In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plp
import seaborn as sns
```

In [2]:
```
visa_df=pd.read_csv('Visadataset - Visadataset.csv')
visa_df
```

Out[2]:

	case_id	continent	education_of_employee	has_job_experience	requires_job_traini
0	EZYV01	Asia	High School	N	
1	EZYV02	Asia	Master's	Y	
2	EZYV03	Asia	Bachelor's	N	
3	EZYV04	Asia	Bachelor's	N	
4	EZYV05	Africa	Master's	Y	
...	...	...	...	...	
25475	EZYV25476	Asia	Bachelor's	Y	
25476	EZYV25477	Asia	High School	Y	
25477	EZYV25478	Asia	Master's	Y	
25478	EZYV25479	Asia	Master's	Y	
25479	EZYV25480	Asia	Bachelor's	Y	

25480 rows × 12 columns

In [3]:
```
education_of_employee=visa_df['education_of_employee'].unique()
education_of_employee
```

Out[3]:
```
array(['High School', "Master's", "Bachelor's", 'Doctorate'], dtype=objec
t)
```

In [4]:
```
ee=[]
for i in education_of_employee:
    e_of_e=visa_df['education_of_employee']==i
    value=len(visa_df[e_of_e])
    ee.append(value)
    print(f'{i}:{value}')
```

```
High School:3420
Master's:9634
Bachelor's:10234
Doctorate:2192
```

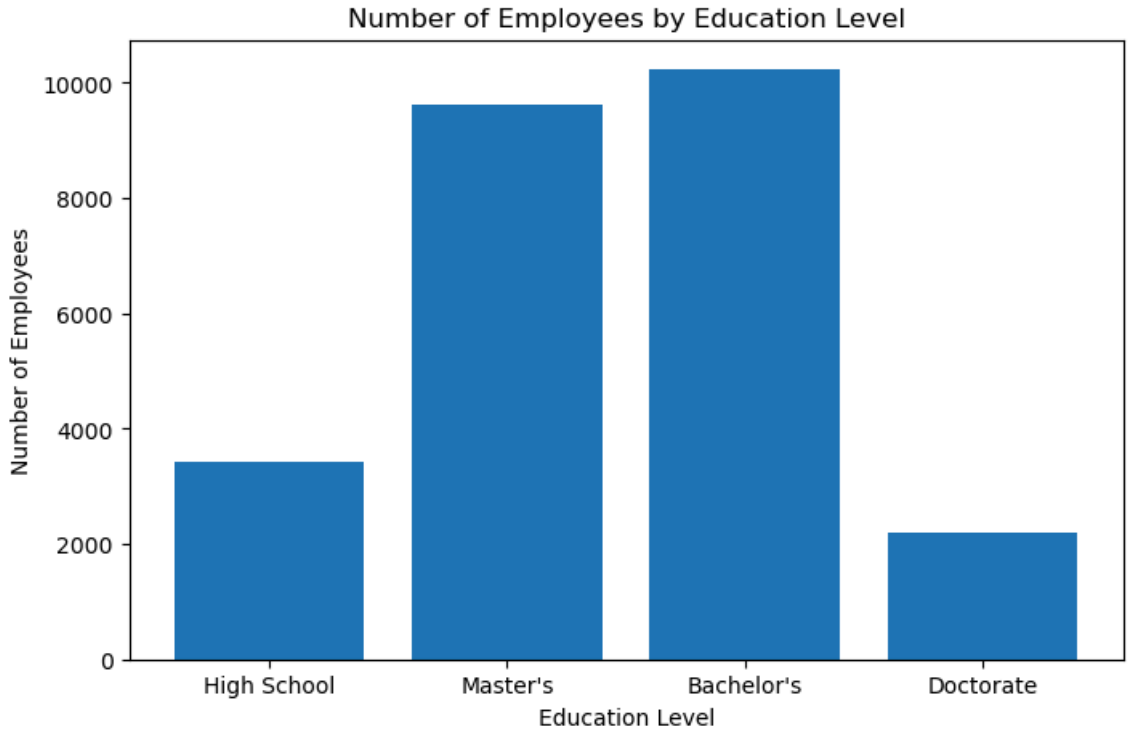
In [26]:
```
education_of_employee_df=pd.DataFrame(zip(education_of_employee,ee),columns
education_of_employee_df
```

Out[26]:

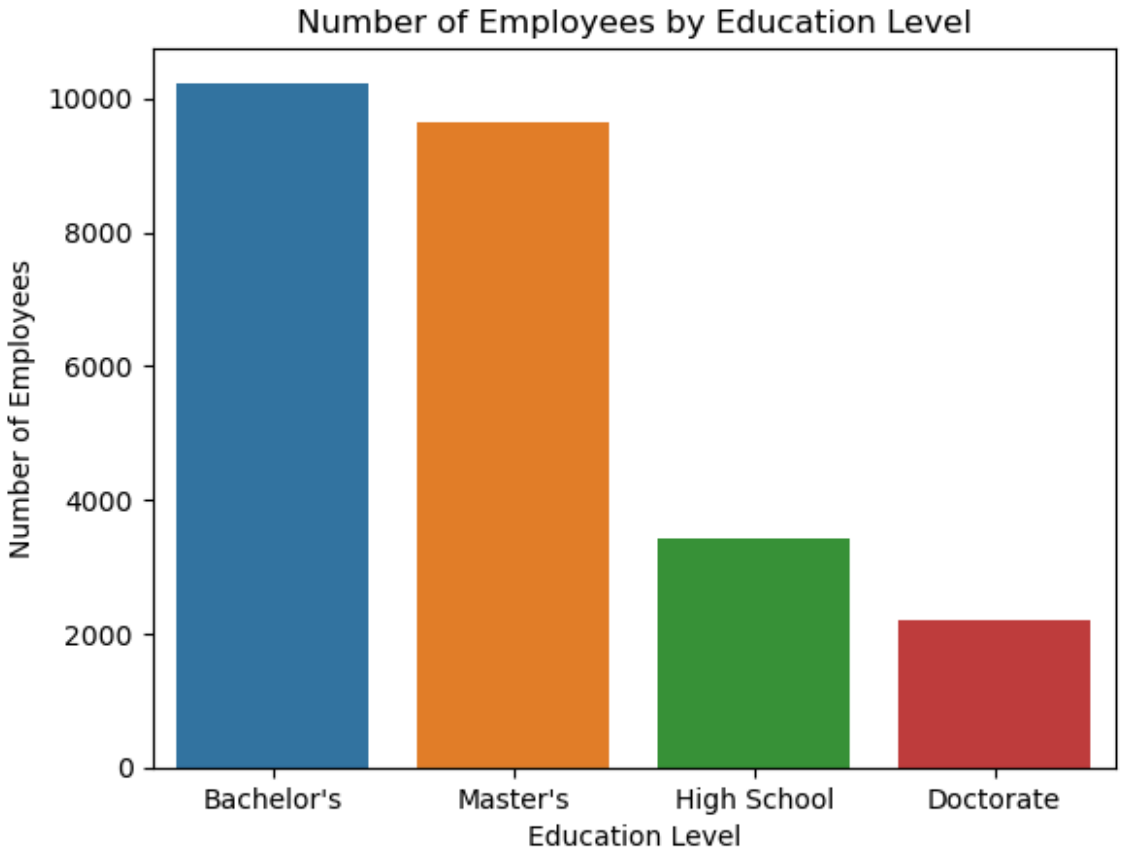
	education_of_employee	count
0	High School	3420
1	Master's	9634
2	Bachelor's	10234
3	Doctorate	2192

In [27]:
```
education_of_employee_df.to_csv("education_of_employee_df.csv",index=False)
```

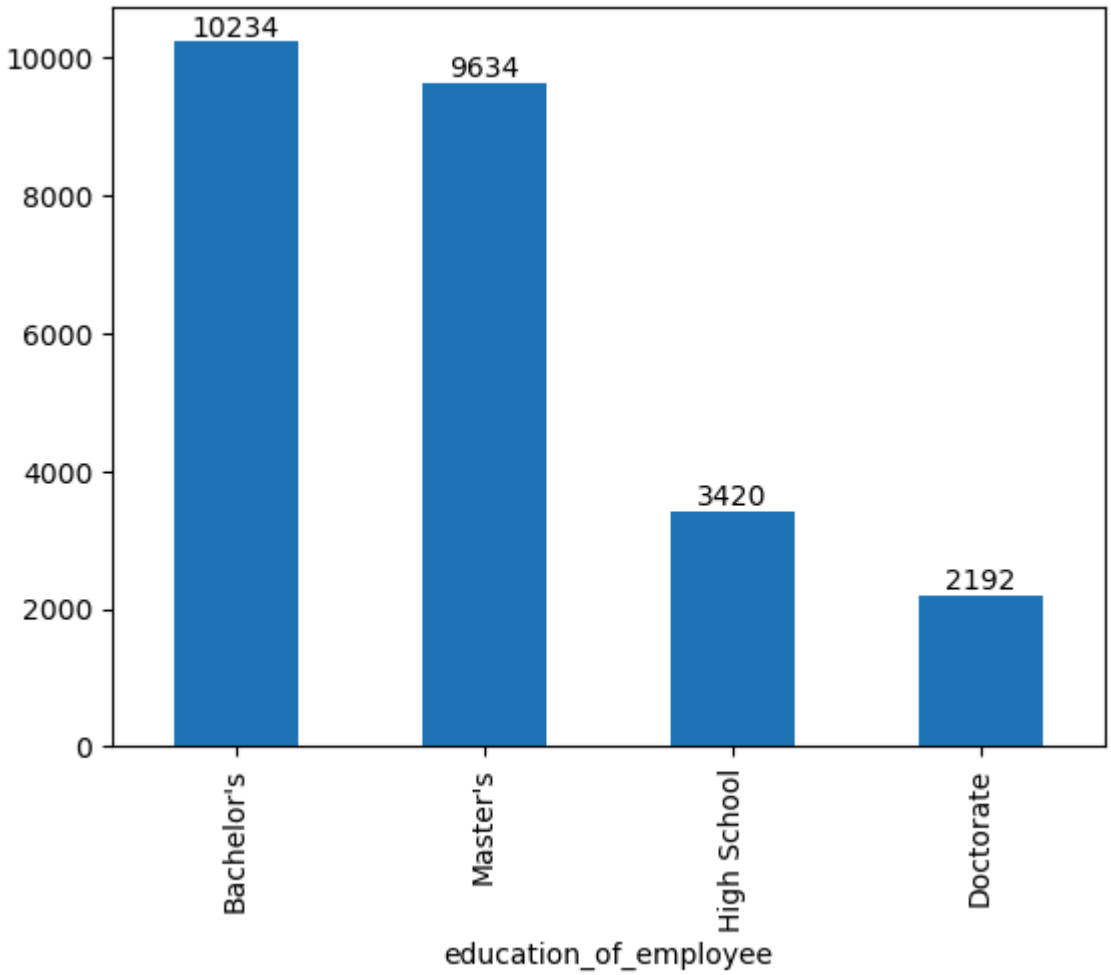
In [36]:
```
plp.figure(figsize=(8,5))
plp.bar('education_of_employee','count',data=education_of_employee_df)
plp.xlabel('Education Level')
plp.ylabel('Number of Employees')
plp.title('Number of Employees by Education Level')
plp.show()
```



In [37]:
```
e=visa_df['education_of_employee'].value_counts()
keys=e.keys()
sns.countplot(data=visa_df,x='education_of_employee',order=keys)
plp.xlabel('Education Level')
plp.ylabel('Number of Employees')
plp.title('Number of Employees by Education Level')
plp.show()
```



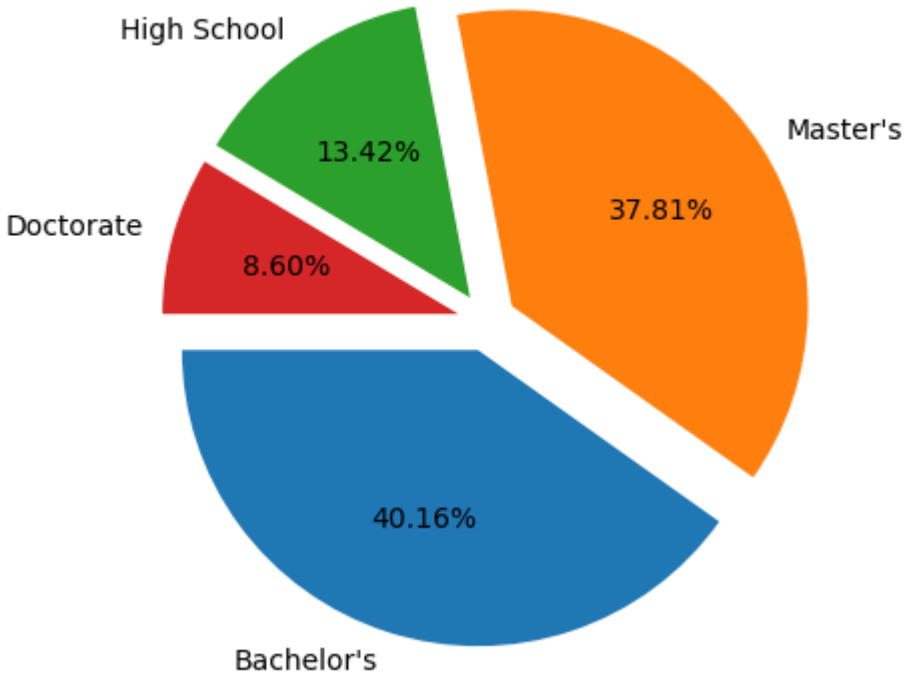
In [52]:
```
ax=e.plot(kind='bar')
ax.bar_label(ax.containers[0])
plp.show()
```



In [41]:
```
cdf=visa_df['education_of_employee'].value_counts()
keys=cdf.keys()
values=cdf.values
keys,values
```

Out[41]:
```
(Index(['Bachelor's', 'Master's', 'High School', 'Doctorate'], dtype='obje
ct', name='education_of_employee'),
 array([10234,  9634,  3420,  2192], dtype=int64))
```

In [50]:
```
plp.pie(values,explode=[0.1,0.1,0.1,0.1],labels=keys,autopct="%0.2f%%",star
plp.show()
```



In [7]:
```
region_of_employment=visa_df['region_of_employment'].unique()
region_of_employment
```

Out[7]:
```
array(['West', 'Northeast', 'South', 'Midwest', 'Island'], dtype=object)
```

In [ ]:
```

```

In [8]:
```
re=[]
for i in region_of_employment:
    r_of_e=visa_df['region_of_employment']==i
    value=len(visa_df[r_of_e])
    re.append(value)
    print(f'{i}:{value}')
```

```
West:6586
Northeast:7195
South:7017
Midwest:4307
Island:375
```

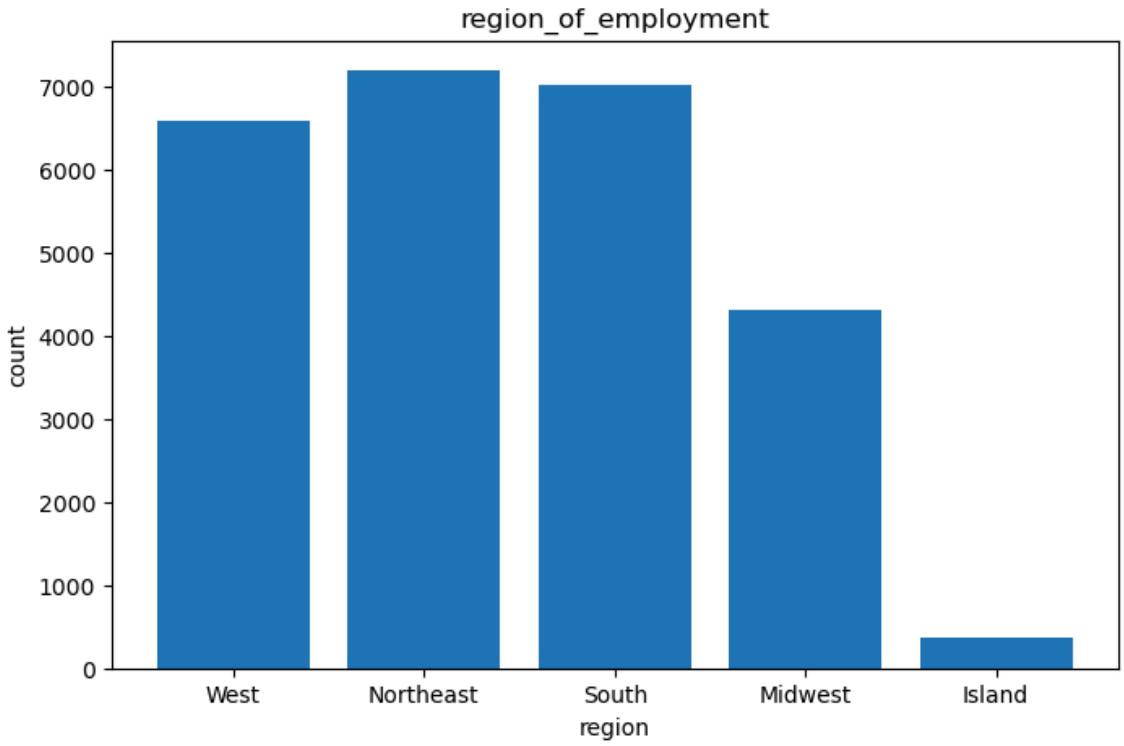
In [9]:
```
region_of_employment_df=pd.DataFrame(zip(region_of_employment,re),columns=[
region_of_employment_df
```

Out[9]:

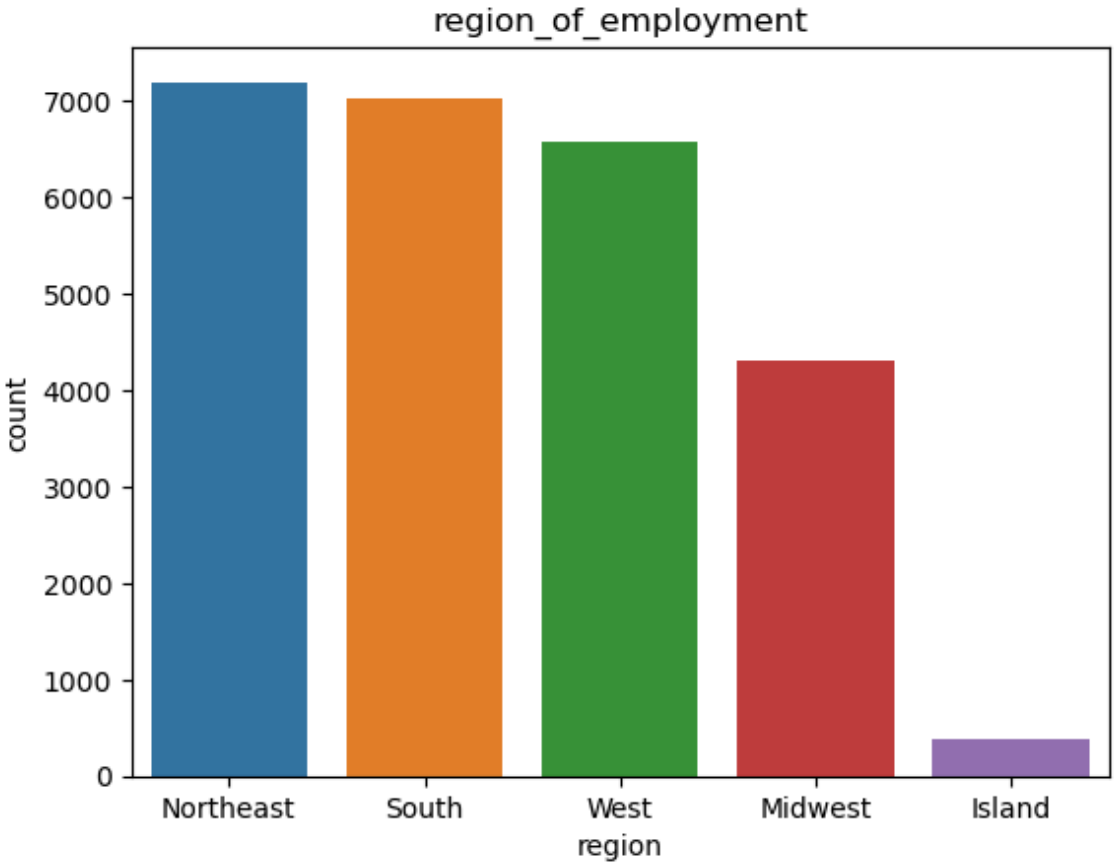
	region_of_employment	count
0	West	6586
1	Northeast	7195
2	South	7017
3	Midwest	4307
4	Island	375

In [10]:
```
region_of_employment_df.to_csv('region_of_employment_df.csv',index=False)
```

In [58]:
```
plp.figure(figsize=(8,5))
plp.bar('region_of_employment','count',data=region_of_employment_df)
plp.xlabel('region')
plp.ylabel('count')
plp.title('region_of_employment')
plp.show()
```



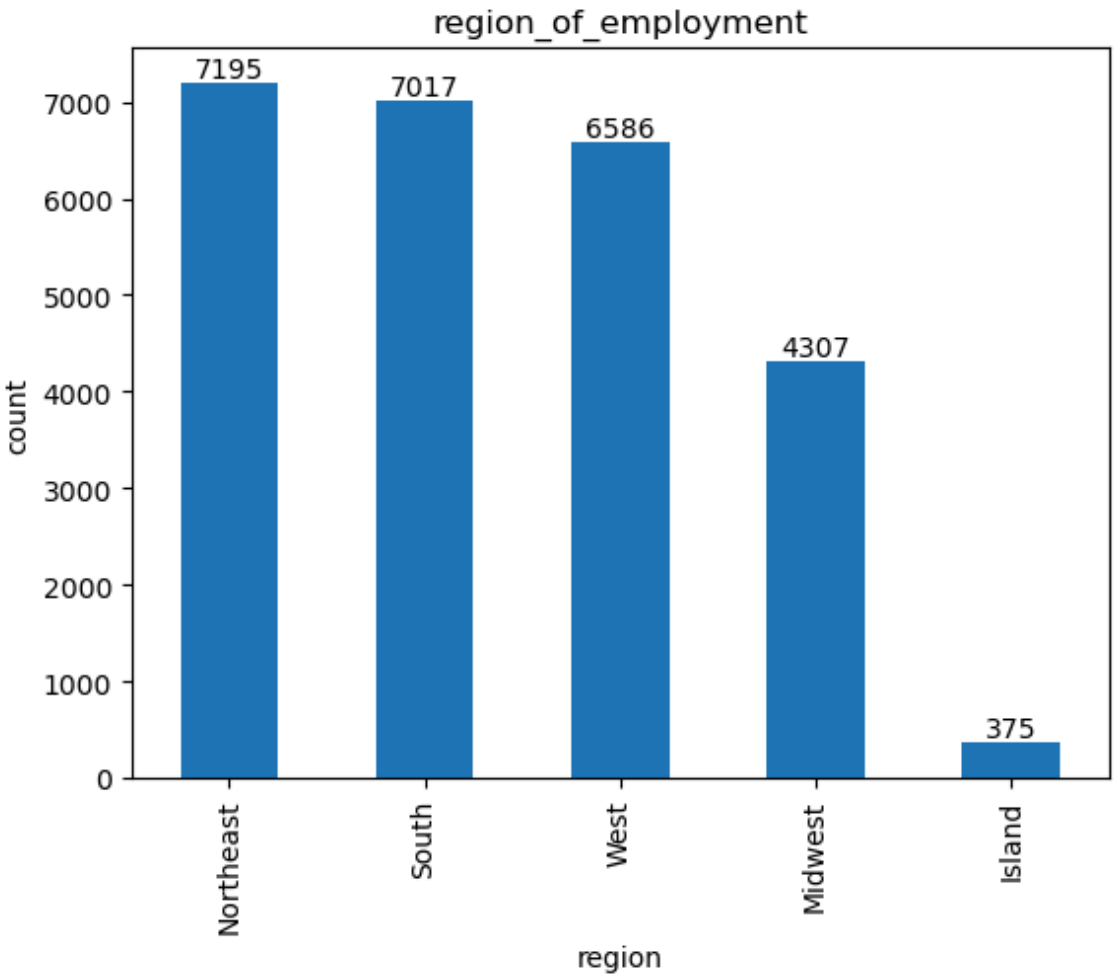
In [73]:
```
a=visa_df['region_of_employment'].value_counts()
key=a.keys()
values=a.values
sns.countplot(data=visa_df,x='region_of_employment',order=key)
plp.title('region_of_employment')
plp.xlabel('region')
plp.show()
```



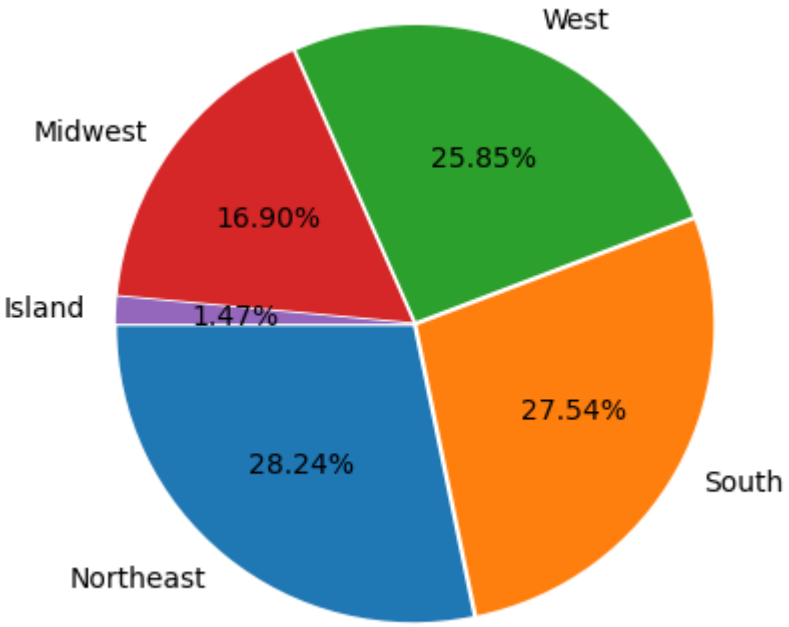
In [74]:
```
key,values
```

Out[74]:
```
(Index(['Northeast', 'South', 'West', 'Midwest', 'Island'], dtype='objec
t', name='region_of_employment'),
 array([7195, 7017, 6586, 4307,  375], dtype=int64))
```

In [67]:
```
b=a.plot(kind='bar')
b.bar_label(b.containers[0])
plp.xlabel('region')
plp.ylabel('count')
plp.title('region_of_employment')
plp.show()
```



In [77]:
```
plp.pie(values,explode=[0.01,0.01,0.01,0.01,0.01],labels=key,autopct='%0.2f
plp.show()
```



In [11]:
```
uniqe_case_status=visa_df['case_status'].unique()
uniqe_case_status
```

Out[11]:
```
array(['Denied', 'Certified'], dtype=object)
```

In [12]:
```
c_c_s=[]
for i in uniqe_case_status:
    c_t=visa_df['case_status']==i
    value=len(visa_df[c_t])
    c_c_s.append(value)
    print(f'{i}:{value}')
```

```
Denied:8462
Certified:17018
```

In [13]:
```
case_status_df=pd.DataFrame(zip(uniqe_case_status,c_c_s),columns=['case_sta
case_status_df
```

Out[13]:

	case_status	count
0	Denied	8462
1	Certified	17018

In [14]:
```
case_status_df.to_csv('case_status_df.csv',index=False)
```

In [ ]:
```

```