In [1]:
```
import pandas as pd #install pandas using pip
import os
import numpy as np #numpy module is used to perform sum on rows/columns
import matplotlib.pyplot as ltp #to plot graphs
```

In [2]:
```
pwd  #current directory
```

Out[2]:
```
'C:\\Users\\susan\\Desktop\\Data anlaysis on forest fires'
```

In [3]:
```
dataset = pd.read_csv(r"D:\Data anlaysis on forest fires\archive\amazon.csv")   #importing the csv files
```

In [4]:
```
dataset.head()  #checking if the csv file is imported using head()-returns first 5 rows of the dataset
```

Out[4]:

| | year | state | month | number | date |
|---|---|---|---|---|---|
| **0** | 1998 | Acre | Janeiro | 0.0 | 1998-01-01 |
| **1** | 1999 | Acre | Janeiro | 0.0 | 1999-01-01 |
| **2** | 2000 | Acre | Janeiro | 0.0 | 2000-01-01 |
| **3** | 2001 | Acre | Janeiro | 0.0 | 2001-01-01 |
| **4** | 2002 | Acre | Janeiro | 0.0 | 2002-01-01 |

In [5]:
```
# headers
print(dataset.columns)
```

```
Index(['year', 'state', 'month', 'number', 'date'], dtype='object')
```

In [6]:
```
# simple analysis plot
# 1.total fires reported by year
# pivot_table function

pivot1 = pd.pivot_table(dataset,values = "number",index = ["year"], aggfunc=np.sum)
pivot1
```

Out[6]:

| | number |
|---|---|
| **year** | |
| **1998** | 20013.971 |
| **1999** | 26882.821 |
| **2000** | 27351.251 |
| **2001** | 29071.612 |
| **2002** | 37390.600 |
| **2003** | 42760.674 |
| **2004** | 38453.163 |
| **2005** | 35004.965 |
| **2006** | 33832.161 |
| **2007** | 33037.413 |
| **2008** | 29378.964 |
| **2009** | 39117.178 |
| **2010** | 37037.449 |
| **2011** | 34633.545 |
| **2012** | 40084.860 |
| **2013** | 35146.118 |
| **2014** | 39621.183 |
| **2015** | 41208.292 |
| **2016** | 42212.229 |
| **2017** | 36685.624 |

In [7]:
```
#translating the month names
month_map = {'Janeiro': 'January', 'Fevereiro': 'February', 'Março': 'March', 'Abril': 'April', 'Maio': 'May',
             'Junho': 'June', 'Julho': 'July', 'Agosto': 'August', 'Setembro': 'September', 'Outubro': 'October',
             'Novembro': 'November', 'Dezembro': 'December'}
#mapping translated months
dataset['month'] = dataset['month'].map(month_map)
```

In [8]:
```
print(dataset['month'])  #checking if it is translated
```

```
0         January
1         January
2         January
3         January
4         January
          ...
6449     December
6450     December
6451     December
6452     December
6453     December
Name: month, Length: 6454, dtype: object
```

In [9]:
```
#total no.fires by month
pivot2 = pd.pivot_table(dataset,values = "number",index = ["month"], aggfunc=np.sum)
pivot2
```

Out[9]:

| | number |
|---|---|
| **month** | |
| **April** | 28188.770 |
| **August** | 88050.435 |
| **December** | 57535.480 |
| **February** | 30848.050 |
| **January** | 47747.844 |
| **July** | 92326.113 |
| **June** | 56010.675 |
| **May** | 34731.363 |
| **November** | 85508.054 |
| **October** | 88681.579 |
| **September** | 58578.305 |

In [10]:
```
#total no.fires by state
pivot3 = pd.pivot_table(dataset,values = "number",index = ["state"], aggfunc=np.sum)
pivot3
```

Out[10]:

| | number |
|---|---|
| **state** | |
| **Acre** | 18464.030 |
| **Alagoas** | 4644.000 |
| **Amapa** | 21831.576 |
| **Amazonas** | 30650.129 |
| **Bahia** | 44746.226 |
| **Ceara** | 30428.063 |
| **Distrito Federal** | 3561.000 |
| **Espirito Santo** | 6546.000 |
| **Goias** | 37695.520 |
| **Maranhao** | 25129.131 |
| **Mato Grosso** | 96246.028 |
| **Minas Gerais** | 37475.258 |
| **Paraiba** | 52435.918 |
| **Par** | 24512.144 |
| **Pernambuco** | 24498.000 |
| **Piau** | 37803.747 |
| **Rio** | 45160.865 |
| **Rondonia** | 20285.429 |
| **Roraima** | 24385.074 |
| **Santa Catarina** | 24359.852 |
| **Sao Paulo** | 51121.198 |
| **Sergipe** | 3237.000 |
| **Tocantins** | 33707.885 |

In [ ]:
```
```