**Instrucciones para ejecutar Web Scraping del RUNT**

**Paso 1: Preparación del Entorno**

Instalación de Python:

Asegúrate de tener Python instalado. Puedes descargar la última versión desde python.org.

Para no presentar problemas de compatibilidad en las librerías instalar Python 3.9.

Durante la instalación, marca la casilla que dice "Agregar Python al PATH".

Instalación de Dependencias:

Abre el "Símbolo del sistema" o "Terminal" en Visual Studio Code o tu PC.

Ejecuta el siguiente comando para instalar las bibliotecas necesarias:

pip install selenium pandas matplotlib opencv-python-headless

Es posible que solicite instalar otras librerías si nunca han sido instaladas en tu entorno, para hacerlo seguir la misma metodología.

**Paso 2: Configuración del Entorno de Ejecución**

Descargar e Instalar chromedriver.exe:

Descarga el controlador de Chrome (chromedriver.exe) desde este enlace.

Descomprime el archivo descargado y coloca chromedriver.exe en una carpeta en tu sistema en la siguiente ruta C:\WebScraping\chromedriver.exe).

**Paso 3: Abrir el Código**

Descarga y abre el script de Python y las funciones necesarias desde este enlace "C:\Users\ctamayo\OneDrive - Renting Colombia S.A\Documentos - Analítica\GERENCIA\_CAPACIDADES\_ANALITICAS\Planes\_de\_Choque\RUNT" y guárdalos en una carpeta fácil de recordar en tu sistema.

**Paso 4: Configuración del Código**

Modificar Rutas y Constantes:

1. Abre el script de Python en un editor de código (Preferiblemente VSCode.)
2. Modifica las rutas de los archivos Excel y las constantes según tus necesidades:

placas = pd.read\_excel(r"ruta\al\archivo\placas.xlsx")

df\_existente = pd.read\_excel(r"ruta\al\archivo\runt.xlsx")

df\_existente\_error = pd.read\_excel(r"ruta\al\archivo\error.xlsx")

Archivo placas: El web scraping hace la búsqueda en el RUNT por medio de las placas aquí diligenciadas.

Archivo runt: Es el archivo donde se escribe la información capturada del web scraping.

Archivo error: Archivo que almacena información de los errores.

(Tener en cuenta que si vas a realizar la extracción en varias ejecuciones debes descartar del listado de placas las ya existentes en el archivo runt para que estas no sean buscadas y escritas nuevamente)

Aquí una muestra de todas las rutas que deben modificarse:









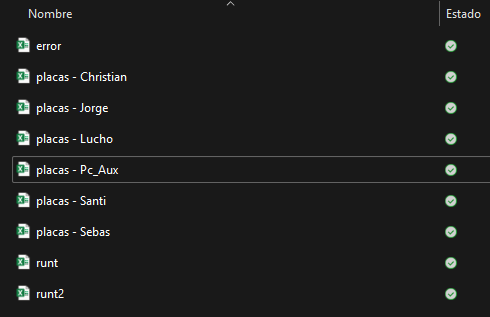








*Las placas que corresponden a cada uno ya se encuentran distribuidas entre los miembros del equipo en archivos de Excel dentro de la carpeta insumos bajo el nombre de “placas - (nombre integrante)”, por favor usar la que corresponde a tu nombre y modifica el nombre de esa específicamente solo por “placas” localmente.*

**

**Paso 5: Script functions:**

En la carpeta encontraras el script functions, esta aloja parte de las funciones del web scraping, con este lo único que debes hacer es modificar la ruta de inicio según la ruta que tengas para la carpeta en tu equipo y posteriormente guardar los cambios:



Nota: Los demás scripts y archivos no es necesario modificarlos.

**Paso 6: Ejecución del Código**

Ejecutar el Script:

Ejecutar el script en la parte superior derecha para Visual Studio Code.

Para otras opciones en general abre el "Símbolo del sistema" o "Terminal".

Navega hasta la carpeta donde guardaste el script usando el comando cd ruta\de\la\carpeta.

Ejecuta el script usando el siguiente comando:

RUNT.py

*Nota: Recomiendo abrir en youtube un video de pantalla negra 24 horas y dejarlo activo mientras se realiza la ejecución para que tu equipo no se bloquee por inactividad.*

**Paso 7: Monitoreo de la Ejecución**

Revisar el Registro:

El script generará registros de la ejecución en el archivo consulta\_runt.log. Puedes abrirlo con un editor de texto para obtener información sobre el progreso y posibles errores.

Luego de ejecutar el código y extraer la información satisfactoriamente comparte el archivo Excel “runt” para realizar la respectiva consolidación.