

The image is a composite. On the right, a close-up of a person's face wearing a blue surgical mask. On the left, a large, detailed green virus particle with many spikes. In the bottom right corner, there is a blue network diagram with nodes and connecting lines.

Analysis of the **new crown epidemic**  
defense capabilities of CA county

---

# BACKGROUND



**Corona Virus Disease** (Corona Virus Disease 2019, COVID-19), referred to as "new coronary pneumonia", named "2019 coronavirus disease" by the World Health Organization, refers to the pneumonia caused by the 2019 new coronavirus infection.

Since December 2019, some hospitals in **Wuhan City**, Hubei Province have successively discovered multiple cases of pneumonia of unknown cause with a history of exposure to the South China seafood market, which were confirmed to be acute respiratory infectious diseases caused by the 2019 novel coronavirus infection.

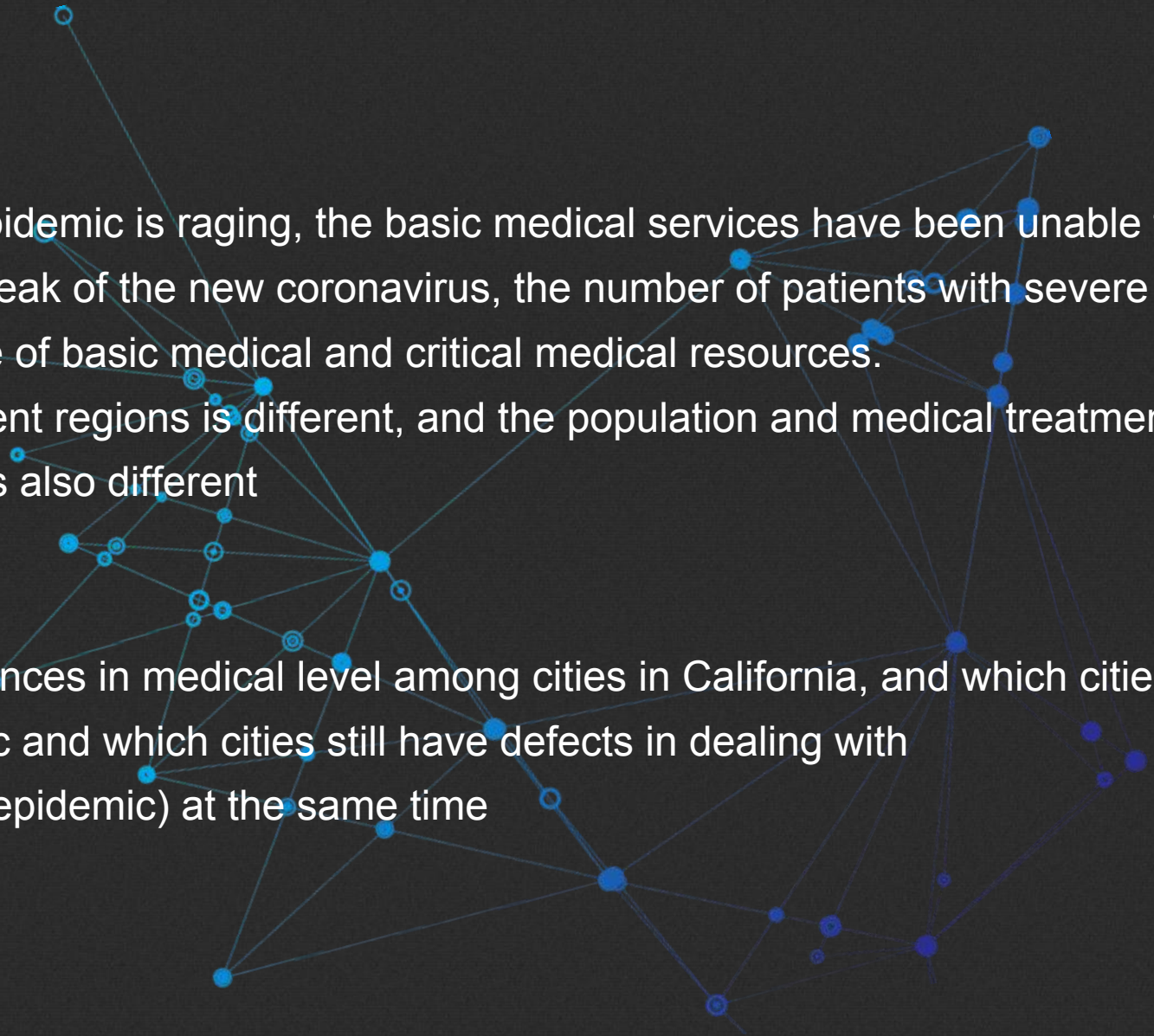
On September 3, the co-chair of the WHO's new crown pneumonia epidemic response assessment expert group announced the list of members of the expert group at a briefing for member states, and Zhong Nanshan was selected. As of 16:59 CET on October 4, 2022 (22:59 Beijing time), the number of confirmed cases worldwide increased by **171,391** from the previous day to **615,777,700**; the number of deaths increased by **662** to **6,527,192**.



# PROBLEM

Today, when the new coronavirus epidemic is raging, the basic medical services have been unable to meet the daily needs. After the outbreak of the new coronavirus, the number of patients with severe diseases has surged, followed by the shortage of basic medical and critical medical resources. The economic development of different regions is different, and the population and medical treatment reserve levels are different, so as to Ability is also different

This article mainly reveals the differences in medical level among cities in California, and which cities have good ability to cope with the new epidemic and which cities still have defects in dealing with extreme epidemic (such as the new epidemic) at the same time



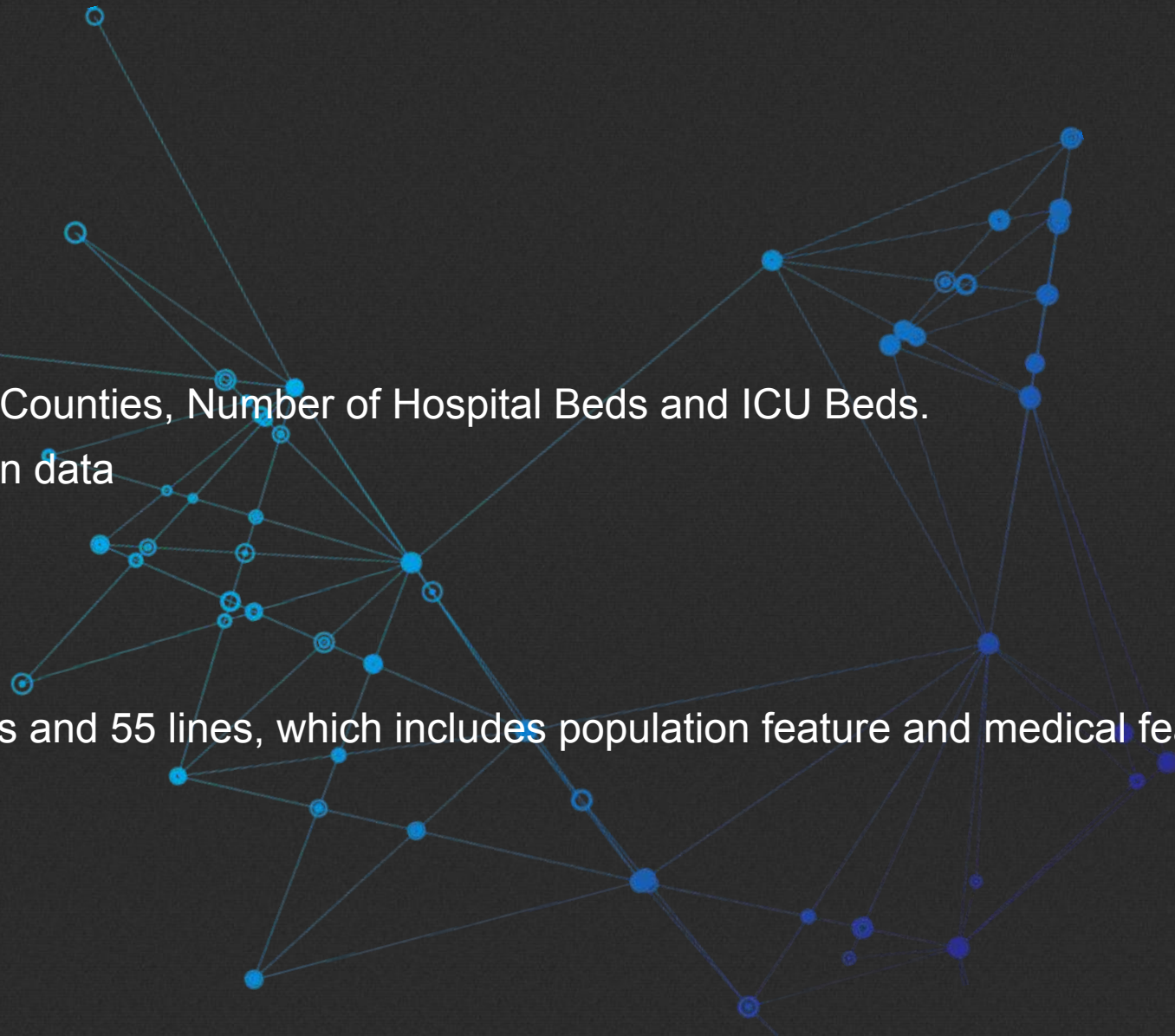
# DATA

## DATA DESCRIPTION

this data include two source :

1. California State Dataset with Counties, Number of Hospital Beds and ICU Beds.
2. California Counties Population data

the final data contains 9 features and 55 lines, which includes population feature and medical feature





# DATA

## DATA CLEANING

Data cleaning, as the name implies, is to filter and remove the duplicate and redundant data in the data to be used; Complete the missing parts and correct or delete the incorrect data. Finally, the data can be further processed and used.

this is my final data,include 8 feature eg:county/all\_hospital\_beds/icu\_available\_beds/lat/Ing/population/bed\_pr\_100\_peo/  
icu\_bed\_per\_100\_people

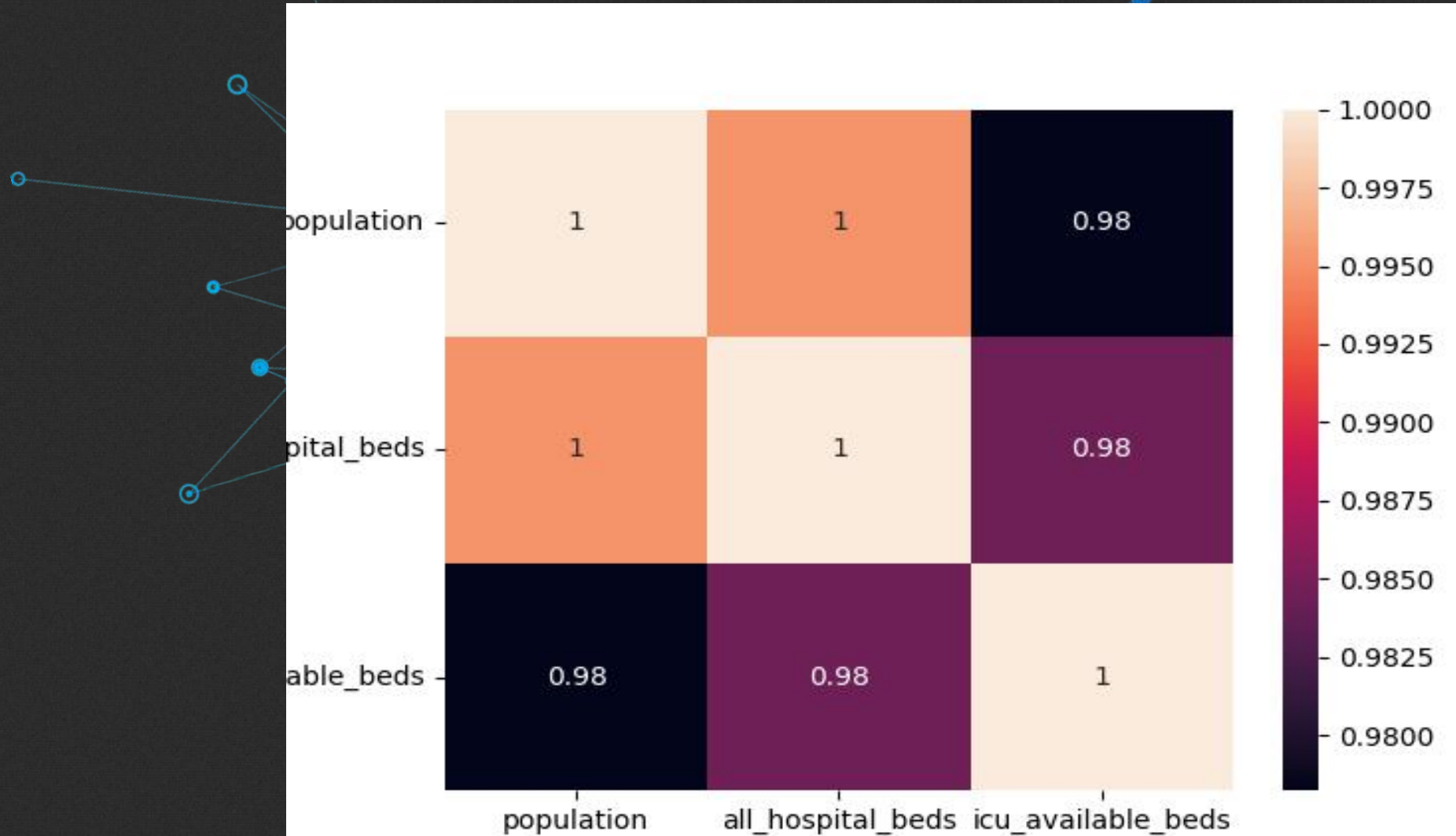
### FINAL DATA

```
cal_df['Bed_per_100_people'] = (cal_df['all_hospital_beds']/cal_df['population'])*100  
cal_df['ICU_Bed_per_100_people'] = (cal_df['icu_available_beds']/cal_df['population'])*100  
cal_df.head()
```

|   | county    | all_hospital_beds | icu_available_beds | lat     | Ing       | population | Bed_per_100_people | ICU_Bed_per_100_people |
|---|-----------|-------------------|--------------------|---------|-----------|------------|--------------------|------------------------|
| 0 | Alameda   | 2496.0            | 94.0               | 37.6469 | -121.8889 | 1656754    | 0.150656           | 0.005674               |
| 1 | Amador    | 54.0              | 2.0                | 38.4464 | -120.6511 | 38429      | 0.140519           | 0.005204               |
| 2 | Butte     | 414.0             | 11.0               | 39.6669 | -121.6007 | 225817     | 0.183334           | 0.004871               |
| 3 | Calaveras | 33.0              | 2.0                | 38.2046 | -120.5541 | 45514      | 0.072505           | 0.004394               |
| 4 | Colusa    | 48.0              | 0.0                | 39.1775 | -122.2370 | 21454      | 0.223735           | 0.000000               |

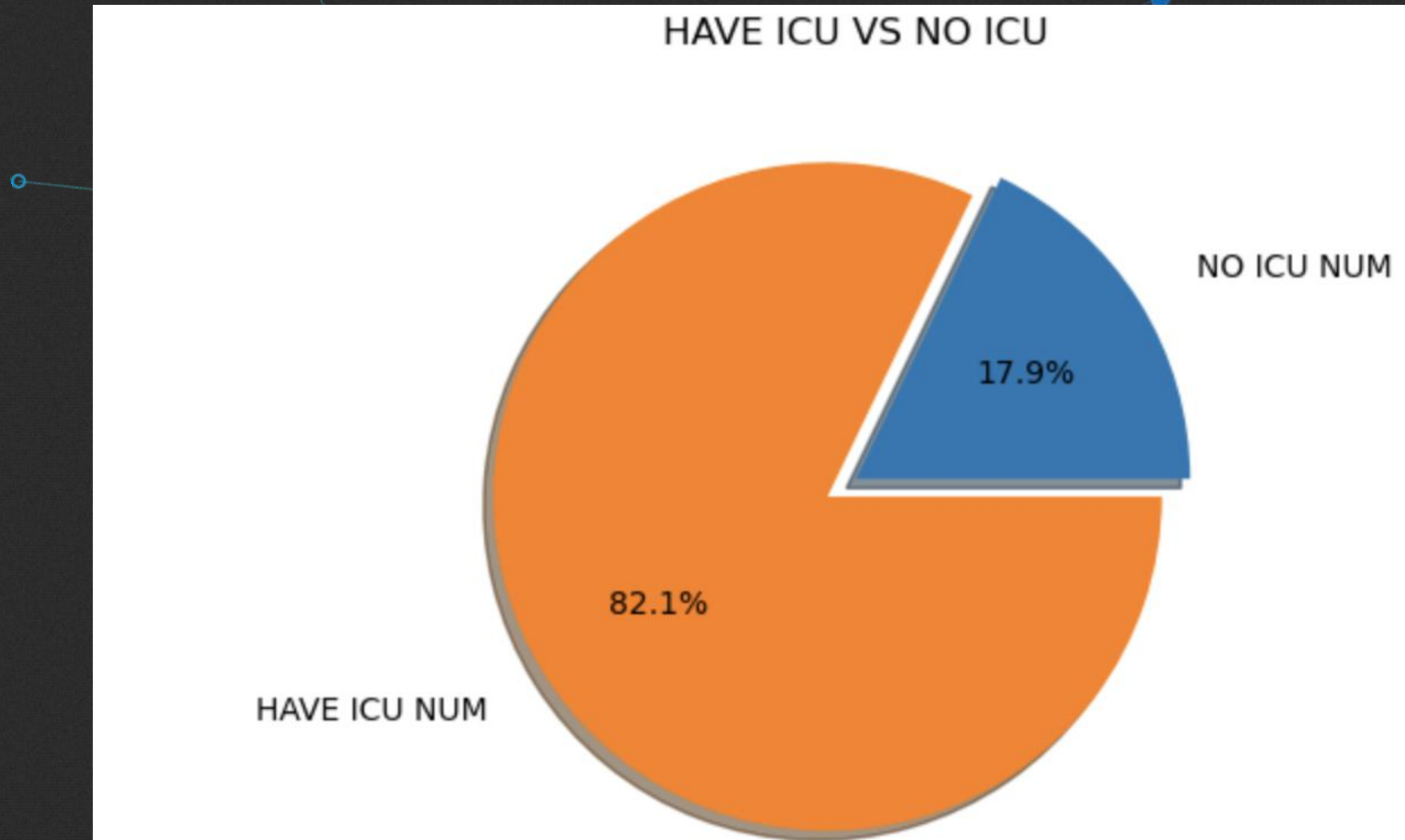


From the perspective of the correlation between population and beds, the number of ordinary beds in each city is strongly related to the number of population, while the correlation between ICU beds is weaker than that of ordinary beds.



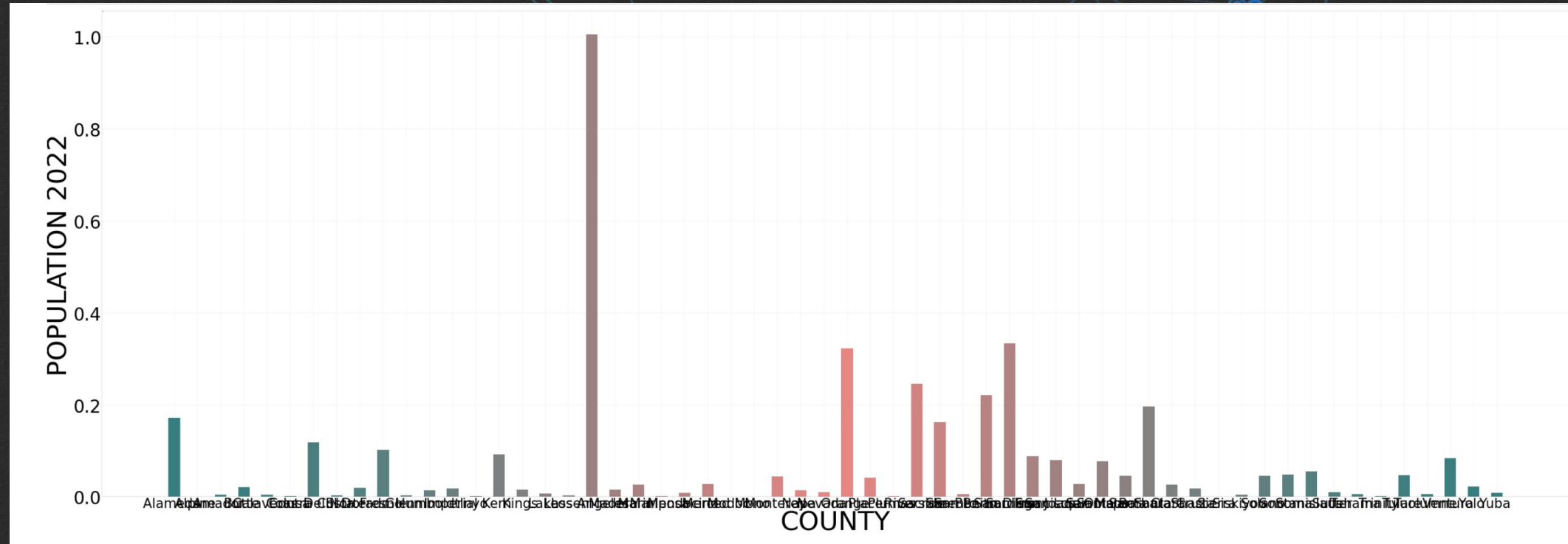


Judging from the distribution of ICU beds in each city, 17.9% of the cities are not equipped with ICU beds, which also means that these cities will be in extreme danger in case of an extreme epidemic.



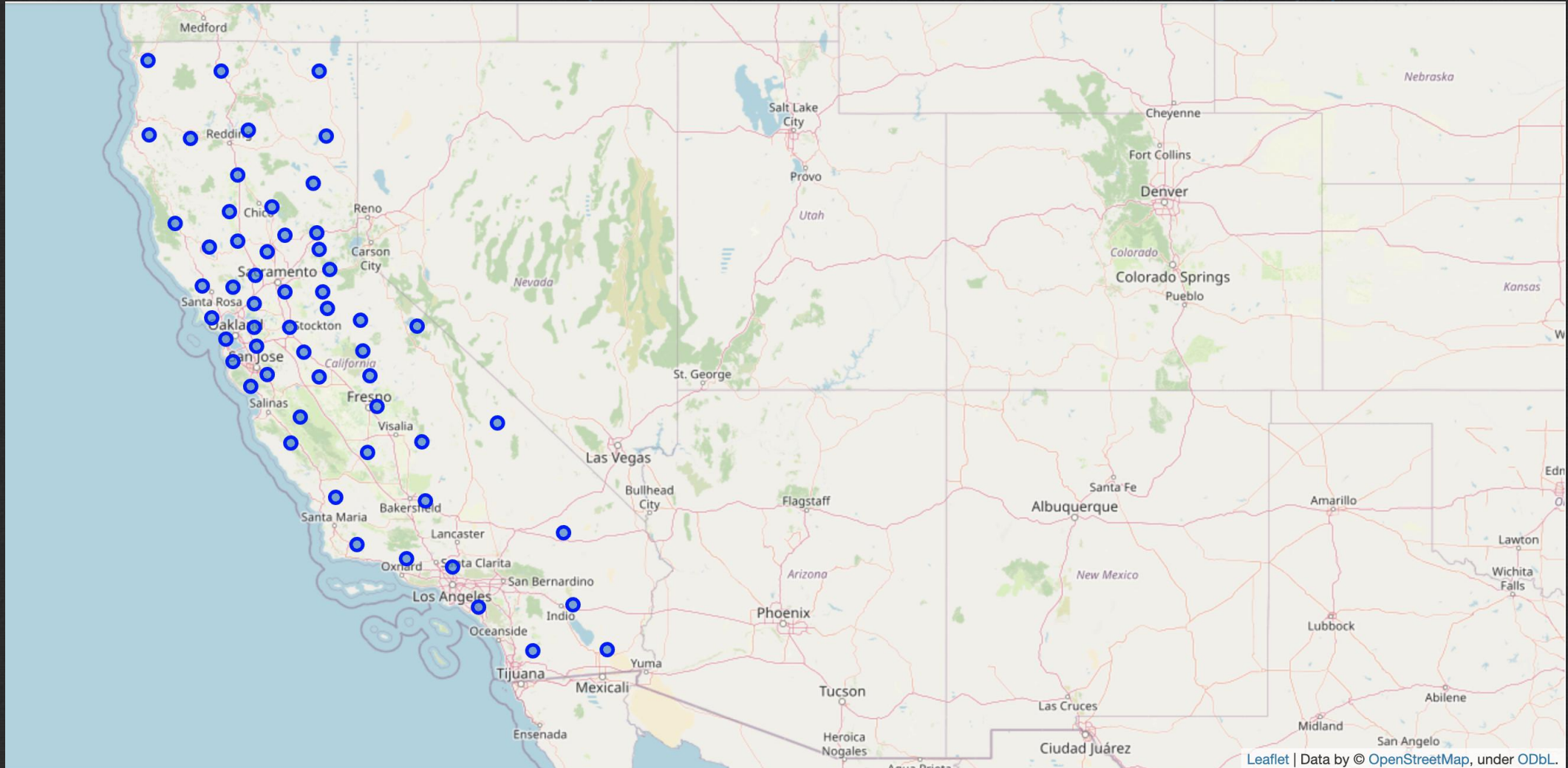


Therefore, once a new epidemic occurs in these cities, effective treatment will not be available



Judging from the distribution of cities in California, some cities in Long Tail are relatively remote and not equipped with effective ICU beds.

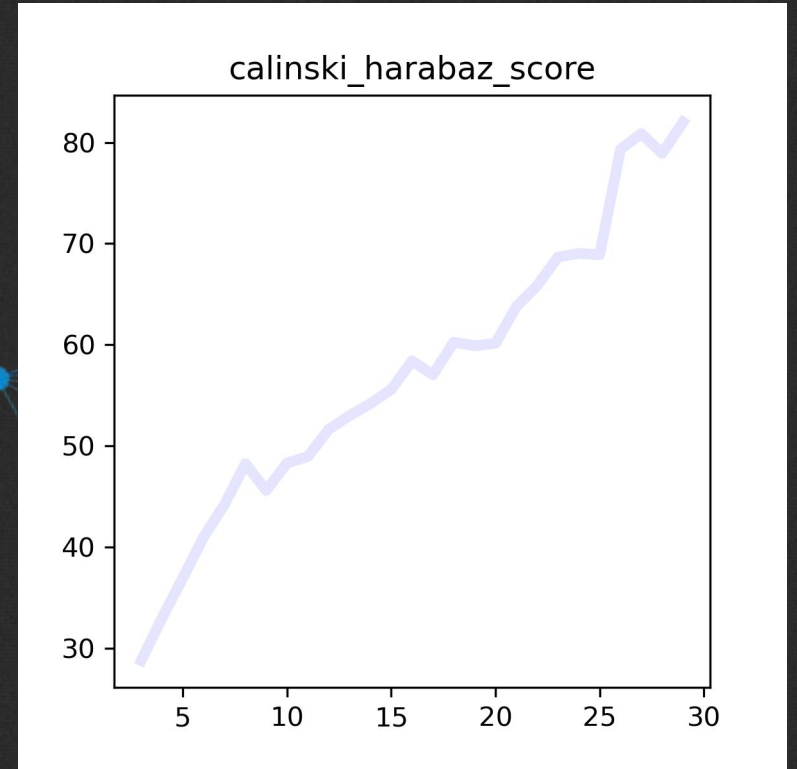
Once problems occur, it will be a hidden danger to transfer patients in time and provide effective treatment





# METHODOLOGY

In this study, **Kmeans algorithm** is used to cluster cities, and **CH algorithm** is used to analyze the optimal classification combination to ensure the scientificity and interpretability of the model



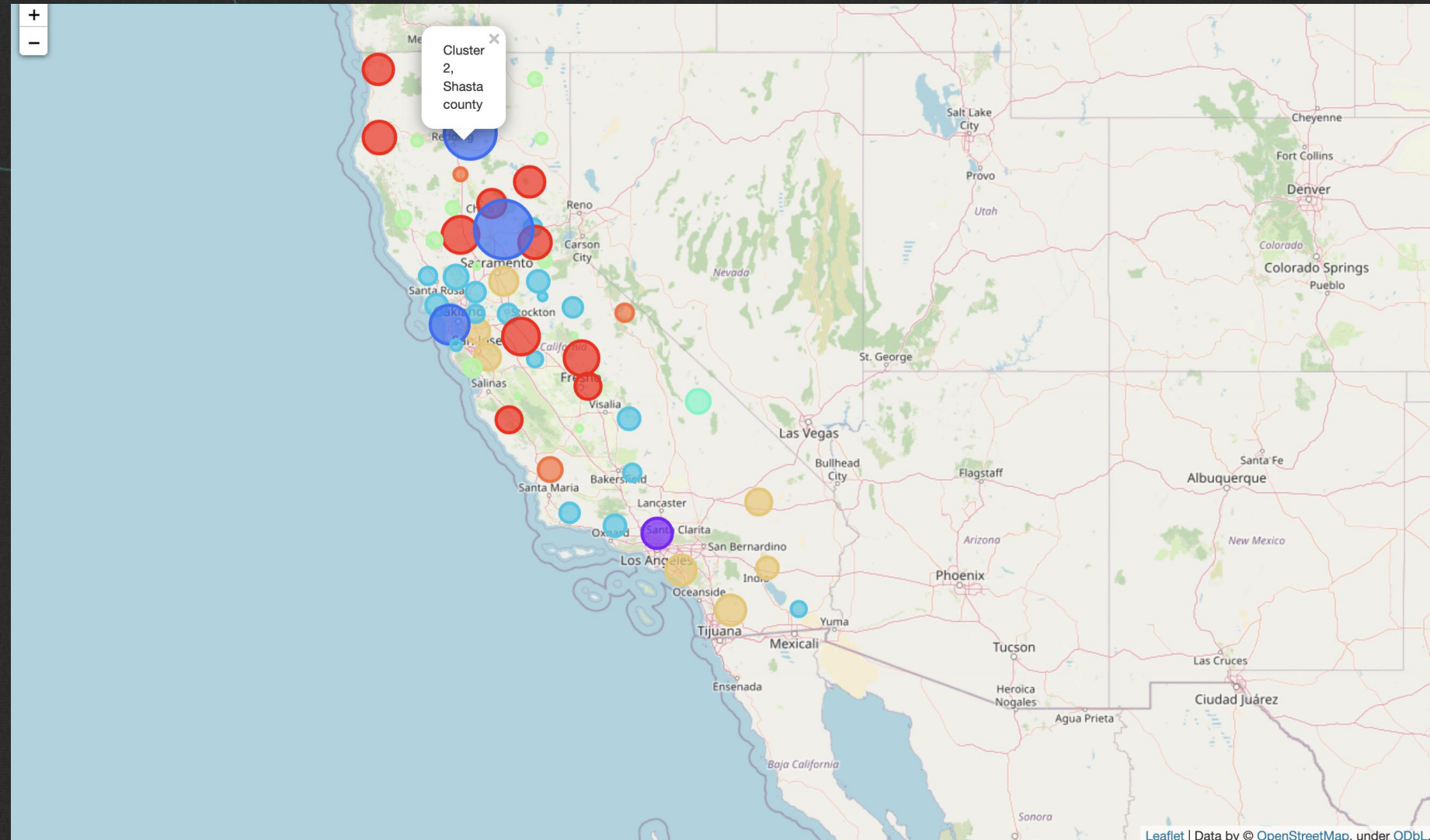
## FIND BEST K VALUE

```
k_value=list()
for k in range(3,30):
    kmeans_model = KMeans(n_clusters=k)
    kmeans_model.fit(clusters_df)
    creator_clust = kmeans_model.labels_
    cluster_centers = kmeans_model.cluster_centers_
    #print('打印分数')
    k_value_tmp=calinski_harabaz_score(clusters_df,creator_clust)
    k_value.append(k_value_tmp)
print(k_value)
```

```
[28.785725348612882, 33.481694983286864, 37.03214115680891, 40.79819594882085, 44.221826258972904, 48.377538064544495, 47.0388448245688, 48.38614731793899, 50.606255328185, 52.06490156461377, 52.82646549156907, 54.540288620233845, 57.19286642766877, 56.39189852863463, 57.54043936924321, 60.204781144818554, 58.7736659316077, 63.6424598907, 51255, 63.492291971373945, 65.24994291671456, 65.05011167628061, 67.81986775975388, 72.10229807859584, 72.80435703908405, 74.19047682984905, 80.90946409463652, 78.38862432461518]
```

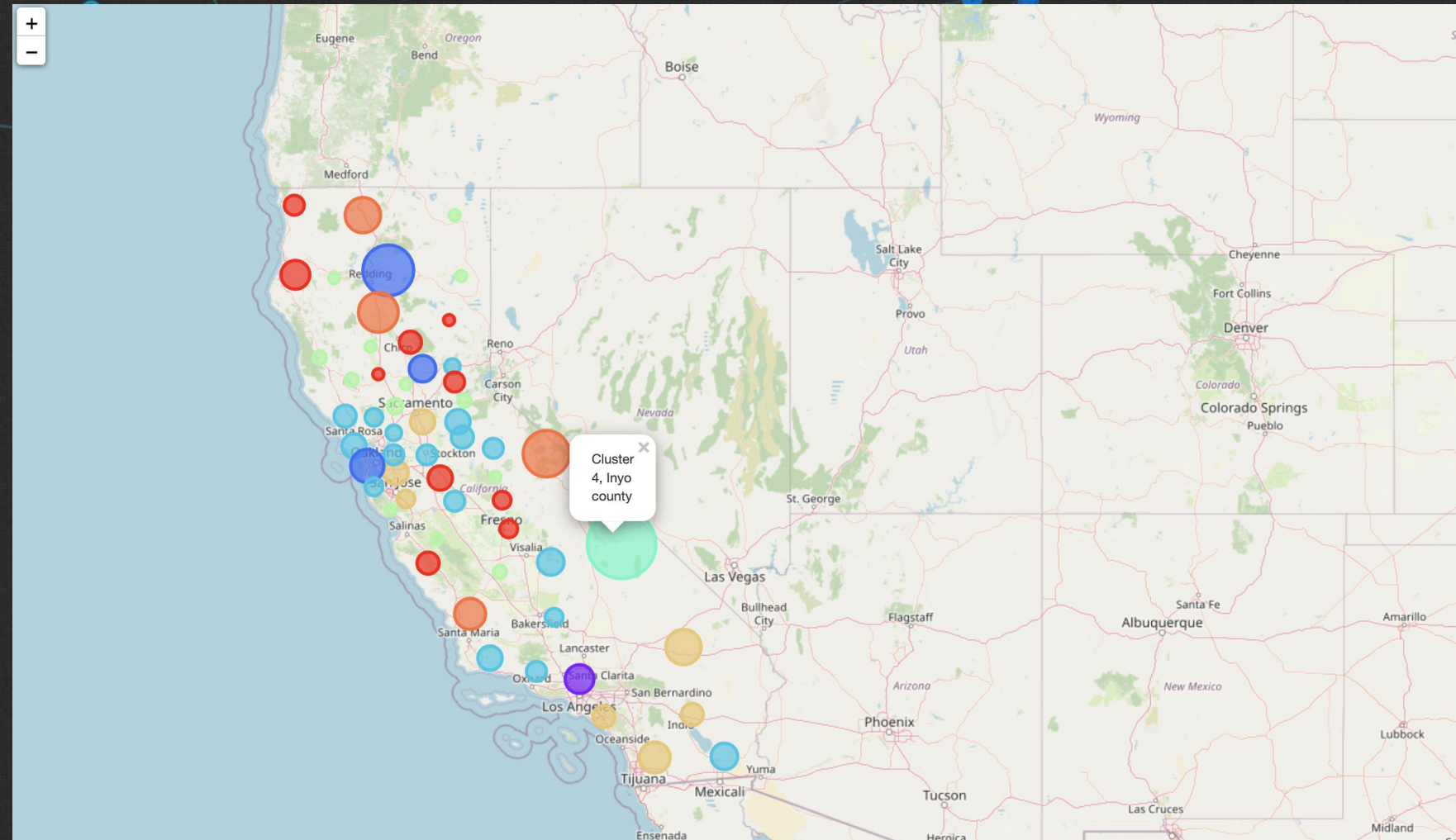
It can be seen from the eight clusters that Cluster 3 has the highest proportion of ordinary beds. However, it is found that the population base of the other two cities is small except San Francisco, which to some extent indicates that San Francisco has the strongest ability to cope with ordinary epidemic

# RESULT





In cluster 5, Yinyou County has the highest proportion of ICU beds, which also shows that when dealing with extreme epidemics, such as new coronavirus, Yinyou County has better processing capacity, and can also help neighboring cities





# CONCLUSION

The purpose of this study is to explore the capacity of cities in California, which has a high population density, in dealing with general medical care and new outbreaks.

The study presents the distribution of common sickbeds and ICU sickbeds in different regions, and the per capita sharing rate of ICU in cities with different population bases

Among them, kmeans model is used to classify the comprehensive population characteristics and medical characteristics of California cities, and finally eight categories are obtained.

Among them, SAN FAN (25%、0.9%) / INYO(16%、2.2%) in classification 3/5 has good resistance to general medical and extreme epidemic respectively.

In classification 6, 13 cities represented by Kings have a per capita share of common sickbeds of 7%, while the share of ICU sickbeds is close to 0%.

Therefore, once an epidemic occurs in these 13 cities, they will face great danger





THANK YOU

---