Weather & Traffic

Leann Kim and Anna Clarkson

OpenWeather API

Parameters:

- Latitude
- Longitude
- API Key

OpenWeather API:

https://api.openweathermap.org/data/2.5/weather?lat=45.506333&lon=-122.750923&appid=57af971f37e6a076513e7383b3028125

```
{
  "coord": {
    "lon": -122.7509,
    "lat": 45.5063
  },
  "weather": [
    {
      "id": 802,
      "main": "Clouds",
      "description": "scattered clouds",
      "icon": "03n"
    }
  ],
  "base": "stations",
  "main": {
    "temp": 280.76,
    "feels_like": 279.16,
    "temp_min": 280.76,
    "temp_max": 280.76,
    "pressure": 1020,
    "humidity": 98,
    "sea_level": 1020,
    "grnd_level": 1009
  },
  "visibility": 10000,
  "wind": {
    "speed": 2.48,
    "deg": 201,
    "gust": 8.65
  },
  "clouds": {
    "all": 30
  },
  "dt": 1714290949,
  "sys": {
    "country": "US",
    "sunrise": 1714309357,
    "sunset": 1714360436
  },
  "timezone": -25200,
  "id": 7262498,
  "name": "West Haven-Sylvan",
  "cod": 200
}
```

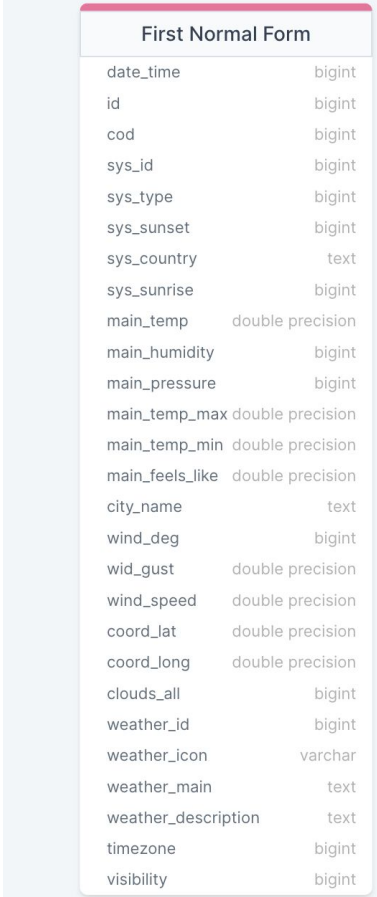

- Weather_main: group of weather parameters (clear, mist, rain, clouds)
- Weather_description: weather condition within each group
  - Example - weather_main: rain, weather_description: light rain
- Weather_icon: weather icon ID

- Temperatures are in Kelvin, which was later transformed into Fahrenheit

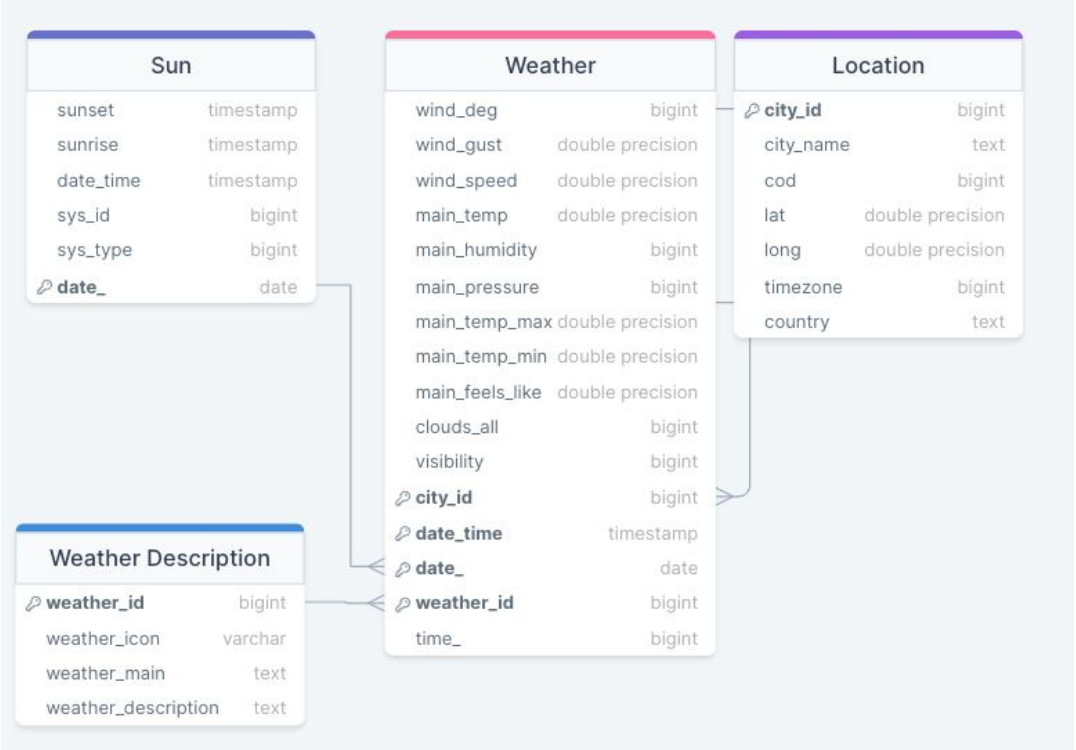
- Wind_speed and wind_gust are in meters per second, which was later transformed into miles per hour

- Dt, sunrise, and sunset are in Unix UTC time, which was later transformed into PDT
  - Unix time: seconds that have elapsed from January 1, 1970

OpenWeather ERD Diagram: First Normal Form



OpenWeather ERD Diagram: Third Normal Form



location_table

city_id	city_name	cod	lat	long	timezone	country
5760068	"West Slope"	200	45.5063	-122.7509	-25200	"US"
7262498	"West Haven-Sylvan"	200	45.5063	-122.7509	-25200	"US"

sun_table

sunset	sunrise	date_time	sys_id	sys_type	date_
2024-04-06 19:45:29	2024-04-06 06:40:33	2024-04-06 00:59:02	2005180	2	2024-04-06
2024-04-07 19:46:46	2024-04-07 06:38:41	2024-04-07 02:00:04	2005180	2	2024-04-07
2024-04-08 19:48:04	2024-04-08 06:36:51	2024-04-08 02:00:02	2005180	2	2024-04-08
2024-04-09 19:49:21	2024-04-09 06:35:00	2024-04-09 02:00:03	2005180	2	2024-04-09
2024-04-10 19:50:38	2024-04-10 06:33:11	2024-04-10 02:00:02	2005180	2	2024-04-10
2024-04-11 19:51:56	2024-04-11 06:31:22	2024-04-11 02:00:04	2005180	2	2024-04-11
2024-04-12 19:53:13	2024-04-12 06:29:33	2024-04-12 02:00:03	2005180	2	2024-04-12

weather_desc

weather_id	weather_icon	weather_main	weather_description
500	10d	Rain	light rain
501	10d	Rain	moderate rain
701	50n	Mist	mist
800	01d	Clear	clear sky
801	02n	Clouds	few clouds
802	03n	Clouds	scattered clouds
803	04d	Clouds	broken clouds
804	04d	Clouds	overcast clouds

weather_table

wind_deg	wind_gust	wind_speed	main_temp	main_humidity	main_pressure	main_temp_max	main_temp_min	main_feels_like	clouds_all	visibility	city_id	date_time	date_	weather_id	time_
261	4.92	2.24	277.93	90	1017	279.28	276.29	276.02	100	10000	5760068	2024-04-06 00:59:02	2024-04-06	804	00:59:02
209	4.02	2.24	277.94	90	1017	279.28	276.45	276.04	55	10000	5760068	2024-04-06 02:00:04	2024-04-06	803	02:00:04
302	4.02	1.79	277.25	93	1018	278.97	275.26	275.7	78	10000	5760068	2024-04-06 05:00:05	2024-04-06	803	05:00:05
233	7.6	3.58	277.63	91	1020	278.84	276.22	274.63	75	10000	5760068	2024-04-06 08:00:10	2024-04-06	803	08:00:10
160		3.6	279.13	88	1021	280.37	277.95	276.44	100	10000	5760068	2024-04-06 11:00:05	2024-04-06	804	11:00:05
229	10.73	5.36	280.61	85	1021	281.61	278.76	277.41	100	10000	7262498	2024-04-06 14:00:04	2024-04-06	500	14:00:04
214	8.49	4.47	281.32	87	1020	282.09	279.85	278.68	100	10000	5760068	2024-04-06 17:00:02	2024-04-06	501	17:00:02
20		3.09	280.14	92	1022	281.45	278.49	277.99	100	10000	5760068	2024-04-06 20:00:05	2024-04-06	804	20:00:05
211	3.58	1.34	279.39	95	1023	280.12	278.21	278.72	97	10000	5760068	2024-04-06 23:00:09	2024-04-06	804	23:00:09
120		1.54	278.66	95	1023	279.74	277.58	277.63	100	10000	5760068	2024-04-07 02:00:04	2024-04-07	701	02:00:04
256	2.24	1.34	278.48	95	1024	279.72	277.05	277.69	100	10000	5760068	2024-04-07 05:00:10	2024-04-07	701	05:00:10
320		1.54	278.74	96	1024	279.91	277.35	277.72	100	1207	5760068	2024-04-07 08:00:03	2024-04-07	701	08:00:03
177	3.13	1.34	281.27	87	1026	282.73	279.9	280.85	84	10000	5760068	2024-04-07 11:00:03	2024-04-07	803	11:00:03
244	5.81	3.13	284.98	65	1024	286.51	282.21	283.92	100	10000	7262498	2024-04-07 14:00:03	2024-04-07	804	14:00:03
197	4.92	2.68	284.43	68	1023	286.29	281.34	283.39	100	10000	5760068	2024-04-07 17:00:20	2024-04-07	804	17:00:20
124	2.24	0.89	281.37	84	1024	282.62	279.39	281.37	100	10000	5760068	2024-04-07 20:00:02	2024-04-07	500	20:00:02
0	1.34	0.45	279.28	90	1025	280.95	277.25	279.28	98	10000	5760068	2024-04-07 23:00:03	2024-04-07	804	23:00:03
270	1.79	0.89	278.27	91	1024	280.39	275.94	278.27	100	10000	5760068	2024-04-08 02:00:02	2024-04-08	804	02:00:02

TomTom API

Parameters:

- Latitude
- Longitude
- Unit (MPH)

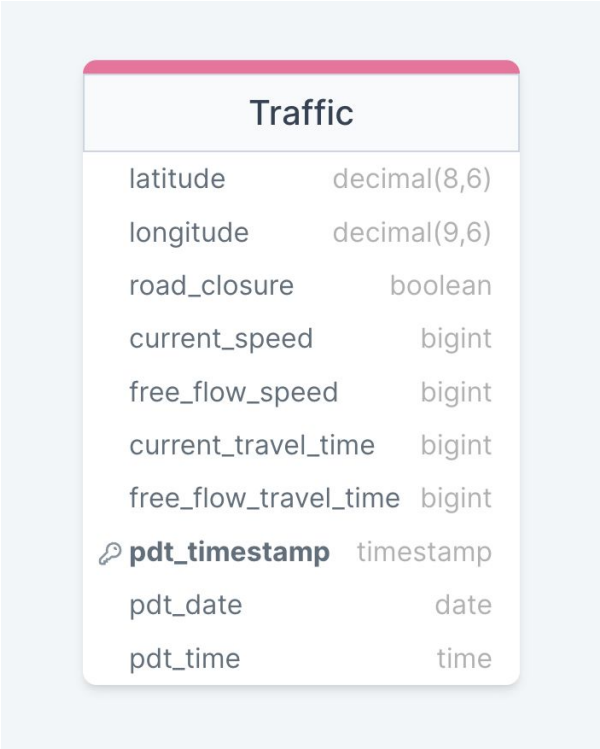
TomTom API:

https://api.tomtom.com/traffic/services/4/flowSegmentData/relative0/10/json?point=45.50800%2C-122.74291&unit=MPH&openLr=false&key=y5wyIm25CVxcnpKzK6IHPGYBkez8pbtI

```
{
  "flowSegmentData": {
    "frc": "FRC1",
    "currentSpeed": 62,
    "freeFlowSpeed": 62,
    "currentTravelTime": 96,
    "freeFlowTravelTime": 96,
    "confidence": 1,
    "roadClosure": false,
    "coordinates": {
      "coordinate": [
        {
          "latitude": 45.5061908648218,
          "longitude": -122.771864075445
        },
        {
          "latitude": 45.5060858817999,
          "longitude": -122.771299650662
        },
        {
          "latitude": 45.5060055012911,
          "longitude": -122.770813292519
        },
        {
          "latitude": 45.5059081387034,
          "longitude": -122.770237263466
        },
        {
          "latitude": 45.5058753690754,
          "longitude": -122.770047501406
        },
        {
          "latitude": 45.5058008157807,
          "longitude": -122.769586071416
        },
        {
          "latitude": 45.5057188073946,
          "longitude": -122.769012100875
        },
        {
          "latitude": 45.5082255281179,
          "longitude": -122.738713798418
        },
        {
          "latitude": 45.5082280486939,
          "longitude": -122.738376190315
        }
      ]
    },
    "@version": "traffic-service-flow 1.0.109"
  }
}
```


- FRC: Functional Road Class
  - 1 is “Major road, less important than a motorway”
- CurrentSpeed: current average speed at given coordinate
- FreeFlowSpeed: free flow speed under ideal conditions
- CurrentTravelTime: travel time in seconds between points
- FreeFlowTravelTime: travel time in seconds under ideal conditions
- Confidence: quality of provided speeds and travel times
  - 1 is full confidence

TomTom ERD Diagram: Third Normal Form



traffic

latitude	longitude	road_closure	current_speed	free_flow_speed	current_travel_time	free_flow_travel_time	pdt_timestamp	pdt_date	pdt_time
45.506191	-122.771864	FALSE	32	62	184	96	2024-04-06 13:00:03.284435	2024-04-06	13:00:03
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-06 17:00:02.863136	2024-04-06	17:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-06 21:00:02.413292	2024-04-06	21:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-07 01:00:02.808758	2024-04-07	01:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-07 05:00:05.767189	2024-04-07	05:00:05
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-07 09:00:03.100492	2024-04-07	09:00:03
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-07 13:00:02.608632	2024-04-07	13:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-07 17:00:03.174529	2024-04-07	17:00:03
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-07 21:00:02.648842	2024-04-07	21:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-08 01:00:02.20589	2024-04-08	01:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-08 05:00:02.669352	2024-04-08	05:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-08 09:00:03.293348	2024-04-08	09:00:03
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-08 13:00:02.924212	2024-04-08	13:00:02
45.506191	-122.771864	FALSE	50	60	119	98	2024-04-08 17:00:02.578251	2024-04-08	17:00:02
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-08 21:00:03.124723	2024-04-08	21:00:03
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-09 01:00:03.721236	2024-04-09	01:00:03
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-09 05:00:02.337253	2024-04-09	05:00:02
45.506191	-122.771864	FALSE	26	61	228	97	2024-04-09 09:00:03.054347	2024-04-09	09:00:03
45.506191	-122.771864	FALSE	62	62	96	96	2024-04-09 13:00:02.969291	2024-04-09	13:00:02
45.506191	-122.771864	FALSE	32	61	188	97	2024-04-09 17:00:02.672772	2024-04-09	17:00:02

Merged Data

- Joined weather and weather_description table by weather_id = final_weather
- Joined the final_weather table with the traffic table by a newly added binned datetime column = weather_traffic
    - Initially only had two matching hours per day (5am and 5pm) due to variance in our cron scheduling
    - Binned our respective date/time columns and ended up matching on seven different hours per day!
        - 1am, 4am, 7am, 1pm, 4pm, and 7pm (in 24 hour clock format)
        - Went from having 44 merged rows to about 130!

Serving: Endpoints

API: data503api-production-563d.up.railway.app
Endpoints:

- /weathertraffic: all info from our merged weather and traffic table
- /weather_main: distinct weather groups (clear, mist, rain, clouds)
- /weather_desc: distinct weather descriptions, plus the weather groups
- /avgspd_weathermain: average speed by weather group
- /avgspd_weatherdesc: average speed by weather description
- /avgspd: average speed by hour and day
- /windsp: wind speed (in mph) by day and hour
- /avgsptemp_byhour: average speed and temp by hour
- /avgsptemp_byhour_day: average speed and temp by hour and day
- /weather_icon: distinct weather icons

/weathertraffic


```
{
  "weather_id": 804,
  "wind_deg": 173,
  "wind_gust": 3.15,
  "wind_speed": 1.54,
  "main_temp": 279.79,
  "main_humidity": 96,
  "main_pressure": 1018,
  "main_temp_max": 279.79,
  "main_temp_min": 279.79,
  "main_feels_like": 278.92,
  "clouds_all": 97,
  "visibility": 10000,
  "city_id": 5760068,
  "date_time": "2024-04-27T05:00:03",
  "date_": "2024-04-27",
  "time_": "05:00:03",
  "weather_icon": "04d",
  "weather_main": "Clouds",
  "weather_description": "overcast clouds",
  "binned_datetime_bin": "2024-04-27T04:00:00",
  "latitude": 45.506191,
  "longitude": -122.771864,
  "road_closure": false,
  "current_speed": 62,
  "free_flow_speed": 62,
  "current_travel_time": 96,
  "free_flow_travel_time": 96,
  "pdt_timestamp": "2024-04-27T05:00:03.162308",
  "binned_time": "04:00:00"
},
{
  "weather_id": 804,
  "wind_deg": 194,
  "wind_gust": 7.15,
  "wind_speed": 2.59,
  "main_temp": 280.92,
  "main_humidity": 91,
  "main_pressure": 1018,
  "main_temp_max": 280.92,
  "main_temp_min": 280.92,
  "main_feels_like": 279.26,
  "clouds_all": 100,
  "visibility": 10000,
  "city_id": 5760068,
  "date_time": "2024-04-27T08:00:02",
  "date_": "2024-04-27",
  "time_": "08:00:02",
  "weather_icon": "04d",
  "weather_main": "Clouds",
  "weather_description": "overcast clouds",
```


/avgspd_weatherdesc

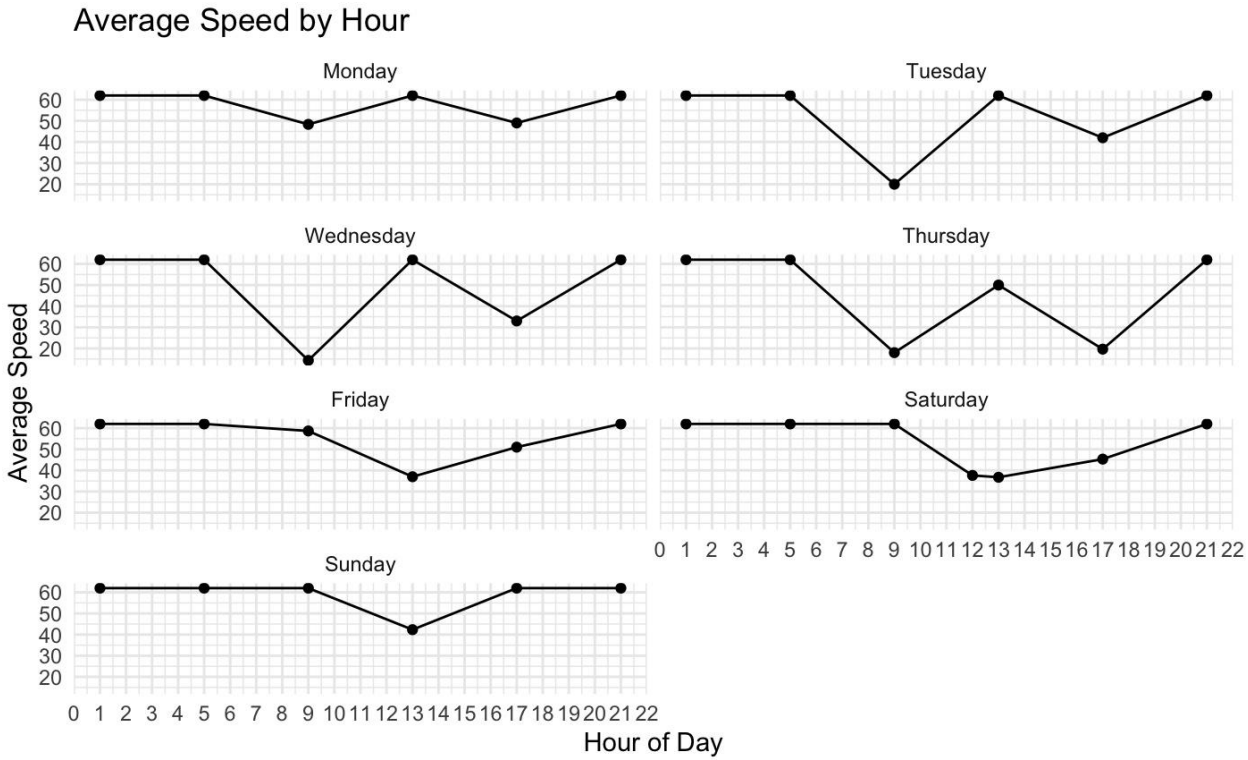

```
[
  {
    "weather_description": "mist",
    "average_speed": 62
  },
  {
    "weather_description": "few clouds",
    "average_speed": 57.8571428571429
  },
  {
    "weather_description": "broken clouds",
    "average_speed": 56.2857142857143
  },
  {
    "weather_description": "overcast clouds",
    "average_speed": 53.9152542372881
  },
  {
    "weather_description": "clear sky",
    "average_speed": 52.05
  },
  {
    "weather_description": "light rain",
    "average_speed": 50
  },
  {
    "weather_description": "moderate rain",
    "average_speed": 38
  },
  {
    "weather_description": "scattered clouds",
    "average_speed": 31.5
  }
]
```


/avgsptemp_byhour_day


```
[
  {
    "hour_of_day": 1,
    "day_of_week": "Sunday",
    "average_speed": 62
  },
  {
    "hour_of_day": 4,
    "day_of_week": "Sunday",
    "average_speed": 62
  },
  {
    "hour_of_day": 7,
    "day_of_week": "Sunday",
    "average_speed": 62
  },
  {
    "hour_of_day": 13,
    "day_of_week": "Sunday",
    "average_speed": 42.3333333333333
  },
  {
    "hour_of_day": 16,
    "day_of_week": "Sunday",
    "average_speed": 62
  },
  {
    "hour_of_day": 19,
    "day_of_week": "Sunday",
    "average_speed": 62
  },
  {
    "hour_of_day": 1,
    "day_of_week": "Monday",
    "average_speed": 62
  },
  {
    "hour_of_day": 4,
    "day_of_week": "Monday",
    "average_speed": 62
  },
  {
    "hour_of_day": 7,
    "day_of_week": "Monday",
    "average_speed": 48.3333333333333
  },
```

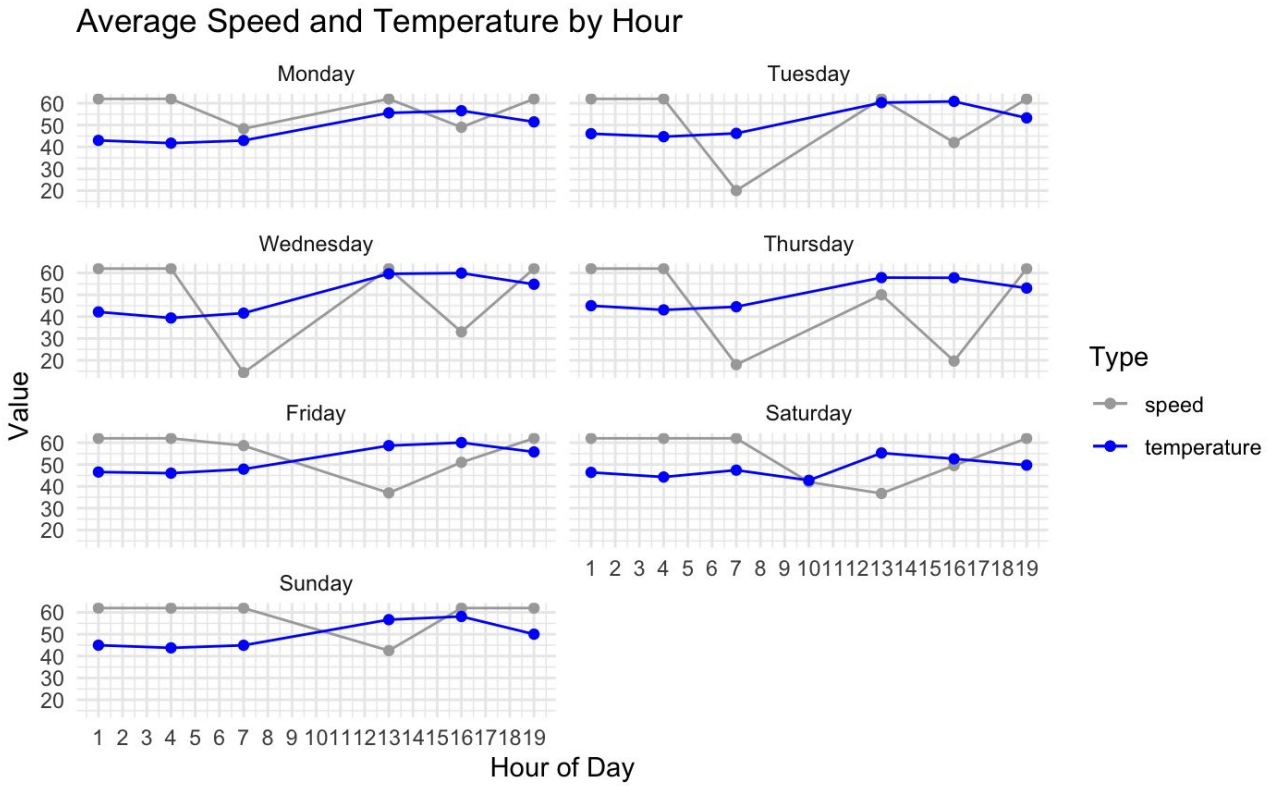
The Big Question:

Does rain affect traffic speeds?



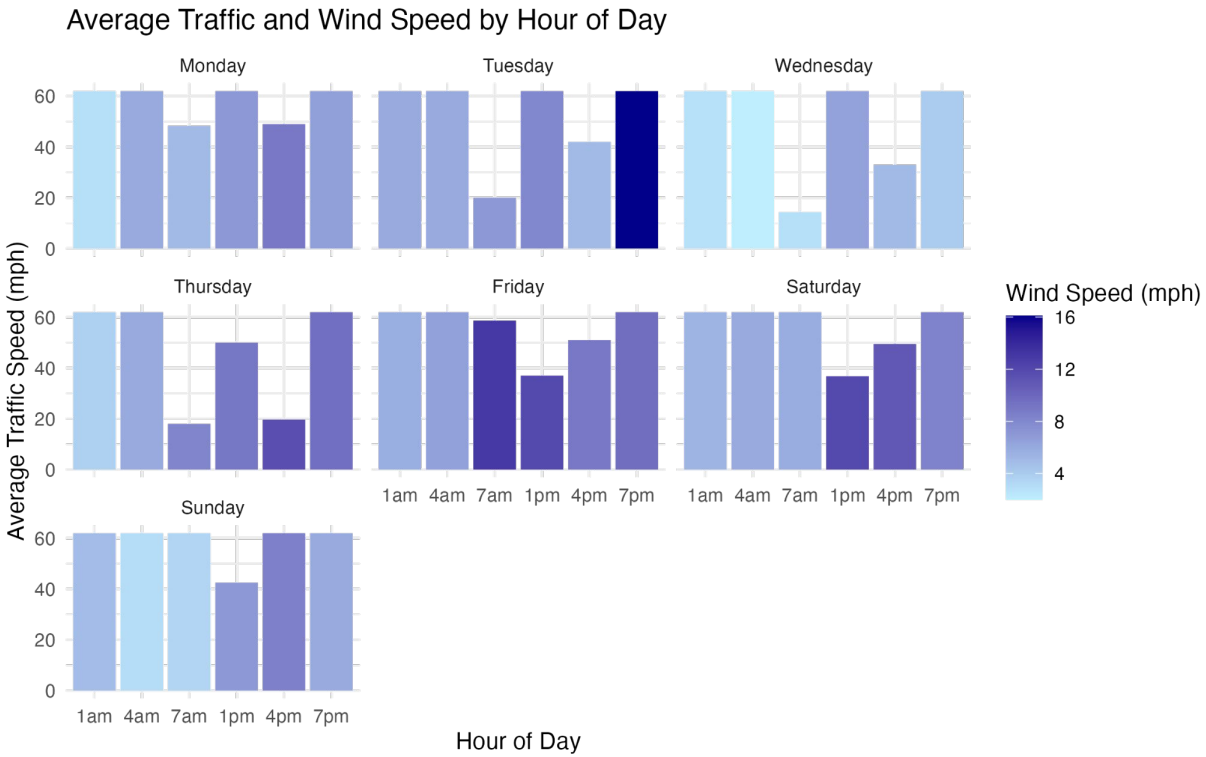
Key Insights:

- Average speeds decrease around 7am and 4pm on the weekdays
- Average speeds decrease after 12pm on weekends



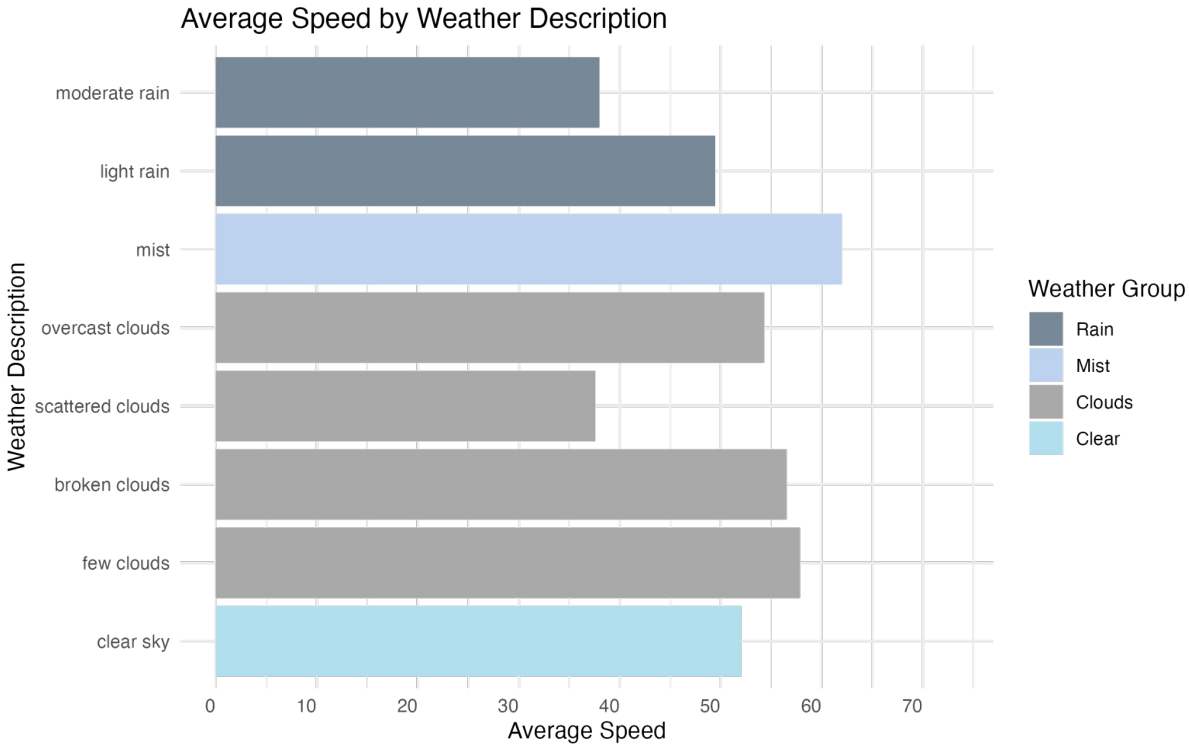
Key Insights:

- Temperatures increase as the day progresses
- No obvious correlation with traffic speed



Key Insights:

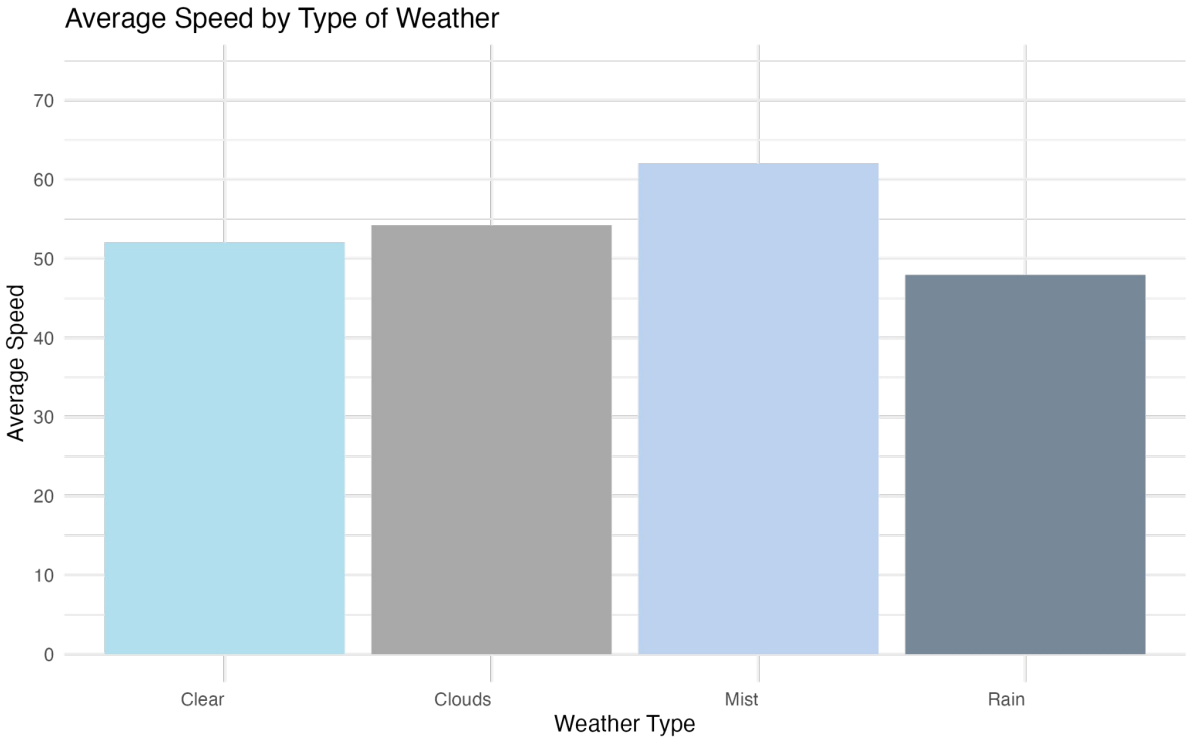
- No obvious correlation between average traffic speed and average wind speed
- It looks cool!



Key Insights:

- Average speed for scattered clouds is suspiciously low
  - There were only 5 instances of scattered clouds
- Average speed for moderate rain is low
  - There were only 2 instances of moderate rain
- Average speed for mist is high
  - There were only 3 instances of mist

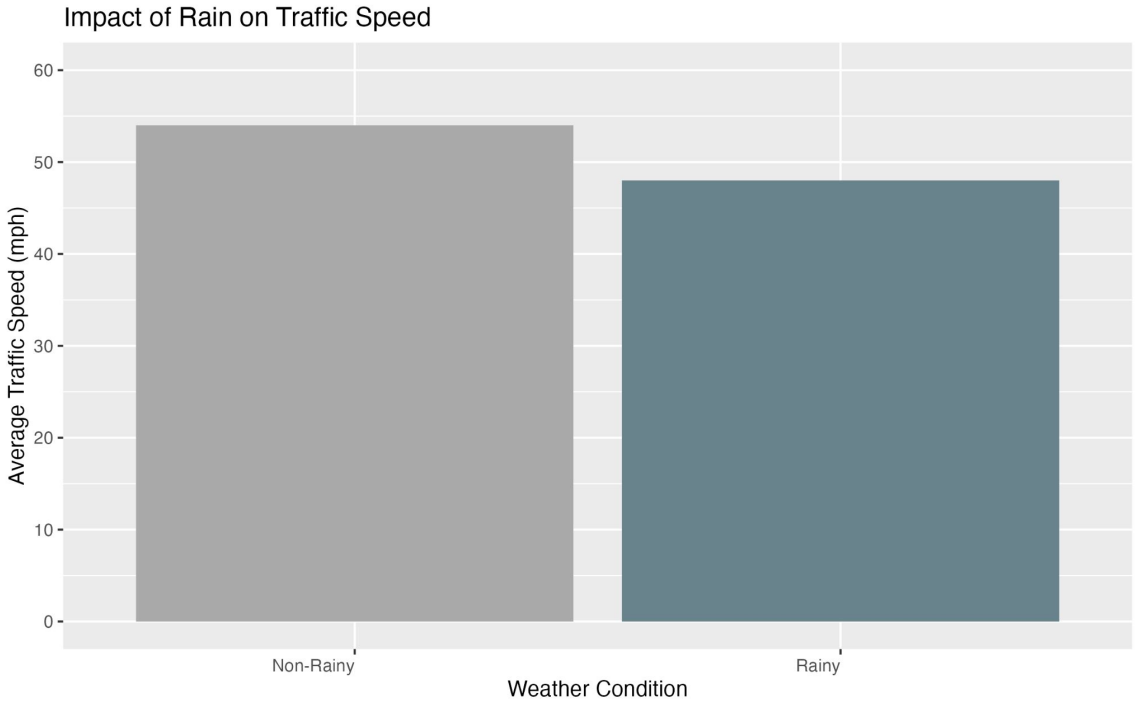
What if we take a wider look and broaden it to the weather groups?



Key Insights:

- Rain has a lower average speed compared to the other groups
  - Not many instances of Rain (about 15)
- There were more instances of Clouds than any of the other groups combined (about 95 instances)
- Mist still averaged high

What if we take a wider look?



- Grouping:
  - Clouds and Clear under Non-Rainy
  - Mist and Rain under Rainy
- Average speed under rainy conditions is lower by 8 mph
  - 116 instances of non-rainy conditions and only 18 instances of rainy conditions

Answered: Does Weather Affect Traffic Speeds?

If we just look at our graphs, we can say that rain does affect traffic speed.

BUT...

There are a lot of confounding variables and lack of data that affects our results, so we can't say for sure. For example, traffic speed is heavily affected by time of day (i.e., when people commute), and local events. Plus, our data isn't representative of all the different weather groups.

Issues & Further Research

- APIs not running on the same cron schedule
- Lack of data for certain weather conditions
    - Collect data over a longer period of time (6 months to a year)
- Scale the groups the same or balance the classes (take analysis into Python)