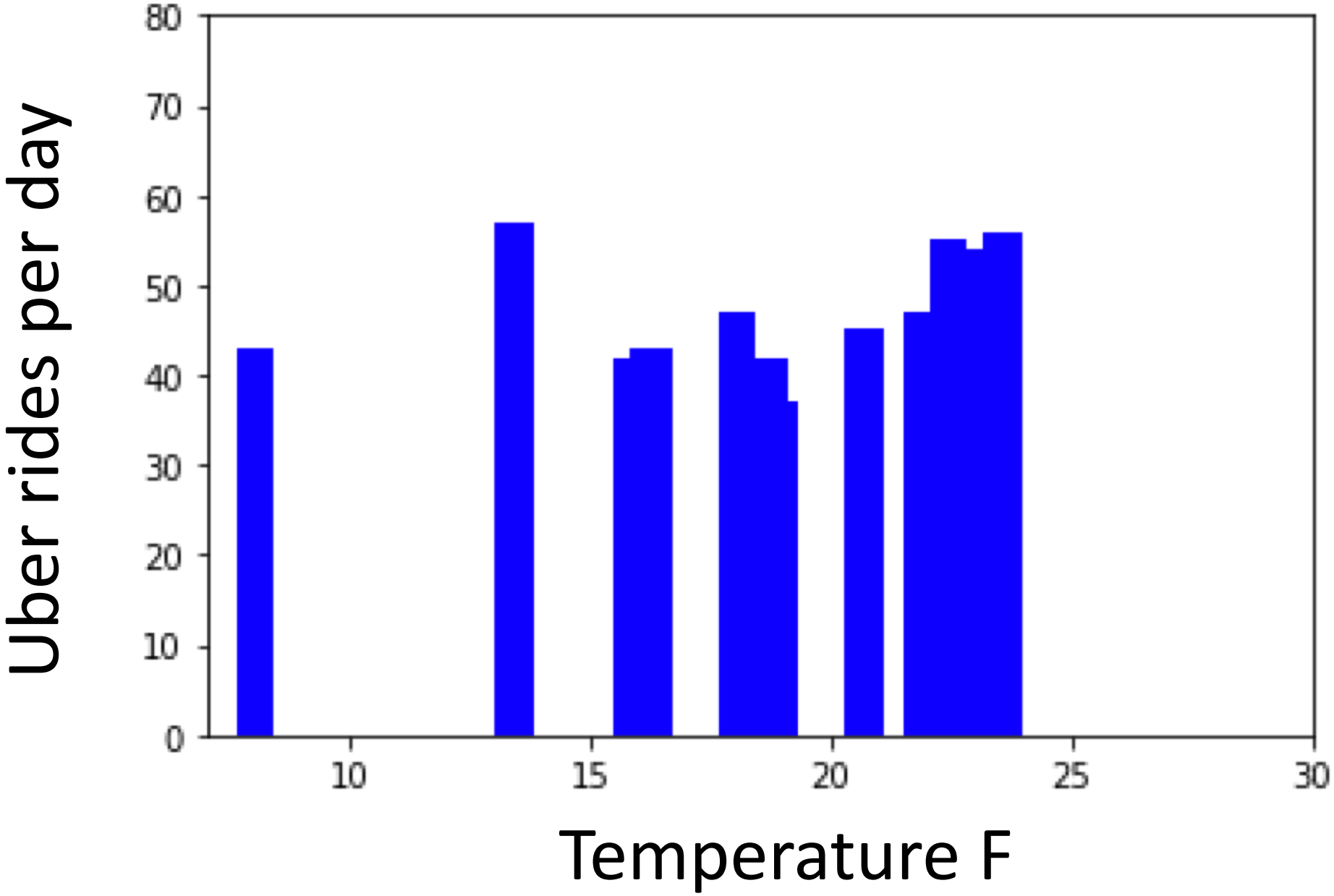

06:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F


07:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

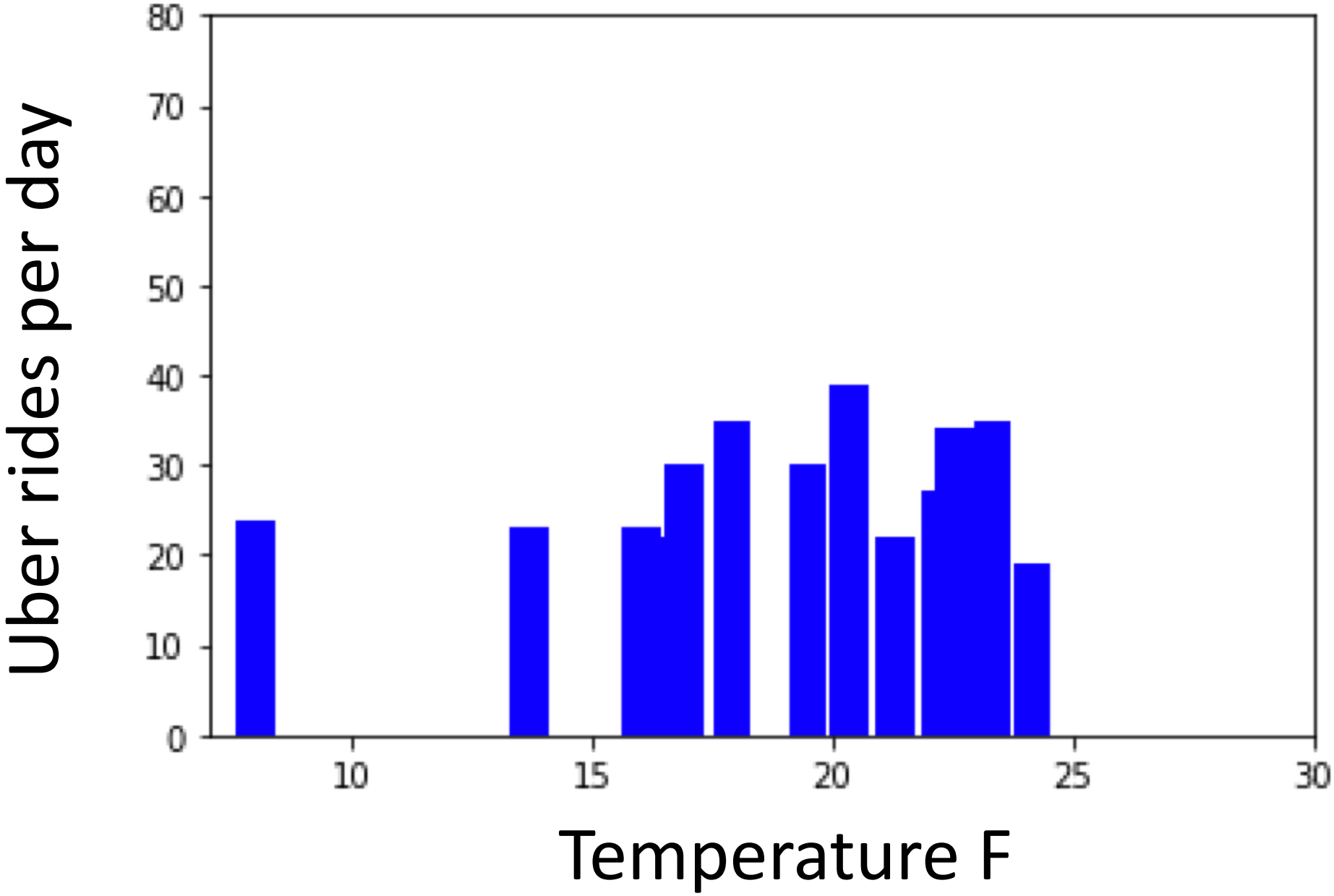

08:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F


09:00
Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

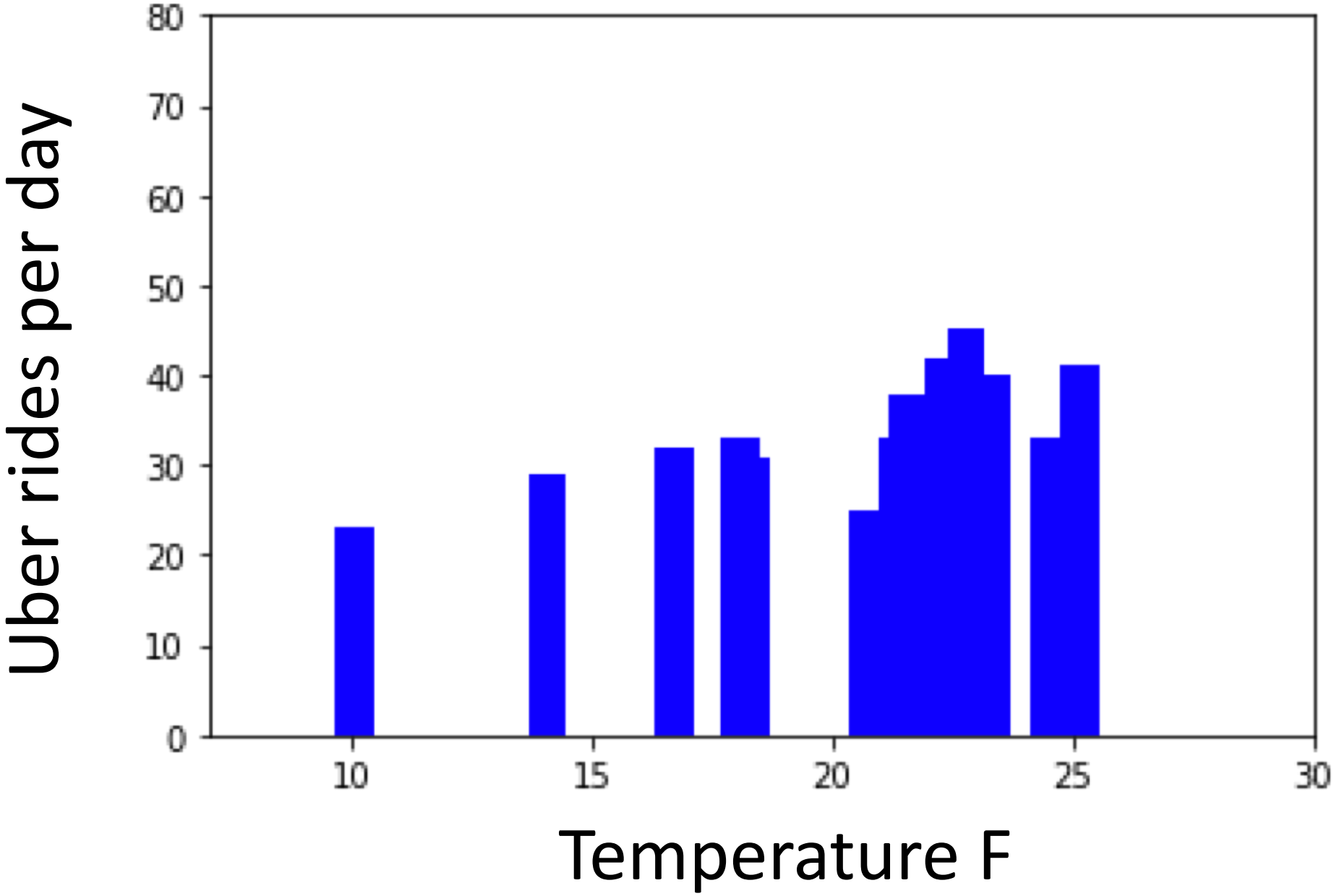
10:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

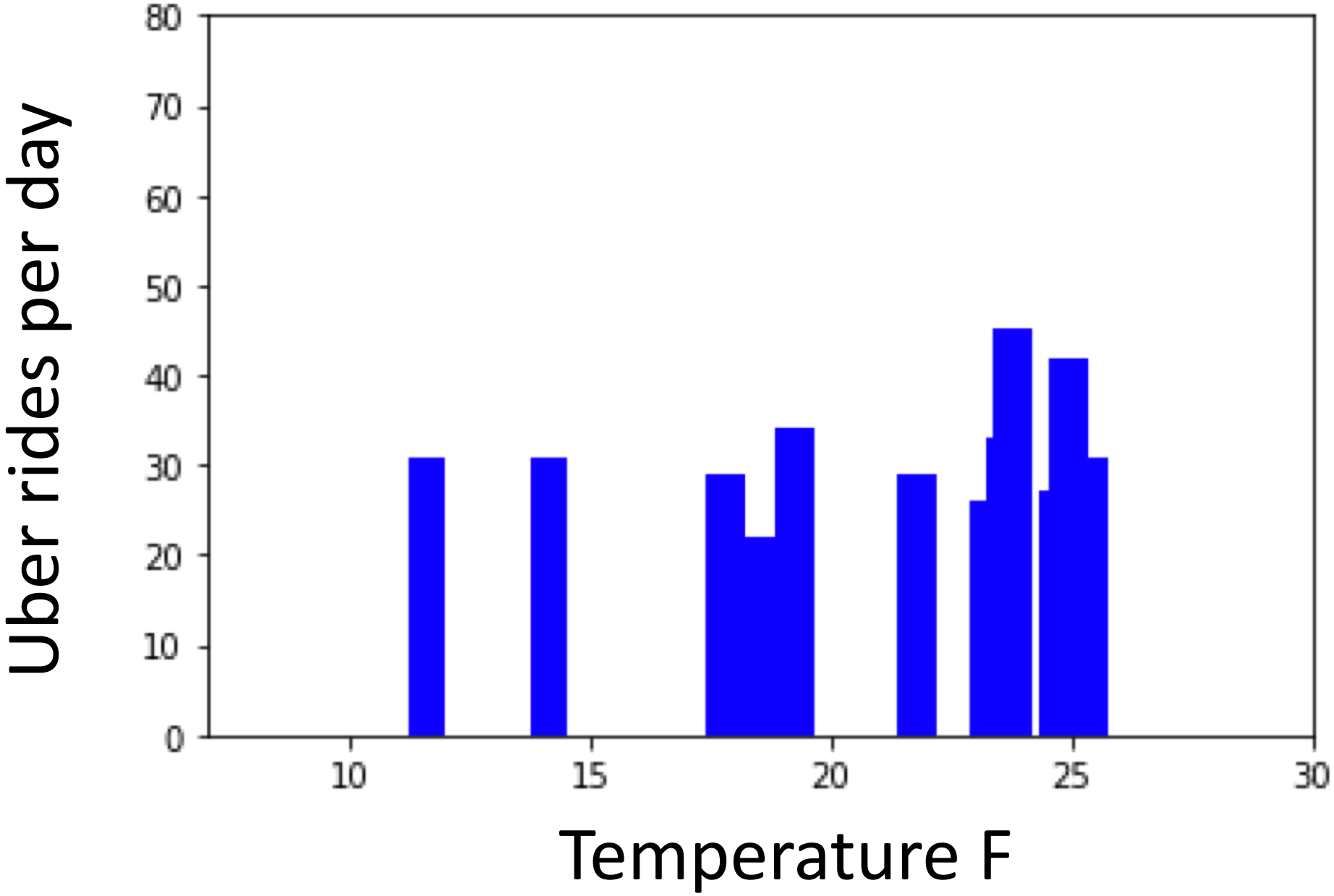
11:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

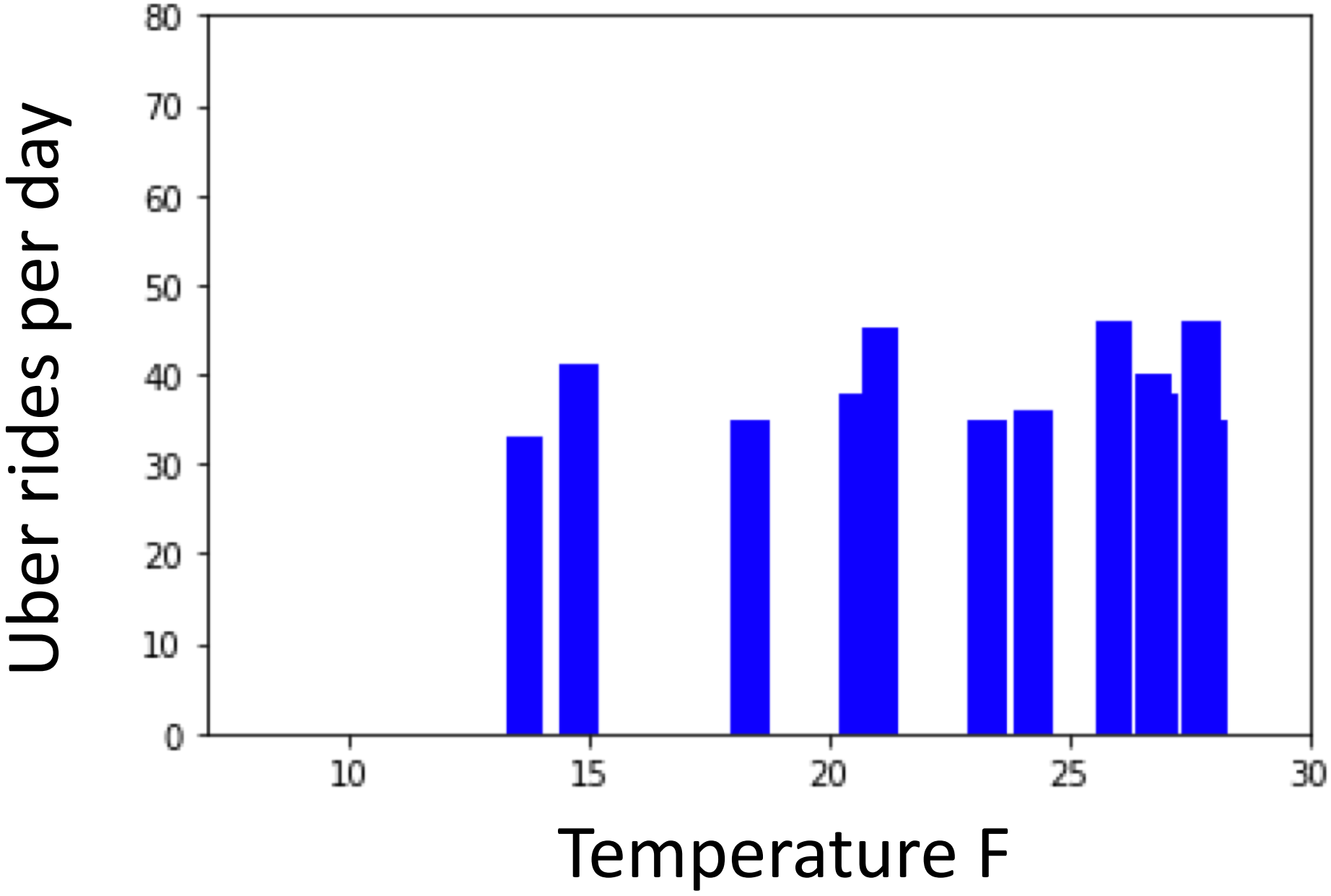
12:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

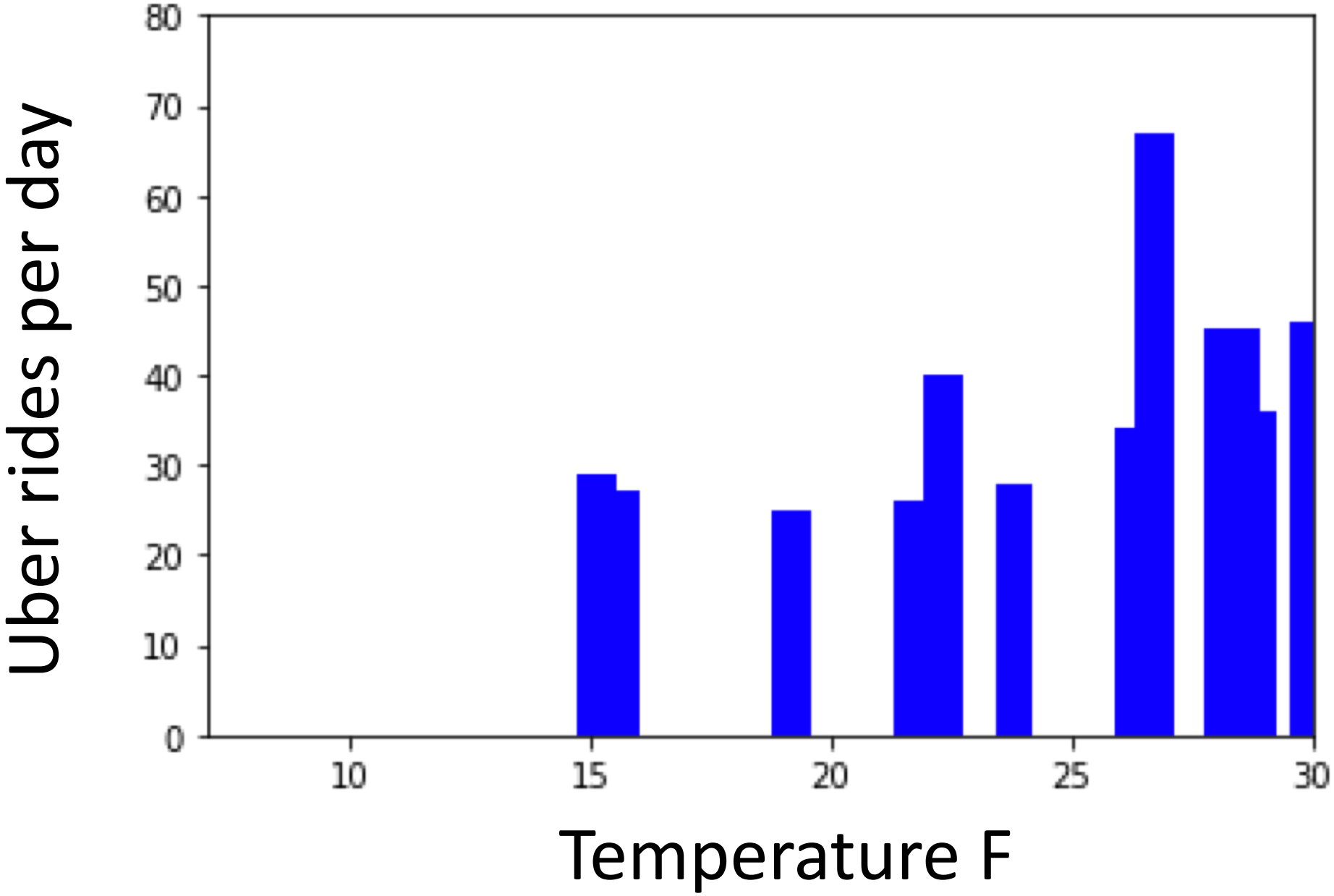
13:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

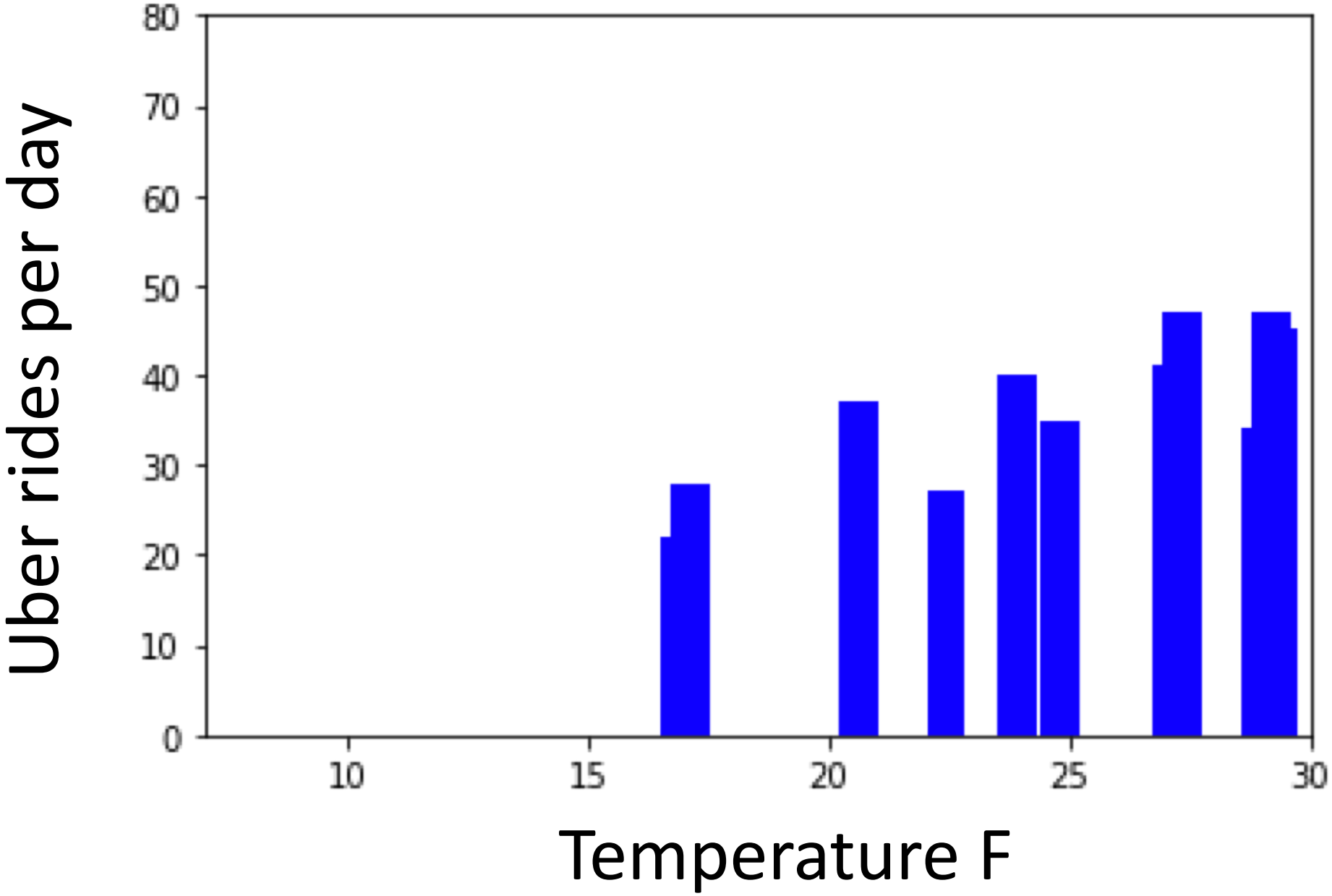
14:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

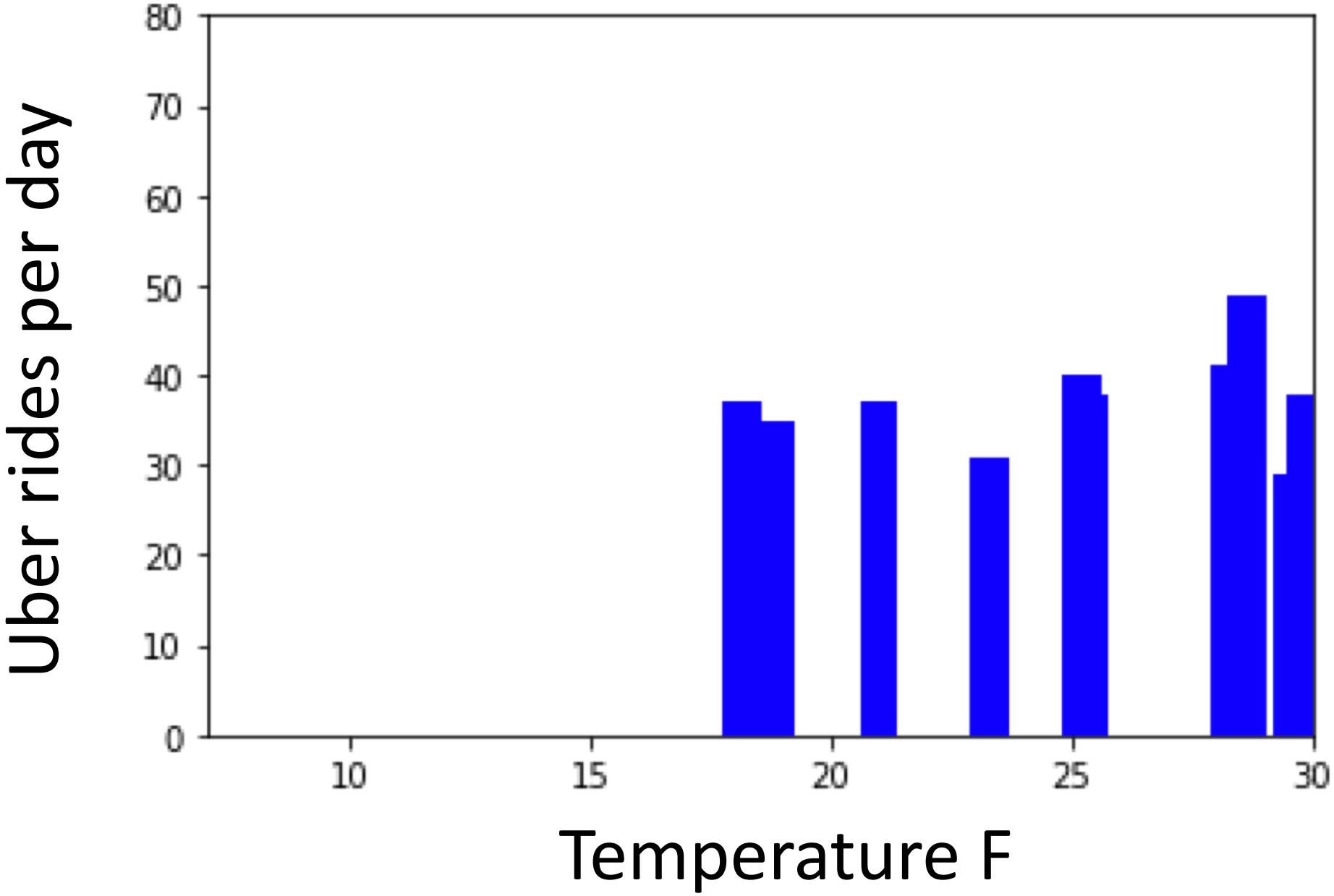
15:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

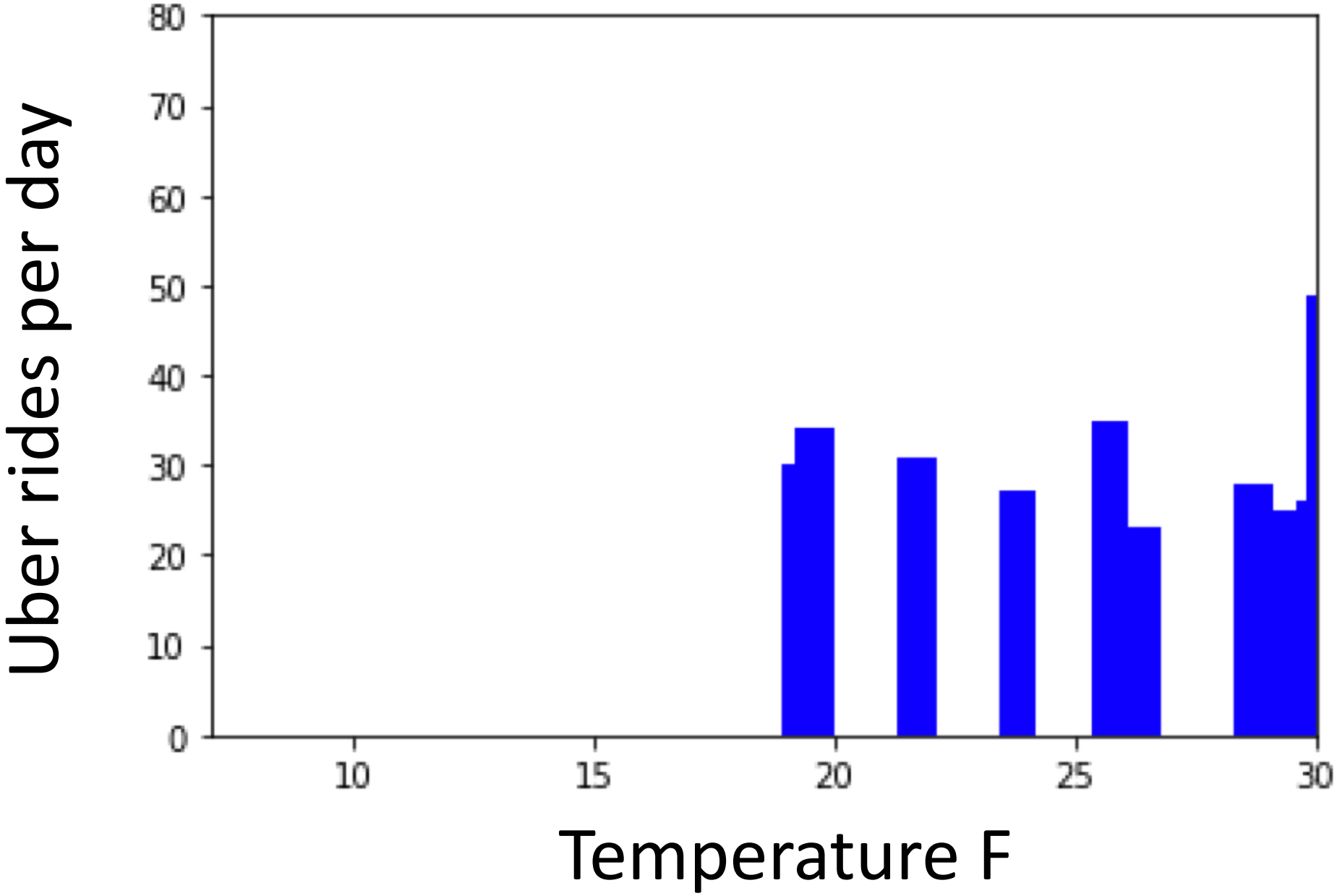
16:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

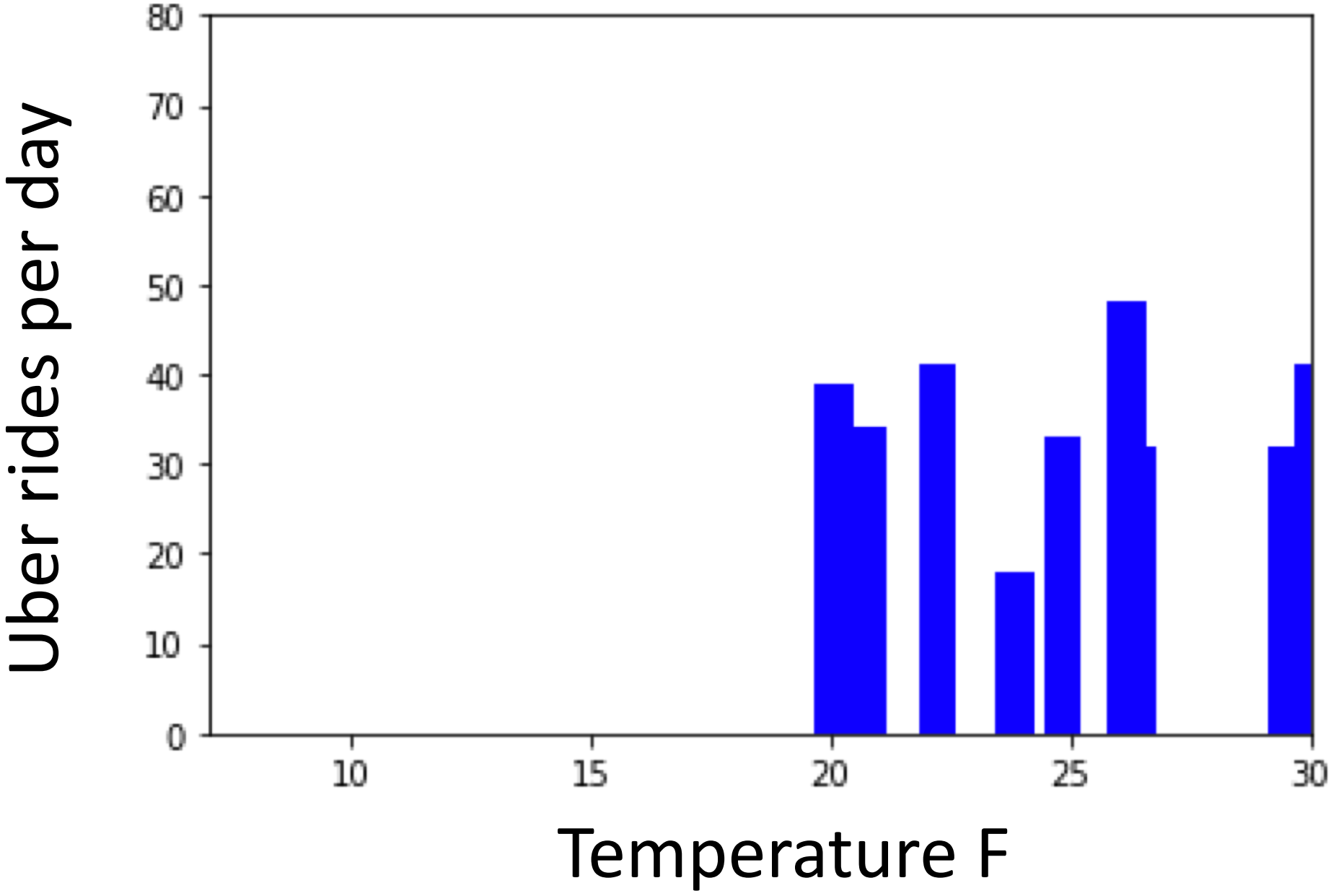
17:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

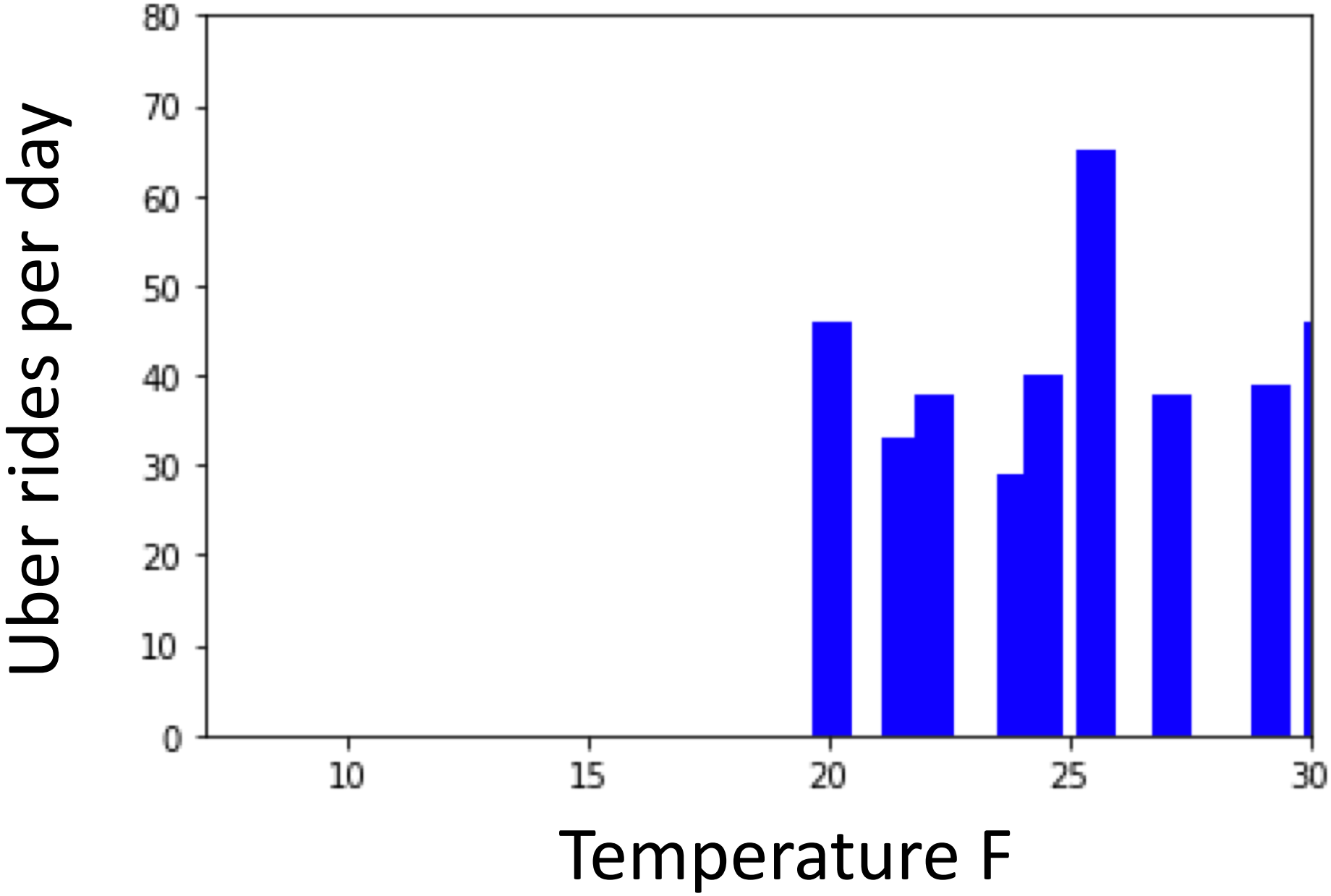
18:00


Uber rides per day
0
10
20
30
40
50
60
70
80
10
15
20
25
30
Temperature F

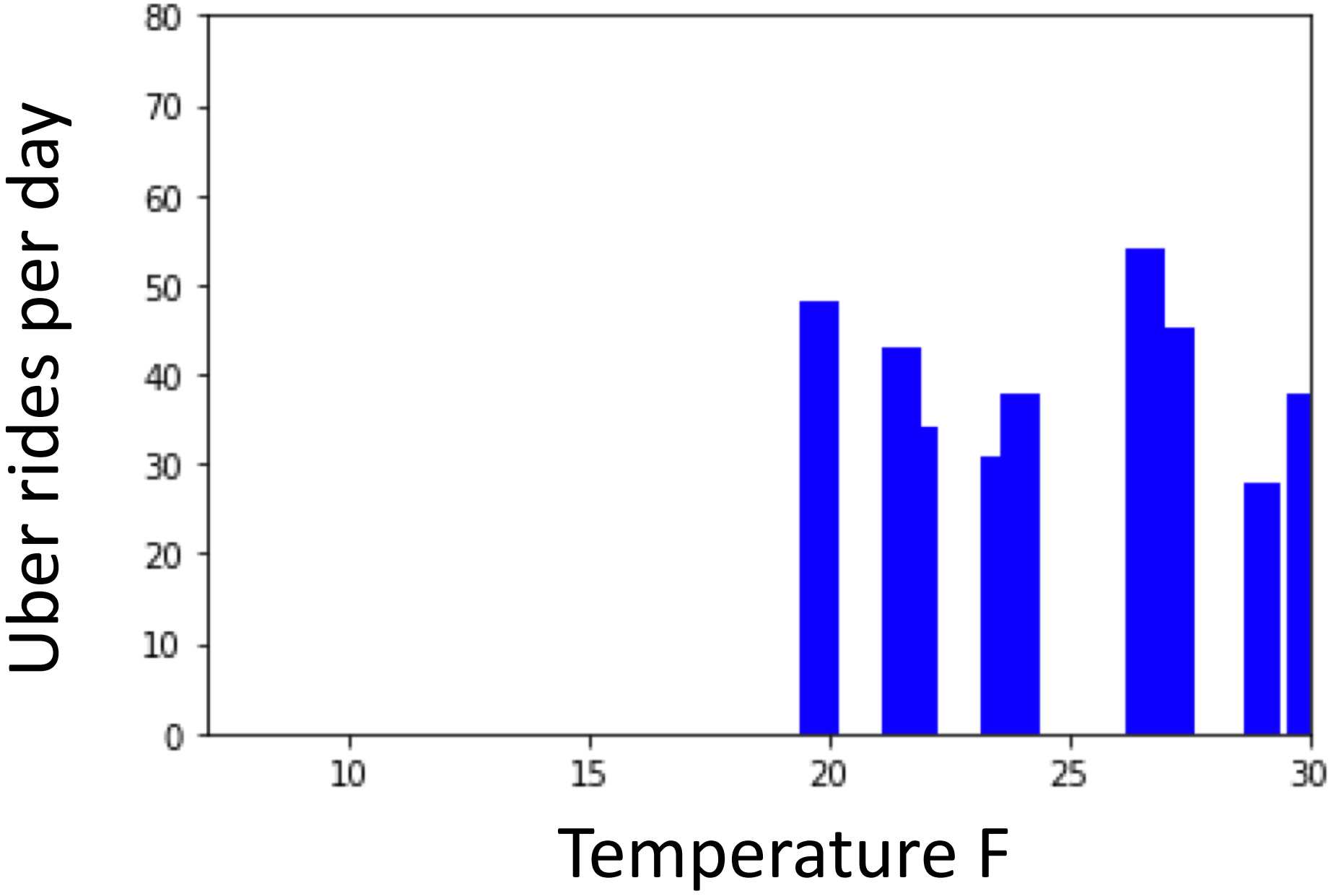
19:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

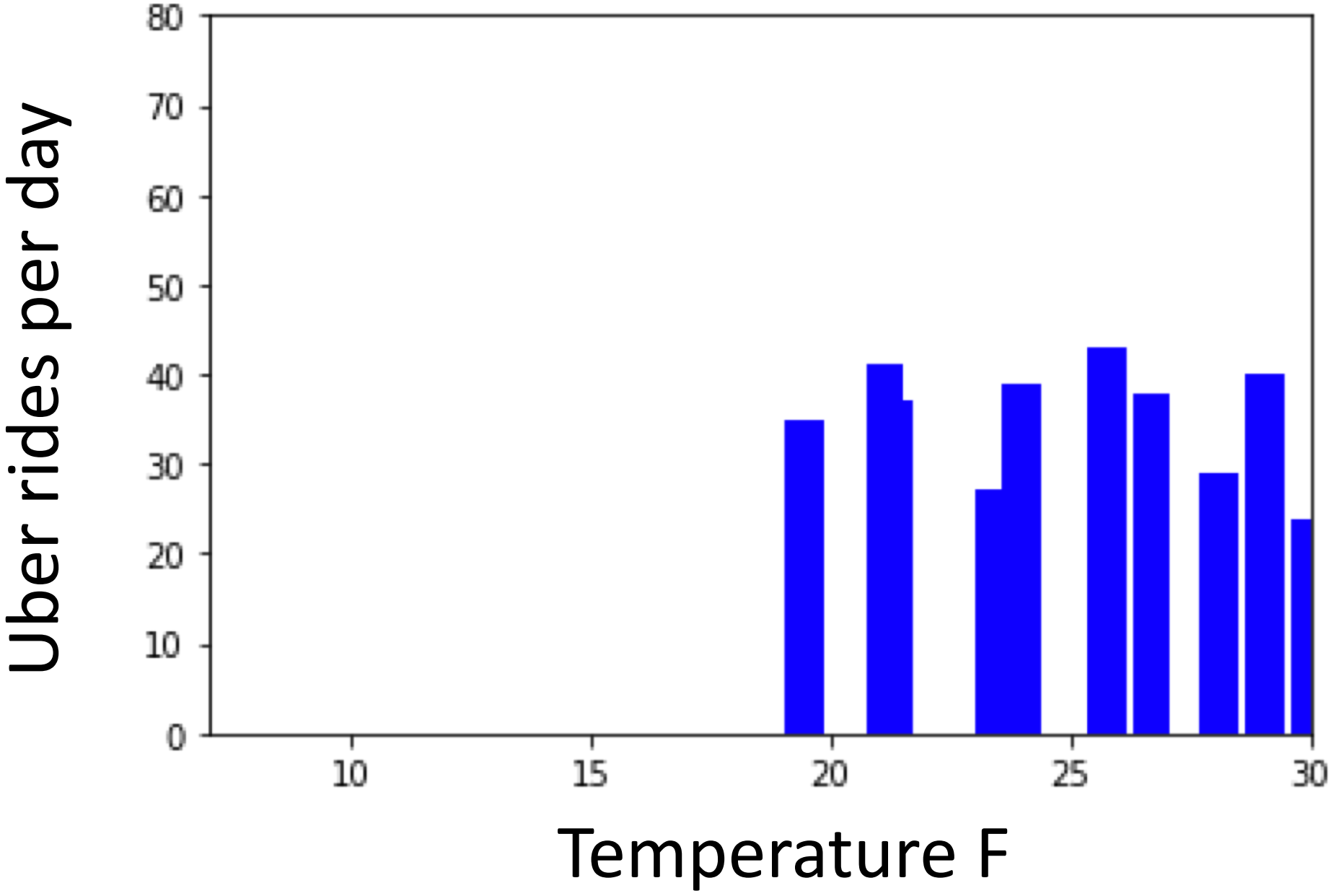
20:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

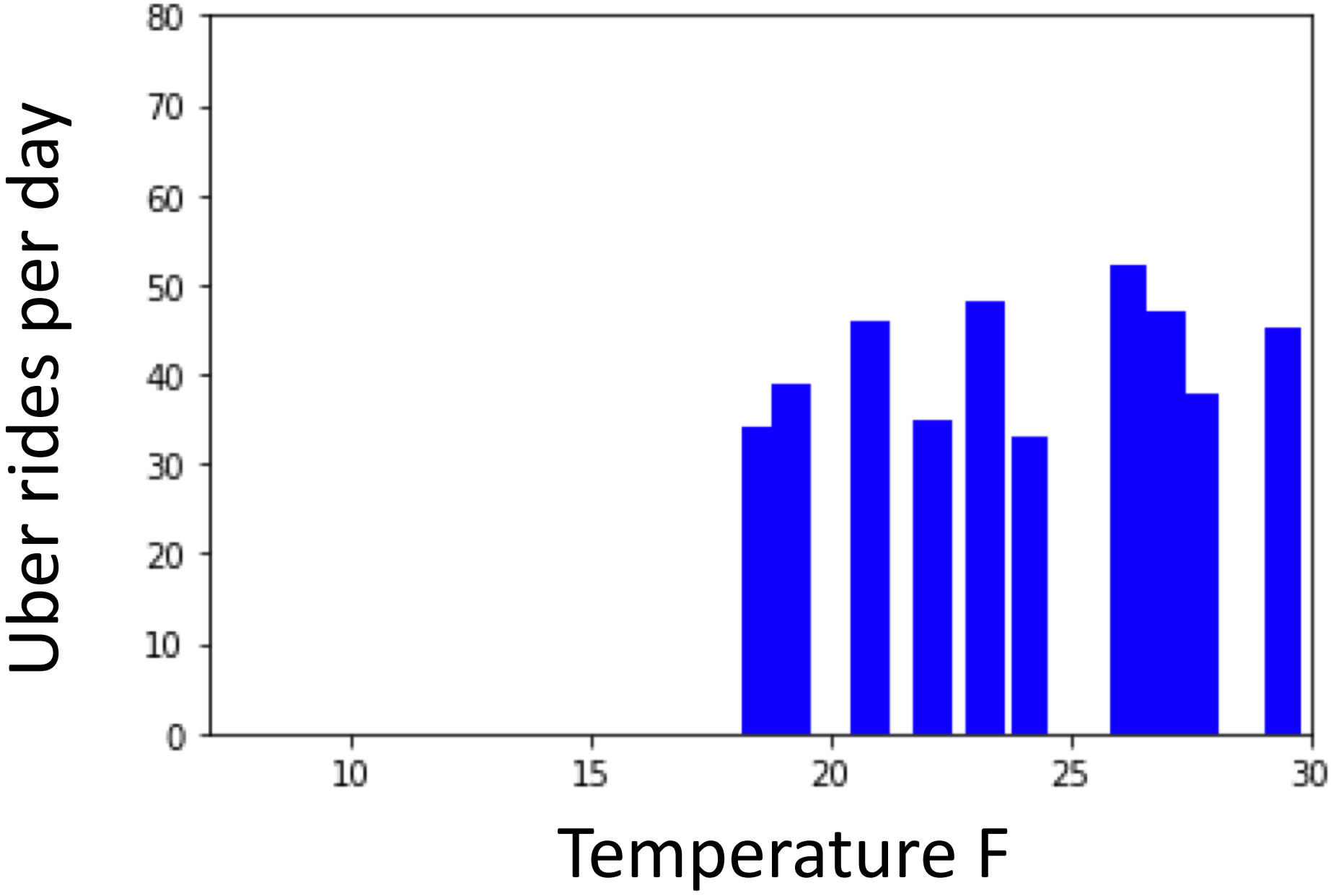
21:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

22:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

23:00


Uber rides per day
80
70
60
50
40
30
20
10
0
10
15
20
25
30
Temperature F

Temperature vs Number of Rides

	observed	expected
COLD	38.787129	41.517429
WARM	40.921170	41.517429
HOT	44.949309	41.517429

Power Divergence Result:
Statistic = 0.4717985698490337 3
P-value = 0.7898602207587081

COLD
below 60 F

WARM
60 F to 80 F

HOT
ABOVE 80 F

Conclusions from Temperature Analysis

- The P-Value is much greater that 0.05. Thus this is Null Hypothesis.
- Different results could be expected if comparing at data through the whole year. The time window July-Sep doesn't contain drastic changes of temperature.
- Different results could expected in non-urban locations and or locations with such high number of taxis, large city train, and city busses network.
- When running the Chi-square test for specific times of the data and/or specific days of the week the P-Value didn't have any significant changes. Always much greater than 0.05. E.g. when comparing noon only (lunch time)

observed	expected
35.000000	35.586957
35.701299	35.586957

Power Divergence Result:
Statistic = 0.010048403330588
P-value = 0.9201524214556676