

# Applied Data Science Capstone -Final Project-

By  
Kavan Zommers

# Introduction

- We will be taking a deeper look at each of the zip codes throughout the State of Rhode Island, determining the top venues in each, the number of gyms, grocery stores, restaurants, and hotels, and the distance to the nearest major airport. This will help to determine the most appropriate area to stay in when traveling there for work on a weekly basis.

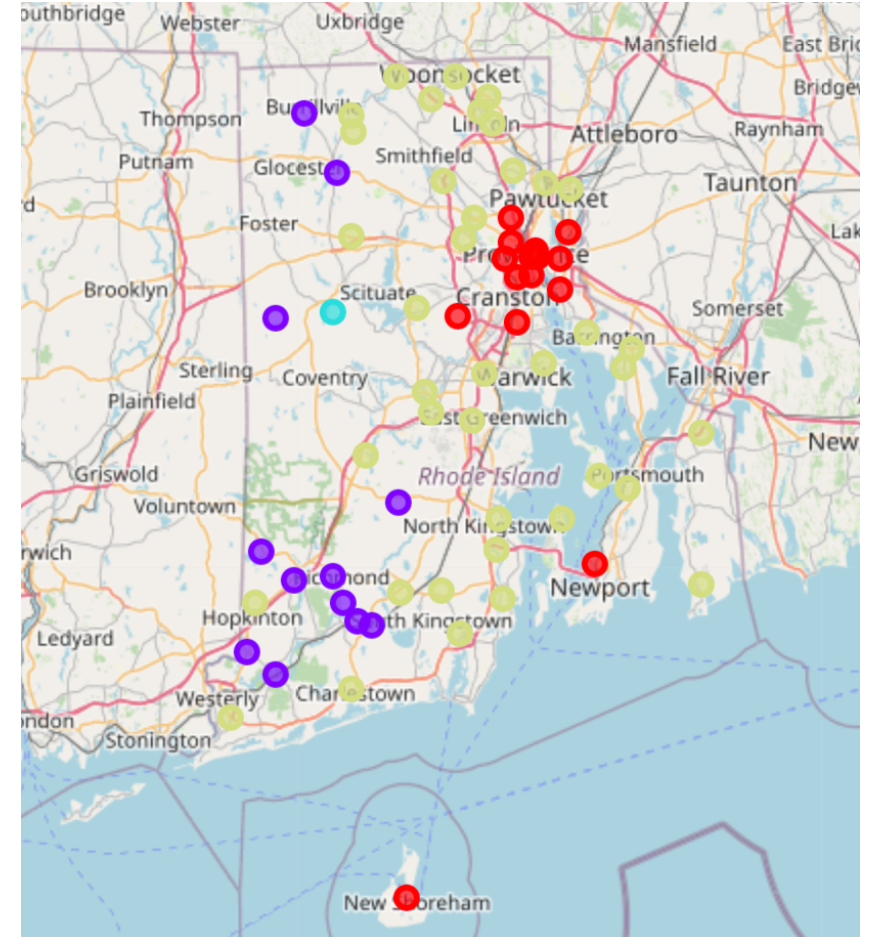


# Data

	PostalCode	City	Latitude	Longitude	Proximity to Airport [mi]
0	02885	Warren	41.728366	-71.270210	8.198661
1	02809	Bristol	41.710403	-71.281610	7.645372
2	02806	Barrington	41.744013	-71.331062	5.256922
3	02921	Cranston, Providence	41.760810	-71.503702	4.699649
4	02919	Johnston	41.839760	-71.495033	8.794294
5	02917	Smithfield	41.895466	-71.522751	12.893345
6	02916	East Providence, Bristol	41.845920	-71.356046	9.309410
7	02915	East Providence, Bristol	41.788613	-71.366901	5.569499
8	02914	East Providence, Bristol	41.818357	-71.365960	7.368240
9	02912	Providence	41.826846	-71.401021	7.344975
10	02911	North Providence	41.859267	-71.483031	9.852921
11	02909	Providence	41.819257	-71.440402	6.709654
12	02908	Providence	41.834972	-71.432172	7.768633
13	02907	Providence	41.800784	-71.424344	5.411361
14	02906	Elmwood, Providence	41.821167	-71.396328	7.014517
15	02905	Providence	41.801115	-71.405824	5.555688
16	02904	Providence	41.859097	-71.434372	9.436165

# Methodology

A K-means clustering algorithm was applied to the postal code and venue data to determine the appropriate clusters with similar characteristics.



```
print('The best locations to stay in RI will be found in Cluster',cluster_info['Score'].idxmax())
```

The best locations to stay in RI will be found in Cluster 3

**Now that the highest-scoring cluster has been identified, we will figure out which zip code within that cluster has the closest proximity to an airport to determine the city in which we should stay.**

```
RI_cluster3 = RI_data.merge(cluster3,on = 'PostalCode',how = 'inner')
```

```
print('When traveling to Rhode Island, based on your preferences you should plan on staying in',RI_cluster3.iloc[RI_cluster3['Proximity to Airport [
```

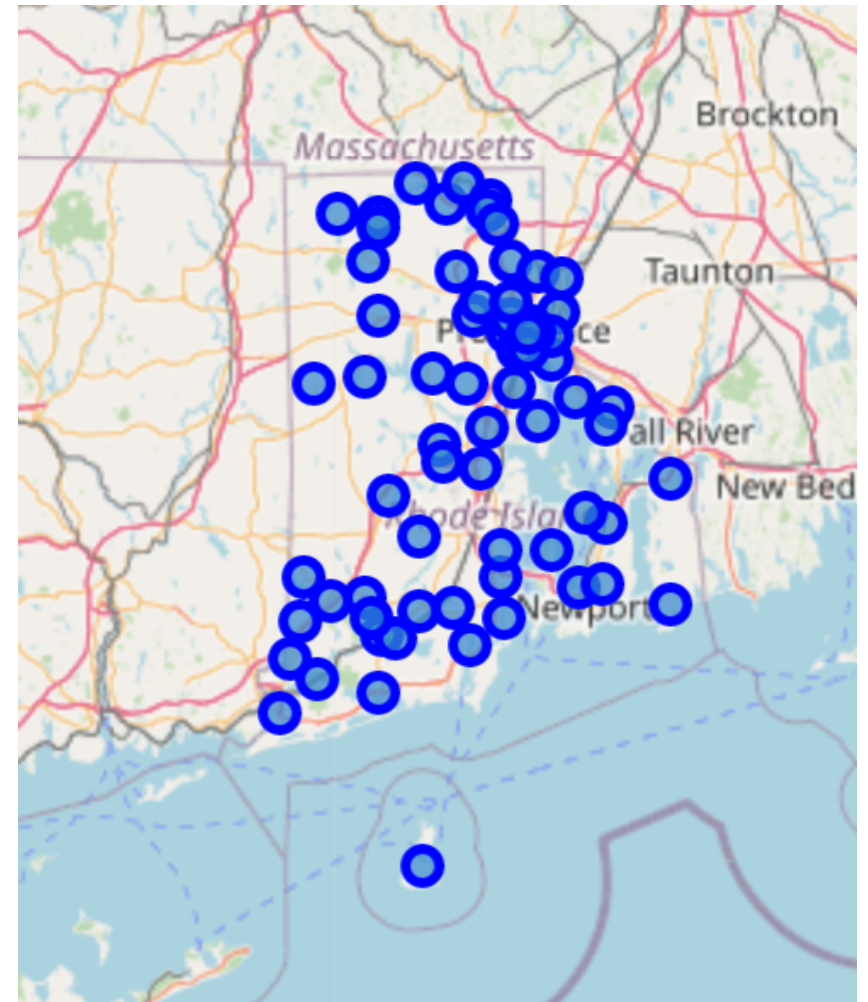
When traveling to Rhode Island, based on your preferences you should plan on staying in Warwick, Bristol

**Warwick / Bristol it is!**

# Results

## Discussion

One of the initial observations made when plotting the zip codes on the map is that there is a zip code being captured for an island just off of the coast. As this would not be a feasible location to stay, it may have been appropriate to remove this code from the data used for clustering.



# Conclusion

- Based on the preferred venues, the most appropriate area to stay in when traveling to Rhode Island is Warwick / Bristol!