In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import plotly.express as px
import seaborn as sns
```

In [2]:
```
pd.set_option('display.max_columns',200)
```

In [3]:
```
df=pd.read_csv(r"C:\Users\berid\OneDrive\Desktop\mydata\globalterrorismdb_0718dist.csv",encoding_errors='ignore')
```

```
C:\Users\berid\AppData\Local\Programs\Python\Python39\lib\site-packages\IPython\core\interactiveshell.py:3251: DtypeWarning: Columns (4,6,31,33,61,62,63,76,79,90,92,94,96,114,115,121) have mixed types.Specify dtype option on import or set low_memory=False.
  exec(code_obj, self.user_global_ns, self.user_ns)
```

In [4]:
```
df['iyear']=df.iyear.astype(int)
```

In [5]:
```
df['imonth']=df['imonth'].replace(0,1,regex=False)
```

In [6]:
```
df.head()
```

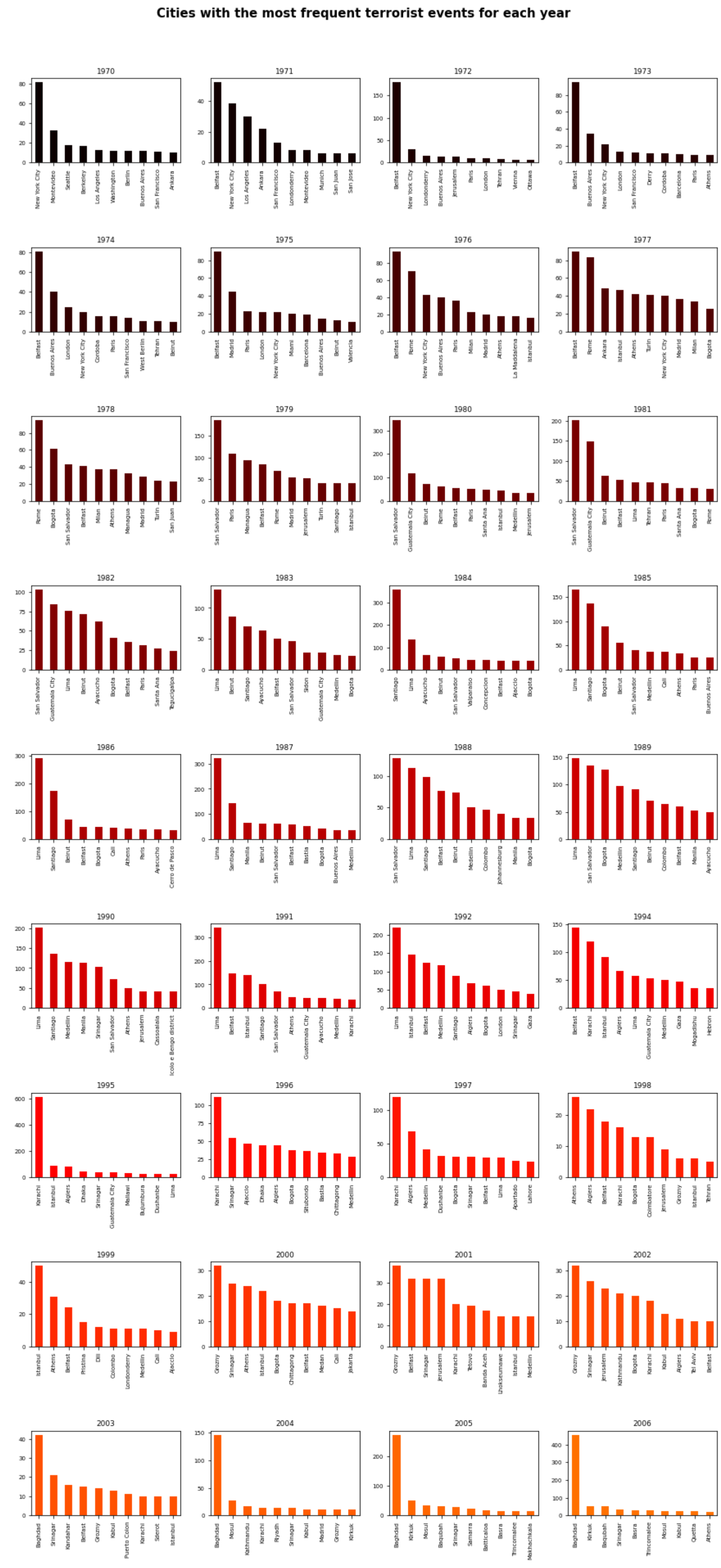
Out[6]:

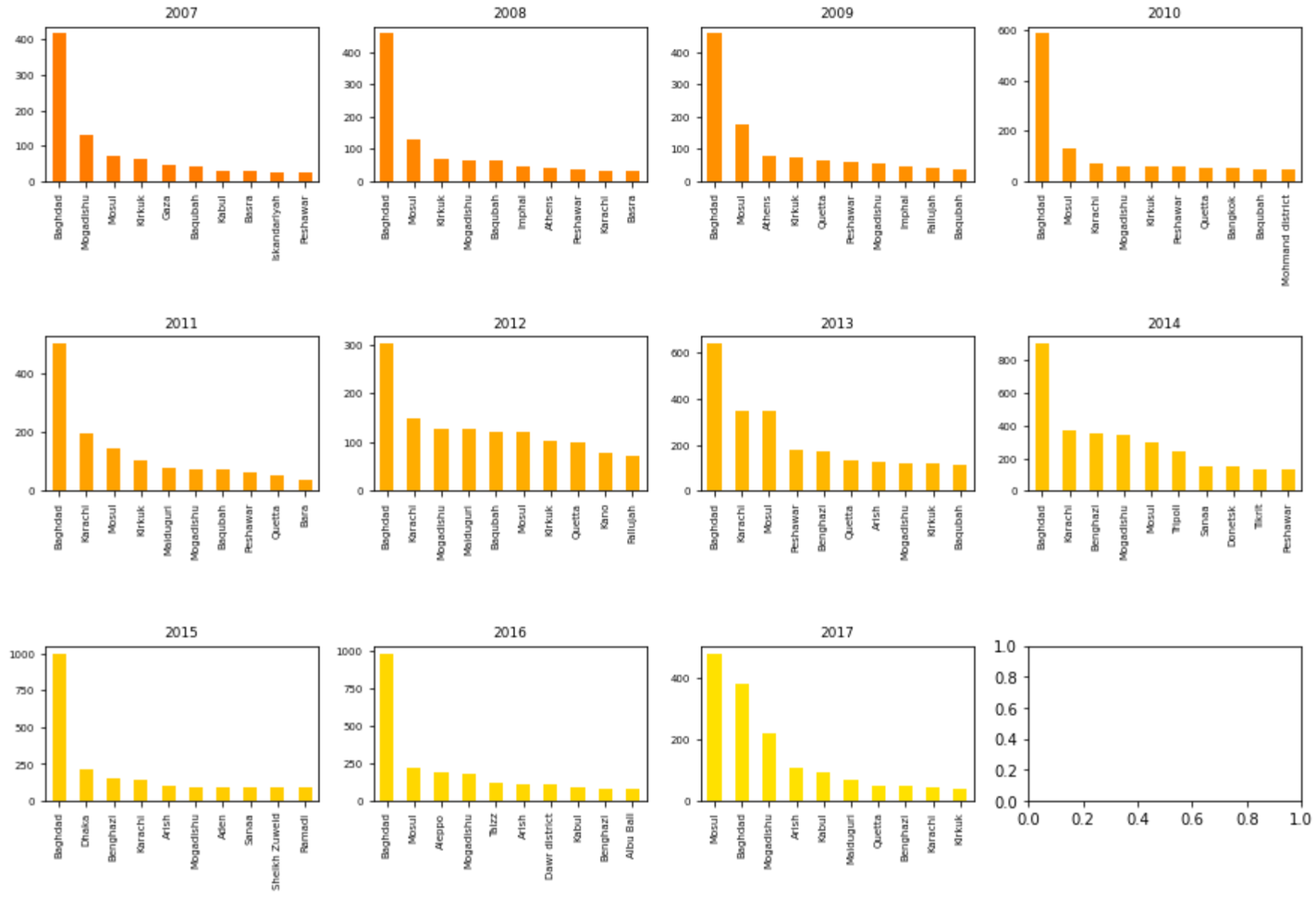
| | eventid | iyear | imonth | iday | approxdate | extended | resolution | country | country_txt | region | region_txt | provstate | city | latitude | longitude | specificity | vicinity | location | summary | crit1 | crit2 | crit3 | doubtt |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 197000000001 | 1970 | 7 | 2 | NaN | 0 | NaN | 58 | Dominican Republic | 2 | Central America & Caribbean | NaN | Santo Domingo | 18.456792 | -69.951164 | 1.0 | 0 | NaN | NaN | 1 | 1 | 1 | |
| **1** | 197000000002 | 1970 | 1 | 0 | NaN | 0 | NaN | 130 | Mexico | 1 | North America | Federal | Mexico city | 19.371887 | -99.086624 | 1.0 | 0 | NaN | NaN | 1 | 1 | 1 | |
| **2** | 197001000001 | 1970 | 1 | 0 | NaN | 0 | NaN | 160 | Philippines | 5 | Southeast Asia | Tarlac | Unknown | 15.478598 | 120.599741 | 4.0 | 0 | NaN | NaN | 1 | 1 | 1 | |
| **3** | 197001000002 | 1970 | 1 | 0 | NaN | 0 | NaN | 78 | Greece | 8 | Western Europe | Attica | Athens | 37.997490 | 23.762728 | 1.0 | 0 | NaN | NaN | 1 | 1 | 1 | |
| **4** | 197001000003 | 1970 | 1 | 0 | NaN | 0 | NaN | 101 | Japan | 4 | East Asia | Fukouka | Fukouka | 33.580412 | 130.396361 | 1.0 | 0 | NaN | NaN | 1 | 1 | 1 | - |

## cities with the most terrorist events for each year

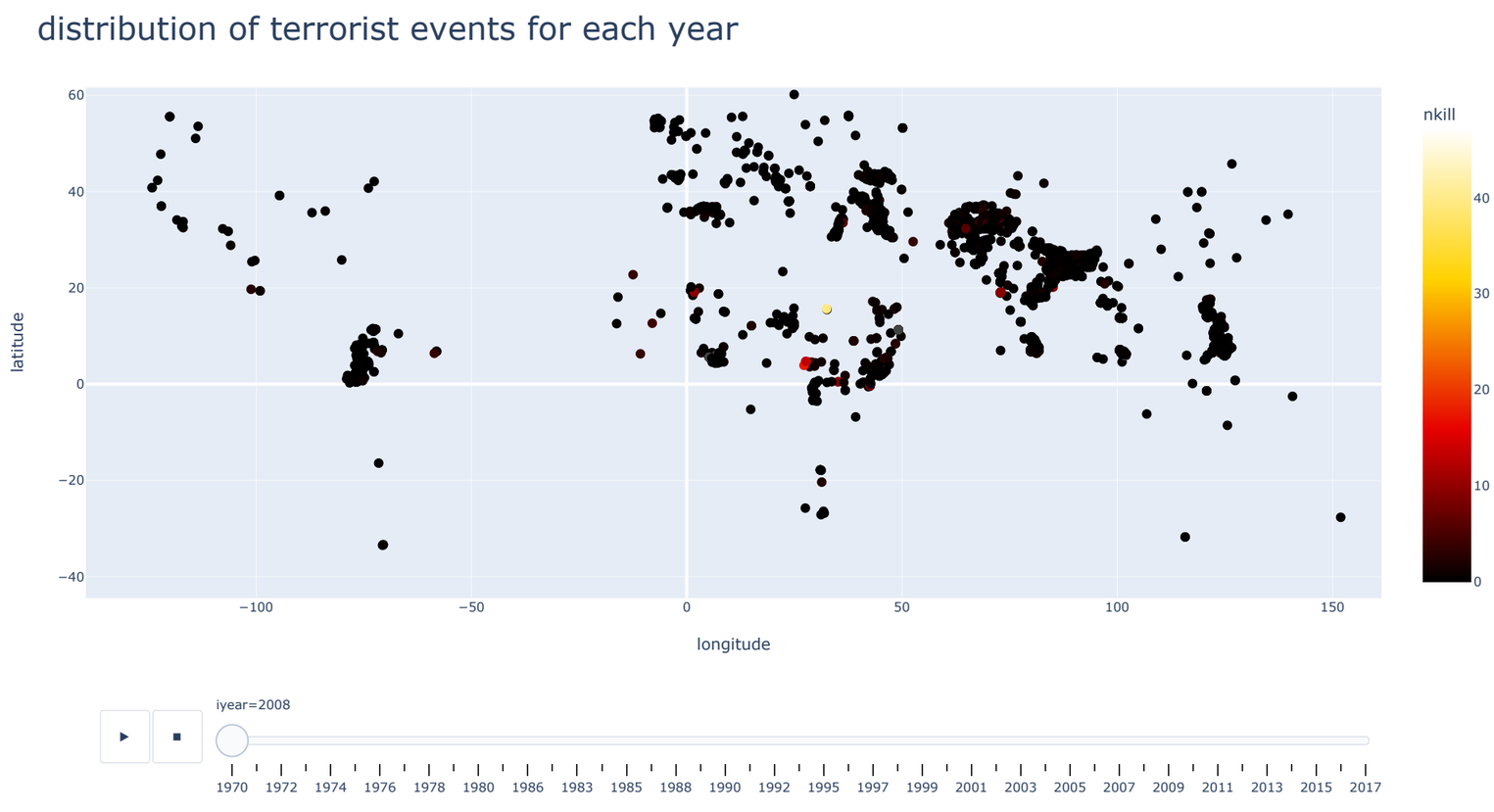
In [7]:
```
years=sorted(df.iyear.unique())
fig,axes=plt.subplots(12,4,figsize=(15,40))
colors=list(plt.get_cmap('hot')(np.linspace(0,0.7,47)))

for year,ax,color in zip(years,axes.ravel(),colors):
    grouped=df[(df.iyear==year)&(df.city.str.contains('known',regex=True,case=False)==False)].groupby('city')['eventid'].count().reset_index().sort_values('eventid',ascending=False)
    grouped.plot(kind='bar',ax=ax,x='city',y='eventid',xlabel='',legend=False,fontsize=7,color=color)
    ax.set_title(str(year),size=9)
plt.suptitle('Cities with the most frequent terrorist events for each year',fontweight='bold',size=15)
plt.subplots_adjust(hspace=1,top=0.95)
plt.show()
```






In [8]:
```
fig=px.scatter(df,x='longitude',y='latitude',
               animation_frame='iyear',
                hover_name='city',
               color='nkill',
                #size='nkill',
                color_continuous_scale=px.colors.sequential.Hot
               )
fig.update_layout(title_text='distribution of terrorist events for each year',font={'size':8},titlefont={'size':20})
fig.layout.updatemenus[0].buttons[0].args[1]["frame"]["duration"] = 2000
fig.show()
```

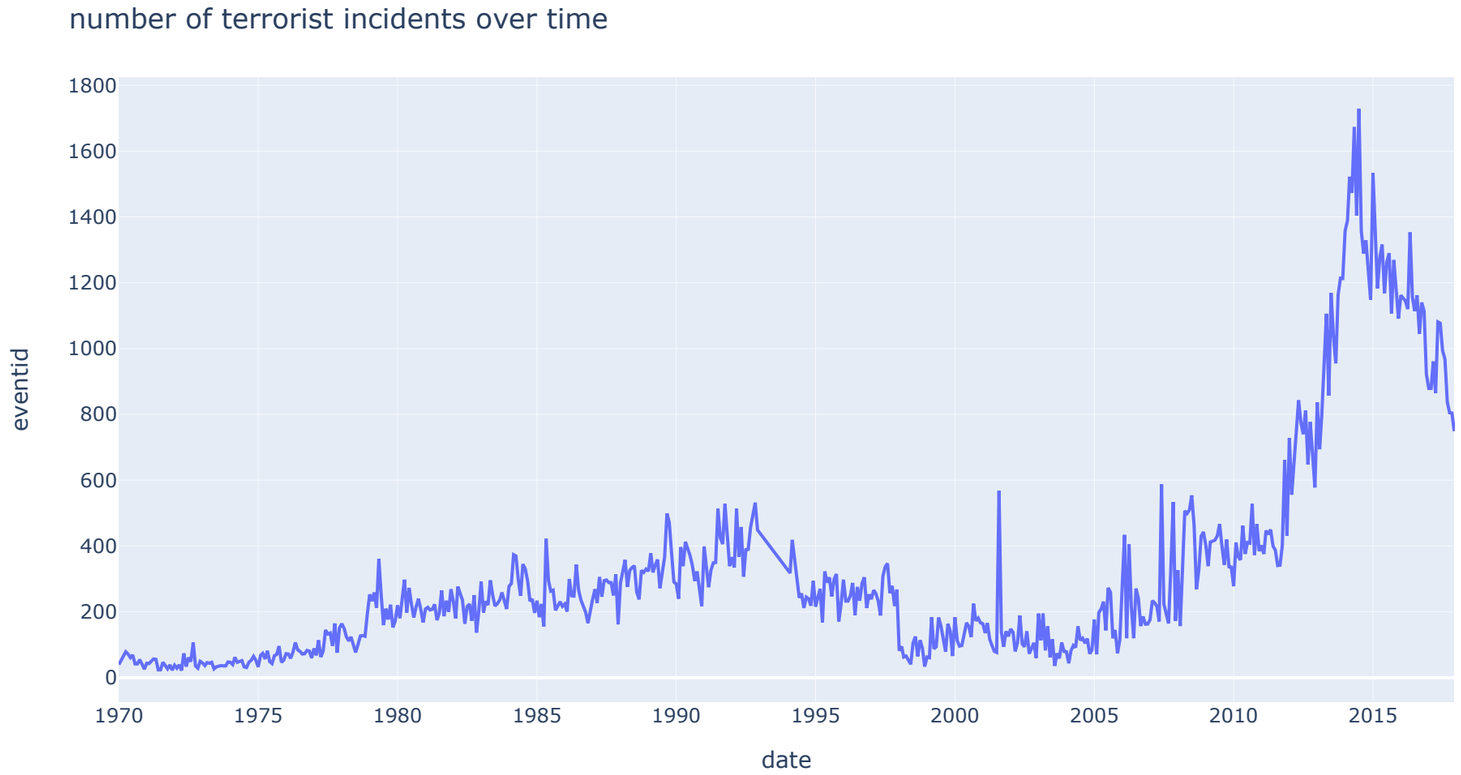



## how number of terrorist events is been increased over time

In [9]:
```
import datetime as dt
```

In [10]:
```
grouped=df.groupby(['iyear','imonth'])['eventid'].count().reset_index().sort_values(['iyear','imonth'])
grouped['date']=grouped.imonth.astype(str)+'-'+grouped.iyear.astype(str)
grouped['date']=pd.to_datetime(grouped['date'])
```

In [11]:
```
fig=px.line(grouped,x='date',y='eventid')
fig.update_layout(title_text='number of terrorist incidents over time')
```

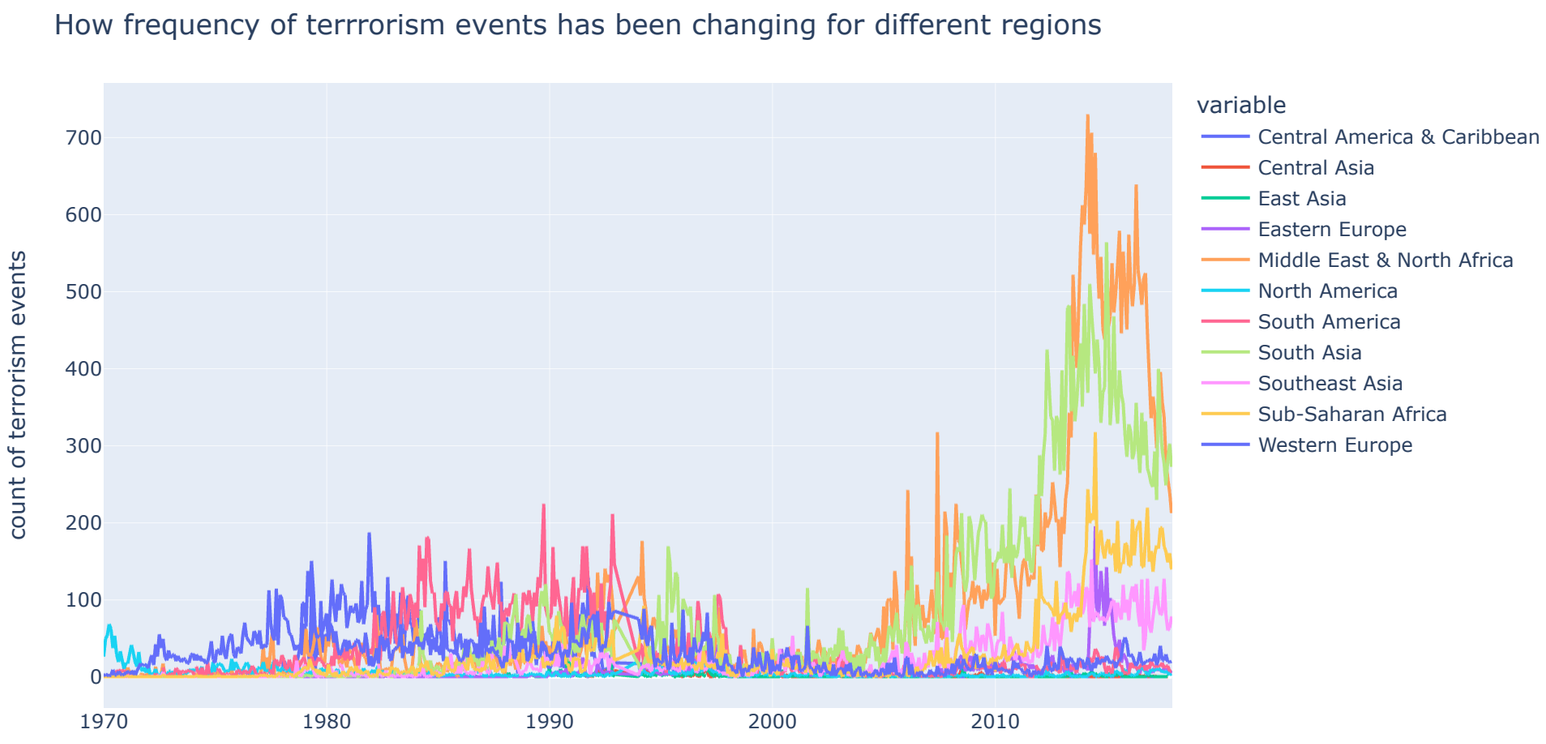


# How terrorism events frequency has been changing in different regions of the world

In [12]:
```
import datetime
```

In [13]:
```
grouped=df.groupby(['region_txt','iyear','imonth'])['eventid'].count().reset_index().sort_values(['iyear','imonth'])
grouped['date']=grouped.iyear.astype(str)+'-'+grouped.imonth.astype(str)
grouped['date']=pd.to_datetime(grouped.date)
pivoted=grouped.pivot('date','region_txt','eventid').reset_index()
```

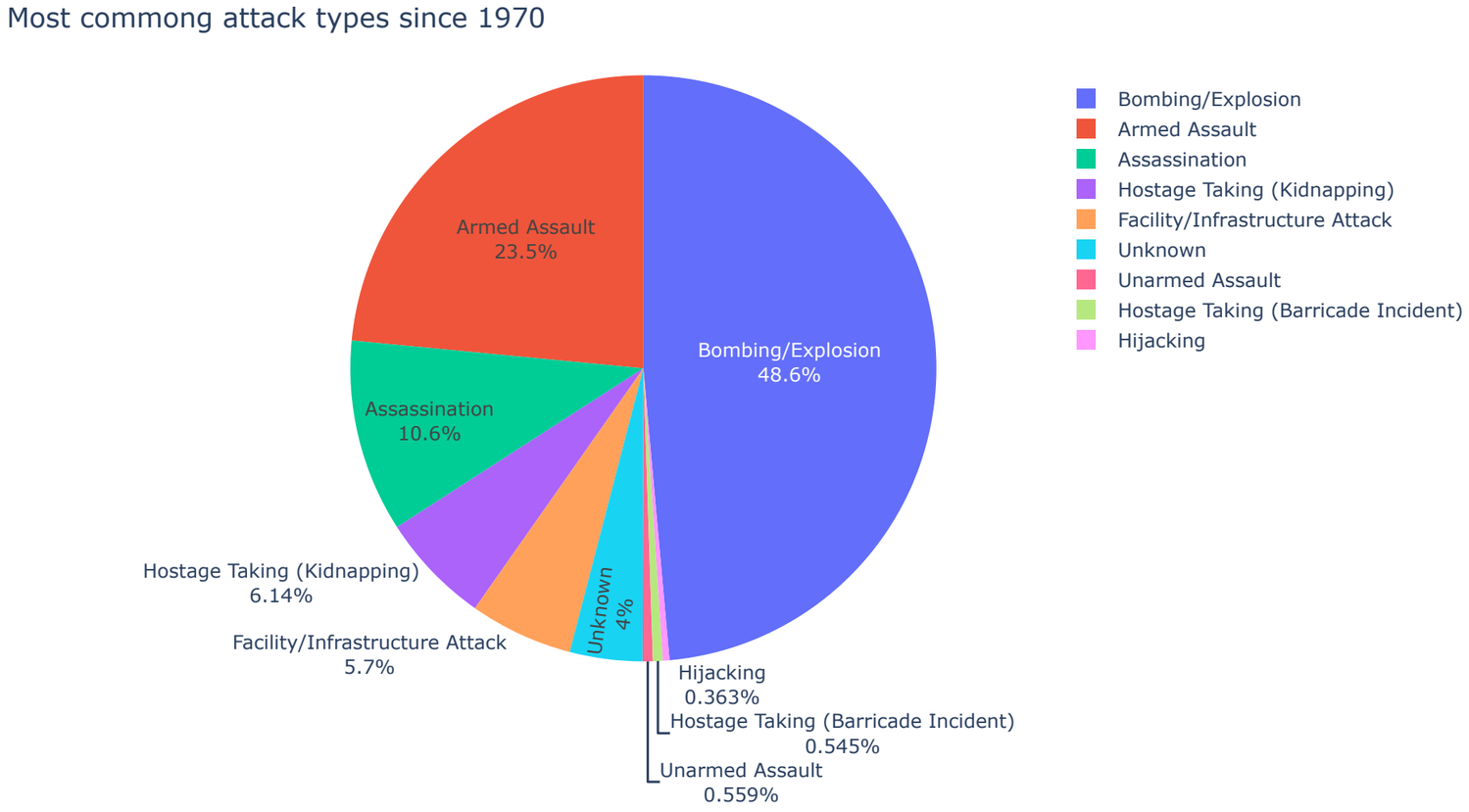
In [14]:
```
fig=px.line(pivoted,x='date',y=pivoted.columns[2:])
fig.update_layout(title_text='How frequency of terrrorism events has been changing for different regions')
fig.update_yaxes(title_text='count of terrorism events')
fig.update_xaxes(title_text=None)
fig.show()
```



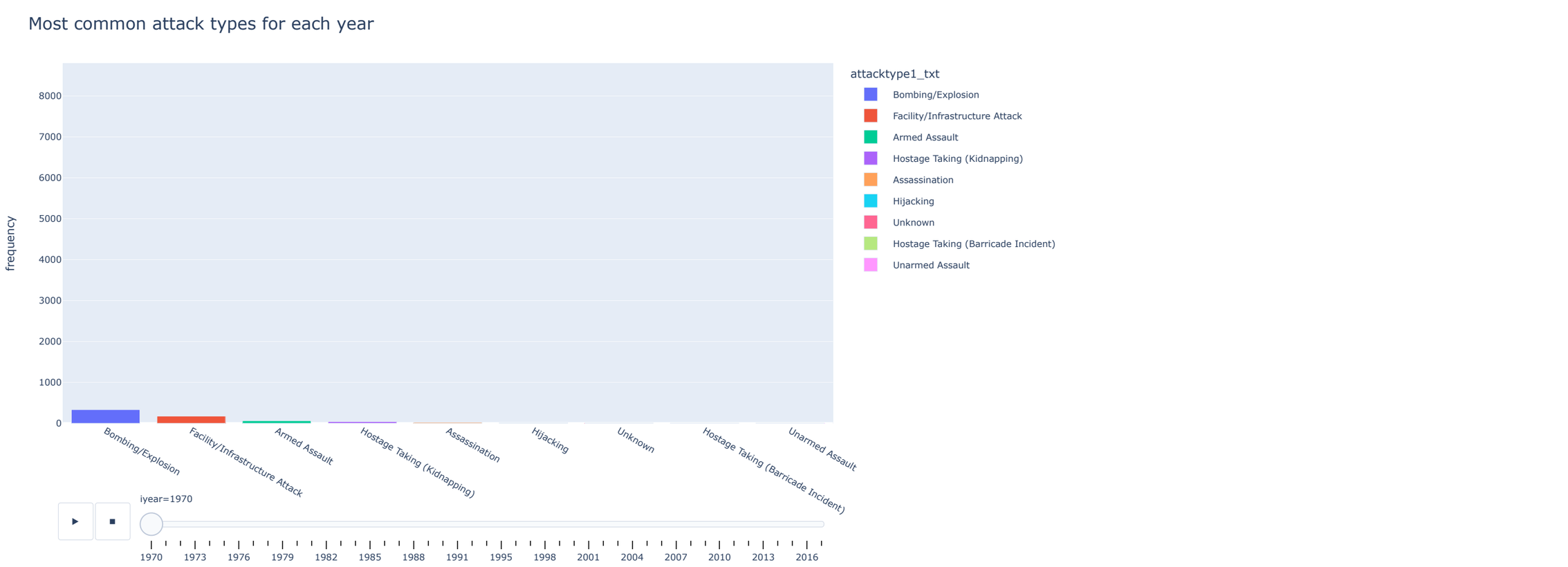

# most common attack type for each year

In [15]:
```
grouped=df.groupby(['iyear','attacktype1_txt'])['eventid'].count().reset_index()
```

In [16]:
```
fig1=px.pie(grouped,values='eventid',names='attacktype1_txt')
fig1.update_traces(#textposition='inside',
                   textinfo='label+percent')
fig1.update_layout(title_text='Most commong attack types since 1970')
```



In [17]:
```
fig2=px.bar(grouped.sort_values(['iyear','eventid'],ascending=[True,False]),
            x='attacktype1_txt',y='eventid',
            animation_frame='iyear',
            color='attacktype1_txt',
           range_y=(0,grouped.eventid.max()))
fig2.update_layout(title_text='Most common attack types for each year',font=dict(size=8),titlefont={'size':15})
fig2.update_yaxes(title_text='frequency')
fig2.update_xaxes(title_text=None)
fig2.show()
```




In [ ]: