

PREREQUISITOS

Python 3.x +
PIP
Virtual ENV
Code Editor

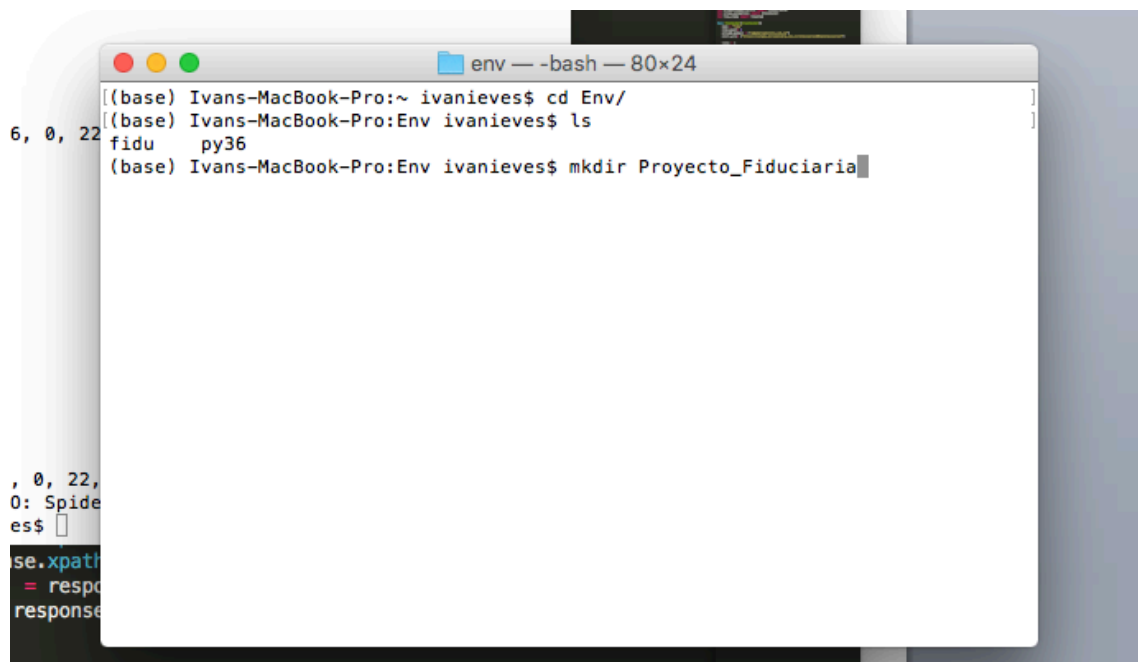
Solución en el Repositorio: <https://github.com/ivanieves/env>

Primer Paso

Instalar entorno virtual (virtualenv vía librería PIP)

Un entorno virtual es una herramienta para mantener dependencias requeridas por diferentes proyectos en lugares separados. virtualenv crea un folder aislado el cual contiene todos los ejecutables necesarios para usar los paquetes que un proyecto de Python demanda.

para esto, nos dirigimos a la Terminal, y en ella navegamos a la carpeta donde usualmente guardamos los entornos virtuales. Con la orientación mkdir creamos el directorio y luego nos aseguramos de habilitar el entorno virtual.



```
env — -bash — 80x24
(base) Ivans-MacBook-Pro:~ ivanieves$ cd Env/
(base) Ivans-MacBook-Pro:Env ivanieves$ ls
fidu  py36
(base) Ivans-MacBook-Pro:Env ivanieves$ mkdir Proyecto_Fiduciaria
```

imagen 01. Puede encontrar el archivo de texto con la creación del entorno virtual (adjunto en: env/Soporte)

Paso Dos

Instalar y Comenzar Scrapy Project

Scrapy es una plataforma colaborativa de código libre que corre en Python para extraer datos de páginas web. Dado que es un framework, Scrapy tiene una serie de herramientas poderosas para hacer el "scraping" o extraer información de webs de manera fácil y eficiente.

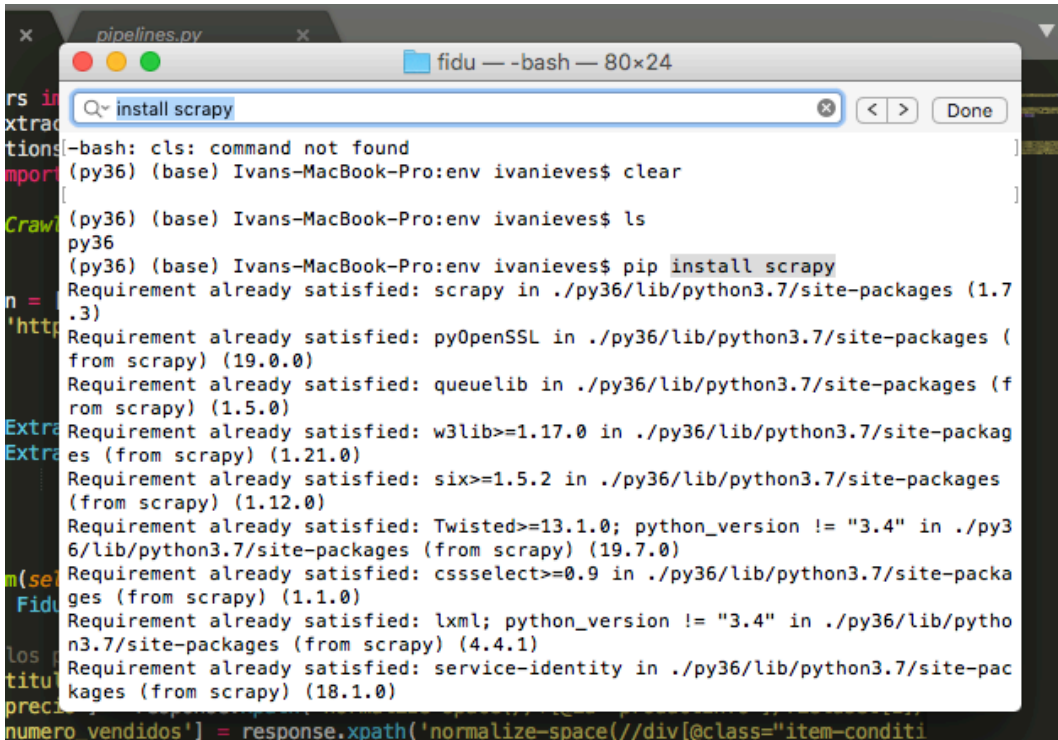


imagen 02.

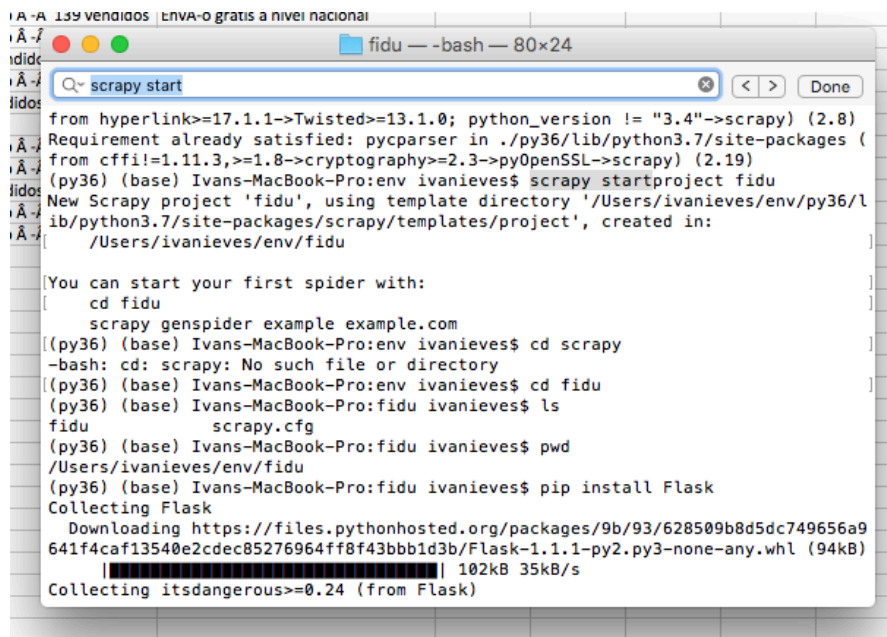


imagen 03.

Paso Tres

Editor de Texto (Sublime Text3)

En este punto, se deja un poco de lado la Terminal. Nos dirigimos a la carpeta que con la creación del proyecto Scrapy se acaba de crear, y la cargamos en el editor de texto de código de preferencia. En este caso Sublime Text. Aquí se modificarán los .py

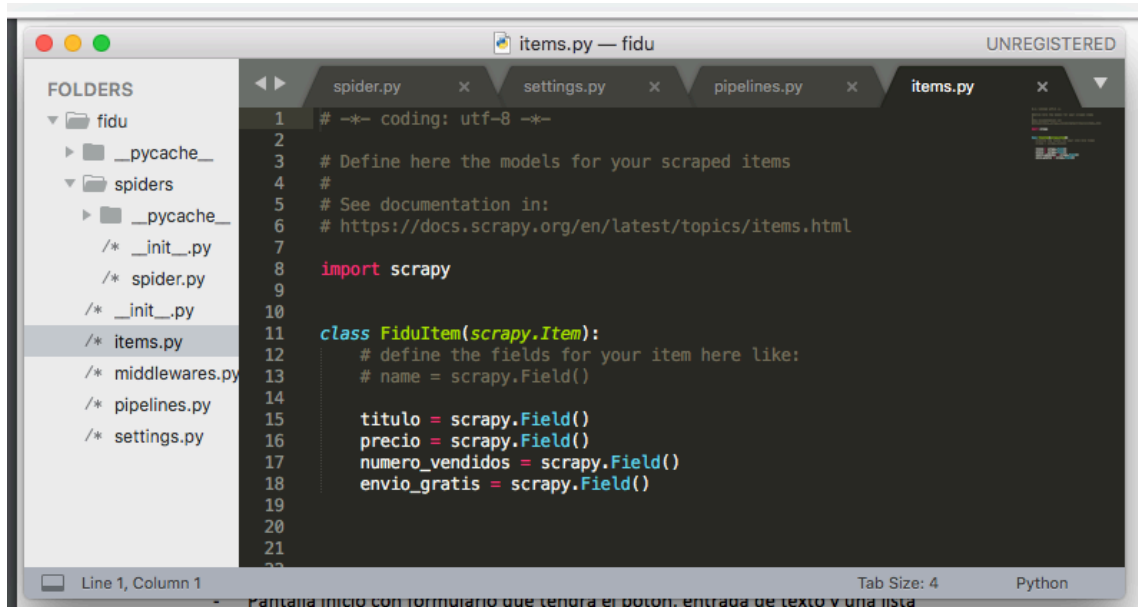


Imagen 04. Ítems.py se crean los ítems según los campos que queramos extraer de la web

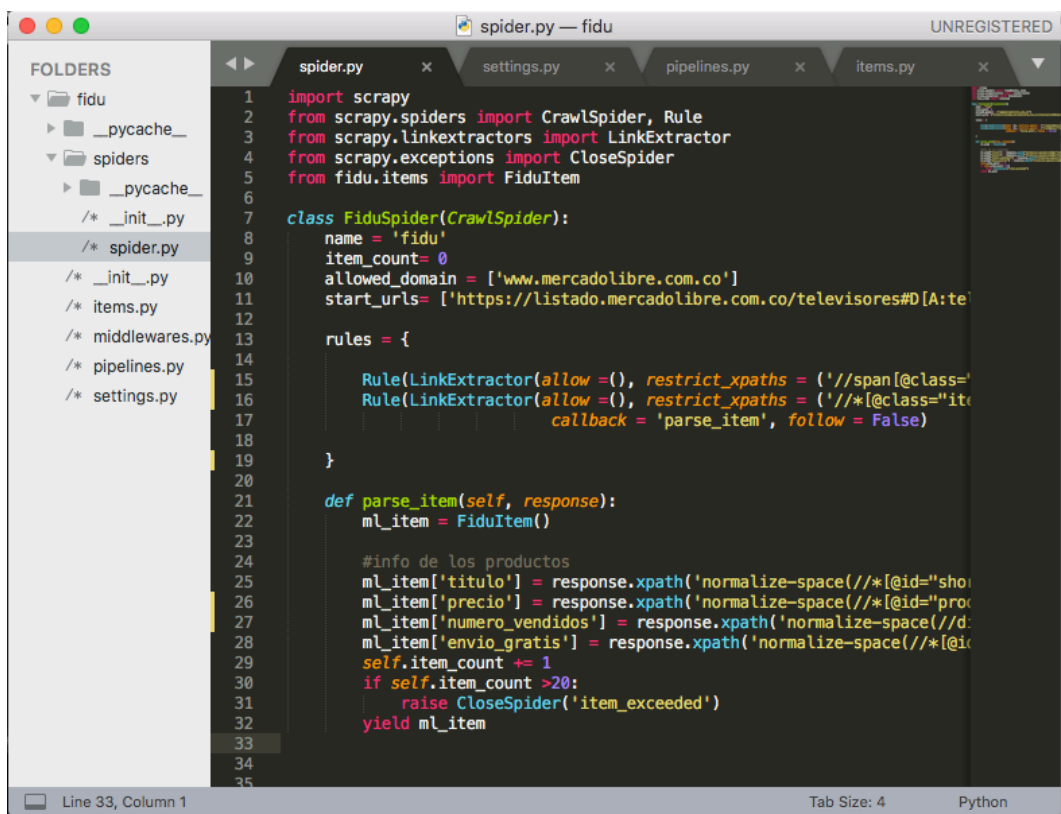


Imagen 05. Se crea el spider en spider.py en la carpeta spiders

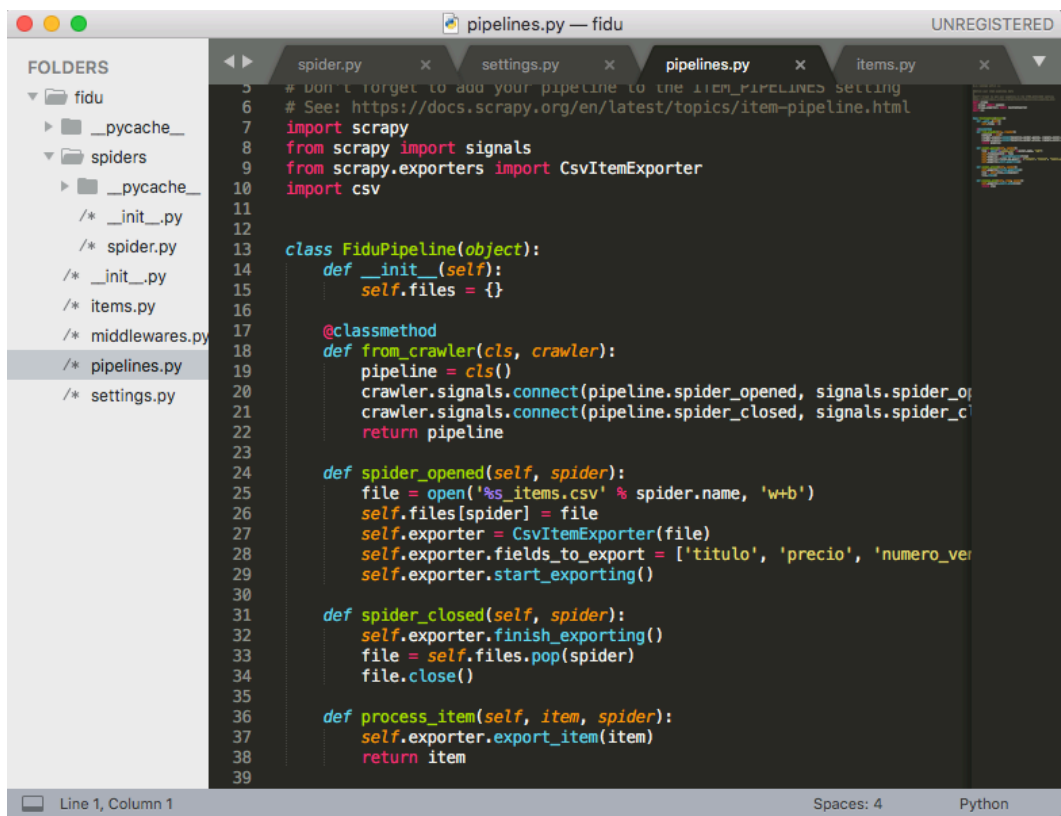


Imagen 06. Pipeline para exportar a csv

Paso Cuarto

Correr scrapy desde la Terminal

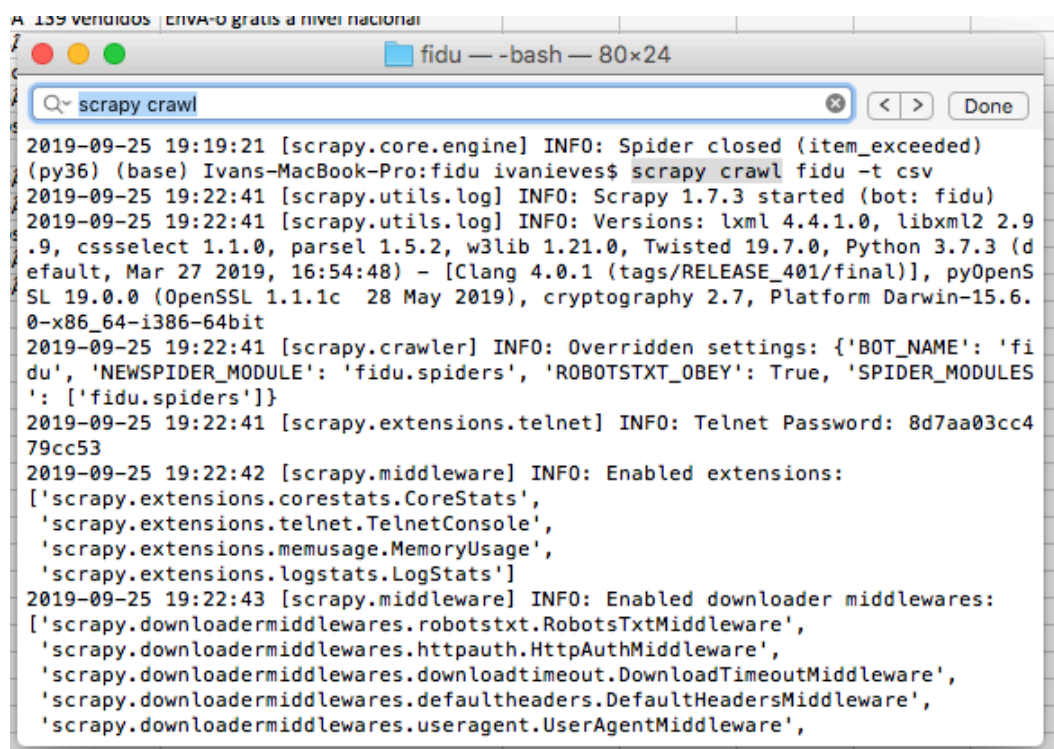


Imagen 07.

Volvemos a la Terminal esta vez para correr el scrapy modificado, que dados los cambios realizados en los archivos .py va a mercadolibre y trae la información requerida. En este caso el número de registros a investigar y el artículo buscado no tienen dinámica front End. En vez de esto, están declarados de manera muy sencilla en el archivo spider.py que creamos en la carpeta spiders de nuestro entorno.

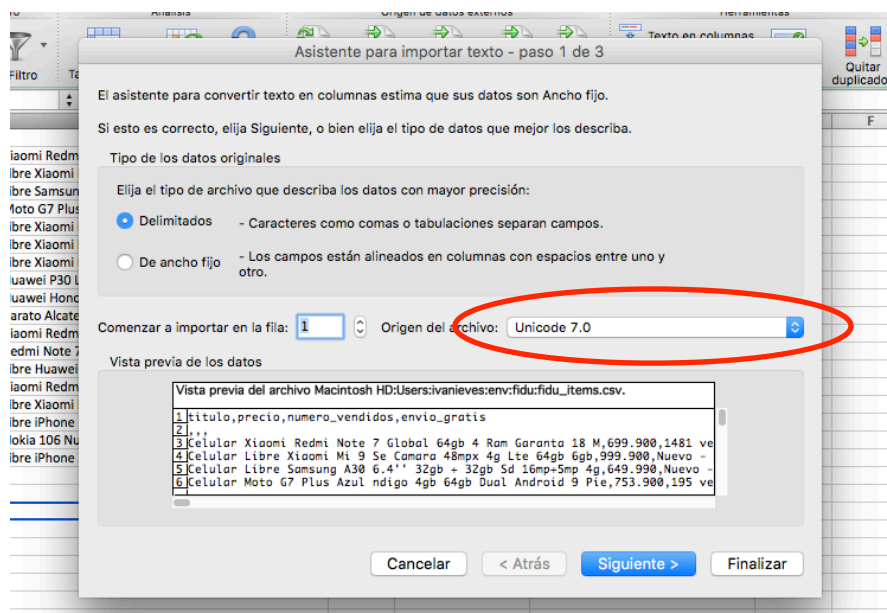
Por Último

Visualización de Nuestra Consulta Mercadolibre

Siguiendo los pasos anteriormente descritos, se va a generar un archivo con nombre fidu_items.csv en la carpeta fidu, a su vez dentro de la carpeta fidu alojada en nuestro repositorio de entornos virtuales env. Este archivo lo podemos cargar a Excel mediante la opción: importación de datos desde un archivo de texto.

titulo	precio	numero_vendidos	envio_gratis
Celular Xiaomi Redmi Note 7 Global 64gb 4 Ram Garanta 18 M	699.9	1481 vendidos	Envío gratis a nivel nacional
Celular Libre Xiaomi Mi 9t Versin Global 128gb 6ram Negro		69 vendidos	Envío gratis a nivel nacional
Celular Libre Xiaomi Mi A3 64gb 48mp 4030mah 4g Lte	699.9	Nuevo - 192 vendidos	Lo retiro en el domicilio del vendedor
Celular Libre Samsung A30 6.4" 32gb + 32gb Sd 16mp+5mp 4g	649.99	Nuevo - 119 vendidos	Envío gratis a nivel nacional
Celular Moto G7 Plus Azul ndigo 4gb 64gb Dual Android 9 Pie	753.9	195 vendidos	Envío gratis a nivel nacional
Celular Huawei P30 Lite 128gb / 4ram 3cam+ Forro Y Vidrio 5d		Nuevo - 179 vendidos	Lo retiro en el domicilio del vendedor
Celular Libre Xiaomi Mi 9t 64gb 48mp Snapdragon 730 4g Lte	1.099.900	Nuevo - 217 vendidos	Lo retiro en el domicilio del vendedor
Celular Libre Xiaomi Mi 9 Se Camara 48mpx 4g Lte 64gb 6gb	999.9	Nuevo - 114 vendidos	Lo retiro en el domicilio del vendedor
Celular Xiaomi Redmi Note 7 64gb Negro	669.9	237 vendidos	Envío gratis a nivel nacional
Celular Libre Xiaomi Redmi Note 7 32gb 48mpx 4g Lte	549.9	Nuevo - 200 vendidos	Lo retiro en el domicilio del vendedor
Celular Nokia 106 Nuevo - Radio Fm Reloj 2019		Nuevo - 144 vendidos	Envío a nivel nacional
Celular Libre Huawei Y9 2019 6.5" 64gb/3gb 16+2mp 13+2mp	648.9	Nuevo - 353 vendidos	Envío gratis a nivel nacional
Celular Libre iPhone 7 128gb 4.7 12 Mpx		76 vendidos	Lo retiro en el domicilio del vendedor
Celular Xiaomi Redmi Note 7 64gb Ram 48mpx 4g Lte	649.9	Nuevo - 109 vendidos	Lo retiro en el domicilio del vendedor
Celular Redmi Note 7 64gb /48mp 4ram + Forro Y Vidrio 5d		Nuevo - 60 vendidos	Envío gratis a nivel nacional
Celular Libre iPhone Xs Max 64gb Ram 4gb 12mp/7mp Garanta	4.099.900	Nuevo - 19 vendidos	Lo retiro en el domicilio del vendedor
Celular Xiaomi Mi 8 Lite 64gb 12+5mp/24mp 4gb Ram+ Vidrio 5d	589.9	Nuevo - 87 vendidos	Lo retiro en el domicilio del vendedor

Nota: Los datos van delimitados por comas. Para una mejor visualización de estos, se debe elegir como origen de archivo: Unicode 7.0



Conclusiones Adicionales

La realización de este proyecto se abordó mediante la instalación de módulos en Python. Para el Web Scrapping, se usó la plataforma colaborativa libre: [scrapy](#) de la biblioteca de código [PIP](#). En el repositorio encontrará en: env/fidu/fidu, los archivos .py modificados y en: env/fidu/fidu/spiders, el spider.py que llama los datos desde mercadolibre.com

Adicionalmente encontrará este documento junto a otros archivos .txt adjuntos con los códigos de la Terminal de los pasos Uno, Dos y Cuatro en: env/Soporte