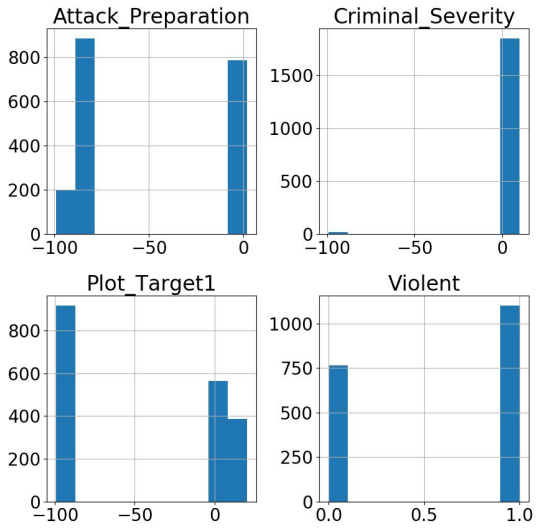
# Capstone 1
# Julia Chapman

The National Consortium for the Study of Terrorism and Responses to Terrorism put together “Profiles of Individual Radicalization in the United States (PIRUS)”. This dataset covers information on over 1,800 individuals who participated in acts under far right, far left, Islamist, or single issue ideologies in the USA from 1948-2016.

Questions:

- Can I predict whether an individual was part of a Violent vs Non-Violent crime?
- Is the violence changing over time?
- What are the top 3 contributing factors?

Description of data:




In [17]: df.shape
Out[17]: (1865, 145)

In [18]: df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1865 entries, 0 to 1864
Columns: 145 entries, Subject_ID to Standing
dtypes: float64(23), int64(103), object(19)
memory usage: 2.1+ MB

Minimum Viable Product (MVP):

- Train a logistics regression models to predict whether an act was violent or not.

MVP+, MVP++, MVP+++

- Investigate the highest contributing factors.
- Find violence correlation to time
- Compare to census or other crime data for top 3 characteristics and determine statistical difference in populations