Task 2:

Columns: The dataset typically contains 95 or 169 columns depending on the specific version. You can see this in the "Dataset Shape" output.

Geometries: Polygon and MultiPolygon.

Geometry Column: It represents the spatial boundaries (shape) of the geographic features.

```
...  Dataset Shape: (177, 169)
          featurecla  scalerank  LABELRANK                     SOVEREIGNT SOV_A3  \
     0  Admin-0 country          1          6                        Fiji    FJI
     1  Admin-0 country          1          3  United Republic of Tanzania    TZA
     2  Admin-0 country          1          7               Western Sahara    SAH
     3  Admin-0 country          1          2                       Canada    CAN
     4  Admin-0 country          1          2     United States of America    US1

        ADM0_DIF  LEVEL                TYPE TLC                        ADMIN  ...  \
     0         0      2  Sovereign country    1                        Fiji  ...
     1         0      2  Sovereign country    1  United Republic of Tanzania  ...
     2         0      2      Indeterminate    1               Western Sahara  ...
     3         0      2  Sovereign country    1                       Canada  ...
     4         1      2            Country    1     United States of America  ...

           FCLASS_TR     FCLASS_ID     FCLASS_PL FCLASS_GR  FCLASS_IT  \
     0          None          None          None      None       None
     1          None          None          None      None       None
     2  Unrecognized  Unrecognized  Unrecognized      None       None
     3          None          None          None      None       None
     4          None          None          None      None       None

           FCLASS_NL FCLASS_SE  FCLASS_BD FCLASS_UA  \
     0          None      None       None      None
     1          None      None       None      None
     2  Unrecognized      None       None      None
     3          None      None       None      None
     4          None      None       None      None

                                                 geometry
     0  MULTIPOLYGON (((180 -16.06713, 180 -16.55522, ...
     1  POLYGON ((33.90371 -0.95, 34.07262 -1.05982, 3...
     2  POLYGON ((-8.66559 27.65643, -8.66512 27.58948...
     3  MULTIPOLYGON (((-122.84 49, -122.97421 49.0025...
     4  MULTIPOLYGON (((-122.84 49, -120 49, -117.0312...

     [5 rows x 169 columns]
```
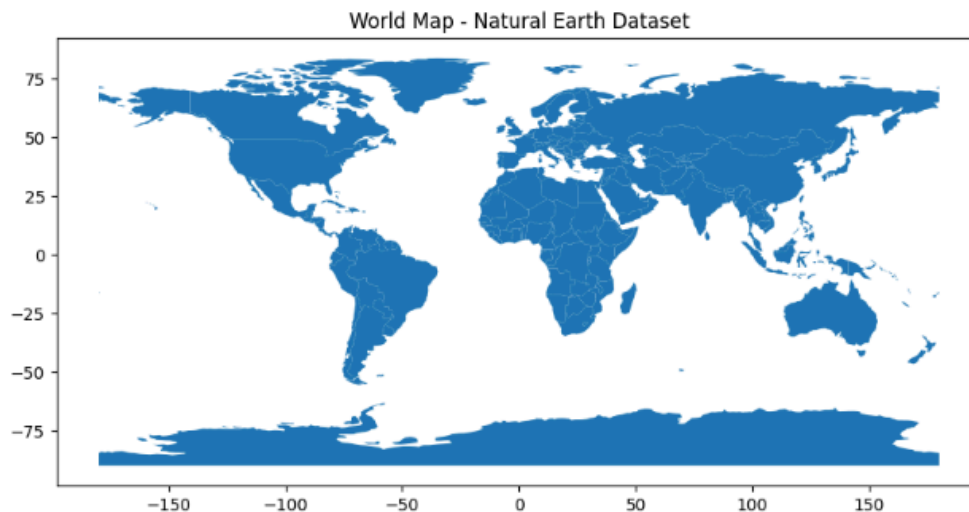
World Map - Natural Earth Dataset

Task 3:

EPSG:4326: It represents WGS84, the standard geographic coordinate system using latitude and longitude.

Importance of CRS: It ensures that spatial data is accurately projected on a 2D surface. Without it, distances and areas will be distorted.

```
Current CRS: EPSG:4326
```

Task 4:

EPSG:4326 Accuracy: It uses degrees. Because the Earth is curved, a degree of longitude varies in physical length depending on latitude, making area calculations mathematically incorrect.

EPSG:3857 Units: It is based on meters.

```
...                         ADMIN      area_m2
0                            Fiji  2.128334e+10
1  United Republic of Tanzania  9.522552e+11
2               Western Sahara  1.171023e+11
3                       Canada  5.216648e+13
4     United States of America  2.186228e+13

Area Reprojection ✅ Complete
```

Task 5:Centroid: The geometric center of a polygon.Clustering: Centroids provide a single

(x,y)

 coordinate per country, which can be used as input features for algorithms like K-Means.Separate X and Y: Machine learning models require distinct numerical input features for each dimension.

```
...                         ADMIN             x             y
0                            Fiji  1.824878e+07 -1.958098e+06
1  United Republic of Tanzania  3.869296e+06 -7.003071e+05
2               Western Sahara -1.348403e+06  2.794163e+06
3                       Canada -1.079779e+07  1.044422e+07
4     United States of America -1.329713e+07  6.667416e+06
```

Top 10 Largest Countries by Area