hw1_rl3089

February 3, 2020

1 HW1 Applied Machine Learning

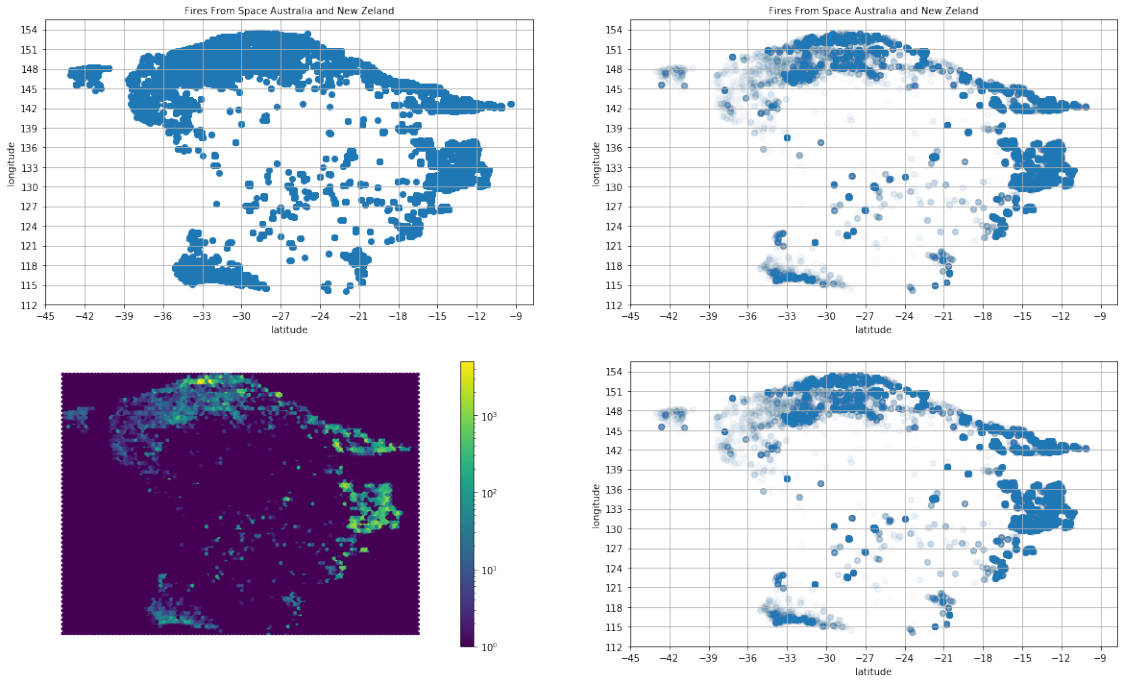
[5]:
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
fire=pd.read_csv("fire_archive_V1_96617.csv")
```

Task 1: Density Plots #1.1 Plot the longitude vs latitude several ways within a single figure (each in its own axes):

[48]:
```
fig, axes = plt.subplots(2, 2,figsize=(20, 12))
x=fire['latitude']
y=fire['longitude']

#first graph
axes[0,0].scatter(x,y)
axes[0,0].set_xlabel("latitude")
axes[0,0].set_ylabel("longitude")
axes[0,0].grid()
axes[0,0].set_title('Fires From Space Australia and New Zeland',size=10)

my_x_ticks = np.arange(-45, -7, 3)
my_y_ticks = np.arange(112, 155, 3)
axes[0,0].set_xticks(my_x_ticks)
axes[0,0].set_yticks(my_y_ticks)

#second graph
axes[0,1].scatter(x,y,alpha=.01)
axes[0,1].set_xlabel("latitude")
axes[0,1].set_ylabel("longitude")
axes[0,1].grid()
axes[0,1].set_title('Fires From Space Australia and New Zeland',size=10)


axes[0,1].set_xticks(my_x_ticks)
axes[0,1].set_yticks(my_y_ticks)

#third graph
```

```
X=[x.tolist(),y.tolist()]
im3=axes[1,0].hexbin(x, y, bins='log')
axes[1,0].axis("off")
plt.colorbar(im3,ax=axes[1,0])


#fourth graph
fire1=fire.sample(frac=0.2)
x1=fire1['latitude']
y1=fire1['longitude']

axes[1,1].scatter(x,y,alpha=.01)
axes[1,1].set_xlabel("latitude")
axes[1,1].set_ylabel("longitude")
axes[1,1].grid()
#axes[1,1].set_title('Fires From Space Australia and New Zeland',size=10)
axes[1,1].set_xticks(my_x_ticks)
axes[1,1].set_yticks(my_y_ticks)
```

[48]:
```
[<matplotlib.axis.YTick at 0x11fb7afd0>,
 <matplotlib.axis.YTick at 0x11fb7acf8>,
 <matplotlib.axis.YTick at 0x121ce23c8>,
 <matplotlib.axis.YTick at 0x11bf3e048>,
 <matplotlib.axis.YTick at 0x11bf3e470>,
 <matplotlib.axis.YTick at 0x119ca39e8>,
 <matplotlib.axis.YTick at 0x11ec7c2e8>,
 <matplotlib.axis.YTick at 0x11ec7ca58>,
 <matplotlib.axis.YTick at 0x119ca31d0>,
 <matplotlib.axis.YTick at 0x11bf3eeb8>,
 <matplotlib.axis.YTick at 0x11ec8dc18>,
 <matplotlib.axis.YTick at 0x11ec8dd30>,
 <matplotlib.axis.YTick at 0x11fba1f60>,
 <matplotlib.axis.YTick at 0x11fba1358>,
 <matplotlib.axis.YTick at 0x11fba1a58>]
```



2.2

The anomalies (measurements) most located in :

#1 Latitude (-33,-27) with Longitude(151,154)

#2 Latitude (-18,-12) with Longitude(129,137)

#3 Latitude (-18,-12) with Longitude(141,144)

Task 2: Visualizing class membership

2.0

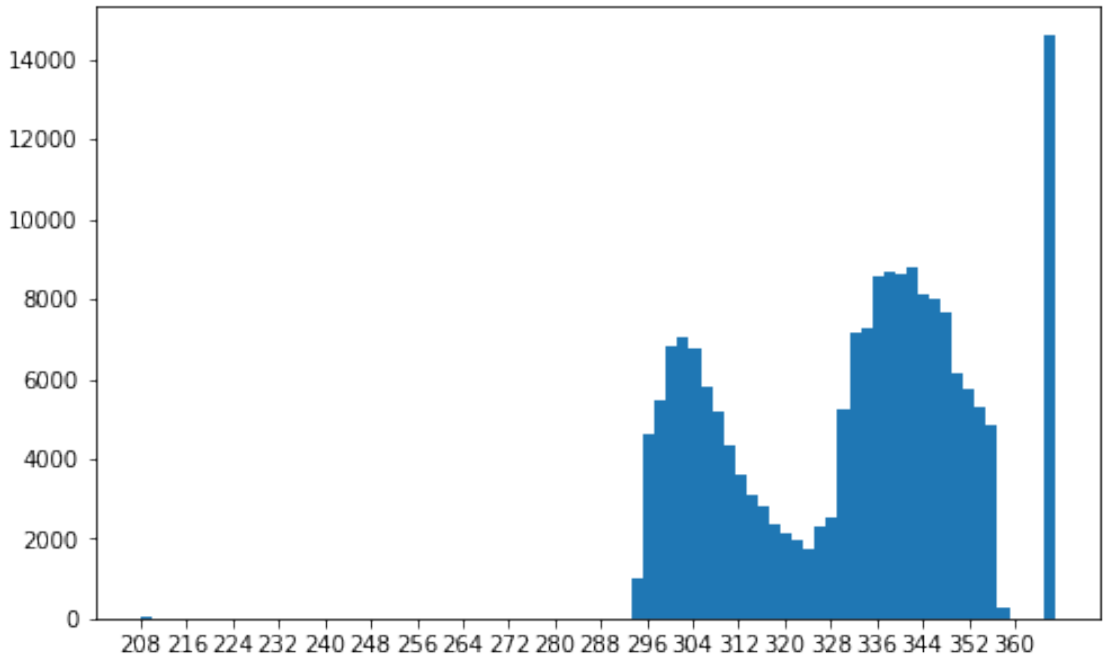
Visualize the distribution of Brightness temperature I-4 as a histogram (with appropriate settings).

Through the graph below, we are certain on the value of 267, it is saturated.

[50]:
```
bright=fire["bright_ti4"]
plt.figure(figsize=(8,5))
plt.hist(bright,bins=80)
plt.xticks(np.arange(min(bright),max(bright),8))
```

[50]:
```
([<matplotlib.axis.XTick at 0x124b458d0>,
  <matplotlib.axis.XTick at 0x124b45208>,
  <matplotlib.axis.XTick at 0x123cab400>,
  <matplotlib.axis.XTick at 0x11ed6e208>,
  <matplotlib.axis.XTick at 0x11ed6e668>,
  <matplotlib.axis.XTick at 0x11ed6eac8>,
  <matplotlib.axis.XTick at 0x11ed6ed68>,
  <matplotlib.axis.XTick at 0x11ed622e8>,
  <matplotlib.axis.XTick at 0x11ed62860>,
  <matplotlib.axis.XTick at 0x11ed62dd8>,
  <matplotlib.axis.XTick at 0x11ed5a390>,
```

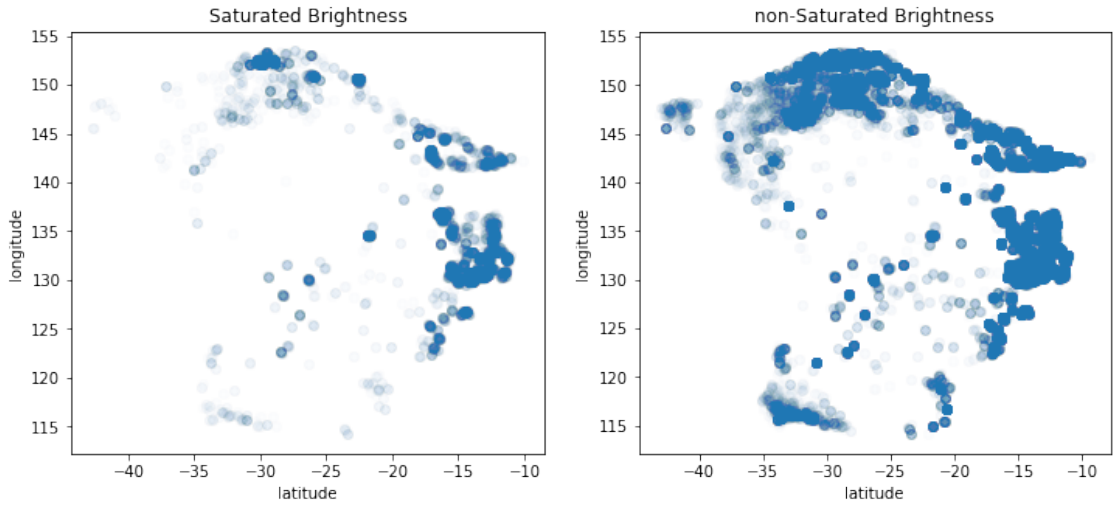
```
 <matplotlib.axis.XTick at 0x11ed5a908>,
 <matplotlib.axis.XTick at 0x11ed5ae80>,
 <matplotlib.axis.XTick at 0x11ed4d438>,
 <matplotlib.axis.XTick at 0x11ed5a780>,
 <matplotlib.axis.XTick at 0x11ed627b8>,
 <matplotlib.axis.XTick at 0x11ed4d860>,
 <matplotlib.axis.XTick at 0x11ed4de48>,
 <matplotlib.axis.XTick at 0x11ed38400>,
 <matplotlib.axis.XTick at 0x11ed38978>],
<a list of 20 Text xticklabel objects>)
```



2.1 Do a small multiples plot of whether the brightness is saturated

[55]:
```
fig, axe = plt.subplots(1, 2,figsize=(12,5))
fire1=fire.loc[fire['bright_ti4']==367]
axe[0].scatter(fire1['latitude'],fire1['longitude'],alpha=.01)
axe[0].set_title("Saturated Brightness")
axe[0].set_xlabel("latitude")
axe[0].set_ylabel("longitude")

fire2=fire.loc[fire['bright_ti4']<367]
axe[1].scatter(fire2['latitude'],fire2['longitude'],alpha=.01)
axe[1].set_title("non-Saturated Brightness")
axe[1].set_xlabel("latitude")
axe[1].set_ylabel("longitude")
```

[55]:
```
Text(0, 0.5, 'longitude')
```



Answer:

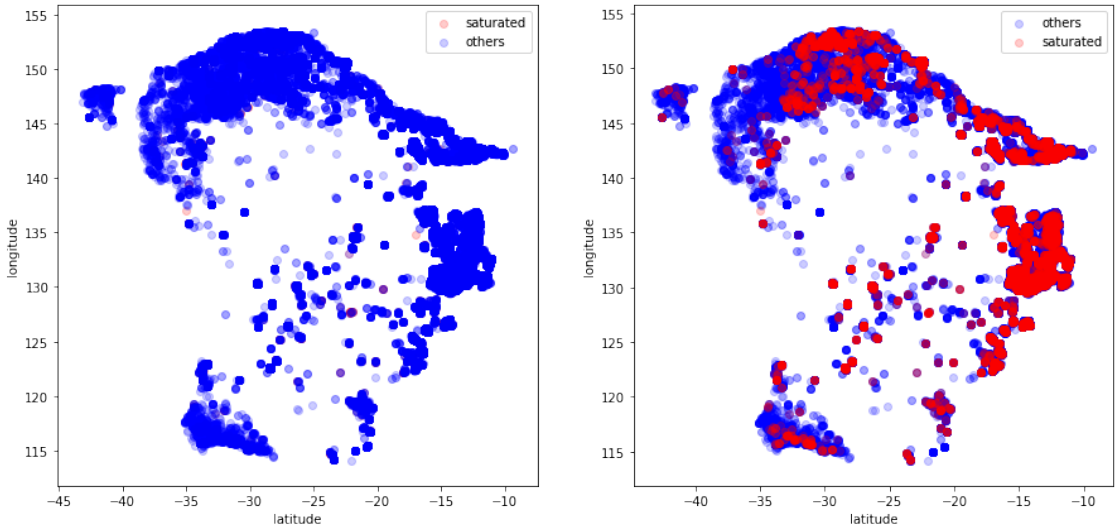
we can find that in the area whose longitude is below 125, brightness is rarely saturated. However, several groups of non-saturated brightness exists in this area.

Saturated brightness located upon logitude 145 is less aggregated than non-saturated brightness

2.2 Plot both groups in the same axes with different colors.

[53]:
```
fig, axe = plt.subplots(1, 2,figsize=(15,7))
fire1=fire.loc[fire['bright_ti4']==367]
fire2=fire.loc[fire['bright_ti4']<367]

axe[0].
 ↪scatter(fire1['latitude'],fire1['longitude'],c="r",label="saturated",alpha=0.
 ↪2)
axe[0].
 ↪scatter(fire2['latitude'],fire2['longitude'],c="b",label="others",alpha=0.2)
axe[0].set_xlabel("latitude")
axe[0].set_ylabel("longitude")
axe[0].legend(loc='best')


axe[1].
 ↪scatter(fire2['latitude'],fire2['longitude'],c="b",label="others",alpha=0.2)
axe[1].
 ↪scatter(fire1['latitude'],fire1['longitude'],c="r",label="saturated",alpha=0.
 ↪2)
axe[1].set_xlabel("latitude")
axe[1].set_ylabel("longitude")
axe[1].legend(loc='best')
```

[53]:
```
<matplotlib.legend.Legend at 0x121f81d68>
```



Answer:

The first graph is mainly covered by the blue points—"non-saturated". Those "saturated" points can be telled in the second graph, as its quantity is far smaller that of "non-saturated" ones.

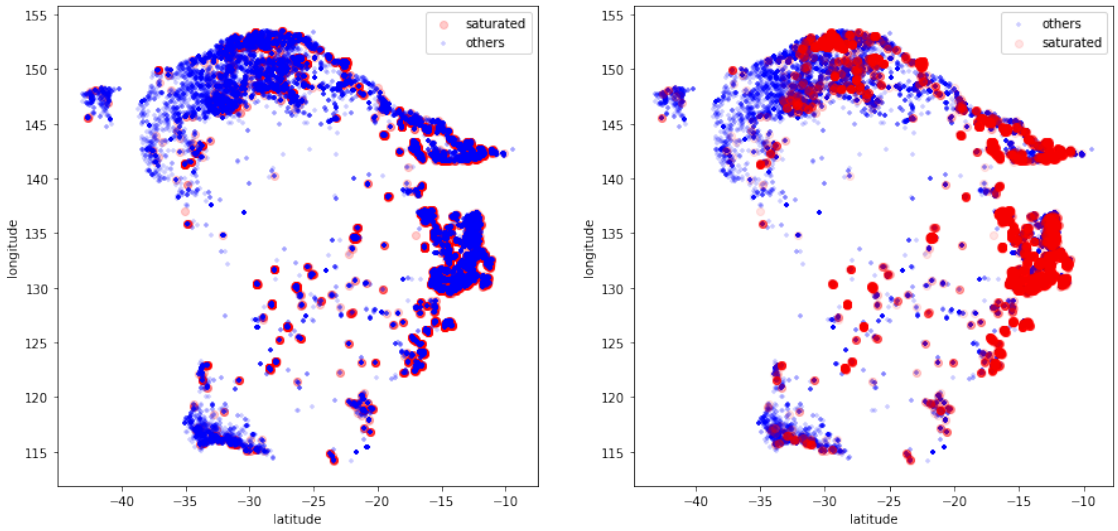
By plotting both groups in the same axes, we are able to tell their location differences much easier.

2.3 Can you find a better way to compare the two distributions?

[57]:
```
fig, axe = plt.subplots(1, 2,figsize=(15,7))
fire1=fire.loc[fire['bright_ti4']==367]
fire2=fire.loc[fire['bright_ti4']<367]

axe[0].
 ↪scatter(fire1['latitude'],fire1['longitude'],c="r",label="saturated",alpha=0.
 ↪2)
axe[0].
 ↪scatter(fire2['latitude'],fire2['longitude'],c="b",label="others",alpha=0.
 ↪2,marker='+',s=20)
axe[0].set_xlabel("latitude")
axe[0].set_ylabel("longitude")
axe[0].legend(loc='best')


axe[1].
 ↪scatter(fire2['latitude'],fire2['longitude'],c="b",label="others",alpha=0.
 ↪2,marker='+',s=20)
axe[1].
 ↪scatter(fire1['latitude'],fire1['longitude'],c="r",label="saturated",alpha=0.
 ↪1)
axe[1].set_xlabel("latitude")
axe[1].set_ylabel("longitude")
```

```
axe[1].legend(loc='best')
```

[57]: `<matplotlib.legend.Legend at 0x11daddf28>`



Answer:

By adjusting the marker size and changed the “others” symbol to “+”, we can better tell the diferences.