



Programación I

Proyecto Integrador

Proyecto Alfa

GHJJKZ6

“Países y Empresas que mas contaminan a nivel mundial”

Colaboradores:

- Juchniewicz Federico
- Zelarayán Román
- Gutiérrez Emma Vilma
- Hanun Romia
- Kanneman Samuel

Profesores:

Silvia Perotti
Hector Prado

Tema a desarrollar:

La propuesta llevada a cabo en el presente trabajo es tomada conjuntamente por el equipo de colaboradores, la misma se centra en la problemática de la contaminación a nivel mundial por parte de los países y sus empresas.

Librerías Utilizadas

- bs4 BeautifulSoup
- Requests
- Pandas
- Pathlib

Para el seguimiento del proyecto se hace uso de la técnica de Scrum en Trello, donde se toma decisión conjunta de los roles.

Roles:

Product Owner: Juchniewicz Federico

Scrum Master: Zelarayán Román

Team:

Gutierrez Emma Vilma

Hanun Romina

Jabie Salomon

Kanneman Samuel

Paginas utilizadas:

<https://www.liderempresarial.com/las-10-empresas-que-mas-contaminan-el-mundo/>

<https://datosmacro.expansion.com/energia-y-medio-ambiente/emisiones-co2?anio=2021>

<https://www.iqair.com/es/world-most-polluted-countries>

Origen de Datos para el Proyecto:

Comenzamos con la búsqueda de información sobre el tema a desarrollar, verificamos que, sobre las páginas web, se pueda llevar a cabo el Web Scraping para extraer datos de las mismas.

Nuestra búsqueda se centró en información sobre la contaminación a nivel global con la producción de plásticos contaminantes, recopilamos información sobre las principales empresas, emisiones de CO₂ fósil por país: incluyen fuentes del uso de combustibles fósiles (combustión, quema), procesos industriales (cemento, acero, productos químicos y urea) y uso de productos; así como también los países y regiones con mayor contaminación del mundo, basado en datos de la concentración de PM_{2.5} media anual $\mu\text{g}/\text{m}^3$.

La materia particulada o PM (por sus siglas en inglés) 2.5, son partículas muy pequeñas en el aire que tiene un diámetro de 2.5 micrómetros (aproximadamente 1 diezmilésimo de pulgada) o menos de diámetro. Esto es menos que el grosor de un cabello humano. La materia particulada, uno de los seis criterios de contaminantes del aire de la U.S. EPA, es una mezcla que puede incluir sustancias químicas orgánicas, polvo, hollín y metales.

Las partículas PM_{2.5}, por tanto, se pueden acumular en el sistema respiratorio y están asociadas, cada vez con mayor consistencia científica, con numerosos efectos negativos sobre la salud, como el aumento de las enfermedades respiratorias y la disminución del funcionamiento pulmonar.

La combustión de combustibles es la principal razón del incremento de las emisiones de CO₂. Los combustibles implicados en dicha producción de CO₂ son:

- La combustión de combustibles fósiles sólidos: como el carbón.
- La combustión de combustibles líquidos, como el petróleo, la gasolina, el diésel y el queroseno.
- La combustión de combustibles gaseosos como el gas natural.

La excesiva emisión de CO₂, hace que la producción de CO₂ en la atmosfera se haya disparado lo provoca la contaminación del aire, el tan nombrado calentamiento global y la acidificación de los océanos.

A través del compromiso y labor de 4,734 voluntarios en 55 países que realizaron un total de 575 auditorías a distintas marcas, catalogaron alrededor de 5 mil marcas y recolectaron 346,494 piezas de plástico con etiquetas visibles de las marcas. En 2020, las empresas que ocuparon los primeros lugares a nivel global en la producción de plásticos contaminantes son: Coca Cola, Pepsico y Nestlé. Las que ocuparon los primeros lugares en México son: Coca Cola, Pepsico y Bimbo. Las marcas han ocupado estas posiciones desde el 2018, se presenta el listado de empresas y marcas que generan más contaminación de plástico de uso único.

Objetivos:

Mediante el conocimiento del cursado de Programación de TSCDIA y curso Cisco de Python, realizamos:

- ✓ Creación de repositorio en GitHub.
- ✓ Técnicas ágiles de Scrum en Trello.
- ✓ Desde Visual Code realizamos la creación y programación de los archivos.py para el proyecto.
- ✓ Programamos incluyendo librerías para la obtención de datos de las páginas relacionadas a la contaminación de países y Empresas a nivel global.
- Se obtienen tres tablas:
 - Empresas que mayor contaminación de plásticos en el mundo.
 - Países con mayor emisión de Dióxido de Carbono.
 - Países y regiones con mayor contaminación del mundo en calidad del aire (datos históricos entre 2018-2021).
- ✓ Utilizar la técnica de Scrum en Trello para el seguimiento del proyecto.
- ✓ Una vez realizada la extracción de datos exportar a csv, donde se realizan gráficas de resultados obtenidos.
- ✓ Crear Base de Datos "db_paises.sql"
- ✓ Crear Tabla "paisest_co2"
- ✓ Realizar una consulta a tabla "paisest_co2", donde filtramos los países que en 2021 aumentarán un 10% las emisiones per cápita de CO2 con respecto al año anterior.

Desarrollo del trabajo

Comenzamos importando las librerías.

Creamos los archivos.py, que necesitaremos para el código y así extraer datos de las páginas con la información objetida del Web Scraping para recolectarlos en archivos csv y así, posteriormente, mostrarlos por consola.

A partir del primer link: <https://www.liderempresarial.com/las-10-empresas-que-mas-contaminan-el-mundo/>

- Definimos el header para la compatibilidad de la página;
- Usamos Request para enviar solicitud a la página y así obtener el código HTML;
- Creamos la variable soup para localizer los elementos de la página web con BeautifulSoup, donde asignamos la variable para obtener el código y definir el parser;
- Inicializamos contadores;
- Creamos 2 listas, donde almacenaremos los datos;
- Buscamos etiquetas donde se encuentran los datos para agregar nombres de empresas y descripción, se agregan a una de las listas;
- Creamos dataframe;
- Exportamos al archive CSV para comparación de datos y valores.

| | NOMBRE | DESCRIPCION |
|--|---------------------------------|---|
| 0 | 1. The Coca Cola Company | Por tercer año consecutivo la empresa se ha po... |
| 1 | 2. Pepsico | Al igual que Coca Cola esta marca se ha posici... |
| 2 | 3. Nestlé | Esta marca sigue figurando por segundo año con... |
| 3 | 4. Unilever | Englobando las marcas de Persil Cornetto y Sun... |
| 4 | 5. Mondelēz International | Formando parte nuevamente de este listado la e... |
| 5 | 6. Mars | En un total de 32 países se recolectaron 678 r... |
| 6 | 7. P&G (Proctel & Gamble) | La empresa de bienes de consumo tiene entre su... |
| 7 | 8. Phillip Morris International | La mayor empresa tabacalera del mundo con marc... |
| 8 | 9. Colgate-Palmolive | Con presencia en 24 de los 55 países esta marc... |
| 9 | 10. Perfetti Van Melle | Responsable de las marcas de Chupa Chups Mento... |
| PS C:\Users\negro\Downloads\ISPC\PI\GHJJKZ6\PROYECTO ALFA\GHJJKZ6> | | |

Con el segundo link: <https://datosmacro.expansion.com/energia-y-medio-ambiente/emisiones-co2?anio=2021>

- Realizamos web scraping;
- Cargamos librerías;
- Definimos el header para la compatibilidad de la página;
- Verificamos si la página responde;
- Ubicamos datos para obtener las tablas;
- Mediante “pandas” ubicamos las tablas que contenga la página web;
- Creamos dataframe filtrando las tablas que contengan la información que necesitamos;
- Eliminamos columnas que no necesitamos;
- Exportamos a CSV.

| | Países | CO2 Totales Mt | CO2 Kg/1000\$ | CO2 t per capita | Var. |
|---|----------------|----------------|---------------|------------------|--------|
| 0 | España | 231914 | 13 | 499 | 9.18% |
| 1 | Alemania | 665884 | 15 | 806 | 5.78% |
| 2 | Reino Unido | 335363 | 11 | 495 | 4.59% |
| 3 | Francia | 382427 | 18 | 458 | 7.88% |
| 4 | Italia | 319669 | 13 | 541 | 8.38% |
| 5 | Portugal | 38677 | 11 | 380 | -1.94% |
| 6 | Estados Unidos | 4.752,879 | 23 | 1424 | 5.78% |
| 7 | Japón | 1.884.604 | 21 | 859 | 3.14% |

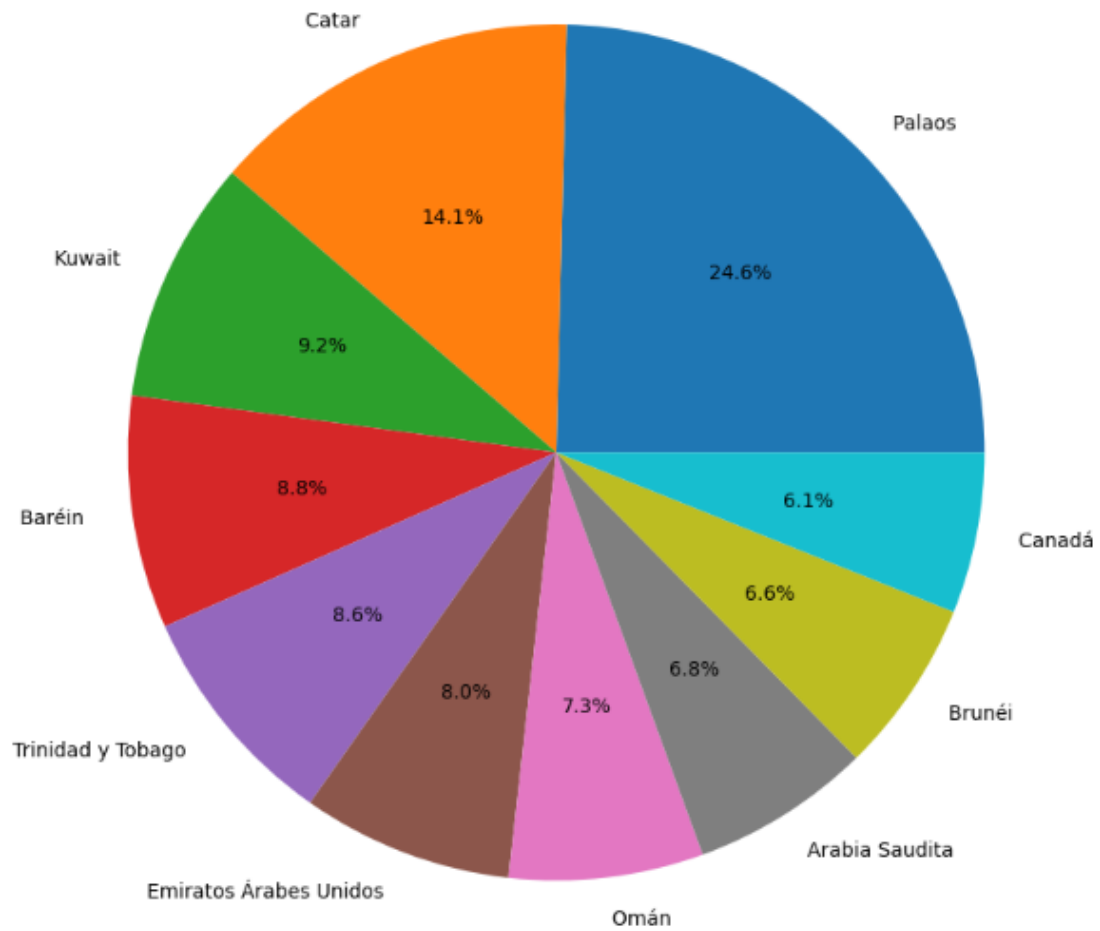
Con el tercer link se procede de la misma forma

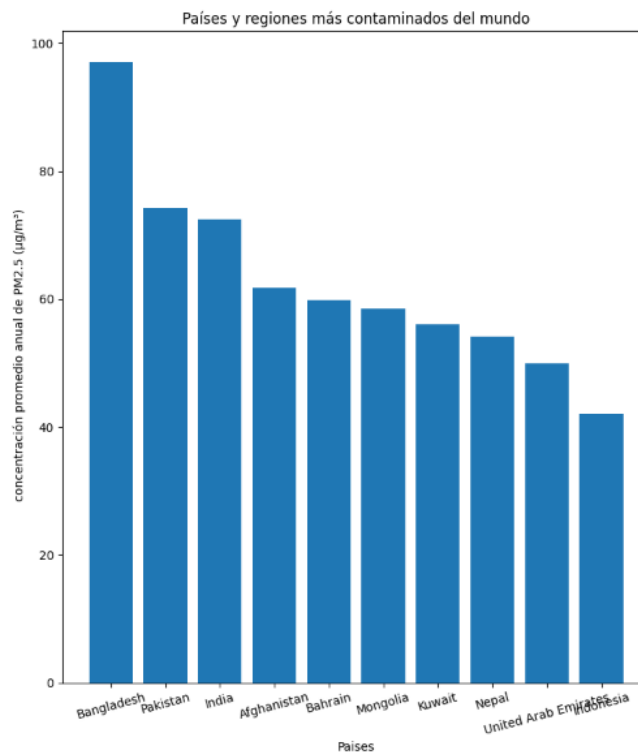
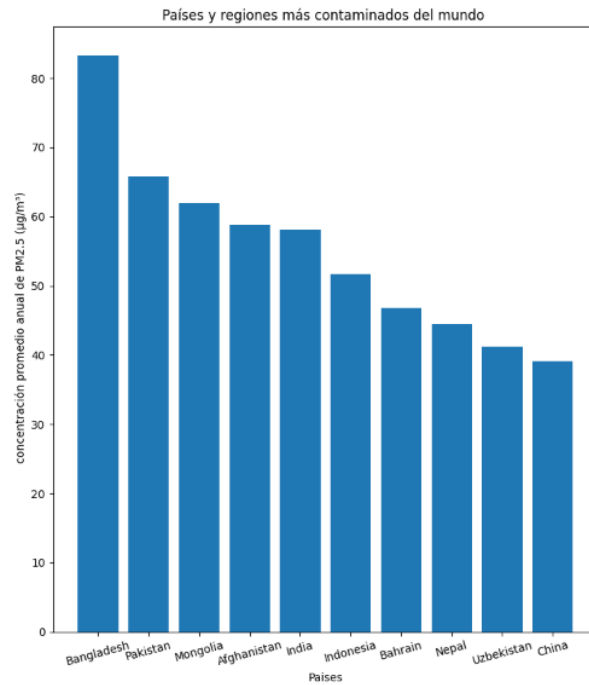
| | Rank | Country/Region | 2021 | 2020 | 2019 | 2018 | Population |
|---|------|----------------|------|------|------|------|------------|
| 0 | 1 | Bangladesh | 76.9 | 77.1 | 83.3 | 97.1 | 164689383 |
| 1 | 2 | Chad | 75.9 | - | - | - | 16425859 |
| 2 | 3 | Pakistan | 66.8 | 59 | 65.8 | 74.3 | 228092331 |
| 3 | 4 | Tajikistan | 59.4 | 38.9 | - | - | 9537642 |
| 4 | 5 | India | 58.1 | 51.9 | 58.1 | 72.5 | 1380004385 |
| 5 | 6 | Oman | 53.9 | 44.4 | - | - | 5106622 |
| 6 | 7 | Kyrgyzstan | 50.8 | 43.5 | 33.2 | - | 6524191 |
| 7 | 8 | Bahrain | 49.8 | 39.7 | 46.8 | 59.8 | 1781583 |
| 8 | 9 | Iraq | 49.7 | - | - | - | 48222503 |
| 9 | 10 | Nepal | 46.8 | 39.2 | 44.5 | 54.1 | 29136888 |

P5 C:\Users\negro\Downloads\ISPC\PI\GHJJKZ6\PROYECTO ALFA\GHJJKZ6>

Creación de gráficas

Países con mayor emisión de CO2 per capita





Data Base

Mediante el archivo .py con el código conectamos a la base de datos que previamente creamos mediante XAMPP, mediante método cursor filtramos los datos que queremos obtener, en éste caso filtramos los países que en 2021 aumentaron un 10% las emisiones per cápita de dióxido de carbono con respecto al año anterior ordenandolo de manera descendente.