Where Should There Be More Electric Charging Hubs in New York State?

Cecilia Pershyn

CIS512

New York State is a leading state in clean and renewable energy, with a commitment to have 50% of the state's electricity supplied by renewable sources by 2030. Thanks to impressive growth in solar, wind, hydro and biomass power in New York State, plug-in electric vehicles (EVs) have particularly low greenhouse gas emissions here.

EVs are a cheap and clean alternative to gas-fueled vehicles. However, small battery storage capacity limits the amount of electric fuel that can be carried on board the vehicle. In order to make the use of EVs more convenient, and help extend vehicle mileage, there is a need to expand the electric power grid and create more public fueling stations.

In this study I will look at vehicle registrations in NYS to determine where electric vehicles are registered. I will look at where charging ports are located in the state and compare to the zip codes with high electric vehicle registrations to determine which zip codes have the highest demand for new charging ports.

Data Sources:

Data on Charging Stations https://data.ny.gov/d/7rrd-248n?category=Energy-Environment&view_name=Electric-Vehicle-Charging-Stations-in-New-York (https://data.ny.gov/d/7rrd-248n?category=Energy-Environment&view_name=Electric-Vehicle-Charging-Stations-in-New-York)

Data on Electric Vehicle Registration https://data.ny.gov/d/uu25-czyc?category=Transportation&view_name=Electric-Vehicles-per-County (https://data.ny.gov/d/uu25-czyc?category=Transportation&view_name=Electric-Vehicles-per-County)

Bibliography:

How New York Uses Renewable Energy. (n.d.). Retrieved from http://www.dec.ny.gov/energy/83070.html (http://www.dec.ny.gov/energy/83070.html)
Date Visited: 10/29/18

How New York Uses Renewable Energy. (n.d.). Retrieved from http://www.dec.ny.gov/energy/83070.html (http://www.dec.ny.gov/energy/83070.html)
Date Visited: 10/29/18

In [60]:
```
%matplotlib inline

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns; sns.set()

#use this function later in notebook to change NaN in "complete" data frame
def fix_count(x):
    if x=='NaN':
        x=0
    elif x=='':
        x=0
    elif type(x) == type(8.8):
        x=0
    return int(x)
#This method did not work.  I attempted to read the csv using a converter, b
#the count, where the NaN occurs, does not happen with the original csv file
#I combine the dataframes.  So I will work on converting the NaN later.

sns.set(style="darkgrid")
```

Electric Vehicle Registration Data

Import and read data table

In [33]:
```
vehicle_reg = pd.read_csv('Vehicle__Snowmobile__and_Boat_Registrations.csv')
```

In [34]:
```
vehicle_reg.head()
```

Out[34]:

	Record Type	VIN	Registration Class	City	State	Zip	County	Model Year	Mal
0	VEH	8995	PAS	BUFFALO	NY	14207	ERIE	1913	DE/I
1	VEH	607SR2131A	PAS	N SYRACUSE	NY	13212	ONONDAGA	1977	CI
2	VEH	537LS7D46CT083476	PAS	NEWFANE	NY	14108	NIAGARA	2012	AZUI
3	VEH	53TBH2MC1BE900166	PAS	ROCKVILLE CTR	NY	11570	NASSAU	2011	THIN
4	VEH	53G1B4A47DB000347	PAS	BROOKLYN	NY	11223	KINGS	2013	COI

Create a new table with only the relevant information to my study

In [35]:
```
veh = vehicle_reg[['City','State','Zip','County']]
veh[:5]
```

Out[35]:

	City	State	Zip	County
0	BUFFALO	NY	14207	ERIE
1	N SYRACUSE	NY	13212	ONONDAGA
2	NEWFANE	NY	14108	NIAGARA
3	ROCKVILLE CTR	NY	11570	NASSAU
4	BROOKLYN	NY	11223	KINGS

There are 9222 registered electric vehicles in NYS

In [36]:
```
len(veh)
```

Out[36]:
```
9222
```

Create a data series showing the number of electric vehicles per zip code in New York State

In [64]:
```
veh['Number of Vehicles'] = ''
reg_car = veh.groupby(['Zip'])
count_car = reg_car.agg({'Number of Vehicles':'count'}).sort_values(['Number
count_car[:5]
```

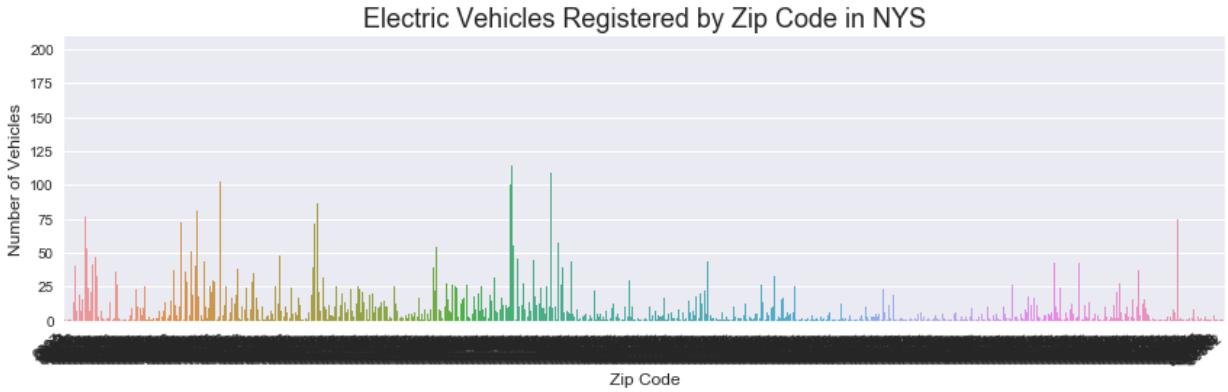
```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy (http://pandas.pydata.or
g/pandas-docs/stable/indexing.html#indexing-view-versus-copy)
  """Entry point for launching an IPython kernel.
```

Out[64]:

	Number of Vehicles
Zip	
10583	200
11746	114
11791	109
10580	103
11743	101

In [38]:
```
ax = sns.countplot(x = "Zip", data = veh)
ax.set_xticklabels(ax.get_xticklabels(),
                   rotation = 40,
                   ha='right')
ax.figure.set_size_inches(12,4)
ax.axes.set_title("Electric Vehicles Registered by Zip Code in NYS",fontsize
ax.set_xlabel("Zip Code",fontsize=12)
ax.set_ylabel("Number of Vehicles",fontsize=12)
#ax_seg = plt.legend(bbox_to_anchor = (1.05, 1), loc = 2, borderaxespad = 0.
plt.tight_layout()
plt.show()
#The best way to display this would be in GIS
```



In [40]:
```
vehicle_reg_zip = vehicle_reg.groupby(['Zip'])
vrzip = vehicle_reg_zip.size().sort_values(ascending = False)
```

In [41]:
```
vrzip = vrzip[:25]
vrzip
```

Out[41]:
```
Zip
10583    200
11746    114
11791    109
10580    103
11743    101
11050     87
10538     81
10013     77
14850     75
10514     73
11030     72
11576     72
14534     67
10504     67
10011     65
10023     60
11797     58
11747     55
11545     54
10014     53
11733     52
10528     51
11568     50
11201     49
10956     48
dtype: int64
```

Charging Hubs in New York State

Import and read data

In [42]:
```
charge_hub = pd.read_csv('Electric_Vehicle_Charging_Stations_in_New_York.csv
charge_hub.head()
```

Out[42]:

	Fuel Type Code	Station Name	Street Address	Intersection Directions	City	State	ZIP	Plus4	Station Phone	Status Code	...
0	ELEC	Hudson Valley Community College - TEC-SMART Bu...	345 Hermes Rd	NaN	Malta	NY	12020	NaN	518-629-7075	E	...
1	ELEC	EDISONPARKFAST	451 9th Ave	LOC #250 #2 LOT#250; ChargePoint America Program	New York	NY	10018	NaN	888-758-4389	E	...
2	ELEC	CARCHARGING	350 W 50th St	ICON MERCURY; Icon parking see attendant for a...	New York	NY	10019	NaN	888-758-4389	E	...
3	ELEC	CARCHARGING	310 W 39th St	ICON 310 W 39TH; Icon Parking systems see vale...	New York	NY	10018	NaN	888-758-4389	E	...
4	ELEC	EDISONPARKFAST	50 W 44th St	LOC #100 LEVEL3; Located in basement level of ...	New York	NY	10036	NaN	888-758-4389	E	...

5 rows × 32 columns

Create a new table with only the relevant information to my study

In [43]:
```
hub = charge_hub[['City','ZIP','Latitude','Longitude',]]
hub[:4]
```

Out[43]:

	City	ZIP	Latitude	Longitude
0	Malta	12020	42.971670	-73.774182
1	New York	10018	40.754512	-73.995938
2	New York	10019	40.763245	-73.988648
3	New York	10018	40.755440	-73.992027

There are 1030 charging ports in New York State

In [44]:
```
len(hub)
```

Out[44]:
```
1030
```

Create a data series showing the number of charging hugs per zip code in New York State

In [45]:
```
hub['Number of Charge Hubs'] = ''
ch_hub = hub.groupby(['ZIP'])
count_hub = ch_hub.agg({'Number of Charge Hubs':'count'}).sort_values(['Numbe
count_hub[:5]
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy (http://pandas.pydata.or
g/pandas-docs/stable/indexing.html#indexing-view-versus-copy)
  """Entry point for launching an IPython kernel.
```

Out[45]:

ZIP	Number of Charge Hubs
10019	34
10016	22
12401	21
10021	17
10022	17

In [46]:
```
charge_hub_zip = charge_hub.groupby(['ZIP'])
chzip = charge_hub_zip.size().sort_values(ascending = False)
```

In [47]:
```
chzip = chzip[:25]
chzip
```

Out[47]:
```
ZIP
10019    34
10016    22
12401    21
10022    17
10021    17
10023    16
14850    16
10065    15
12866    15
10028    14
10011    14
10024    13
11201    12
10025    12
10003    12
10036    11
10014     9
14221     9
10013     9
10075     8
11790     8
10001     8
11901     7
12205     7
11215     7
dtype: int64
```

Combine count_car and count_hub dataframes using concat, name the resulting table "complete"

In [72]:
```
frames = [count_car, count_hub]
complete = pd.concat(frames)
complete
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: FutureWar
ning: Sorting because non-concatenation axis is not aligned. A future ver
sion
of pandas will change to not sort by default.

To accept the future behavior, pass 'sort=True'.

To retain the current behavior and silence the warning, pass sort=False
```

Out[72]:

	Number of Charge Hubs	Number of Vehicles
10583	NaN	200.0
11746	NaN	114.0
11791	NaN	109.0
10580	NaN	103.0
11743	NaN	101.0
11050	NaN	87.0
10538	NaN	81.0
10013	NaN	77.0
14850	NaN	75.0
10514	NaN	73.0
11030	NaN	72.0
11576	NaN	72.0
10504	NaN	67.0
14534	NaN	67.0
10011	NaN	65.0
10023	NaN	60.0
11797	NaN	58.0
11747	NaN	55.0
11545	NaN	54.0
10014	NaN	53.0
11733	NaN	52.0
10528	NaN	51.0
11568	NaN	50.0
11201	NaN	49.0
10128	NaN	48.0

	Number of Charge Hubs	Number of Vehicles
10804	NaN	48.0
10956	NaN	48.0
10024	NaN	47.0
11768	NaN	47.0
11040	NaN	46.0
...	...	...
14216	1.0	NaN
14222	1.0	NaN
14414	1.0	NaN
14428	1.0	NaN
14445	1.0	NaN
14454	1.0	NaN
14485	1.0	NaN
14487	1.0	NaN
14526	1.0	NaN
14527	1.0	NaN
14534	1.0	NaN
14551	1.0	NaN
14569	1.0	NaN
14605	1.0	NaN
14607	1.0	NaN
14612	1.0	NaN
14615	1.0	NaN
14624	1.0	NaN
14626	1.0	NaN
14627	1.0	NaN
14642	1.0	NaN
14702	1.0	NaN
14731	1.0	NaN
14760	1.0	NaN
14817	1.0	NaN
14818	1.0	NaN
14837	1.0	NaN
14853	1.0	NaN
14870	1.0	NaN

	Number of Charge Hubs	Number of Vehicles
14881	1.0	NaN

1420 rows × 2 columns

In [78]:
```
#attempting to use replace to get rid of NaN
complete.replace('NaN',0)
```

Out[78]:

	Number of Charge Hubs	Number of Vehicles
10583	NaN	200.0
11746	NaN	114.0
11791	NaN	109.0
10580	NaN	103.0
11743	NaN	101.0
11050	NaN	87.0
10538	NaN	81.0
10013	NaN	77.0
14850	NaN	75.0
10514	NaN	73.0
11030	NaN	72.0
11576	NaN	72.0
10504	NaN	67.0
14534	NaN	67.0
10011	NaN	65.0
10023	NaN	60.0
11797	NaN	58.0
11747	NaN	55.0
11545	NaN	54.0
10014	NaN	53.0
11733	NaN	52.0
10528	NaN	51.0
11568	NaN	50.0
11201	NaN	49.0
10128	NaN	48.0
10804	NaN	48.0
10956	NaN	48.0
10024	NaN	47.0
11768	NaN	47.0
11040	NaN	46.0
...	...	...
14216	1.0	NaN

	Number of Charge Hubs	Number of Vehicles
14222	1.0	NaN
14414	1.0	NaN
14428	1.0	NaN
14445	1.0	NaN
14454	1.0	NaN
14485	1.0	NaN
14487	1.0	NaN
14526	1.0	NaN
14527	1.0	NaN
14534	1.0	NaN
14551	1.0	NaN
14569	1.0	NaN
14605	1.0	NaN
14607	1.0	NaN
14612	1.0	NaN
14615	1.0	NaN
14624	1.0	NaN
14626	1.0	NaN
14627	1.0	NaN
14642	1.0	NaN
14702	1.0	NaN
14731	1.0	NaN
14760	1.0	NaN
14817	1.0	NaN
14818	1.0	NaN
14837	1.0	NaN
14853	1.0	NaN
14870	1.0	NaN
14881	1.0	NaN

1420 rows × 2 columns

In [67]:
```
#attempt #2: use fillna() to replace NaN in dataframe with 0
total = complete['Number of Charge Hubs'].fillna(value=0, inplace=True)
total
#The kernel runs, but there is no visible output.
```

In [81]:
```
#YES! finally got it :)
#This code kills two birds with one stone, since it also converts to interge
complete = complete.fillna(0).astype(int)
complete
```

Out[81]:

	Number of Charge Hubs	Number of Vehicles
10583	0	200
11746	0	114
11791	0	109
10580	0	103
11743	0	101
11050	0	87
10538	0	81
10013	0	77
14850	0	75
10514	0	73
11030	0	72
11576	0	72
10504	0	67
14534	0	67
10011	0	65
10023	0	60
11797	0	58
11747	0	55
11545	0	54
10014	0	53
11733	0	52
10528	0	51
11568	0	50
11201	0	49
10128	0	48
10804	0	48
10956	0	48
10024	0	47
11768	0	47
11040	0	46
...	...	...

	Number of Charge Hubs	Number of Vehicles
14216	1	0
14222	1	0
14414	1	0
14428	1	0
14445	1	0
14454	1	0
14485	1	0
14487	1	0
14526	1	0
14527	1	0
14534	1	0
14551	1	0
14569	1	0
14605	1	0
14607	1	0
14612	1	0
14615	1	0
14624	1	0
14626	1	0
14627	1	0
14642	1	0
14702	1	0
14731	1	0
14760	1	0
14817	1	0
14818	1	0
14837	1	0
14853	1	0
14870	1	0
14881	1	0

1420 rows × 2 columns

Next Steps

1. Combine count_car and count_hub data series into one data frame to compare the number of cars in each zip code and the number of charging hubs. This will help give an idea of demand

for hubs. (*Completed*)

2. Map number of cars per zip code and the locations of charging hubs using geocoding and ArcGIS

In [ ]: