```
In [51]:
          import pandas as pd
          import numpy as np
          visadf = pd.read_csv(r"C:\Users\Anuja_PC\OneDrive\Documents\dataFiles\Visadatase
In [2]:
          visadf
Out[2]:
                     case_id continent education_of_employee has_job_experience requires_job_1
               0
                     EZYV01
                                   Asia
                                                    High School
                                                                                 Ν
               1
                     EZYV02
                                   Asia
                                                        Master's
               2
                     EZYV03
                                   Asia
                                                      Bachelor's
                                                                                 Ν
               3
                     EZYV04
                                   Asia
                                                      Bachelor's
                                                                                 Ν
               4
                     EZYV05
                                  Africa
                                                       Master's
                                                                                  Υ
          25475 EZYV25476
                                   Asia
                                                      Bachelor's
                                                                                  Υ
          25476 EZYV25477
                                   Asia
                                                    High School
          25477 EZYV25478
                                   Asia
                                                       Master's
                                                                                  Υ
          25478 EZYV25479
                                                        Master's
                                   Asia
          25479 EZYV25480
                                   Asia
                                                      Bachelor's
                                                                                  Υ
         25480 rows × 12 columns
```

#### Converting case\_id into numerical data

## Converting continent into numerical data

```
In [42]: visadf = pd.read_csv(r"C:\Users\Anuja_PC\OneDrive\Documents\dataFiles\Visadatase
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
visadf['continent']=le.fit_transform(visadf['continent'])
valueList = visadf['continent'].values
```

```
labels = le.inverse_transform(visadf['continent'])
          dictContinent=dict(zip(labels,valueList))
          dictContinent
Out[42]: {'Asia': 1,
           'Africa': 0,
           'North America': 3,
           'Europe': 2,
           'South America': 5,
           'Oceania': 4}
In [43]: visadfContinent = visadf['continent']
          visadfContinent
Out[43]: 0
                    1
          1
                    1
          2
                    1
          3
                    1
          4
                    0
          25475
                   1
          25476
                   1
          25477
                   1
          25478
                   1
          25479
          Name: continent, Length: 25480, dtype: int32
In [44]:
          visadf ## converted continent into numeric data
Out[44]:
                     case_id continent education_of_employee has_job_experience requires_job_1
              0
                     EZYV01
                                     1
                                                   High School
                                                                               Ν
              1
                     EZYV02
                                                      Master's
              2
                     EZYV03
                                     1
                                                     Bachelor's
                                                                               Ν
              3
                     EZYV04
                                     1
                                                     Bachelor's
                                                                               Ν
              4
                     EZYV05
                                     0
                                                      Master's
                                                                                Υ
          25475 EZYV25476
                                     1
                                                     Bachelor's
                                                                                Υ
          25476 EZYV25477
                                                   High School
          25477 EZYV25478
                                     1
                                                      Master's
                                                                                Υ
          25478 EZYV25479
                                     1
                                                      Master's
          25479 EZYV25480
                                     1
                                                     Bachelor's
                                                                                Υ
         25480 rows × 12 columns
```

Converting education\_of\_employee into numerical data

```
In [45]: visadf = pd.read_csv(r"C:\Users\Anuja_PC\OneDrive\Documents\dataFiles\Visadatase
         from sklearn.preprocessing import LabelEncoder
         le = LabelEncoder()
         visadf['education_of_employee']=le.fit_transform(visadf['education_of_employee']
         valueList = visadf['education_of_employee'].values
         labels = le.inverse_transform(visadf['education_of_employee'])
         dictEducation_of_employee=dict(zip(labels,valueList))
         dictEducation_of_employee
Out[45]: {'High School': 2, "Master's": 3, "Bachelor's": 0, 'Doctorate': 1}
In [47]: visadfEduOfEmployee = visadf['education_of_employee']
         visadfEduOfEmployee
Out[47]: 0
                  2
                  3
         2
                  0
         3
                  0
         4
                  3
         25475
                  0
         25476
                  2
         25477 3
         25478
                3
         25479
         Name: education_of_employee, Length: 25480, dtype: int32
        visadf ## converted education_of_employee into numeric data
```

Out[48]:		case_id	continent	education_of_employee	has_job_experience	requires_job_1
	0	EZYV01	Asia	2	N	
	1	EZYV02	Asia	3	Υ	
	2	EZYV03	Asia	0	N	
	3	EZYV04	Asia	0	N	
	4	EZYV05	Africa	3	Υ	
	•••					
	25475	EZYV25476	Asia	0	Υ	
	25476	EZYV25477	Asia	2	Υ	
	25477	EZYV25478	Asia	3	Υ	
	25478	EZYV25479	Asia	3	Υ	
	25479	EZYV25480	Asia	0	Υ	
	25480 rd	ows × 12 colu	umns			
	4					•
	Conver	ting has job	experienc	e into numerical data		

## Converting has\_job\_experience into numerical data

```
In [52]: visadf = pd.read_csv(r"C:\Users\Anuja_PC\OneDrive\Documents\dataFiles\Visadatase
    con=visadf['has_job_experience']=='Y'
    visadf['has_job_experience']=np.where(con,0,1)
    visadf #converted has_job_experience column to numeric
```

52]:		case_id	continent	education_of_employee	has_job_experience	requires_job_
	0	EZYV01	Asia	High School	1	
	1	EZYV02	Asia	Master's	0	
	2	EZYV03	Asia	Bachelor's	1	
	3	EZYV04	Asia	Bachelor's	1	
	4	EZYV05	Africa	Master's	0	
	•••					
254	175	EZYV25476	Asia	Bachelor's	0	
254	176	EZYV25477	Asia	High School	0	
254	177	EZYV25478	Asia	Master's	0	
254	178	EZYV25479	Asia	Master's	0	
254	179	EZYV25480	Asia	Bachelor's	0	
2548	30 rc	ows × 12 colu	ımns			

Out[55]:

## Converting requires\_job\_training into numerical data

```
In [55]: visadf = pd.read_csv(r"C:\Users\Anuja_PC\OneDrive\Documents\dataFiles\Visadatase
    con=visadf['requires_job_training']=='Y'
    visadf['requires_job_training']=np.where(con,0,1)
    visadf #converted requires_job_training column to numeric
```

	case_id	continent	education_of_employee	has_job_experience	requires_job_t
0	EZYV01	Asia	High School	N	
1	EZYV02	Asia	Master's	Υ	
2	EZYV03	Asia	Bachelor's	N	
3	EZYV04	Asia	Bachelor's	N	
4	EZYV05	Africa	Master's	Υ	
•••					
25475	EZYV25476	Asia	Bachelor's	Υ	
25476	EZYV25477	Asia	High School	Υ	
25477	EZYV25478	Asia	Master's	Υ	
25478	EZYV25479	Asia	Master's	Υ	
25479	EZYV25480	Asia	Bachelor's	Υ	
25480 re	ows × 12 coli	umns			

# Converting region\_of\_employment into numerical data

Out[57]:		case_id	continent	education_of_employee	has_job_experience	requires_job_1
	0	EZYV01	Asia	High School	N	
	1	EZYV02	Asia	Master's	Υ	
	2	EZYV03	Asia	Bachelor's	N	
	3	EZYV04	Asia	Bachelor's	N	
	4	EZYV05	Africa	Master's	Υ	
	•••					
	25475	EZYV25476	Asia	Bachelor's	Υ	
	25476	EZYV25477	Asia	High School	Υ	
	25477	EZYV25478	Asia	Master's	Υ	
	25478	EZYV25479	Asia	Master's	Υ	
	25479	EZYV25480	Asia	Bachelor's	Υ	
	25480 rd	ows × 12 colu	umns			
	4					•

## Converting unit\_of\_wage into numerical data

Out[61]:		case_id	continent	education_of_employee	has_job_experience	requires_job_1			
	0	EZYV01	Asia	High School	N				
	1	EZYV02	Asia	Master's	Υ				
	2	EZYV03	Asia	Bachelor's	N				
	3	EZYV04	Asia	Bachelor's	N				
	4	EZYV05	Africa	Master's	Υ				
	•••								
	25475	EZYV25476	Asia	Bachelor's	Υ				
	25476	EZYV25477	Asia	High School	Υ				
	25477	EZYV25478	Asia	Master's	Υ				
	25478	EZYV25479	Asia	Master's	Υ				
	25479	EZYV25480	Asia	Bachelor's	Υ				
	25480 rd	ows × 12 colu	umns						
	4			_		<b>&gt;</b>			
In [62]:	visadf	- nd nead	csv(n"C•\II	cans\Anuia PC\OneDnive	\Documents\dataEile	as\Visadatase			
TII [02].	<pre>visadf = pd.read_csv(r"C:\Users\Anuja_PC\OneDrive\Documents\dataFiles\Visadatase con=visadf['full_time_position']=='Y'</pre>								
	visadf	['full_time	_position'	]=np.where(con,0,1) s_job_training column	to numeric				
Out[62]:	visadf	['full_time #converted	_position' d require	]=np.where(con,0,1)		requires_job_1			
Out[62]:	visadf	['full_time #converted	_position' d require	]=np.where(con,0,1) s_job_training column		requires_job_1			
Out[62]:	visadf visadf	['full_time #converted case_id	_position' d require continent	]=np.where(con,0,1) s_job_training column education_of_employee	has_job_experience	requires_job_1			
Out[62]:	visadf visadf	['full_time #converted case_id	position' d require  continent  Asia	]=np.where(con,0,1) s_job_training column education_of_employee  High School	has_job_experience	requires_job_1			
Out[62]:	visadf visadf 0	case_id  EZYV01  EZYV02	position' d required  continent  Asia  Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's	has_job_experience  N  Y	requires_job_1			
Out[62]:	visadf visadf 0 1	case_id  EZYV01  EZYV02  EZYV03	position' d require  continent  Asia  Asia  Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's	has_job_experience  N  Y  N	requires_job_1			
Out[62]:	visadf visadf 0 1 2	case_id EZYV01 EZYV02 EZYV03 EZYV04	position' d require  continent  Asia  Asia  Asia  Asia  Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Bachelor's	has_job_experience  N  Y  N  N	requires_job_1			
Out[62]:	visadf visadf 0 1 2 3	case_id EZYV01 EZYV02 EZYV03 EZYV04 EZYV05	position' d required  continent  Asia  Asia  Asia  Asia  Africa	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Bachelor's  Master's	has_job_experience  N  Y  N  N  Y	requires_job_t			
Out[62]:	visadf visadf 0 1 2 3 4	case_id  EZYV01  EZYV02  EZYV03  EZYV04  EZYV05	position' d required  continent  Asia  Asia  Asia  Asia  Africa	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Bachelor's  Master's  Master's	has_job_experience  N Y N N V V	requires_job_1			
Out[62]:	0 1 2 3 4 25475	case_id EZYV01 EZYV02 EZYV03 EZYV04 EZYV05	position' d required continent  Asia Asia Asia Asia Africa Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Bachelor's  Master's   Bachelor's	has_job_experience  N Y N N Y  N Y  N Y	requires_job_1			
Out[62]:	0 1 2 3 4  25475 25476	case_id EZYV01 EZYV02 EZYV03 EZYV04 EZYV05 EZYV25476	position'd required continent  Asia Asia Asia Asia Africa Asia Asia Africa Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Bachelor's  Master's   Bachelor's  High School	has_job_experience  N Y N N Y  N Y Y Y Y  T Y  T Y Y	requires_job_1			
Out[62]:	0 1 2 3 4  25475 25476	case_id EZYV01 EZYV02 EZYV03 EZYV04 EZYV05 EZYV25476 EZYV25477 EZYV25478	position'd required continent  Asia Asia Asia Asia Africa Asia Asia Asia Africa Asia Asia Asia	education_of_employee  High School  Master's  Bachelor's  Master's  Master's  Master's  High School  Master's  Master's  Master's  Master's	has_job_experience  N Y N N Y Y Y Y Y Y	requires_job_1			
Out[62]:	visadf visadf 0 1 2 3 4  25475 25476 25477 25478	'full_time #converted   case_id   EZYV01   EZYV02   EZYV03   EZYV04   EZYV05     EZYV25476   EZYV25477   EZYV25478   EZYV25479	position'd required  continent  Asia  Asia  Asia  Africa   Asia  Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Master's   Bachelor's  High School  Master's  Master's  High School  Master's	has_job_experience  N Y N N N Y  Y Y Y  Y Y Y Y	requires_job_1			
Out[62]:	visadf visadf 0 1 2 3 4  25475 25476 25477 25478	'full_time #converted	position'd required  continent  Asia  Asia  Asia  Africa   Asia  Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Master's   Bachelor's  High School  Master's  Master's  High School  Master's	has_job_experience  N Y N N N Y  Y Y Y  Y Y Y Y	requires_job_t			
	0 1 2 3 4 25475 25476 25477 25478 25479	case_id  EZYV01  EZYV02  EZYV03  EZYV04  EZYV05   EZYV25476  EZYV25477  EZYV25477  EZYV25479  EZYV25480  Dows × 12 column	position'd required continent  Asia Asia Asia Africa Asia Asia Asia Asia Asia Asia Asia Asia	]=np.where(con,0,1) s_job_training column  education_of_employee  High School  Master's  Bachelor's  Master's   Bachelor's  High School  Master's  Bachelor's  High School  Master's  Bachelor's  Bachelor's	has_job_experience  N Y N N N Y  Y Y Y Y Y Y Y Y				
Out[62]:	0 1 2 3 4 25475 25476 25477 25478 25479 visadf	case_id  EZYV01  EZYV02  EZYV03  EZYV04  EZYV05   EZYV25476  EZYV25477  EZYV25477  EZYV25479  EZYV25480  Dows × 12 column	position'd required continent  Asia Asia Asia Africa Asia Asia Asia Asia Asia Asia Csv(r"C:\U	=np.where(con,0,1)   s_job_training column    education_of_employee	has_job_experience  N Y N N N Y  Y Y Y Y Y Y Y Y				

visadf['full\_time\_position']=np.where(con,0,1)
visadf #converted full\_time\_position column to numeric

Out[63]:

	case_id	continent	education_of_employee	has_job_experience	requires_job_t
0	EZYV01	Asia	High School	N	
1	EZYV02	Asia	Master's	Υ	
2	EZYV03	Asia	Bachelor's	N	
3	EZYV04	Asia	Bachelor's	N	
4	EZYV05	Africa	Master's	Υ	
•••					
25475	EZYV25476	Asia	Bachelor's	Υ	
25476	EZYV25477	Asia	High School	Υ	
25477	EZYV25478	Asia	Master's	Υ	
25478	EZYV25479	Asia	Master's	Υ	
25479	EZYV25480	Asia	Bachelor's	Υ	
25480 r	ows × 12 col	umns			
4					