UBER

Abstract

Uber is a San Francisco based company founded in 2009. It is currently one of the fastest growing companies worldwide. The idea behind Uber is a car sharing, and using technology to connect potential passengers and drivers by using smart phone apps. It is popularity is surging because of its flexibility for the drivers and the convenience for the passengers. Although it gets a lot criticism in terms of market share and legal issues towards taxi drivers and the local authorities, it still continues to grow therefore attracts more people.

The main contribution of this paper is analyzing the data available for Uber in 2014 and answering some questions regarding correlation between the rider patterns and weather conditions. Secondly, it tries to show the relation between passenger and driver preferences. Moreover, this paper is also making assumptions about the economic and sociological impacts of Uber in the society by analyzing the data available.

1.Introduction

Uber Technologies Inc. is an American international transportation network company headquartered in San Francisco, California. The company develops, markets and operates the Uber mobile app, which allows consumers with smartphones to submit a trip request which is then routed to Uber drivers who use their own cars.[1]

Uber was founded as "UberCab" by Travis Kalanick and Garrett Camp in 2009 and the app was released the following June. Beginning in 2012, Uber expanded internationally in 58 companies. In 2014, it experimented with carpooling features and made other updates. By late-2015, Uber was estimated to be worth more then $50 B. [1]

Unlike a traditional taxi company, Uber does not operate its own cars. Instead provides a platform where individual drivers can log-in, use their own cars and set their own schedules. The passengers also use the smarts phones to use the app and then platform connects these two, and charges passengers for the ride. Uber gets 10-20% from each ride. There is no money exchange between the driver and the passenger.

In 2012, Uber introduced a dynamic component to pricing known as the surge multiplier. As the name suggests, fare prices are multiplied by the surge multiplier; typically, the multiplier is 1, but during times of high demand it increases. Uber has stated that there are two goals of this system: first, higher profits may increase supply by incentivizing drivers to come online. Second, higher prices may reduce demand by discouraging price-elastic customers. [2]

In this paper I aim to analyze publicly available data for Uber rides in New York in 2014, and try to answer following questions:

- What is the daily rush hours for trips
- How does usage change weekly base
- Does Uber mostly used by business people for work/home trips
- Compare week days to weekend trips
- Where does rider goes in weekdays and weekends (with some assumption)

2.Related Work

There are some papers that have studied the car sharing practice.

1. An Analysis of the Labor Market for Uber's Driver-Partners in the United States
   https://s3.amazonaws.com/uber-static/comms/PDF/Uber_Driver-Partners_Hall_Kreuger_2015.pdf

2. Analyzing Uber Ride Sharing GPS Data
   http://blogs.mathworks.com/loren/2014/09/06/analyzing-uber-ride-sharing-gps-data/

3. Making a Bayesian Model to Infer Uber Rider Destinations
   http://www.datascienceweekly.org/newsletters/data-science-weekly-newsletter-issue-41

The second related work is an analyze about San Francisco. It shows us the weekends in San Francisco is the busiest time for Uber. However, in my paper New York City, Sunday is the slowest day. In other words New York City datasets shows the Uber usage is mostly for work days, not much for weekend which is definitely opposite of San Francisco datasets. Another difference between these two paper is in data sizes. In San Francisco paper, the data has one-year period around 1.1 million record of rides, however in my analyze New York City has six months data period around 4.5 million in 6 months.

In San Francisco dataset usage peaks in early hours of the day and afternoon is the slowest time but, New York City datasets usage also peaks after 5am and mostly get the top after 3 pm.

3. Description of Data Sources

In this study, I used the public trip dataset for Uber for New York City for date range from Apr-2014 to Sep-2014 and it has 4,451,746 trip records. [3] The dataset has been anonymized by removing names and trip-end points. Date and time of day are still intact. The dataset has the information of: Trip start Date/Time, Pickup Latitude, Pickup Longitude, and Base Company.

The Base Company is unique for New York City. In many cities where Uber operates, it is purely digital service. However, New York City requires Uber to operate from bases to dispatch rides. The data set has five different base companies (Weiter, Hinter, Schmecken, Danach and Unter), all owned by Uber. [4]

In order to analyze the data, I used Python for programing, python data analysis library pandas for data frame, python interactive visualization library Bokeh and Basemap matplotlib toolkit for plotting data on geographically referenced map projections.

The data for the weather conditions in 2014, especially for specific events, was very hard to find. I derived 2014 rainfall reports data from National Weather Service Forecast Office's website. [5] I also used day by day description of the weather events happened in 2014 from Weather for Central New York website. [6] I used this data to explain and match the abnormalities in my graphs.

4. Results

After analyzing my data, I generated many graphs. However, I found a few of them useful for my analysis. Below I have explained them one by one:

Figure 1:

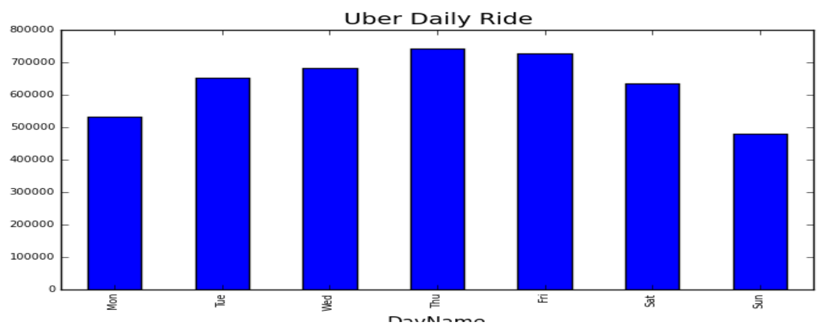
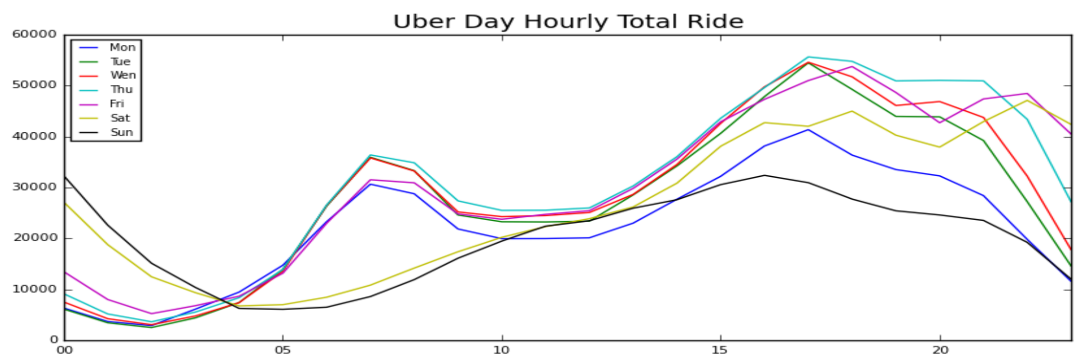


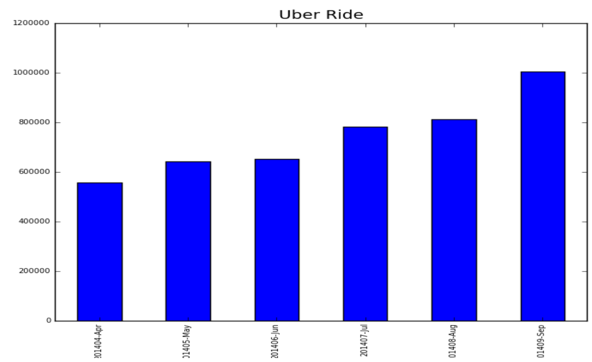
Figure 2:



The Figure 1 shows how all the data set is distributed according to days, and the Figure 2 shows how the data is distributed hours by day of week. As we mentioned above (San Francisco, related work two) I expected to see higher rides during the weekend. However, in my New York City graph Sunday is the slowest day. The busiest days are Thursday and Friday.

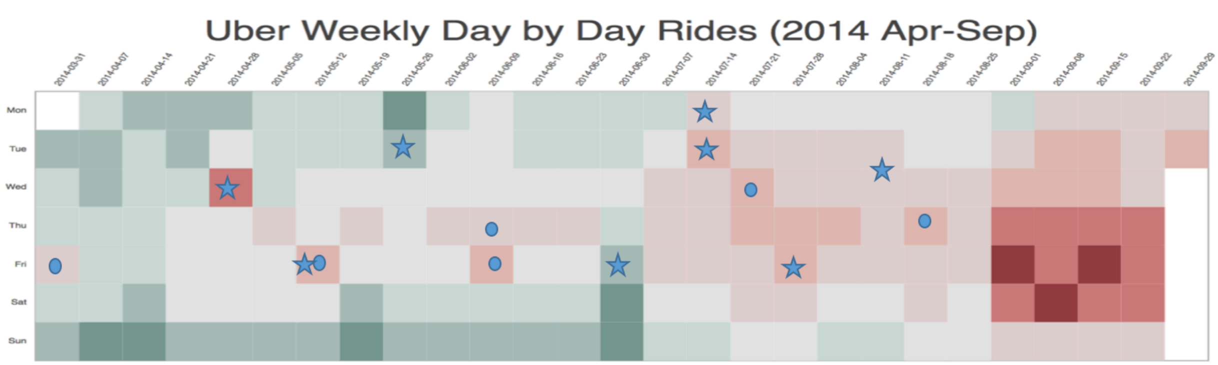
From Figure 2 we can see riders work and social patterns. For example, during the weekdays at early rush hours is busy time, whereas during the weekend at those times the lowest time of the day. Again, as we see from the graph Friday and Saturday riders tend to pick up a ride at late evening, however this is not the case when the next day is the work day. Around the afternoon rush hour (after 3pm), all days except Sunday has a steady increase.

Figure 3:



This figure shows us the monthly change in rider numbers. As we see there was a steady increase starting July, then another major shift in September. When I looked for the reasons for these changes I found out that Uber made 20% limited time price cut on July 7th. [7] Before this cut the number of rides were very low (around 10K) for especially last 2 days. However, after the cut, steady increase started (above 20K) and never went back to under 20K. Uber announced that it makes 20% summer price cut permanent on September 25th. [8] This also explains the ongoing increase in the rides. In September Uber had another increase. I assume this was a result of the start of the public schools , college and Universities fall semester.

Figure 4:



Rainfall Reports (in) Highest number for data range	Date	Ride
2.11	5/16/14	32,020
1.9	8/21/14	31,686
1.2	6/25/14	25,973
1.14	6/12/14	29,375
1.14	6/13/14	30,704
1.07	7/23/14	33,367
1.01	4/4/14	26,353

Figure 5

Significant Weather Events	Date	Ride
Heavy Rain Event	4/30/14	35,718
Heavy Rain Event	5/16/14	32,020
Straight Line Wind Damage	**5/27/14**	**16,773**
Hurricane Arthur	**7/4/14**	**13,921**
Severe Thunderstorms and Flooding	7/14/14	26,923
Severe Thunderstorms and Flooding	7/15/14	33,791
Flash Flooding	8/1/14	31,686
Historic LI Flash Flooding	8/12/14	29,137
Historic LI Flash Flooding	8/13/14	29,259
Severe Weather and Flash Flooding	8/31/14	22,781

Figure 6

The Figure 4 shows weekly day-by-day rides. It enables us to see the changes in the number of rides by colors. Darker colors mean more rides. I first spotted the days with the different colors, and then in figure 5 [5] and figure 6 [6]. As we can see on April 30th, there was a heavy rain event record. Also, many dates that are darker color on the Figure 4 have higher rainfall report number. Higher number means higher, severe weather, and all resulted in higher number of rides for our data range. The bad weather tends to result in more rides. However, there are exceptions like July 4 when Hurricane Arthur was in effect, and May 27 when heavy winds were in effect; there were less rides.

Figure 7:

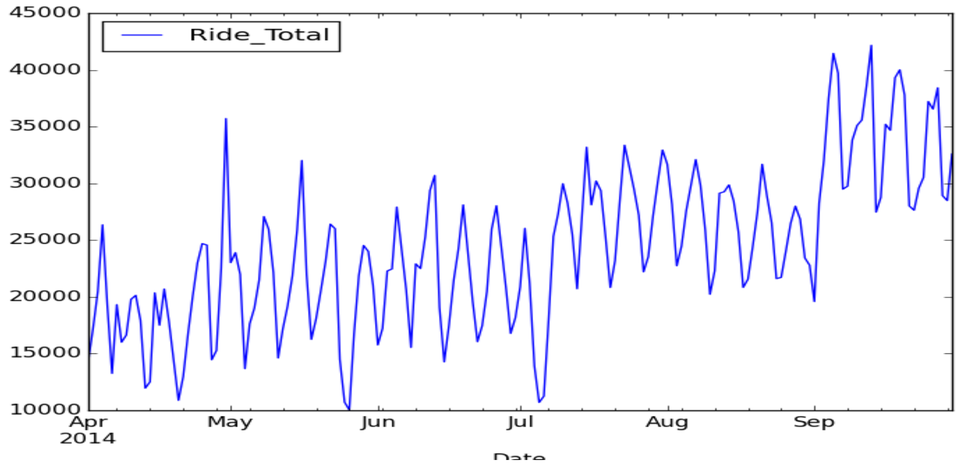
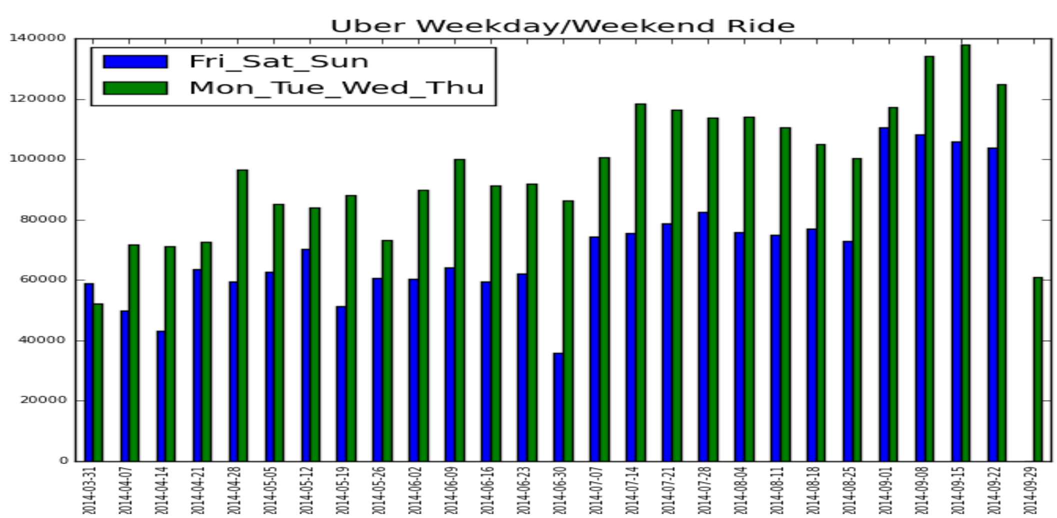


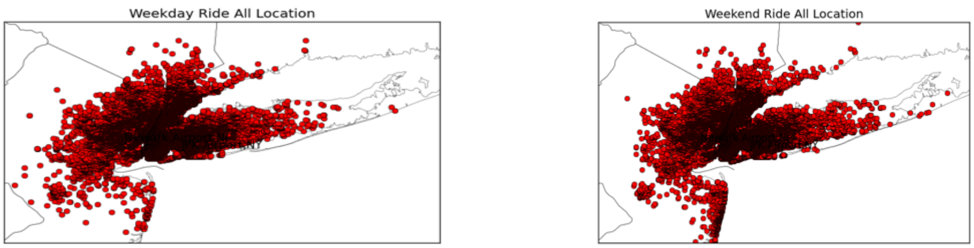
Figure 7 is another graph where we can see the differences in daily rides. May 27 and July 4 events that we mentioned on Figure 6 are visible in here too. Another trend we can see from this graph is after the Uber's price cut at July 7th a steady increase starts in rides, also after September there is another shift in rides mostly due to the opening of the school season.

Figure 8:



This figure shows the weekly rider patters. As we see from the figure 8, weekdays are higher then weekend rides.

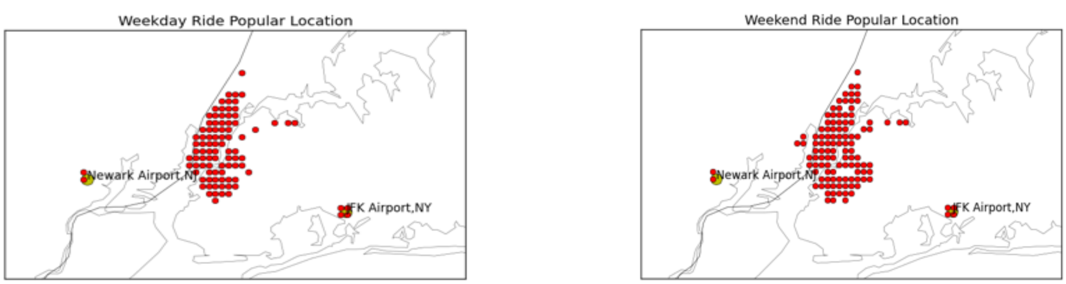
Figure 9:



When I plot the longitude and latitude data, I got messy point clouds like in this Figure 9. As you see it is hard to see the details. In general, there is not much difference in these two graphs, only in weekday figure there are more dots scattered around, whereas in weekend figure dots are usually close to the center.

In Figure 10, I tried to omit the excessive data by rounding longitude and latitude numbers by two digits to show nearby dots in the same category. By doing this I grouped the dots, and made the graph clearer. I grouped data by weekdays and weekends to see the differences for rides. I found the total rides for each group and calculated the percentage for each record. I dropped the records that has a percentage smaller than % 0.1. When I checked the total for rest of the record by weekdays and weekend, I had still % 91 of all dataset.

Figure 10:



These two figures are clearly shows us that weekday and weekend ride behavior are almost similar. Also from this figures we can see that in both graphs popular destinations are similar, like airport.

5. Discussion

As I mentioned in my proposal my aim was to answer a few questions by using the available data. As a result of reviewing the publicly available Uber data, and generating and analyzing the figures, I was able to answer almost all of my questions.

Q1: What is the daily rush hours for trips? : According to Figure 2, morning peak time is around after 6am, and afternoon is after 3pm. Because the trend is going upwards during this time, meaning there are more request for rides. As we can see, Uber exhibits regular trends, supply and demand show daily patterns that peak around rush hour.

Q2: How does usage change weekly base? : Figure 8 answers this question. I was expecting to see the more rides during the weekend because of the night time activities in New York, however the graph shows more activity during the weekdays.

Q3: Does Uber mostly used by business people for work/home trips? : I can only predict the answer for this question, because I could not find any data to give complete answer. However, because we are able to see the rush hour trends from the graphs easily, I can assume that the riders are the business people or students whose regular hours are 8am and 5pm.

Q4: Compare week days to weekend trips : I expected to see more rides during the weekend for New York City, however weekday rides were higher for all data, which is opposite of San Francisco. I can assume that this is because New York City is a business center.

Q5: Where does rider goes in weekdays and weekends (with some assumption)? : I could not find any clear difference to specify this. For example, in both times airports were popular locations.

Uber has a big social and economic impacts. Because it allows drivers work on their own flexible schedule, many people have a part-time job opportunity during their available times. This allows people to work off times, add additional income to their families. [9]

5. Conclusion and Future Work

In this paper my aim was to answer my questions, and I could be able to do so. My paper shows a correlation between weather conditions and rider request. Uber will need more driver in their platform during this time. This may help Uber drivers predict their income by following forecast.

Not enough data for detailed analyze. With data for several years there will be better analysis.

Uber has a promising future. New features of Uber, like Uberpool, Uber doing food deliveries, Uber cash in India etc. will benefit it even more. The main challenge that facing Uber now and in the future is I believe is the legal issues its facing.

Google's investment in Uber may be the start of a Nexus car (self-driving vehicles) program. [10] If so this will be the new era in transportation business.

6 References

1. McAlone, N. (2015, September 13). “Here's how Uber got its start and grew to become the most valuable startup in the world”. *Business Insider*. Retrieved December 12, 2015, from http://www.businessinsider.com/history-of-uber-and-its-rise-to-become-the-most-valuable-startup-in-the-world-2015-9

2. Mislove, A. (2015, October 28). Peeking Beneath the Hood of Uber. Retrieved December 13, 2015, from http://www.ccs.neu.edu/home/leonchen/papers/uber-imc15.pdf

3. Dataset, Uber trip record. Retrieved December 12, 2015, from https://public.opendatasoft.com/explore/dataset/uber-trip-nyc-14/?tab=export

4. Fischer, B. (2015, January 6). “NYC suspends Uber bases until they hand over trip records”. *Business Journal*. Retrieved December 12, 2015, from http://www.bizjournals.com/newyork/blog/techflash/2015/01/nyc-suspends-uber-bases-until-they-hand-over-trip.html

5. National Weather Service Forecast Office (New York, NY): Significant Weather Events Archive. Retrieved December 12, 2015, from http://www.erh.noaa.gov/okx/stormtotals.html

6. Rainfall Reports: Report for Year 2014. Retrieved December 12, 2015, from http://www.cnyweather.com/wxraindetail.php?year=2014

7. Furman, P. (2014, July 7). “Uber slashes prices by 20% in NYC for limited time, saying it's now cheaper than yellow taxi cabs”. *New York Daily News*. Retrieved December 12, 2015, from http://nydn.us/1lP4ctV

8. “Uber makes 20 percent summer price cut permanent”. (2014, September 25). *New York Post*. Retrieved December 12, 2015, from http://nypost.com/2014/09/25/uber-makes-20-percent-summer-price-cut-permanent/

9. White, G. (2015, June 8). “In the Sharing Economy, No One's an Employee”. The *Atlantic.* Retrieved December 12, 2015.

10. Abrams, S. (2013, August 9). “Is Google's investment in Uber the start of a Nexus car program?”. *Roundandtrack.* Retrieved December 12, 2015, from http://www.roadandtrack.com/car-culture/features/a5318/go-news-google-investment-in-uber-might-mean-nexus-cars/

Rainfall Reports (in)

Report for Year 2014

Data last updated 12/11/2015 10:10 PM.

Other Years:

2015 Go

Daily Detail · Monthly Summary · Seasonal Summary

Temperature | Rain | Wind | Wind Run | Wind Direction | Snow

Snow Depth | Barometric Pressure | Degree Days | Sunshine Hours | Max Solar

Solar kWh | UV | Dew Point | Wet Bulb | ET | Humidity

Day	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.00	0.33	0.00	0.00	0.37	---	0.42	0.00	0.00	0.00	0.00	0.04
2	0.00	0.00	0.05	0.00	0.02	---	0.12	0.07	0.69	0.00	0.00	0.05
3	0.00	0.00	0.00	0.00	0.04	0.46	0.99	0.86	0.00	0.03	0.00	0.08
4	0.01	0.24	0.00	1.01	0.29	0.00	0.23	0.01	0.00	0.41	0.00	0.00
5	0.10	0.24	0.02	0.06	0.00	0.00	0.00	0.00	0.00	0.01	0.05	0.05
6	0.05	0.57	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.00	0.28	0.47
7	0.00	0.00	0.00	0.52	0.00	0.00	0.67	0.02	0.01	0.00	0.03	0.00
8	0.06	0.00	0.00	0.33	0.00	0.00	0.36	0.00	0.00	0.22	0.00	0.01
9	0.09	0.07	0.00	0.00	0.22	0.11	0.34	0.00	0.00	0.00	0.11	0.33
10	0.02	0.16	0.00	0.00	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.68
11	0.52	0.06	0.00	0.11	0.00	0.84	0.00	0.00	0.21	0.00	0.00	0.17
12	0.00	0.00	0.81	0.00	0.00	1.14	0.41	0.57	0.02	0.00	0.06	0.06
13	0.10	0.00	0.05	0.11	0.31	1.14	0.26	0.95	0.43	0.00	0.04	0.02
14	0.41	0.93	0.00	0.33	0.26	0.01	0.02	0.02	0.00	0.00	0.10	0.04
15	0.00	0.08	0.12	0.65	0.00	0.00	0.02	0.16	0.00	2.25	0.00	0.00
16	0.00	0.05	0.00	0.00	2.11	0.00	0.00	0.21	0.33	0.84	0.00	0.14
17	0.00	0.02	0.00	0.00	0.01	0.81	0.00	0.10	0.00	0.03	0.86	0.13
18	0.07	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.30	0.00	0.04
19	0.03	0.02	0.06	0.03	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.04
20	0.06	0.40	0.16	0.00	0.00	0.00	0.00	0.42	0.00	0.01	0.01	0.00
21	0.00	0.60	0.06	0.00	0.00	0.00	0.00	1.90	0.73	0.07	0.46	0.00
22	0.00	0.00	0.00	0.26	0.12	0.00	0.00	0.00	0.05	0.04	0.20	0.00
23	0.00	0.00	0.00	0.08	0.08	0.00	1.07	0.00	0.00	0.10	0.14	0.16
24	0.00	0.05	0.00	0.00	0.15	0.52	0.01	0.00	0.00	0.03	0.17	0.61
25	0.00	0.03	0.00	0.20	0.00	1.20	0.00	0.00	0.00	0.01	0.00	0.10
26	0.00	0.00	0.00	0.17	0.00	0.01	0.02	0.00	0.00	0.17	0.26	0.00
27	0.06	0.07	0.00	0.01	0.00	0.00	0.59	0.00	0.00	0.01	0.03	0.00
28	0.00	0.08	0.51	0.00	0.00	0.00	0.67	0.00	0.00	0.06	0.10	0.41
29	0.00		0.61	0.61	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00
30	0.00		0.58	0.32	0.00	0.29	0.26	0.00	0.00	0.07	0.00	0.00
31	0.00		0.00		0.00		0.51	0.32		0.00		0.00
Rain Days	13	18	11	16	13	12	18	14	9	20	16	20
Month Total	1.58 in	4.00 in	3.03 in	4.80 in	4.51 in	6.61 in	6.97 in	5.64 in	2.50 in	4.75 in	2.90 in	3.63 in
YTD Total	1.58 in	5.58 in	8.61 in	13.41 in	17.92 in	24.53 in	31.50 in	37.14 in	39.64 in	44.39 in	47.29 in	50.92 in

Color Key												
< 0.25	0.25 - 0.50	0.50 - 0.75	0.75 - 1.00	1.00 - 1.25	1.25 - 1.50	1.50 - 1.75	1.75 - 2.00	2.00 - 2.25	2.25 - 2.50	2.50 - 2.75	2.75 - 3.00	3.00>

erh.noaa.gov

National Weather Service Forecast Office

New York, NY

Home News Organization

Local Forecast by "City, St" or ZIP code

City, St Go

Current Hazards
- Local Outlook
- National Outlook
- Precipitation (HPC)
- Severe (SPC)
- Tropical (NHC)
- Submit a Storm Report

Current Conditions
- Observations
- Rivers & Lakes
- AHPS

Radar and Satellite
- Local Radar
- Nationwide Radar
- Satellite Products

Forecasts
- Activity Planner
- Local Forecasts
- --Text Forecasts
- Air Quality
- Aviation
- Fire Weather
- Hydrology
- Marine
- Space Weather
- Surf Zone
- Wireless
- Experimental
- RSS/XML Feeds

Climate
- Local
- National
- More...

Significant Weather Events Archive

2014

- Jan 3rd Snow Event
- Jan 7th Wind Event
- Jan 18th Snow Squalls
- Jan 21-22nd Snow Event
- Jan 25th Snow Squalls
- Jan 29th Light Snow Event
- Feb 3rd Snow Event
- Feb 5th Snow and Ice Event
- Feb 9th Light Snow (evening)
- Feb 13-14th Nor'Easter
- Feb 15th Snow Event
- Feb 18th Light Snow Event
- Mar 13th Wind Advisory
- Mar 31st Morning Snow Band - CT/Long Island
- Apr 29-May 1st Heavy Rain Event
- May 15-17th Heavy Rain Event
- May 27th Straight Line Wind Damage Assessment in Southbury CT
- Jul 3th Severe Late Day Thunderstorms and Flooding
- Jul 4th Hurricane Arthur
- Jul 14th Severe Thunderstorms and Flash Flooding
- Jul 15th Severe Thunderstorms and Flash Flooding
- Jul 27th Wolcott Tornado
- Aug 1st Flash Flooding
- Aug 12-13th Historic LI Flash Flooding
- Aug 31st Severe Weather and Flash Flooding
- October 15-16th Heavy Rain Event
- October 22-23rd Rain and Wind Event
- November 1-2nd Rain and Wind Event
- November 14th Light Snow Event
- November 17th Heavy Rain Event
- November 24th Wind Event
- November 26th Winter Storm
- December 9th Nor'Easter