```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df = pd.read_csv('./indian_food.csv')
df.head()
```

```
            name
ingredients  \
0      Balu shahi                    Maida flour, yogurt, oil, sugar

1          Boondi                          Gram flour, ghee, sugar

2  Gajar ka halwa        Carrots, milk, sugar, ghee, cashews, raisins

3          Ghevar  Flour, ghee, kewra, milk, clarified butter, su...

4      Gulab jamun  Milk powder, plain flour, baking powder, ghee,...


          diet  prep_time  cook_time flavor_profile    course
state  \
0  vegetarian         45         25          sweet  dessert  West
Bengal
1  vegetarian         80         30          sweet  dessert
Rajasthan
2  vegetarian         15         60          sweet  dessert
Punjab
3  vegetarian         15         30          sweet  dessert
Rajasthan
4  vegetarian         15         40          sweet  dessert  West
Bengal

   region
0    East
1    West
2   North
3    West
4    East
```

```python
df.describe()
```

```
       prep_time    cook_time
count  255.000000  255.000000
mean    31.105882   34.529412
std     72.554409   48.265650
min     -1.000000   -1.000000
25%     10.000000   20.000000
50%     10.000000   30.000000
```

```
75%       20.000000    40.000000
max      500.000000   720.000000

df.isna().sum()

name                0
ingredients         0
diet                0
prep_time           0
cook_time           0
flavor_profile      0
course              0
state               0
region              1
dtype: int64

df.loc[df['region'].isna(),'region'] = 'North'

df.loc[df['state']== '-1']
```

	name	ingredients
7	Kaju katli	Cashews, ghee, cardamom, sugar
9	Kheer	Milk, rice, sugar, dried fruits
10	Laddu	Gram flour, ghee, sugar
12	Nankhatai	Refined flour, besan, ghee, powdered sugar, yo...
94	Khichdi	Moong dal, green peas, ginger, tomato, green c...
96	Kulfi falooda	Rose syrup, falooda sev, mixed nuts, saffron, ...
98	Lauki ki subji	Bottle gourd, coconut oil, garam masala, ginge...
109	Pani puri	Kala chana, mashed potato, boondi, sev, lemon
111	Papad	Urad dal, sev, lemon juice, chopped tomatoes
115	Rajma chaval	Red kidney beans, garam masala powder, ginger,...
117	Samosa	Potatoes, green peas, garam masala, ginger, dough
128	Dosa	Chana dal, urad dal, whole urad dal, blend ric...
130	Idli	Split urad dal, urad dal, idli rice, thick poh...
144	Masala Dosa	Chana dal, urad dal, potatoes, idli rice, thic...

145	Pachadi	Coconut oil, cucumber, curd, curry leaves, mus...
149	Payasam	Rice, cashew nuts, milk, raisins, sugar
154	Rasam	Tomato, curry leaves, garlic, mustard seeds, h...
156	Sambar	Pigeon peas, eggplant, drumsticks, sambar powd...
158	Sevai	Sevai, parboiled rice, steamer
161	Uttapam	Chana dal, urad dal, thick poha, tomato, butter
162	Vada	Urad dal, ginger, curry leaves, green chilies,...
164	Upma	Chana dal, urad dal, ginger, curry leaves, sugar
231	Brown Rice	Brown rice, soy sauce, olive oil
248	Red Rice	Red pepper, red onion, butter, watercress, oli...

	diet	prep_time	cook_time	flavor_profile	course	state	region
7	vegetarian	10	20	sweet	dessert	-1	-1
9	vegetarian	10	40	sweet	dessert	-1	-1
10	vegetarian	10	40	sweet	dessert	-1	-1
12	vegetarian	20	30	sweet	dessert	-1	-1
94	vegetarian	40	20	spicy	main course	-1	-1
96	vegetarian	45	25	sweet	dessert	-1	-1
98	vegetarian	10	20	spicy	main course	-1	-1
109	vegetarian	15	2	spicy	snack	-1	-1
111	vegetarian	5	5	spicy	snack	-1	-1
115	vegetarian	15	90	spicy	main course	-1	North
117	vegetarian	30	30	spicy	snack	-1	-1
128	vegetarian	360	90	spicy	snack	-1	South
130	vegetarian	360	90	spicy	snack	-1	South

```
144  vegetarian       360       90        spicy       snack    -
1   South
145  vegetarian        10       25          -1  main course    -
1   South
149  vegetarian        15       30        sweet     dessert    -
1   South
154  vegetarian        10       35        spicy  main course    -
1   South
156  vegetarian        20       45        spicy  main course    -
1   South
158  vegetarian       120       30          -1  main course    -
1   South
161  vegetarian        10       20        spicy       snack    -
1   South
162  vegetarian        15       20        spicy       snack    -
1   South
164  vegetarian        10       20        spicy       snack    -
1       -1
231  vegetarian        15       25          -1  main course    -
1       -1
248  vegetarian        -1       -1          -1  main course    -
1       -1
```

```python
x = df.groupby('state')
['ingredients'].apply(list).reset_index(name='ingredients')

# function to get the uniqueue values from each string
def get_unique_ingred(ingred):
    for i in ingred:
        word = i.lower().split(',')
    return set(word)

x['ingredients'] = x['ingredients'].apply(get_unique_ingred)
x
```

```
              state                                ingredients
0                -1  { watercress,  butter,  red onion, red pepper,...
1    Andhra Pradesh                    { rice flour, green moong beans}
2             Assam        {glutinous rice,  black sesame seeds,  gur}
3             Bihar     { filling, sattu,  dough,  atta,  mustard oil}
4      Chhattisgarh  {chana dal,  garam masala powder,  arhar dal, ...
5               Goa  { fennel seeds,  ginger powder,  grated coconu...
6           Gujarat  { sugar,  peas,  baking soda,  grated coconut,...
7           Haryana  { gram flour, besan,  curry leaves,  garam mas...
8   Jammu & Kashmir  { badam,  dried rose petals, cottage cheese,  ...
9         Karnataka  { black sesame seeds,  curry leaves, thin rice...
10           Kerala  { masala,  sesame oil,  whole red beans,  tama...
11   Madhya Pradesh  { arrowroot powder,  all purpose flour, milk p...
12      Maharashtra         { khus khus, gobi,  potato,  coconut,  beans}
13          Manipur  {forbidden black rice,  garlic powder,  olive ...
14     NCT of Delhi  { butter,  greek yogurt, chicken,  cashew nuts...
```

```
15       Nagaland  { chillies,  axone,  water,  pork,  salt, rice}
16         Odisha  {curd,  curry leaves,  cooked rice,  dry chilli}
17         Punjab  { fish fillets,  green bell pepper,  yogurt,  ...
18      Rajasthan  { khus khus, whole wheat flour,  dry coconut, ...
19     Tamil Nadu  { cinnamon, meat curry powder,  tomato,  chick...
20     Telangana   {rose water,  almonds,  milk,  saffron,  white...
21        Tripura  { ginger and garlic,  onions, boiled pork,  ch...
22  Uttar Pradesh  {whole wheat flour,  musk melon seeds,  edible...
23    Uttarakhand                    { coconut,  molu leaf, khoa}
24    West Bengal  { brinjal,  ridge gourd,  banana, green beans,...
```

```python
x.drop(0,inplace=True)
```

```
x
```

```
            state                                ingredients
1   Andhra Pradesh                 { rice flour, green moong beans}
2            Assam        {glutinous rice,  black sesame seeds,  gur}
3            Bihar      { filling,  sattu,  dough,  atta,  mustard oil}
4     Chhattisgarh  {chana dal,  garam masala powder,  arhar dal, ...
5              Goa  { fennel seeds,  ginger powder,  grated coconu...
6          Gujarat  { sugar,  peas,  baking soda,  grated coconut,...
7          Haryana  { gram flour, besan,  curry leaves,  garam mas...
8   Jammu & Kashmir  { badam,  dried rose petals, cottage cheese,  ...
9        Karnataka  { black sesame seeds,  curry leaves, thin rice...
10          Kerala  { masala,  sesame oil,  whole red beans,  tama...
11   Madhya Pradesh  { arrowroot powder,  all purpose flour, milk p...
12     Maharashtra     { khus khus,  gobi,  potato,  coconut,  beans}
13         Manipur  {forbidden black rice,  garlic powder,  olive ...
14    NCT of Delhi  { butter,  greek yogurt, chicken,  cashew nuts...
15        Nagaland  { chillies,  axone,  water,  pork,  salt, rice}
16          Odisha  {curd,  curry leaves,  cooked rice,  dry chilli}
17          Punjab  { fish fillets,  green bell pepper,  yogurt,  ...
18       Rajasthan  { khus khus, whole wheat flour,  dry coconut, ...
19      Tamil Nadu  { cinnamon, meat curry powder,  tomato,  chick...
20       Telangana  {rose water,  almonds,  milk,  saffron,  white...
21         Tripura  { ginger and garlic,  onions, boiled pork,  ch...
22   Uttar Pradesh  {whole wheat flour,  musk melon seeds,  edible...
23     Uttarakhand                    { coconut,  molu leaf, khoa}
24     West Bengal  { brinjal,  ridge gourd,  banana, green beans,...
```

```python
x['ingredients'] = x['ingredients'].apply(' '.join)
x
```

```
            state                                ingredients
1   Andhra Pradesh                 rice flour green moong beans
2            Assam         glutinous rice  black sesame seeds  gur
3            Bihar             filling sattu  dough  atta  mustard oil
4     Chhattisgarh  chana dal  garam masala powder  arhar dal  whi...
5              Goa   fennel seeds  ginger powder  grated coconut  ...
6          Gujarat   sugar  peas  baking soda  grated coconut ridg...
```

```
7             Haryana    gram flour besan  curry leaves  garam masala ...
8     Jammu & Kashmir    badam  dried rose petals cottage cheese  pist...
9           Karnataka    black sesame seeds  curry leaves thin rice fl...
10             Kerala    masala  sesame oil  whole red beans  tamarind...
11     Madhya Pradesh    arrowroot powder  all purpose flour milk powd...
12        Maharashtra              khus khus gobi  potato  coconut  beans
13           Manipur    forbidden black rice  garlic powder  olive oil...
14       NCT of Delhi    butter  greek yogurt chicken  cashew nuts  ga...
15           Nagaland            chillies  axone  water  pork  salt rice
16             Odisha          curd  curry leaves  cooked rice  dry chilli
17             Punjab    fish fillets  green bell pepper  yogurt  biry...
18          Rajasthan    khus khus whole wheat flour  dry coconut  ses...
19         Tamil Nadu    cinnamon meat curry powder  tomato  chicken c...
20          Telangana    rose water  almonds  milk  saffron  white brea...
21            Tripura    ginger and garlic  onions boiled pork  chillies
22      Uttar Pradesh    whole wheat flour  musk melon seeds  edible gu...
23        Uttarakhand                      coconut  molu leaf khoa
24        West Bengal    brinjal  ridge gourd  banana green beans  bit...
```

```python
from sklearn.feature_extraction.text import TfidfVectorizer

tfidf = TfidfVectorizer()

# Fit and transform text column to get TF-IDF matrix
vec = tfidf.fit_transform(x['ingredients'])

# Convert sparse matrix to dense matrix and create DataFrame
df_tfidf = pd.DataFrame(vec.todense(),
columns=tfidf.get_feature_names_out())

# Print results
print(df_tfidf)
```

```
        all   almonds       and     arhar  arrowroot      atta
axone  \
0  0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
1  0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
2  0.000000  0.000000  0.000000  0.000000   0.000000  0.420872
0.000000
3  0.000000  0.000000  0.000000  0.251635   0.000000  0.000000
0.000000
4  0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
5  0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
6  0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
7  0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
```

```
                                          0.000000
8    0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
9    0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
10   0.377578  0.000000  0.000000  0.000000   0.377578  0.000000
0.000000
11   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
12   0.000000  0.332084  0.000000  0.000000   0.000000  0.000000
0.000000
13   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
14   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.460391
15   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
16   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
17   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
18   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
19   0.000000  0.336520  0.000000  0.000000   0.000000  0.000000
0.000000
20   0.000000  0.000000  0.411783  0.000000   0.000000  0.000000
0.000000
21   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
22   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000
23   0.000000  0.000000  0.000000  0.000000   0.000000  0.000000
0.000000

        badam    baking    banana  ...     sugar  tamarind      thin
tomato  \
0    0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000
0.000000
1    0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000
0.000000
2    0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000
0.000000
3    0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000
0.000000
4    0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000
0.000000
5    0.000000  0.384465  0.000000  ...  0.384465  0.000000  0.000000
0.000000
6    0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000
0.000000
```

```
7   0.346459   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
8   0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.441473
0.000000
9   0.000000   0.000000   0.000000   ...   0.000000   0.439448   0.000000
0.000000
10  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
11  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
12  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
13  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
14  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
15  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
16  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
17  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
18  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.405895
19  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
20  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
21  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
22  0.000000   0.000000   0.000000   ...   0.000000   0.000000   0.000000
0.000000
23  0.000000   0.000000   0.351162   ...   0.000000   0.000000   0.000000
0.000000
```

|	urad	water	wheat	white	whole	yogurt
0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	0.251635	0.000000	0.202164	0.222696	0.186238	0.000000
4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.000000	0.000000	0.000000	0.325241	0.000000
10	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000	0.000000	0.000000	0.315331

```
14   0.000000   0.407445   0.000000   0.000000   0.000000   0.000000
15   0.000000   0.000000   0.000000   0.000000   0.000000   0.000000
16   0.000000   0.000000   0.000000   0.000000   0.000000   0.305667
17   0.000000   0.000000   0.294096   0.000000   0.270928   0.000000
18   0.000000   0.000000   0.000000   0.000000   0.000000   0.000000
19   0.000000   0.336520   0.000000   0.336520   0.000000   0.000000
20   0.000000   0.000000   0.000000   0.000000   0.000000   0.000000
21   0.000000   0.000000   0.261250   0.000000   0.240669   0.000000
22   0.000000   0.000000   0.000000   0.000000   0.000000   0.000000
23   0.000000   0.000000   0.000000   0.000000   0.000000   0.000000

[24 rows x 110 columns]
```

```python
from sklearn.cluster import KMeans

wcss = []
for i in range(1,25):
    km = KMeans(n_clusters=i,init='k-means++',random_state=42)
    km.fit(vec)
    wcss.append(km.inertia_)

plt.plot(range(1,25),wcss)
plt.xlabel('No. of clusters')
plt.ylabel('WCSS')
plt.title('WCSS graph')
plt.show()
```

WCSS graph
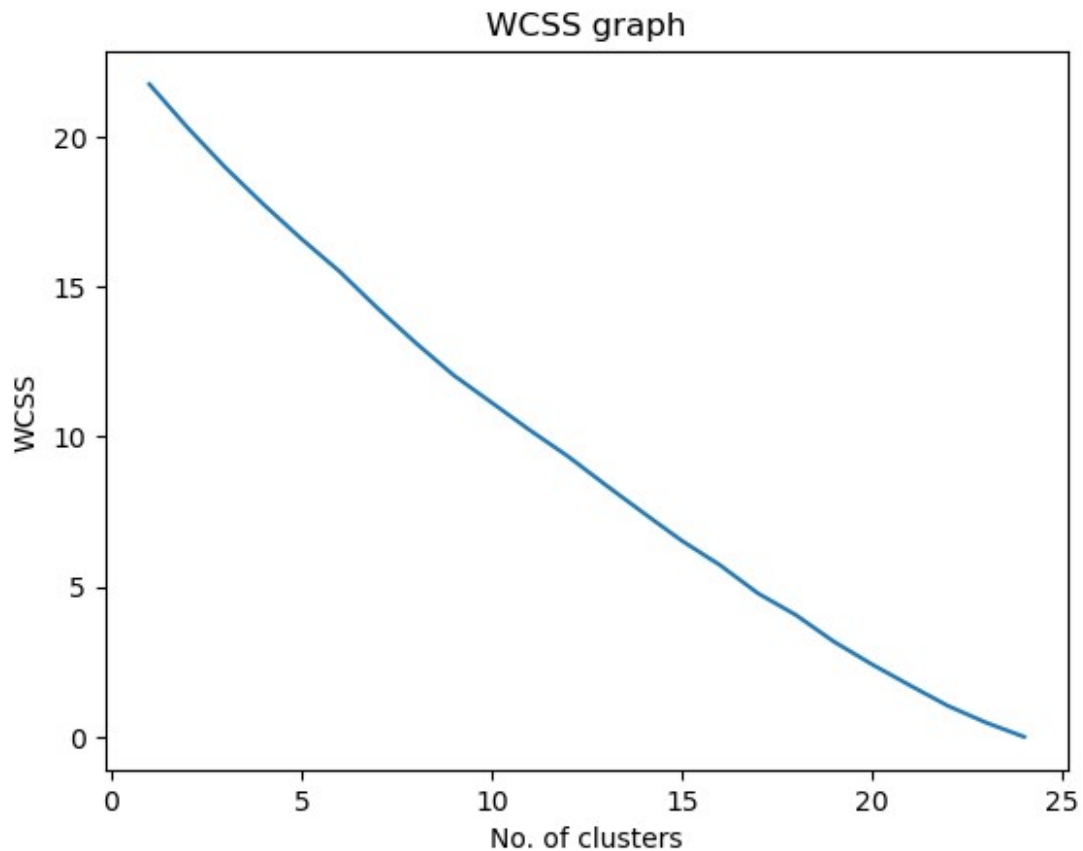

```
no_cluster = 10

kmeans = KMeans(n_clusters=no_cluster,init='k-means++',
random_state=36)
y_pred = kmeans.fit_predict(vec)

x['cluster'] = y_pred
x
```

	state	ingredients
1	Andhra Pradesh	rice flour green moong beans
2	Assam	glutinous rice black sesame seeds gur
3	Bihar	filling sattu dough atta mustard oil
4	Chhattisgarh	chana dal garam masala powder arhar dal whi...
5	Goa	fennel seeds ginger powder grated coconut ...
6	Gujarat	sugar peas baking soda grated coconut ridg...

7	Haryana	gram flour besan curry leaves garam masala ...
8	Jammu & Kashmir	badam dried rose petals cottage cheese pist...
9	Karnataka	black sesame seeds curry leaves thin rice fl...
10	Kerala	masala sesame oil whole red beans tamarind...
11	Madhya Pradesh	arrowroot powder all purpose flour milk powd...
12	Maharashtra	khus khus gobi potato coconut beans
13	Manipur	forbidden black rice garlic powder olive oil...
14	NCT of Delhi	butter greek yogurt chicken cashew nuts ga...
15	Nagaland	chillies axone water pork salt rice
16	Odisha	curd curry leaves cooked rice dry chilli
17	Punjab	fish fillets green bell pepper yogurt biry...
18	Rajasthan	khus khus whole wheat flour dry coconut ses...
19	Tamil Nadu	cinnamon meat curry powder tomato chicken c...
20	Telangana	rose water almonds milk saffron white brea...
21	Tripura	ginger and garlic onions boiled pork chillies
22	Uttar Pradesh	whole wheat flour musk melon seeds edible gu...
23	Uttarakhand	coconut molu leaf khoa
24	West Bengal	brinjal ridge gourd banana green beans bit...

```
    cluster
1       9
2       8
3       1
4       0
5       4
6       3
7       0
8       5
9       8
10      1
```

```
11        5
12        7
13        4
14        0
15        2
16        8
17        9
18        7
19        0
20        5
21        2
22        6
23        1
24        3
```

```python
import geopandas as gpd

out_res = pd.concat([x['state'],x['cluster']], axis=1)
out_res.replace('NCT of Delhi','Delhi',inplace=True)

indian_map = gpd.read_file('./India-State-and-Country-Shapefile-
Updated-Jan-2020-master/India_State_Boundary.shp')

indian_map.replace('Jammu and Kashmir','Jammu & Kashmir',inplace=True)
indian_map.replace('Telengana','Telangana',inplace=True)
indian_map.replace('Tamilnadu','Tamil Nadu',inplace=True)
indian_map.replace('Chhattishgarh','Chhattisgarh',inplace=True)

merged =
indian_map.set_index('State_Name').join(out_res.set_index('state'))
merged
```

```
                                       geometry  \
Andaman & Nicobar                      MULTIPOLYGON (((10341718.474
1449533.161, 1034...
Andhra Pradesh                         POLYGON ((9426056.496
2174632.352, 9426228.484...
Arunachal Pradesh                      POLYGON ((10696175.277
3434232.650, 10696981.8...
Assam                                  POLYGON ((10380499.251
2872443.723, 10380499.2...
Bihar                                  POLYGON ((9362949.333
3188807.607, 9362966.106...
Chandigarh                             POLYGON ((8546255.616
3606050.813, 8546315.400...
Chhattisgarh                           POLYGON ((9275926.808
2765881.317, 9276185.437...
Daman and Diu and Dadra and Nagar Haveli  MULTIPOLYGON (((8122247.822
```

2312434.407, 81223...	
Delhi 3359116.190, 8583476.212...	POLYGON ((8583390.570
Goa 1779394.764, 8223279.301...	POLYGON ((8223217.424
Gujarat 2837315.493, 7915101.603...	POLYGON ((7914780.837
Haryana 3516490.865, 8524451.392...	POLYGON ((8524318.539
Himachal Pradesh 3929291.879, 8548760.706...	POLYGON ((8548682.698
Jammu & Kashmir 3927668.327, 8550332.102...	POLYGON ((8550375.654
Jharkhand 2772949.712, 9762301.816...	POLYGON ((9762288.285
Karnataka 2090389.205, 86086...	MULTIPOLYGON (((8608594.474
Kerala 1436381.747, 8347795.745...	POLYGON ((8347733.191
Ladakh 3927668.327, 8548619.625...	POLYGON ((8550375.654
Lakshadweep 930182.487, 813526...	MULTIPOLYGON (((8135256.290
Madhya Pradesh 3106498.184, 8724579.382...	POLYGON ((8724343.278
Maharashtra 2515416.345, 82809...	MULTIPOLYGON (((8280974.863
Manipur 2960789.340, 10528432.7...	POLYGON ((10527945.945
Meghalaya 3013858.327, 10222165.9...	POLYGON ((10222042.433
Mizoram 2817021.246, 10326465.4...	POLYGON ((10326423.582
Nagaland 3126858.281, 10597031.2...	POLYGON ((10596805.532
Odisha 2579790.782, 95786...	MULTIPOLYGON (((9578537.936
Puducherry 1232399.360, 88784...	MULTIPOLYGON (((8878474.160
Puducherry 1888315.624, 9149005.345...	POLYGON ((9150619.186
Punjab 3830799.529, 8442574.742...	POLYGON ((8442331.679
Rajasthan 3529026.887, 8234599.327...	POLYGON ((8234599.326
Sikkim 3265074.341, 9865469.610...	POLYGON ((9864726.992
Tamil Nadu 1513831.235, 89395...	MULTIPOLYGON (((8939353.702
Telangana	POLYGON ((8720284.876

```
2259244.214, 8720421.528...
Tripura                                    POLYGON ((10260260.337
2818339.599, 10260273.8...
Uttar Pradesh                              POLYGON ((8637489.997
3555885.598, 8637654.287...
Uttarakhand                                POLYGON ((8801802.136
3692833.282, 8802083.049...
West Bengal                                POLYGON ((9800305.279
3151090.311, 9800377.779...


                                           cluster
Andaman & Nicobar                            NaN
Andhra Pradesh                               9.0
Arunachal Pradesh                            NaN
Assam                                        8.0
Bihar                                        1.0
Chandigarh                                   NaN
Chhattisgarh                                 0.0
Daman and Diu and Dadra and Nagar Haveli     NaN
Delhi                                        0.0
Goa                                          4.0
Gujarat                                      3.0
Haryana                                      0.0
Himachal Pradesh                             NaN
Jammu & Kashmir                              5.0
Jharkhand                                    NaN
Karnataka                                    8.0
Kerala                                       1.0
Ladakh                                       NaN
Lakshadweep                                  NaN
Madhya Pradesh                               5.0
Maharashtra                                  7.0
Manipur                                      4.0
Meghalaya                                    NaN
Mizoram                                      NaN
Nagaland                                     2.0
Odisha                                       8.0
Puducherry                                   NaN
Puducherry                                   NaN
Punjab                                       9.0
Rajasthan                                    7.0
Sikkim                                       NaN
Tamil Nadu                                   0.0
Telangana                                    5.0
Tripura                                      2.0
Uttar Pradesh                                6.0
Uttarakhand                                  1.0
West Bengal                                  3.0
```
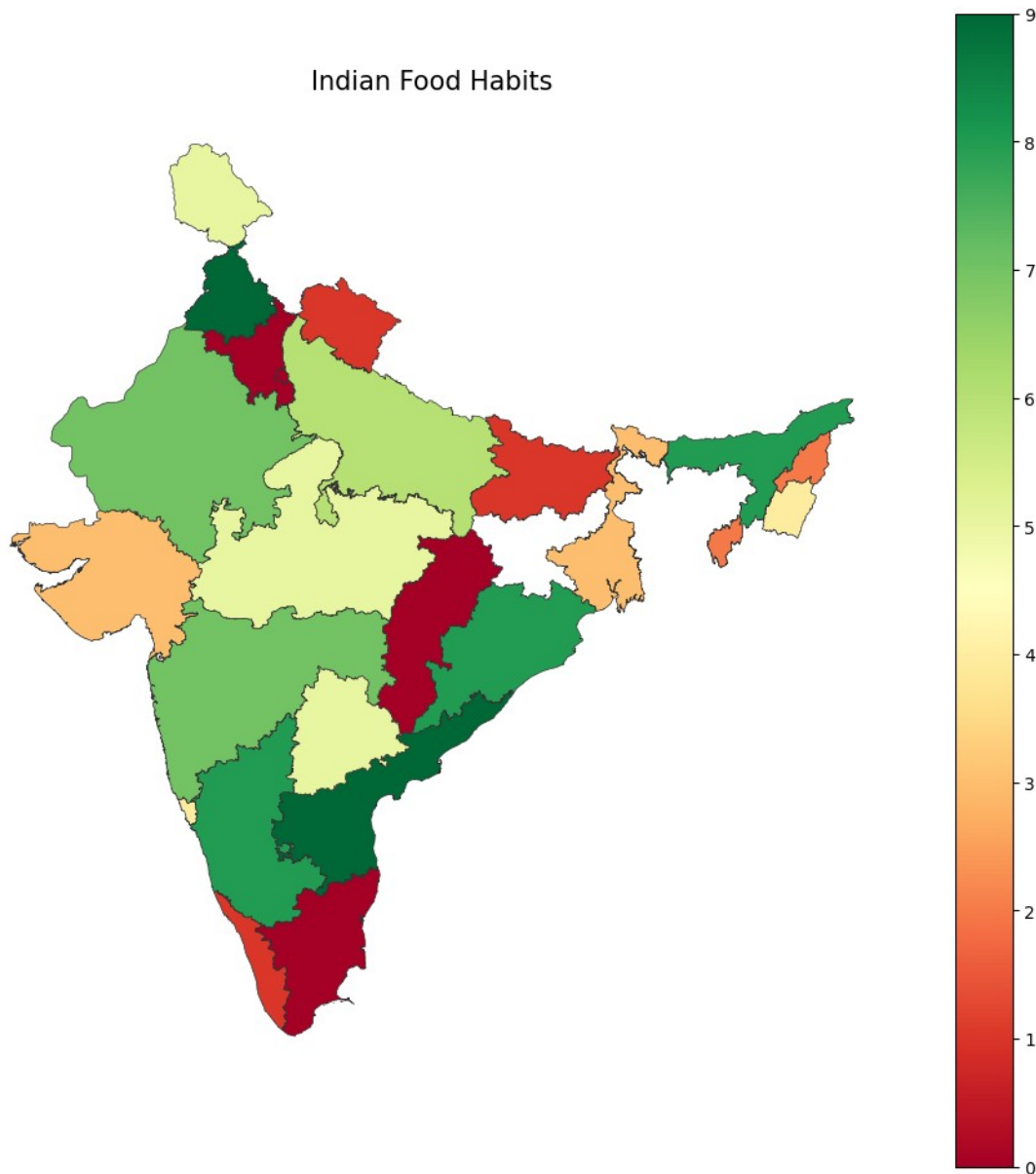
```
fig, ax = plt.subplots(1, figsize=(12, 12))
ax.axis('off')
ax.set_title('Indian Food Habits',
             fontdict={'fontsize': '15', 'fontweight' : '3'})
fig = merged.plot(column='cluster', cmap='RdYlGn', linewidth=0.5,
ax=ax, edgecolor='0.2',legend=True)
```
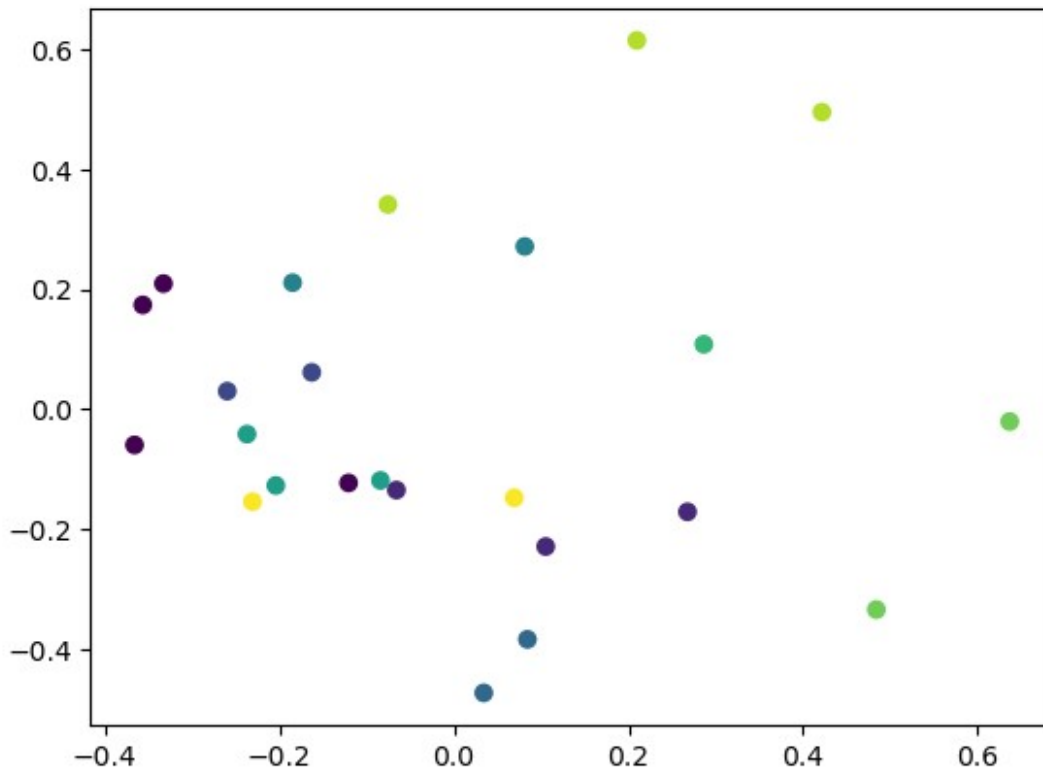

Indian Food Habits

```
from sklearn.decomposition import PCA

pca = PCA(n_components=10)
pca_matrix = pca.fit_transform(vec.toarray())
labels = kmeans.labels_
# Create scatter plot of data points, using cluster labels as color
```

```
plt.scatter(pca_matrix[:, 0], pca_matrix[:, 1], c=labels)
plt.show()
```



```
# Print the state and cluster for each row
for idx, row in x.iterrows():
    print(f"{row['state']} belongs to cluster {row['cluster']}")

# Visualize the clustering result using a bar chart
cluster_counts = x.groupby('cluster')['state'].count()
plt.bar(cluster_counts.index, cluster_counts.values)
plt.xlabel('Cluster')
plt.ylabel('Number of States')
plt.show()
```

```
Andhra Pradesh belongs to cluster 9
Assam belongs to cluster 8
Bihar belongs to cluster 1
Chhattisgarh belongs to cluster 0
Goa belongs to cluster 4
Gujarat belongs to cluster 3
Haryana belongs to cluster 0
Jammu & Kashmir belongs to cluster 5
Karnataka belongs to cluster 8
Kerala belongs to cluster 1
Madhya Pradesh belongs to cluster 5
Maharashtra belongs to cluster 7
```

```
Manipur belongs to cluster 4
NCT of Delhi belongs to cluster 0
Nagaland belongs to cluster 2
Odisha belongs to cluster 8
Punjab belongs to cluster 9
Rajasthan belongs to cluster 7
Tamil Nadu belongs to cluster 0
Telangana belongs to cluster 5
Tripura belongs to cluster 2
Uttar Pradesh belongs to cluster 6
Uttarakhand belongs to cluster 1
West Bengal belongs to cluster 3
```