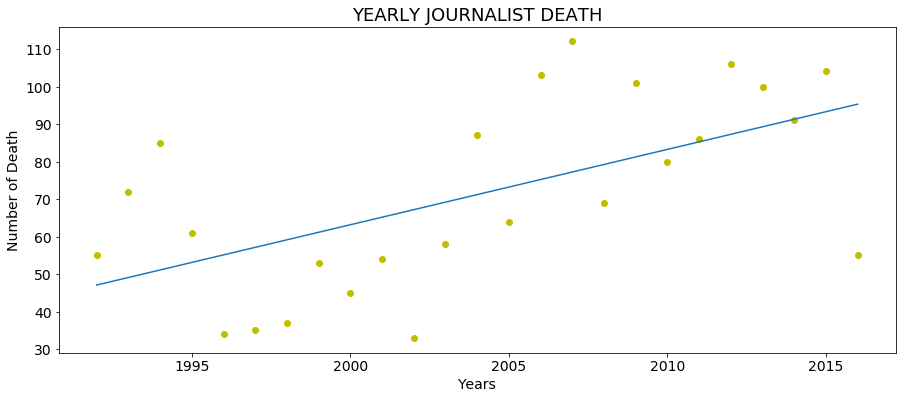
\*Python code at the end

**HAS CONFLICT REPORTING BECOME MORE DANGEROUS?**

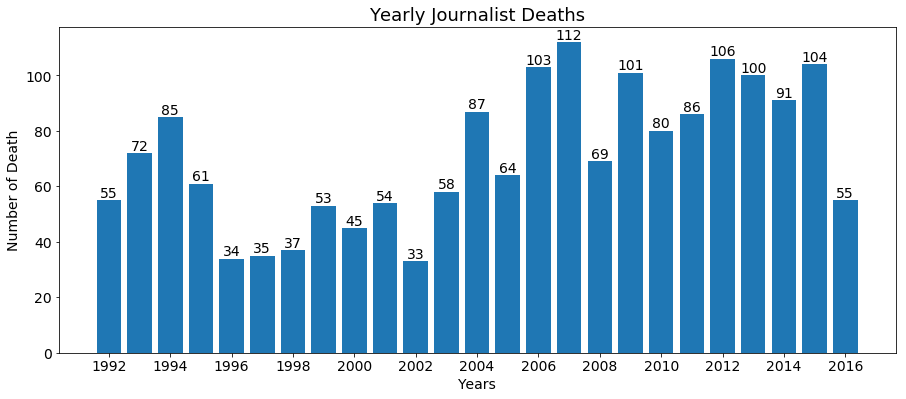
Journalists are afforded some protection against direct attack during a conflict by International Humanitarian Law. As state in Customary Rules of International armed conflicts and the non-international armed conflicts(IAC/NIAC), "Civilian journalists engaged in professional missions in areas of armed conflict must be respected and protected as long as they are not taking a direct part in hostilities."

This data set details account of every journalist killed on duty worldwide was compiled by Committee to Protect Journalists (CPJ), since 1992. CPJ claims to have invested each death. After investigation and verification, classify them as work related if the journalist was killed in line of duty: murder, crossfire, and/or while carrying out a dangerous assignment. Cases involving unclear motives, but with a potential link to journalism, are classified as "unconfirmed" and CPJ continues to investigate.

For now, danger will be defined as number of death per year. Relationship between death of journalist and time is not strong (Rvalue = 0.59), but it is statistically significant (Pvalue = 0). Positive slope indicates an overall increase in death of journalist while on duty over the years but it is not large enough to definitively conclude the null hypothesis.



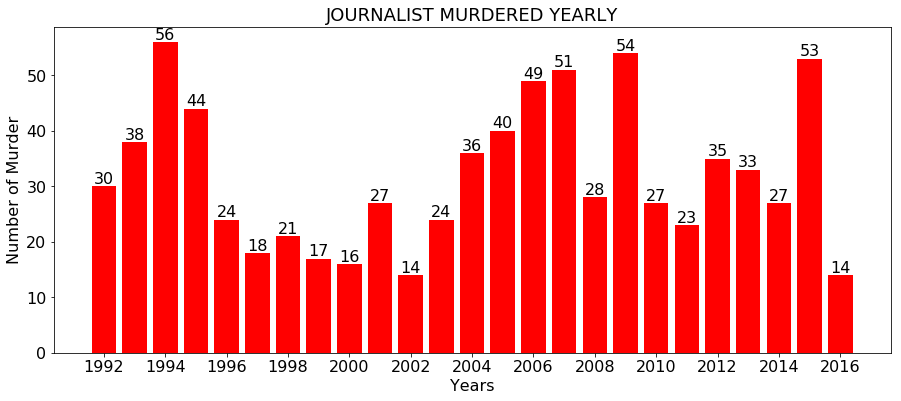
Although journalist have faced danger in every armed conflict, changing nature of warfare might have made them a greater target. Armed conflicts are no longer between conventional state adversaries, it changed not only who fought but also how it is fought. Below is a graph of the same data set, better visualized as bar graph.



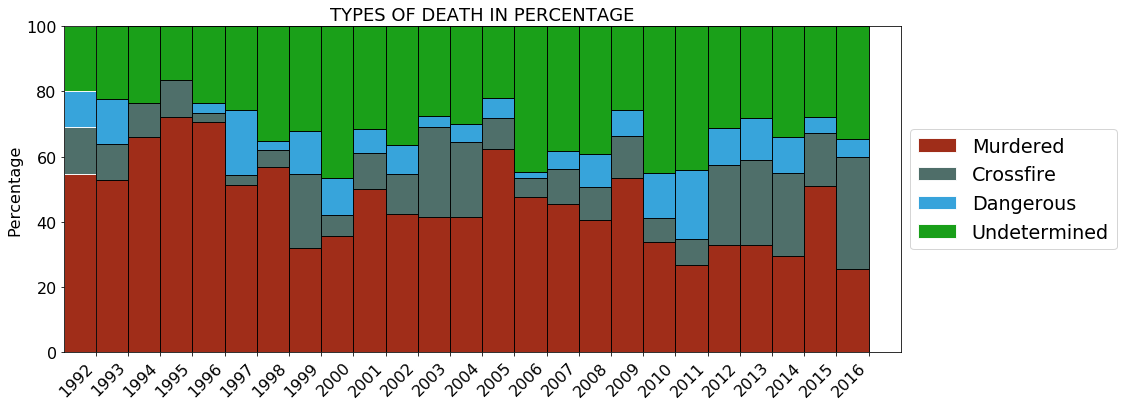
We don't know if journalists are being targeted because small increase in death we observe could be largely due to increase in news outlets, number, size and duration of the conflicts. Here is a total distribution of type of deaths and murder makes up 44 percent of the deaths, the largest contributor.

|  |  |
| --- | --- |
| Type of Death | Percentage |
| Murder | 44.67% |
| Crossfire/Combat-Related | 14.87% |
| Dangerous Assignment | 8.53% |
| Undetermined | 31.93% |

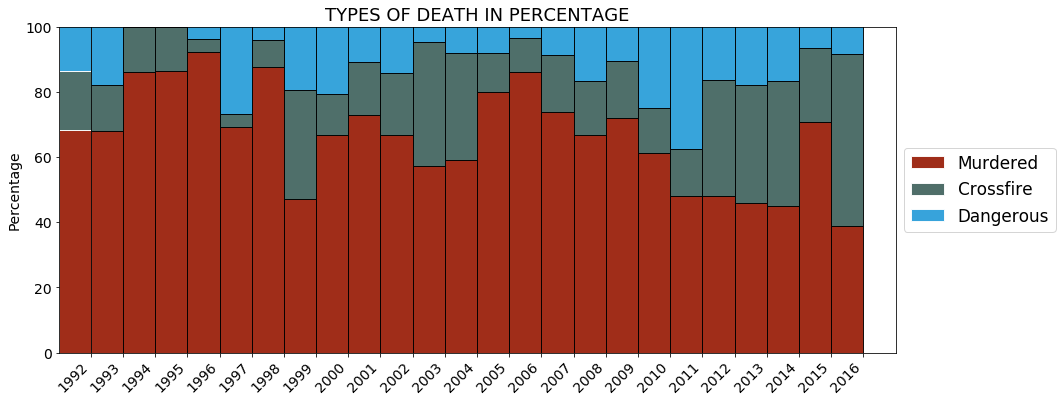
Deaths in categories “Crossfire/Combat-Related” and “Dangerous Assignment” is due to randomness, without purpose or intention, and their values would vary randomly over time. As oppose to specific and intentional targeted killing of journalist categorized by 'Murder' label would give a better indication of increase in danger. I would expect the danger to increase, indicated by increase number of murder because of shift in nature of war to secretive and increase in news outlets, number, size and duration of the conflicts.



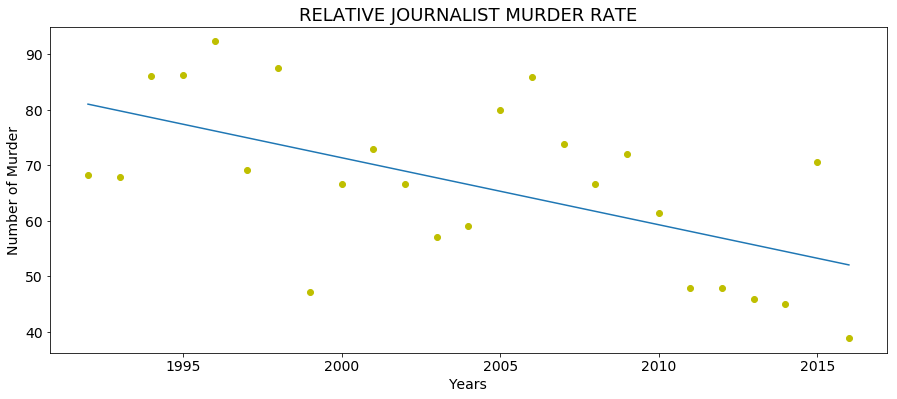
Just looking at the number of murder, we don’t see the predicted trend. This could be due to various varying aspects of the conflicts and it would be better to look the proportional contribution from each category.



The set of data in "Undermined" category is messing up my data, they belong in one of the other three categories, but they are still under investigation and to be determined. To add another prospective for better analysis, a solution would be to distribute "Undetermined" data points into the other categories three proportionately.



Above is bar graph of the three categories with propositional distribution of “undermined”, and below is the linear regression analysis of above data for “Murder” categories.



Relationship between murder of journalist and time is not strong (Rvalue = 0.59), but it is statistically significant (Pvalue = 0). Negative slope (slope = -1.20) indicates an overall decrease in murder of journalist while on duty over the years, but it is not large enough to definitively conclude the null hypothesis.

Recent article by Washington post claims, “[2017 was the most dangerous year ever for journalists. 2018 might be even worse](https://www.washingtonpost.com/news/worldviews/wp/2018/02/01/2017-was-the-most-dangerous-year-ever-for-journalists-2018-might-be-even-worse/?noredirect=on&utm_term=.33537da7fdba),” based on data from Reporters Without Border. The report listed 18 killed world wide but their claims ‘most dangerous year’ was based on number of reporters imprisoned, which all time high of 189 worldwide. Since arrests also fall under the specific and intentional targeting of journalist similar to murder category, so combination of this could be used to examined and gauge the danger for journalist.

```

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

%matplotlib inline

from scipy.stats import linregress

import sklearn

from sklearn.linear\_model import LinearRegression

# Data frame of journalist death since 1992 to 2016

df\_cpj = pd.read\_csv('cpj.csv')

DateRange = np.arange(1992,2017)

count = []

EmptyDF = []

DateRange

for year in DateRange:

somth = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year))]

count = [year, somth['Name'].count()]

EmptyDF.append(count)

df\_2 = pd.DataFrame()

df\_2 = df\_2.append(EmptyDF)

df\_2.columns = ["Year","Count"]

plt.figure(figsize=(15,6))

fit = np.polyfit(df\_2["Year"], df\_2["Count"],1)

fit\_fn = np.poly1d(fit)

plt.title('YEARLY JOURNALIST DEATH')

plt.xlabel("Years")

plt.ylabel('Number of Death')

plt.rc('axes', titlesize=18)

plt.rc('font', size=14)

plt.plot(df\_2["Year"], df\_2["Count"], 'yo', df\_2["Year"], fit\_fn(df\_2["Year"]))

linregress(df\_2["Year"], df\_2["Count"])

| Variable | Value |

|:-------------|:--------:|

|slope |2.00 |

|intercept | -3952.21 |

|rvalue |0.59 |

|pvalue |0.00 |

|stderr |0.58 |

#plt.figure(figsize=(15,6))

#plt.scatter(df\_2["Year"], df\_2["Count"])

#plt.xticks(range(df\_2["Year"].min(), df\_2["Year"].max() +2, 2))

#plt.ylabel("Number of Death")

#plt.xlabel("Years")

#plt.show()

fig, ax = plt.subplots(figsize=(15,6))

rects = ax.bar(df\_2["Year"], df\_2["Count"])

plt.xticks(range(df\_2["Year"].min(), df\_2["Year"].max() + 1, 2))

ax.set\_title('Yearly Journalist Deaths')

ax.set\_xlabel("Years")

ax.set\_ylabel('Number of Death')

def autolabel(rects):

# attach some text labels

for rect in rects:

height = rect.get\_height()

ax.text(rect.get\_x() + rect.get\_width()/2., 1.0\*height,

'%d' % int(height),

ha='center', va='bottom')

plt.rc('axes', titlesize=18)

plt.rc('font', size=14)

autolabel(rects)

plt.show()

TypeMurder = df\_cpj[df\_cpj['Type\_death'] == 'Murder']

TypeCross = df\_cpj[df\_cpj['Type\_death'] == 'Crossfire/Combat-Related']

TypeDanger = df\_cpj[df\_cpj['Type\_death'] == 'Dangerous Assignment']

TypeUnknown = df\_cpj[df\_cpj['Type\_death'] == 'Unknown']

TypeUndetermined = df\_cpj['Name'].count() + TypeUnknown['Name'].count() - (TypeMurder['Name'].count() +

TypeCross['Name'].count() +

TypeDanger['Name'].count() +

TypeUnknown['Name'].count())

m = 100\*TypeMurder['Name'].count()/df\_cpj['Name'].count()

c = 100\*TypeCross['Name'].count()/df\_cpj['Name'].count()

d = 100\*TypeDanger['Name'].count()/df\_cpj['Name'].count()

#uk = TypeUnknown['Name'].count() #/df\_cpj['Name'].count()

u = 100\*TypeUndetermined/df\_cpj['Name'].count()

d = {'DeathType': ['Murder', 'Crossfire','Dangerous Assignment', 'Undetermined'], 'Percentage': [m,c,d,u]}

TypeofDeath = pd.DataFrame(data=d)

TypeofDeath.round({'Percentage': 2})

#uk

DateRange3 = np.arange(1992,2017)

EmptyDF3 = []

for year3 in DateRange3:

somth3 = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3)) & df\_cpj['Type\_death'].str.contains('Murder')]

count3 = [year3, somth3['Name'].count()]

EmptyDF3.append(count3)

df\_3 = pd.DataFrame()

df\_3 = df\_3.append(EmptyDF3)

df\_3.columns = ["Year","Count"]

fig, ax = plt.subplots(figsize=(15,6))

rects = ax.bar(df\_3["Year"], df\_3["Count"], color='r')

plt.xticks(range(df\_3["Year"].min(), df\_3["Year"].max() + 2, 2))

ax.set\_title('JOURNALIST MURDERED YEARLY')

ax.set\_xlabel("Years")

ax.set\_ylabel('Number of Murder')

def autolabel(rects):

# attach some text labels

for rect in rects:

height = rect.get\_height()

ax.text(rect.get\_x() + rect.get\_width()/2., 1.0\*height,

'%d' % int(height),

ha='center', va='bottom')

plt.rc('axes', titlesize=18)

plt.rc('font', size=16)

autolabel(rects)

plt.show()

#df\_2.loc[i,"Count"]

DateRange4 = np.arange(1992,2017)

EmptyDF3 = []

i = 0

for year3 in DateRange4:

TypeMurder = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3)) & df\_cpj['Type\_death'].str.contains('Murder')]

TypeCross = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3)) & df\_cpj['Type\_death'].str.contains('Crossfire/Combat-Related')]

TypeDanger = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3)) & df\_cpj['Type\_death'].str.contains('Dangerous Assignment')]

# totesinyear = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3))]

count3 = [year3, 100\*TypeMurder['Name'].count()/df\_2.loc[i,"Count"],

100\*TypeCross['Name'].count()/df\_2.loc[i,"Count"],

100\*TypeDanger['Name'].count()/df\_2.loc[i,"Count"],

100\*(df\_2.loc[i,"Count"] - TypeMurder['Name'].count() - TypeCross['Name'].count() - TypeDanger['Name'].count())

/df\_2.loc[i,"Count"]]

i += 1

EmptyDF3.append(count3)

dfcpjPer = pd.DataFrame()

dfcpjPer = dfcpjPer.append(EmptyDF3)

dfcpjPer.columns = ["Year","Murdered","Crossfire","Dangerous","Other"]

# Create a figure with a single subplot

f, ax = plt.subplots(1, figsize=(15,6))

# Set bar width at 1

bar\_width = 1

# positions of the left bar-boundaries

bar\_l = [i for i in range(len(dfcpjPer['Murdered']))]

# positions of the x-axis ticks (center of the bars as bar labels)

tick\_pos = [i+(bar\_width/2) for i in bar\_l]

# Create a bar chart in position bar\_1

ax.bar(bar\_l,

# using pre\_rel data

dfcpjPer['Murdered'],

# labeled

label='Murdered',

# with alpha

alpha=0.9,

# with color

color='#961600',

# with bar width

width=bar\_width,

# with border color

edgecolor='white'

)

# Create a bar chart in position bar\_1

ax.bar(bar\_l,

# using mid\_rel data

dfcpjPer['Crossfire'],

# with pre\_rel

bottom=dfcpjPer['Murdered'],

# labeled

label='Crossfire',

# with alpha

alpha=0.9,

# with color

color='#3C5F5A',

# with bar width

width=bar\_width,

# with border color

edgecolor='white'

)

# Create a bar chart in position bar\_1

ax.bar(bar\_l,

# using post\_rel data

dfcpjPer['Dangerous'],

# with pre\_rel and mid\_rel on bottom

bottom=[i+j for i,j in zip(dfcpjPer['Murdered'], dfcpjPer['Crossfire'])],

# labeled

label='Dangerous',

# with alpha

alpha=0.9,

# with color

color='#219AD8',

# with bar width

width=bar\_width,

# with border color

edgecolor='white'

)

# Create a bar chart in position bar\_1

ax.bar(bar\_l,

# using post\_rel data

dfcpjPer['Other'],

# with pre\_rel and mid\_rel on bottom

bottom=[i+j+k for i,j,k in zip(dfcpjPer['Murdered'], dfcpjPer['Crossfire'],dfcpjPer['Dangerous'])],

# labeled

label='Undetermined',

# with alpha

alpha=0.9,

# with color

color='#019600',

# with bar width

width=bar\_width,

# with border color

edgecolor='white'

)

# Set the ticks to be first names

plt.xticks(tick\_pos, dfcpjPer['Year'])

ax.set\_title('TYPES OF DEATH IN PERCENTAGE')

ax.set\_ylabel("Percentage")

ax.set\_xlabel("")

# Let the borders of the graphic

plt.xlim([min(tick\_pos)-bar\_width, max(tick\_pos)+bar\_width])

plt.ylim(-0, 100)

# rotate axis labels

plt.setp(plt.gca().get\_xticklabels(), rotation=45, horizontalalignment='right')

plt.legend(loc='center left', bbox\_to\_anchor=(1, 0.5), fontsize = 'large')

plt.rc('axes', titlesize=18)

plt.rc('font', size=14)

# shot plot

plt.show()

DateRange4 = np.arange(1992,2017)

EmptyDF3 = []

i = 0

for year3 in DateRange4:

TypeMurder = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3)) & df\_cpj['Type\_death'].str.contains('Murder')]

TypeCross = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3)) & df\_cpj['Type\_death'].str.contains('Crossfire/Combat-Related')]

TypeDanger = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3)) & df\_cpj['Type\_death'].str.contains('Dangerous Assignment')]

totesinyear = df\_cpj[df\_cpj['Date'].str.contains('{}'.format(year3))]

totes = TypeMurder['Name'].count() + TypeCross['Name'].count() + TypeDanger['Name'].count()

Typeundetermined = totesinyear['Name'].count() - totes

Murder = Typeundetermined \* (TypeMurder['Name'].count()/totes)

Cross = Typeundetermined \* (TypeCross['Name'].count()/totes)

Danger = Typeundetermined \* (totesinyear['Name'].count()/totes)

count3 = [year3,

100\*(TypeMurder['Name'].count() + Murder)/df\_2.loc[i,"Count"],

100\*(TypeCross['Name'].count() + Cross)/df\_2.loc[i,"Count"],

100\*(TypeDanger['Name'].count() + Danger)/df\_2.loc[i,"Count"], ]

i += 1

EmptyDF3.append(count3)

dfcpjPer = pd.DataFrame()

dfcpjPer = dfcpjPer.append(EmptyDF3)

dfcpjPer.columns = ["Year","Murdered","Crossfire","Dangerous"]

# positions of the left bar-boundaries

bar\_l = [i for i in range(len(dfcpjPer['Murdered']))]

# positions of the x-axis ticks (center of the bars as bar labels)

tick\_pos = [i+(bar\_width/2) for i in bar\_l]

# Create a bar chart in position bar\_1

ax.bar(bar\_l,

# using pre\_rel data

dfcpjPer['Murdered'],

# labeled

label='Murdered',

# with alpha

alpha=0.9,

# with color

color='#961600',

# with bar width

width=bar\_width,

# with border color

edgecolor='white'

)

# Create a bar chart in position bar\_1

ax.bar(bar\_l,

# using mid\_rel data

dfcpjPer['Crossfire'],

# with pre\_rel

bottom=dfcpjPer['Murdered'],

# labeled

label='Crossfire',

# with alpha

alpha=0.9,

# with color

color='#3C5F5A',

# with bar width

width=bar\_width,

# with border color

edgecolor='white'

)

# Create a bar chart in position bar\_1

ax.bar(bar\_l,

# using post\_rel data

dfcpjPer['Dangerous'],

# with pre\_rel and mid\_rel on bottom

bottom=[i+j for i,j in zip(dfcpjPer['Murdered'], dfcpjPer['Crossfire'])],

# labeled

label='Dangerous',

# with alpha

alpha=0.9,

# with color

color='#219AD8',

# with bar width

width=bar\_width,

# with border color

edgecolor='white'

)

# Set the ticks to be first names

plt.xticks(tick\_pos, dfcpjPer['Year'])

ax.set\_title('TYPES OF DEATH IN PERCENTAGE')

ax.set\_ylabel("Percentage")

ax.set\_xlabel("")

# Let the borders of the graphic

plt.xlim([min(tick\_pos)-bar\_width, max(tick\_pos)+bar\_width])

plt.ylim(-0, 100)

# rotate axis labels

plt.setp(plt.gca().get\_xticklabels(), rotation=45, horizontalalignment='right')

plt.legend(loc='center left', bbox\_to\_anchor=(1, 0.5), fontsize = 'large')

plt.rc('axes', titlesize=18)

plt.rc('font', size=14)

# shot plot

plt.show()

plt.figure(figsize=(15,6))

fit = np.polyfit(dfcpjPer["Year"], dfcpjPer['Murdered'],1)

fit\_fn = np.poly1d(fit)

plt.title('RELATIVE JOURNALIST MURDER RATE')

plt.xlabel("Years")

plt.ylabel('Number of Murder')

plt.rc('axes', titlesize=18)

plt.rc('font', size=14)

plt.plot(dfcpjPer["Year"], dfcpjPer['Murdered'], 'yo', dfcpjPer["Year"], fit\_fn(dfcpjPer["Year"]))

linregress(dfcpjPer["Year"], dfcpjPer['Murdered'])

```