0_data_prep_final

December 11, 2025

1 Data Prep

Excel bestanden omzetten naar bruikbaar .csv formaat via onderstaande code:

```python
[6]: import pandas as pd
     import os

     # 1. Definieer je bestandsnamen
     excel_file = '/home/jovyan/work/drone-project/cnas-dataset-raw.xlsx' # Pas naam
      ↪aan indien nodig
     csv_output = 'cnas_drone_proliferation.csv'

     # 2. Check of het bestand bestaat
     if os.path.exists(excel_file):
         # Lees Excel in met Pandas
         # 'sheet_name=0' pakt het eerste tabblad. Pas aan als je data op een ander
      ↪tabblad staat.
         df_pandas = pd.read_excel(excel_file, sheet_name=0)

         # 3. Bekijk de eerste regels (even spieken of het klopt)
         print("Voorbeeld van de data:")
         print(df_pandas.head())

         # 4. Opslaan als CSV
         # index=False zorgt dat je geen rij-nummers (0, 1, 2...) in je CSV krijgt
         df_pandas.to_csv(csv_output, index=False)
         print(f"\nSucces! Bestand opgeslagen als: {csv_output}")
     else:
         print(f"Kan bestand niet vinden: {excel_file}")
```

```
Voorbeeld van de data:
   Entry Number         Region        Seeker Supplier  Paltform Model  \
0             1  North America  United States   Israel          Hunter
1             2   Indo-Pacific      Singapore   Israel  Searcher Mk II
2             3         Europe         France   Israel          Hunter
3             4         Europe    Netherlands   France         Sperwer
4             5   Indo-Pacific      Sri Lanka   Israel           Scout
```

```
     Platform   Type   Year of Interest   Year of Order   Year of First Delivery  \
0              UAV                   NaN           1992.0                    1993.0
1              UAV                   NaN           1993.0                    1994.0
2              UAV                   NaN           1995.0                    1997.0
3              UAV                   NaN           1995.0                    2000.0
4              UAV                   NaN           1995.0                    1998.0

     Year of  Identification, Completion, or Cancellation Status Source 1  \
0                                                   1996.0        Yes    SIPRI
1                                                   1995.0        Yes    SIPRI
2                                                   1997.0        Yes    SIPRI
3                                                   2001.0        Yes    SIPRI
4                                                   1998.0        Yes    SIPRI

   Source 2 Source 3                                                  Comments
0       NaN      NaN  SIPRI Comments: "Deal incl. 7 ground-control s…
1       NaN      NaN                                                      NaN
2       NaN      NaN                      SIPRI Comments: "$50m deal"
3       NaN      NaN        SIPRI Comments: "$82m deal; retired in 2011"
4       NaN      NaN  SIPRI Comments: "Secondhand but modernized bef…

Succes! Bestand opgeslagen als: cnas_drone_proliferation.csv
```

```python
# 1. Definieer je bestandsnamen
excel_file = '/home/jovyan/work/drone-project/DroneWarsData.xlsx' # Pas naam
  aan indien nodig
csv_output = '/home/jovyan/work/drone-project/dronewars.csv'

# 2. Check of het bestand bestaat
if os.path.exists(excel_file):
    # Lees Excel in met Pandas
    # 'sheet_name=0' pakt het eerste tabblad. Pas aan als je data op een ander
  tabblad staat.
    df_pandas = pd.read_excel(excel_file, sheet_name=0)

    # 3. Bekijk de eerste regels (even spieken of het klopt)
    print("Voorbeeld van de data:")
    print(df_pandas.head())

    # 4. Opslaan als CSV
    # index=False zorgt dat je geen rij-nummers (0, 1, 2...) in je CSV krijgt
    df_pandas.to_csv(csv_output, index=False)
    print(f"\nSucces! Bestand opgeslagen als: {csv_output}")
else:
    print(f"Kan bestand niet vinden: {excel_file}")
```

```
Voorbeeld van de data:
   Strike ID        Country Date (MM-DD-YYYY) President Village/local area  \
```

```
0     AFG001   Afghanistan          01-01-2015    Obama              Unknown
1     AFG002   Afghanistan          01-03-2015    Obama              Unknown
2     AFG003   Afghanistan          01-03-2015    Obama              Unknown
3     AFG004   Afghanistan          01-06-2015    Obama              Unknown
4     AFG005   Afghanistan          01-07-2015    Obama              Unknown


        District Province Village/Local Area, District, Province, Country  \
0        Spera    Khost                 Unknown, Spera, Khost, Afghanistan
1        Gayan   Paktika             Unknown, Gayan, Paktika, Afghanistan
2        Spera    Khost                 Unknown, Spera, Khost, Afghanistan
3   Dara I Pech   Kunar      Unknown, Dara I Pech, Kunar, Afghanistan
4      Unknown    Logar          Unknown, Unknown, Logar, Afghanistan


   Lat/Long of Location  Modified Location \nwhen Applicable  …  \
0  32.8078448,68.6456967                                NaN  …
1    32.984041,69.430782                                NaN  …
2  32.8078448,68.6456967                                NaN  …
3    34.872898,71.155620                                NaN  …
4    34.014551,69.192391                                NaN  …


   Maximum total people killed Minimum civilians reported killed  \
0                           0                                  0
1                          18                                  0
2                           7                                  0
3                           0                                  0
4                           3                                  0


   Maximum civilians reported killed Minimum children reported killed  \
0                                  0                                 0
1                                  0                                 0
2                                  0                                 0
3                                  0                                 0
4                                  0                                 0


   Maximum children reported killed Minimum reported injured  \
0                                 0                        0
1                                 0                        0
2                                 0                        0
3                                 0                        0
4                                 0                        0


   Maximum reported injured  \
0                        0
1                        0
2                        0
3                        0
4                        0
```

```
                                   Strike link  \
0  http://www.thebureauinvestigates.com/2015/02/1…
1  http://www.thebureauinvestigates.com/2015/02/1…
2  http://www.thebureauinvestigates.com/2015/02/1…
3  http://www.thebureauinvestigates.com/2015/02/1…
4  http://www.thebureauinvestigates.com/2015/02/1…


                                 Timeline URL  Index
0  http://www.thebureauinvestigates.com/2015/02/1…     77
1                                          NaN      3
2                                          NaN      4
3                                          NaN    120
4                                          NaN      5

[5 rows x 36 columns]


Succes! Bestand opgeslagen als: /home/jovyan/work/drone-project/dronewars.csv
```