

L'esercizio è mirato a prendere confidenza con la produzione di dati nei formati XML, CSV e JSON.

Si esegua la progettazione di un sistema per l'analisi di dati. La progettazione deve prevedere:

1. L'individuazione di un argomento di proprio interesse multi-dominio
2. La produzione di almeno tre dataset rispettivamente nei formati XML, CSV, e JSON.
3. Richiesta 1: Il dataset XML dovrà contenere almeno 10 figli del nodo padre e almeno 2 nidificazioni per ogni figlio, dovranno inoltre essere presenti tag con e senza attributi.
4. Richiesta 2: il dataset JSON dovrà contenere almeno 10 oggetti JSON contenenti diversi tipi di dato (tra cui array e altri oggetti JSON)
5. Richiesta 3: il dataset CSV dovrà essere composto da un header e almeno 10 righe. Inoltre per campi a propria scelta sarà necessario inserire almeno una virgola, un carattere di nuova linea, e i doppi apici.

Procedere con la redazione di un documento che dovrà includere:

- La descrizione del caso di studio scelto
- Analisi e raccolta dei requisiti (generici e poi più specifici)
- Strutturazione dei requisiti in gruppi di frasi omogenee
- Glossario dei termini
- Descrivere nel dettaglio quale sarebbe l'outcome atteso dal sistema (analisi di business / analisi predittive), come potrebbe essere di supporto nelle decisioni ed eventualmente quali 'storie' è possibile ricavare

Produzione di rifiuti urbani

Caso di studio scelto:

Il fine di questa progettazione è quello di analizzare le informazioni sulla produzione di rifiuti urbani, focalizzandoci sulla produzione regione per regione. Inoltre, si valuterà la medesima situazione dal punto di vista del riciclo e si daranno delle considerazioni su come agire sulle regioni che tendono a riciclare di meno.

Nota metodologica

Fonti e perimetro dei dati utilizzati

I dataset utilizzati sono forniti dall'Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA). Essi presentano la quantità totale di rifiuti urbani generati in Italia, disaggregati a livello regionale. Sempre lo stesso ente ci fornisce i dati percentuali sulla raccolta differenziata dei rifiuti urbani su scala regionale, divisi per frazioni merceologiche.

I dataset forniti sono nei formati XML, JSON e CSV. I dati sono classificabili come dati a 4 stelle secondo la classificazione di Berners Lee (disponibilità di file OWL), i metadati sono classificabili come 3 punti (sono associati internamente al dataset). La licenza è di tipo Italian Open Data License 2.0 (IODL 2.0), che consente di riprodurre, distribuire, utilizzare le informazioni a patto che venga riportata la paternità dei dati.

Richiesta 1: DATASET XML

```
<?xml version="1.0" encoding="UTF-8"?>
<rifiuti_urbani_italia>
  <regione>
    <nome_regione>Piemonte</nome_regione>
    <codice_regione>13</codice_regione>
    <abitanti>4.356mil</abitanti>
    <kg_rifiuti>2270ton</kg_rifiuti>

    <rapporto_kg_rifiuti_per_abitante>516</rapporto_kg_rifiuti_per_abitante>
    <!--cifre non reali del rapporto Kg rifiuti/abitanti regione-->
  </regione>
  <regione>
    <nome_regione>Lombardia</nome_regione>
    <codice_regione>10</codice_regione>
    <abitanti>7.456mil</abitanti>
    <kg_rifiuti>4562ton</kg_rifiuti>

    <rapporto_kg_rifiuti_per_abitante>612</rapporto_kg_rifiuti_per_abitante>
  </regione>
  <regione>
    <nome_regione>Veneto</nome_regione>
    <codice_regione>21</codice_regione>
    <abitanti>5.748mil</abitanti>
    <kg_rifiuti>4874ton</kg_rifiuti>

    <rapporto_kg_rifiuti_per_abitante>547</rapporto_kg_rifiuti_per_abitante>
  </regione>
  <regione>
    <nome_regione>Emilia Romagna</nome_regione>
    <codice_regione>06</codice_regione>
    <abitanti>4.987mil</abitanti>
    <kg_rifiuti>4578ton</kg_rifiuti>
```

```
<rapporto_kg_rifiuti_per_abitante>657</rapporto_kg_rifiuti_per_abitante>
>
  </regione>
  <regione>
    <nome_region>Marche</nome_region>
    <codice_region>11</codice_region>
    <abitanti>1.575mil</abitanti>
    <kg_rifiuti>2577ton</kg_rifiuti>

<rapporto_kg_rifiuti_per_abitante>567</rapporto_kg_rifiuti_per_abitante>
>
  </regione>
  <regione>
    <nome_region>Molise</nome_region>
    <codice_region>12</codice_region>
    <abitanti>1.321mil</abitanti>
    <kg_rifiuti>1547ton</kg_rifiuti>

<rapporto_kg_rifiuti_per_abitante>324</rapporto_kg_rifiuti_per_abitante>
>
  </regione>
  <regione>
    <nome_region>Lazio</nome_region>
    <codice_region>08</codice_region>
    <abitanti>8.987mil</abitanti>
    <kg_rifiuti>7.958ton</kg_rifiuti>

<rapporto_kg_rifiuti_per_abitante>612</rapporto_kg_rifiuti_per_abitante>
>
  </regione>
  <regione>
    <nome_region>Campania</nome_region>
    <codice_region>05</codice_region>
    <abitanti>6.476mil</abitanti>
    <kg_rifiuti>5687ton</kg_rifiuti>

<rapporto_kg_rifiuti_per_abitante>758</rapporto_kg_rifiuti_per_abitante>
>
  </regione>
  <regione codice="04">
    <nome_region>Calabria</nome_region>
    <abitanti>5.345mil</abitanti>
```

```
<kg_rifiuti>3928ton</kg_rifiuti>

<rapporto_kg_rifiuti_per_abitante>362</rapporto_kg_rifiuti_per_abitante
>
</regione>
<regione codice="14">
  <nome_regione>Puglia</nome_regione>
  <abitanti>5.476mil</abitanti>
  <kg_rifiuti>4582ton</kg_rifiuti>

<rapporto_kg_rifiuti_per_abitante>541</rapporto_kg_rifiuti_per_abitante
>
</regione>
</rifiuti_urbani_italia>
```

Visualizzazione su browser:

```
▼<rifiuti_urbani_italia>
  ▼<regione>
    <nome_regione>Piemonte</nome_regione>
    <codice_regione>13</codice_regione>
    <abitanti>4.356mil</abitanti>
    <kg_rifiuti>2270ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>516</rapporto_kg_rifiuti_per_abitante>
    <!-- cifre non reali del rapporto Kg rifiuti/abitanti regione -->
  </regione>
  ▼<regione>
    <nome_regione>Lombardia</nome_regione>
    <codice_regione>10</codice_regione>
    <abitanti>7.456mil</abitanti>
    <kg_rifiuti>4562ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>612</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione>
    <nome_regione>Veneto</nome_regione>
    <codice_regione>21</codice_regione>
    <abitanti>5.748mil</abitanti>
    <kg_rifiuti>4874ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>547</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione>
    <nome_regione>Emilia Romagna</nome_regione>
    <codice_regione>06</codice_regione>
    <abitanti>4.987mil</abitanti>
    <kg_rifiuti>4578ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>657</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione>
    <nome_regione>Marche</nome_regione>
    <codice_regione>11</codice_regione>
    <abitanti>1.575mil</abitanti>
    <kg_rifiuti>2577ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>567</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione>
    <nome_regione>Molise</nome_regione>
    <codice_regione>12</codice_regione>
    <abitanti>1.321mil</abitanti>
    <kg_rifiuti>1547ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>324</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione>
    <nome_regione>Lazio</nome_regione>
    <codice_regione>08</codice_regione>
    <abitanti>8.987mil</abitanti>
    <rapporto_kg_rifiuti_per_abitante>612</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione>
    <nome_regione>Campania</nome_regione>
    <codice_regione>05</codice_regione>
    <abitanti>6.476mil</abitanti>
    <kg_rifiuti>5687ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>758</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione codice="04">
    <nome_regione>Calabria</nome_regione>
    <abitanti>5.345mil</abitanti>
    <kg_rifiuti>3928ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>362</rapporto_kg_rifiuti_per_abitante>
  </regione>
  ▼<regione codice="14">
    <nome_regione>Puglia</nome_regione>
    <abitanti>5.476mil</abitanti>
    <kg_rifiuti>4582ton</kg_rifiuti>
    <rapporto_kg_rifiuti_per_abitante>541</rapporto_kg_rifiuti_per_abitante>
  </regione>
</rifiuti_urbani_italia>
```

Richiesta 3: DATASET CSV:

Riciclo di rifiuti urbani;;;

Regione;;;

;Codice regione;abitanti (mil);kg rifiuti urbani riciclati (ton);Kg rifiuti riciclati * abitante

Piemonte;13;4.356;1.152;234

Lombardia;10;7.456;2.458;306

Veneto;21;5.748;1.567;224

Emilia-Romagna;6;4.987;1.354;321

Marche;11;1.575;897;301

Molise;12;1.321;1.125;125

Lazio;8;8.987;2.156;357

Campania;5;6.476;2.110;215

Calabria;4;5.345;987;125

Puglia ;14;5.476;2101;254

“Fonte: Protezione e la Ricerca Ambientale, “““ISPRA””” ;;;

“Presentano la quantità totale di rifiuti urbani generati in Italia \n disaggregati a livello regionale.”

Visualizzazione su excel:

2	Regione					
3		Codice reg	abitanti (m	kg rifiuti urbani riciclati (ton)	Kg rifiuti riciclati * abitante	
4	Piemonte	13	4.356	1.152	234	
5	Lombardia	10	7.456	2.458	306	
6	Veneto	21	5.748	1.567	224	
7	Emilia-Rom	6	4.987	1.354	321	
8	Marche	11	1.575	897	301	
9	Molise	12	1.321	1.125	125	
10	Lazio	8	8.987	2.156	357	
11	Campania	5	6.476	2.110	215	
12	Calabria	4	5.345	987	125	
13	Puglia	14	5.476	2101	254	
14	Fonte: Protezione e la Ricerca Ambientale, “ISPRA”.					
15	Presentano la quantità totale di rifiuti urbani generati in Italia					
16	disaggregati a livello regionale.					
17						

	column 1 🗑	column 2	column 3	column 4	column 5
1	Tabella 10.1: Produzione di rifiuti urbani				
2	Regione				
3		Codice	abitanti (mil)	kg rifiuti (ton)	Kg rifiuti * abitante
4	Piemonte	13	4.356	2.270	516
5	Lombardia	10	7.456	4.562	612
6	Veneto	21	5.748	4.874	547
7	Emilia-Romagna	6	4.987	4.578	657
8	Marche	11	1.575	2577	567
9	Molise	12	1.321	3.344	324
10	Lazio	8	8.987	7.958	612
11	Campania	5	6.476	5.687	758
12	Calabria	4	5.345	3.928	362
13	Puglia	14	5.476	4582	541
14	Fonte: ISPRA				

Richiesta 2: DATASET JSON:

```
{
  "rapporto rifiuti urbani italia/rifiuti riciclati per regione": [
    {
      "nome regione": "Piemonte",
      "codice regione": 13,
      "numero abitanti": 4356000,
      "ton rifiuti": 2270,
      "feedback attendibile": true,
      "rifiuti riciclati per settori merceologici": [
        {
          "settore": "tessile",
          "ton rifiuti riciclati": 120
        },
        {
          "settore": "organico",
          "ton rifiuti riciclati": 212
        },
        {
          "settore": "metalli",
          "ton rifiuti riciclati": 22
        }
      ],
      "nome regione": "Lombardia",
      "codice regione": 10,
      "numero abitanti": 4356000,
      "ton rifiuti": 2270,
      "feedback attendibile": true,
      "rifiuti riciclati per settori merceologici": [
        {
          "settore": "tessile",
          "ton rifiuti riciclati": 120
        },
        {
          "settore": "organico",
          "ton rifiuti riciclati": 212
        },
        {
          "settore": "metalli",
          "ton rifiuti riciclati": 22
        }
      ],
      "nome regione": "Veneto",
      "codice regione": 21,
```

```
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
{"nome regione":"Emilia Romagna",
"codice regione":6,
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
{"nome regione":"Marche",
"codice regione":11,
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
{"nome regione":"Molise",
"codice regione":12,
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
{"nome regione":"Lazio",
"codice regione":8,
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
```



```
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
{"nome regione":"Campania",
"codice regione":5,
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
{"nome regione":"Calabria",
"codice regione":4,
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
{"nome regione":"Puglia",
"codice regione":14,
"numero abitanti":4356000,
"ton rifiuti":2270,
"feedback attendibile":true
"rifiuti riciclati per settori merceologici":[
  {"setteore":"tessile","ton rifiuti riciclati":120},
  {"setteore":"organico","ton rifiuti riciclati":212},
  {"setteore":"metalli","ton rifiuti riciclati":22}
],
]}
```

Analisi e raccolta dei requisiti:

Attributi disponibili per l'analisi:

- **nome regione**
- **codice regione**
- **numero di abitanti**
- **kg rifiuti (ton)**
- **kg rifiuti (ton)/numero abitanti**
- **rifiuti riciclati per settori merceologici**

Nel dataset in **XML**, abbiamo i dati sui rifiuti urbani regionali, con informazioni sul numero di abitanti, tonnellate di rifiuti prodotti e il rapporto tra tonnellate rifiuti prodotti e numero di abitanti.

Nel dataset in **JSON**, i dati forniti analizzano il rapporto tra rifiuti urbani regionali e i rifiuti riciclati per settore merceologico. Contengono anch'essi il numero di abitanti e tonnellate di rifiuti prodotti, con un booleano sull'affidabilità dei dati forniti dalle regioni.

Il dataset in **CSV**, contiene gli stessi attributi dell'XML, ma sotto forma appunto di formato CSV.

Outcome atteso dal sistema:

Grazie ai dati raccolti sarà possibile effettuare le seguenti **considerazioni** relative al tasso di rifiuti regionali riciclati in Italia. Osservando questi dataset relativi alle regioni, sarà possibile valutare su quali aree intensificare la regolamentazione sui rifiuti differenziabili e predisporre delle linee guida al miglioramento di queste ultime.

Grazie a **tecniche predittive** (Data Scientist), si potrà stimare l'evoluzione della situazione per gli anni successivi.