

Proyecto Ecobici Guía Sugerida de Retos o Pasos a llevar a cabo para resolver la problemática

RETO 1

Obtener los conjuntos de datos

Descripción del reto

En la descripción del proyecto se describe como la empresa Ecobici pone a disposición en su sitio web datos abiertos de dos tipos, es necesario revisar cada link proporcionado para cada caso.

Para el caso de los datos históricos obtener los archivos correspondientes al año 2023 en formato CSV y para el caso de los datos en tiempo real realizar una consulta y guardar los resultados en un archivo en formato JSON.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo request última versión disponible
- Librería o módulo json última versión disponible

RETO 2

Agrupar datos de viajes del año 2023

Descripción del reto

Debido a que el sitio Web de Ecobici proporciona los datos históricos por mes se obtuvo un archivo en formato CSV por cada uno de los meses para el año 2023, por lo que deberás agrupar los datos de cada archivo en un sólo archivo llamado 2023.csv realizando un programa en Python que realice ésta tarea.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo csv última versión disponible

RETO 3

Explorar los datos y realizar limpieza de valores nulos y anómalos

Descripción del reto

En éste punto ya contamos con archivos con datos, pero que son desconocidos para nosotros, así que deberás de realizar una exploración de todos los archivos descargados identificando nombres de columnas o valores en cada columna en el caso de archivos csv o las llaves en objetos y sus valores en archivos json con el objetivo que que logres familiarizarte con los datos que has obtenido.

Después deberás de buscar si hay valores nulos o anómalos y decidir si es necesario eliminarlos o transformarlos, obteniendo así un nuevo archivo con datos limpios, realiza ésto por cada archivo obtenido.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo pandas 2.0 o mayor

RETO 4

Transformar datos en formato JSON a CSV

Descripción del reto

Para poder relacionar datos de manera más imple y directa es recomendable que todos los datos estén en un mismo formato y el formato por excelencia es el CSV debido a su similitud con una tabla de datos, por lo que deberás crear un programa en Python para transformar cada unos de los archivos JSON en archivos CSV.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo pandas 2.0 o mayor

RETO 5

Adquirir nuevos datos y complementar los ya existentes

Descripción del reto

En muchas ocasiones los conjuntos de datos obtenidos están incompletos por error o por diseño o simplemente no tienen toda la información que necesitamos, así que llegado a éste punto habrás notado que los datos de las cicloestaciones sólo incluyen datos de geolocalización (longitud y latitud) por lo que necesitas realizar un programa en Python que usando los datos de longitud y latitud de cada cicloestación obtenga los datos de código postal, colonia y alcaldía (opcional) y los adicione a los datos existentes obteniendo así un nuevo archivo en formato CSV de cicloestaciones que puedes llamar cicloestaciones-con-direcciones.csv.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo pandas 2.0 o mayor
- Librería o módulo geopy última versión disponible

RETO 6

Relacionar los datos de viajes y cicloestaciones

Descripción del reto

En éste punto estamos listos para crear un conjunto de datos (dataset) con toda la información útil para poder generar algunas estadísticas, por lo tanto necesitas crear un programa en Python que lea los datos de los viajes y también los datos de las cicloestaciones y los combine en base a una columna común en ambos archivos y el conjunto de datos resultante lo deberás de guardar en el archivo viajes-completo.csv

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo pandas 2.0 o mayor

RETO 7

Obtener algunos de los resultados obtenidos en la encuesta de la SEMOVI

Descripción del reto

Revisa la encuesta realizada por la SEMOVI en el link proporcionado e identifica la mayor cantidad de resultados estadísticos que puedas obtener con el conjunto de datos final, entonces realiza un programa en Python que genere esos resultados pero para el año 2023 y genera un reporte.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo pandas 2.0 o mayor

RETO 8

Generar algunas de las gráficas obtenidas en la encuesta de la SEMOVI

Descripción del reto

Ya con datos y resultados una forma de complementar es generando gráficas que muestren los resultados de forma gráfica y para ello deberás de realizar un programa en Python que haga uso de la librería Matplotlib y Pandas para realizar algunas de las gráficas mostradas en la encuesta de la SEMOVI pero para el año 2023 y genera un nuevo reporte incluyendo ahora las gráficas.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo pandas 2.0 o mayor
- Librería o módulo matplotlib última versión disponible

RETO 9

Análisis de datos adicionales

Descripción del reto

Hasta éste punto se ha obtenido los resultados sugeridos por la encuesta de la SEMOVI, pero no siempre se contará con una encuesta o guía que sugiera los resultados a obtener, por lo que ahora realizarás un programa en Python que obtenga 3 resultados adicionales usando el mismo conjunto de datos, por ejemplo podrías determinar el número de bicis regresadas por un desperfecto por cada mes o podrías mejorar alguno de los resultados de la encuesta, como por ejemplo el porcentaje de viajes por cada hora a lo largo de todo el día en todo lo que va del año 2023. Genera un reporte incluyendo tus resultados y las gráficas correspondientes.

Listado de herramientas que se utilizan para la solución del reto

- Python 3.7 o mayor
- Navegador web Google Chrome o Firefox
- Jupyter Lab 2.0 o mayor
- Librería o módulo pandas 2.0 o mayor
- Librería o módulo matplotlib última versión disponible