Fire Safety Recommendation

Introduction

As per the survey conducted it has been noticed that fire safety for many of the nursing providers has been of concern. There is a concern on how to move with the solution. The annual number of fires in health care properties reported to United States fire departments has followed a moderate downward trend since 2003. After falling between 2003 and 2010, the estimated number of fires has increased each year since 2011.

There is an urgent need of machine learning algorithm to assist the professionals and nursing home providers to make an effective decisions to tackle all fire safety deficiencies.

Data and Methodology Section

Data was collected from https://healthdata.gov/dataset/fire-safety-deficiencies.

I cleaned the data by sorting it based on severity code. Pre-processed the data using StandardScaler.

I applied the K-Means to cluster different part of country based on Fire Safety zone. Formed the clustered dataset.

```
array([[-0.0612209 ,  1.31154847],
       [16.33428994,  1.31154847],
       [-0.0612209 , -1.1337079 ],
       ...,
       [-0.0612209 ,  1.31154847],
       [-0.0612209 , -1.1337079 ],
       [-0.0612209 ,  0.08892029]])
```

 Found the centroids

```
array([[5.01482702e+00, 7.66062603e-03, 3.00000000e+00],
       [3.00000000e+00, 3.21913997e-03, 1.49113743e+00],
       [2.79013635e+00, 2.49492312e-03, 3.00000000e+00],
       [5.01031339e+00, 6.18803646e-03, 2.00000000e+00],
       [1.94769286e+00, 3.53103118e-03, 1.40110120e+00],
       [4.00000000e+00, 2.93754774e-03, 3.00000000e+00],
       [4.00000000e+00, 3.15633123e-03, 2.00000000e+00],
       [5.00695145e+00, 2.75935261e-03, 1.00000000e+00],
       [4.00000000e+00, 3.38367066e-03, 1.00000000e+00]])
```

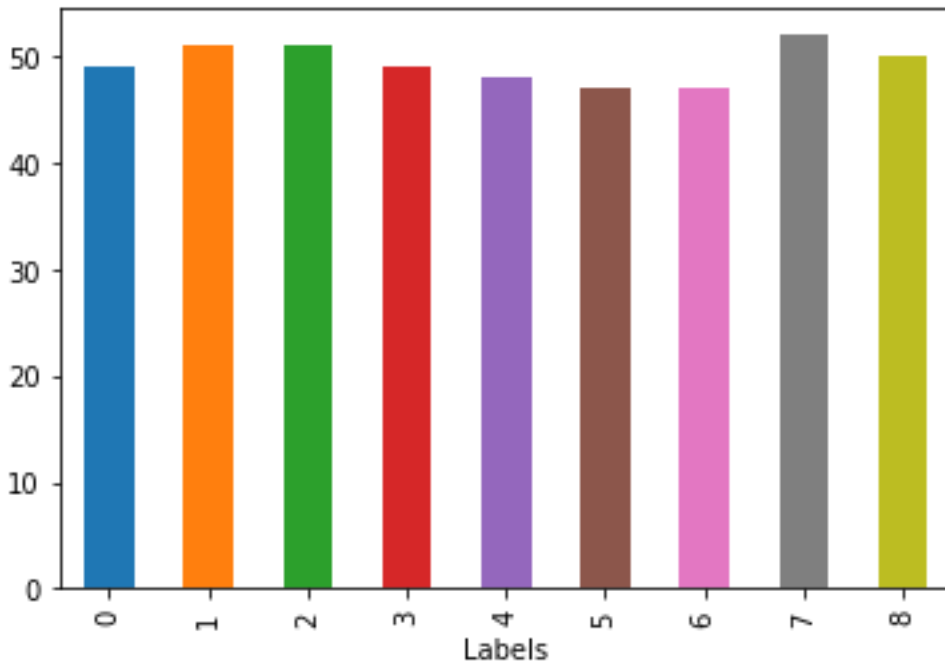Labels was created and attached the Dataframe.

And found the average severity

Labels	Scope Severity Code
0.0	5.014827
1.0	3.000000
2.0	2.790136
3.0	5.010313
4.0	1.947693
5.0	4.000000
6.0	4.000000
7.0	5.006951
8.0	4.000000

Plotted the graph.
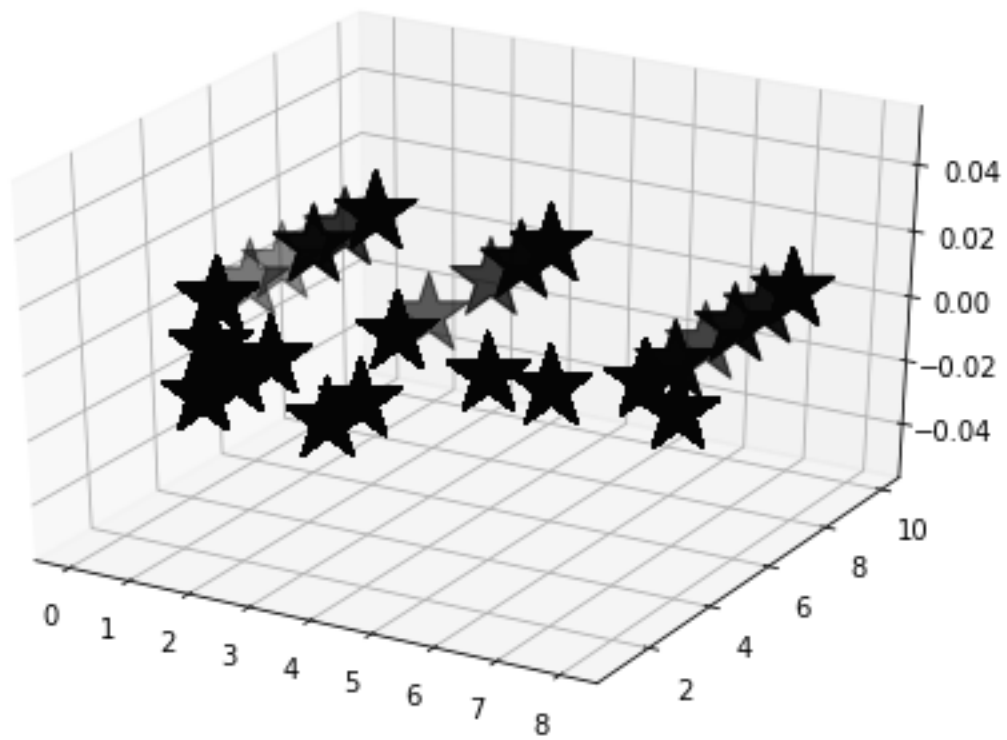
Visualize based on State

After this is done I used FourSquare to know rehabilitation center and found following

```
{'meta': {'code': 200, 'requestId': '5d371bd0a1979300374177fe'},
 'response': {'venues': [{'categories': [{'icon': {'prefix': 'https://ss3.
4sqi.net/img/categories_v2/building/conventioncenter_',
       'suffix': '.png'},
     'id': '4bf58dd8d48988d1ff931735',
     'name': 'Convention Center',
     'pluralName': 'Convention Centers',
     'primary': True,
     'shortName': 'Convention Center'}],
   'hasPerk': False,
   'id': '57d0e244cd10ec59d8f5429f',
   'location': {'address': '6020 Marian Dr',
    'cc': 'US',
    'city': 'Baltimore',
    'country': 'United States',
    'distance': 316,
    'formattedAddress': ['6020 Marian Dr',
     'Baltimore, MD 21215',
     'United States'],
    'labeledLatLngs': [{'label': 'display',
      'lat': 39.346109,
      'lng': -76.711121}],
```

```
      'lat': 39.346109,
      'lng': -76.711121,
      'postalCode': '21215',
      'state': 'MD'},
     'name': 'New Psalmist Retreat & Meeting Center',
     'referralId': 'v-1563892688'}]}}
```

Applied Logistic regression to predict if any deficiency is complaint or not.

```
array(['0', '0', '0', ..., '0', '0', '0'], dtype='<U21')
```

Probability of this prediction was

```
array([[9.99958355e-01, 4.16454238e-05], [9.99958355e-01, 4.16454238e-05],
       [9.99958355e-01, 4.16454238e-05], ..., [9.69241688e-01, 3.07583116e-02],
       [9.69241688e-01, 3.07583116e-02], [9.69241688e-01, 3.07583116e-02]])
```

Conclusion

As we see severity level 7 is occurring many of the states. We should concentrate on that first. Based on cluster we can apply different safety mechanism

To solve the problem, We should cluster different fire deficiency zone and action accordingly. We also did Mapping to show the area. We used Logistic to know the probability of complaint deficiency