# Analyzing Data

## Prison Helicopter Escapes

We begin by importing some helper functions.

## Get the Data

Now, let's get the data from the [List of helicopter prison escapes (https://en.wikipedia.org/wiki/List_of_helicopter_prison_escapes)](https://en.wikipedia.org/wiki/List_of_helicopter_prison_escapes) Wikipedia article.

In [3]:
```
from helper import *
```

In [4]:
```
url='https://en.wikipedia.org/wiki/List_of_helicopter_prison_escapes'
```

In [5]:
```
data=data_from_url(url)
```

Let's print the first three rows

In [6]:
```
for row in data[:3]:
    print(row)
```

```
['August 19, 1971', 'Santa Martha Acatitla', 'Mexico', 'Yes', 'Joel David
Kaplan Carlos Antonio Contreras Castro', "Joel David Kaplan was a New York
businessman who had been arrested for murder in 1962 in Mexico City and wa
s incarcerated at the Santa Martha Acatitla prison in the Iztapalapa borou
gh of Mexico City. Joel's sister, Judy Kaplan, arranged the means to help
Kaplan escape, and on August 19, 1971, a helicopter landed in the prison y
ard. The guards mistakenly thought this was an official visit. In two minu
tes, Kaplan and his cellmate Carlos Antonio Contreras, a Venezuelan counte
rfeiter, were able to board the craft and were piloted away, before any sh
ots were fired.[9] Both men were flown to Texas and then different planes
flew Kaplan to California and Contreras to Guatemala.[3] The Mexican gover
nment never initiated extradition proceedings against Kaplan.[9] The escap
e is told in a book, The 10-Second Jailbreak: The Helicopter Escape of Joe
l David Kaplan.[4] It also inspired the 1975 action movie Breakout, which
starred Charles Bronson and Robert Duvall.[9]"]
['October 31, 1973', 'Mountjoy Jail', 'Ireland', 'Yes', "JB O'Hagan Seamus
TwomeyKevin Mallon", 'On October 31, 1973, an IRA member hijacked a helico
pter and forced the pilot to land in the exercise yard of Dublin\'s Mountj
oy Jail\'s D Wing at 3:40\xa0p.m., October 31, 1973. Three members of the
```

We begin by importing some helper functions

In [7]:
```
index=0
```

In [8]:
```
for row in data:
    data[index]=row[:-1]
    index=index+1
print(data[:3])
```

```
[['August 19, 1971', 'Santa Martha Acatitla', 'Mexico', 'Yes', 'Joel David Ka
plan Carlos Antonio Contreras Castro'], ['October 31, 1973', 'Mountjoy Jail',
'Ireland', 'Yes', "JB O'Hagan Seamus TwomeyKevin Mallon"], ['May 24, 1978',
'United States Penitentiary, Marion', 'United States', 'No', 'Garrett Brock T
rapnellMartin Joseph McNallyJames Kenneth Johnson']]
```

In [9]:
```
for row in data:
    row[0]=fetch_year(row[0])
print(data[:5])
```

```
[[1971, 'Santa Martha Acatitla', 'Mexico', 'Yes', 'Joel David Kaplan Carlos A
ntonio Contreras Castro'], [1973, 'Mountjoy Jail', 'Ireland', 'Yes', "JB O'Ha
gan Seamus TwomeyKevin Mallon"], [1978, 'United States Penitentiary, Marion',
'United States', 'No', 'Garrett Brock TrapnellMartin Joseph McNallyJames Kenn
eth Johnson'], [1981, 'Fleury-Mérogis, Essonne, Ile de France', 'France', 'Ye
s', 'Gérard DupréDaniel Beaumont'], [1981, 'Orsainville Prison, Quebec City',
'Canada', 'No', 'Marina Paquet (hijacker)Giles Arseneault (prisoner)']]
```

In [10]:
```
min_year = min(data, key=lambda x: x[0])[0]
max_year = max(data, key=lambda x: x[0])[0]
```

In [11]:
```
print(min_year)
print(max_year)
```

```
1971
2020
```

In [12]:
```
years=[]
for year in range(min_year, max_year+1):
    years.append([year,0])
```

In [13]:
```
print(years)
```

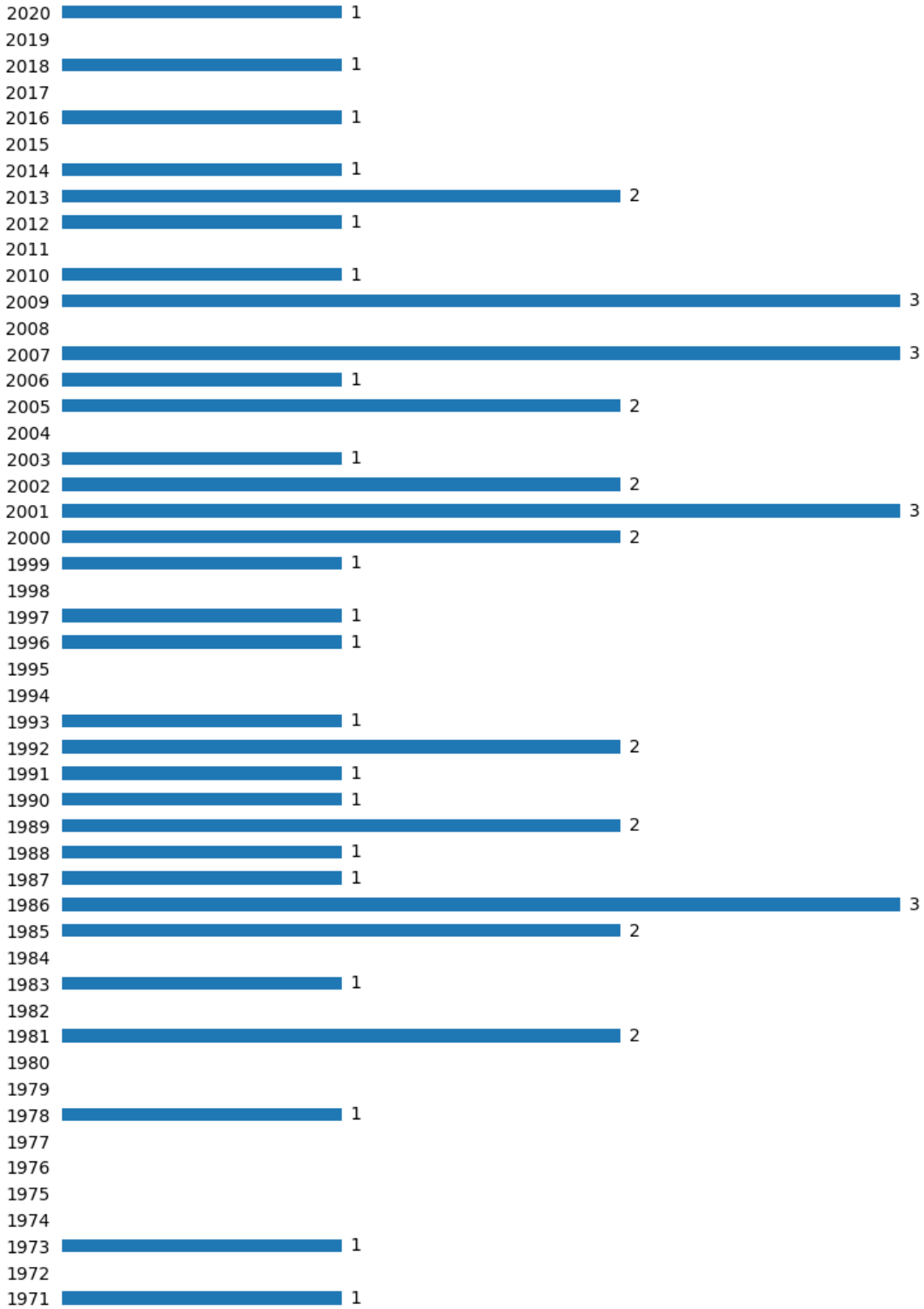
```
[[1971, 0], [1972, 0], [1973, 0], [1974, 0], [1975, 0], [1976, 0], [1977, 0],
[1978, 0], [1979, 0], [1980, 0], [1981, 0], [1982, 0], [1983, 0], [1984, 0],
[1985, 0], [1986, 0], [1987, 0], [1988, 0], [1989, 0], [1990, 0], [1991, 0],
[1992, 0], [1993, 0], [1994, 0], [1995, 0], [1996, 0], [1997, 0], [1998, 0],
[1999, 0], [2000, 0], [2001, 0], [2002, 0], [2003, 0], [2004, 0], [2005, 0],
[2006, 0], [2007, 0], [2008, 0], [2009, 0], [2010, 0], [2011, 0], [2012, 0],
[2013, 0], [2014, 0], [2015, 0], [2016, 0], [2017, 0], [2018, 0], [2019, 0],
[2020, 0]]
```

In [14]:
```
for row in data:
    for year_attempt in years:
        year = year_attempt[0]
        if row[0] == year:
            year_attempt[1] += 1
print(years)
```

```
[[1971, 1], [1972, 0], [1973, 1], [1974, 0], [1975, 0], [1976, 0], [1977, 0],
[1978, 1], [1979, 0], [1980, 0], [1981, 2], [1982, 0], [1983, 1], [1984, 0],
[1985, 2], [1986, 3], [1987, 1], [1988, 1], [1989, 2], [1990, 1], [1991, 1],
[1992, 2], [1993, 1], [1994, 0], [1995, 0], [1996, 1], [1997, 1], [1998, 0],
[1999, 1], [2000, 2], [2001, 3], [2002, 2], [2003, 1], [2004, 0], [2005, 2],
[2006, 1], [2007, 3], [2008, 0], [2009, 3], [2010, 1], [2011, 0], [2012, 1],
[2013, 2], [2014, 1], [2015, 0], [2016, 1], [2017, 0], [2018, 1], [2019, 0],
[2020, 1]]
```

In which year did the most attempts at breaking out of prison with a helicopter occur?

In [20]:
```
%matplotlib inline
barplot(years)
```




The years in which the most helicopter prison break attempts occurred were 1986, 2001, 2007 and 2009, with a total of three attempts each.

ATTEMPTS BY COUNTRY

let's investigate attempts by country

In [19]:
```
countries_frequency = df["Country"].value_counts()
print_pretty_table(countries_frequency)
```

| Country | Number of Occurrences |
|---|---|
| France | 15 |
| United States | 8 |
| Greece | 4 |
| Canada | 4 |
| Belgium | 4 |
| United Kingdom | 2 |
| Australia | 2 |
| Brazil | 2 |
| Ireland | 1 |
| Chile | 1 |
| Puerto Rico | 1 |
| Russia | 1 |
| Mexico | 1 |
| Netherlands | 1 |
| Italy | 1 |

Conclusions

France has highest of attempts of prison break with 15 occurrences.