# Defining a function for processing report strings

def process\_report\_strings(incident\_string, attribute\_names):

# Splitting the incident string by semicolons and alsoremove extra quotes and whitespaces by using strip.

values = [value.strip().strip('"') for value in incident\_string.strip().split(";")]

# Handling missing attribute i.e. considering image as missed attribute

if len(values) < len(attribute\_names): # Checking length to identify if values length are less than the attribute names then i need to append it as "None"

values.append('none') # Adding 'none' if image is missing

# Creating a dictionary from the attribute names and values

return dict(zip(attribute\_names, values))

# Processing each incident and storing results in a list by calling the defined function multiple times for all the given incidents to ensure that all are being added again to a list

report\_list\_incindents = [process\_report\_strings(incident, attribute\_names) for incident in incidents]

# Checking the result by lopping to find all the outputs

for report\_incidents in report\_list\_incindents:

print(report\_incidents)

Report\_list\_incindents

import re

for single\_incident in report\_list\_incindents:

for key in single\_incident:

if isinstance(single\_incident[key], str):

single\_incident[key] = single\_incident[key].replace('\n', '')

match = re.search(r'https?://([^/]+)', single\_incident['url'])

single\_incident['domain'] = match.group(1) if match else 'none'

date\_published = single\_incident['date published']

date\_submitted = single\_incident['date submitted']

year\_published, month\_published, day\_published = map(int, date\_published.split('-'))

year\_submitted, month\_submitted, day\_submitted = map(int, date\_submitted.split('-'))

days\_difference = (year\_submitted - year\_published) \* 360 + (month\_submitted - month\_published) \* 30 + (day\_submitted - day\_published)

single\_incident['time difference'] = days\_difference

for incident in report\_list\_incindents:

print(incident)

Report\_list\_incindents