

## ESERCIZIO nr1

I tre dataset strutturati scelti riguardano rispettivamente:

1. Produzione globale di plastica;
2. Produzione di plastica per settore;
3. Inquinamento di plastica per persona.

Il dataset 1. Ha i seguenti descrittori:

Entity	Code	Year	Global plastics production
--------	------	------	----------------------------

I descrittori indicano l'entità in esame nei confronti della quale viene quantificata la quantità di plastica prodotta; il codice di riferimento; l'anno in cui il quantitativo di plastica è stato prodotto e misurato e, infine, il quantitativo effettivo (desumo in mln di tonnellate dal dataset 2.) di plastica prodotta in misura globale.

Il dataset permette di comprendere come la produzione di plastica sia aumentata negli anni (dal 1950 al 2019 per i dati in esame).

Gli attributi del dataset indicano il nome dell'entità di riferimento, in questo caso il globo; i vari codici di riferimento (in questo caso è lo stesso per ogni rilevazione); gli anni di riferimento dal 1950 al 2019 e il quantitativo di produzione di plastica misurato per ogni singolo anno in esame.

Il dataset 2. Presenta i seguenti descrittori:

Entity	Code	Year	Primary plastic production (million tonnes)
--------	------	------	---

I descrittori indicano l'entità, in questo caso rappresentante il settore industriale per il quale vengono rilevati i dati; i codici di riferimento; l'anno di riferimento per la produzione di plastica per settore industriale e il quantitativo effettivo di plastica prodotta per settore.

Gli attributi contengono le seguenti specifiche: in Entity troviamo i diversi settori industriali (packaging, Textiles, etc) per i quali è stata rilevata la quantità di plastica prodotta; il codice di riferimento valorizzato solo per un Entity; gli anni di riferimento, in questo caso sempre lo stesso "2015"; il quantitativo per ogni settore di plastica prodotta misurato in milioni di tonnellate.

Il dataset 3. Ha i seguenti descrittori:

Entity	Code	Year	Per capita plastic waste (kg/person/day)
--------	------	------	--

I descrittori indicano l'entità, in questo caso rappresentante la nazione per la quale vengono rilevati i dati; i codici di riferimento; l'anno di riferimento per la rilevazione e il quantitativo effettivo di plastica prodotta in kg al giorno per ogni persona.

Gli attributi indicano le nazioni di riferimento con i loro codici identificativi; l'anno di rilevazione (2010) e il quantitativo in kg di inquinamento giornaliero di plastica per persona.

I due dataset non strutturati sono i seguenti:

1. post social in cui compaiono le parole “plastic production” e “plastic waste” diviso per nazioni;
2. articoli di giornali o di notizie di mass media.

L’obiettivo sarebbe quello di identificare in quanti post/articoli di giornale, divisi per singola nazione, è contenuta la tematica “plastic waste” per identificare la sensibilità e l’attenzione che ogni nazione mostra nei confronti del tema in oggetto in proporzione rispetto alle altre nazioni.

Con il primo dataset si analizza quanto il tema sia trattato dall’utente medio; con il secondo si vuole misurare in quali nazioni la tematica sia maggiormente trattata da quotidiani istituzionali che hanno una maggiore incidenza sull’opinione pubblica.