



Entrenamiento Piscine Python para datascience - 2

Tabla de datos

Resumen: Hoy aprenderás a cargar, manipular y mostrar tablas de datos.

Versión: 1.00

Contenido

I	Reglas generales	2
II	Instrucciones específicas del día	3
III	Ejercicio 00	4
IV	Ejercicio 01	5
V	Ejercicio 02	6
VI	Ejercicio 03	7
VII	Presentación y evaluación por pares	9

Capítulo I

Reglas generales

- Debes renderizar tus módulos desde una computadora en el clúster usando una máquina virtual:
 - Puede elegir el sistema operativo que utilizará para su máquina virtual
 - Su máquina virtual debe contar con todo el software necesario para realizar su proyecto. Este software debe estar configurado e instalado.
- O puedes utilizar directamente la computadora en caso de que tengas las herramientas disponibles.
 - Asegúrate de tener el espacio en tu sesión para instalar lo que necesitas para todos los módulos (usa goinfre si tu campus lo tiene)
 - Debes tener todo instalado antes de las evaluaciones.
- Sus funciones no deberían cerrarse inesperadamente (error de segmentación, error de bus, doble liberación, etc.) salvo por comportamientos indefinidos. Si esto sucede, su proyecto se considerará no funcional y recibirá una 0 Durante la evaluación.
- Le animamos a crear programas de prueba para su proyecto, incluso si este trabajo no es suficiente. **No será necesario enviarlo y no será calificado..** Te dará la oportunidad de evaluar fácilmente tu trabajo y el de tus compañeros. Estas pruebas te resultarán especialmente útiles durante tu defensa. De hecho, durante la defensa, eres libre de utilizar tus propias pruebas y/o las pruebas de los compañeros a los que estás evaluando.
- Envía tu trabajo al repositorio git que se te haya asignado. Solo se calificará el trabajo que se encuentre en el repositorio git. Si se le asigna a Deepthought la tarea de calificar tu trabajo, se hará después de las evaluaciones de tus pares. Si ocurre un error en alguna sección de tu trabajo durante la calificación de Deepthought, la evaluación se detendrá.
- Debes utilizar la versión Python 3.10
- Las importaciones de bibliotecas deben ser explícitas, por ejemplo, debe "importar numpy como np". No se permite importar "from pandas import *" y obtendrá 0 en el ejercicio.
- No hay ninguna variable global.
- ¡Por Odín, por Thor! ¡Usa tu cerebro!

Capítulo II

Instrucciones específicas del día

- No hay código en el ámbito global. ¡Utilice funciones!
- Cada programa debe tener su main y no ser un simple script:

```
definición principal():  
    # sus pruebas y su manejo de errores  
  
si __nombre__ == "__principal__":  
    principal()
```


- Cualquier excepción no detectada invalidará los ejercicios, incluso en el caso de un error que se le solicitó probar.
- Puedes utilizar cualquier función incorporada si no está prohibida en el ejercicio.
- Todas tus funciones deben tener una documentación (__doc__)
- Su código debe estar en la norma
 - pip instala flake8
 - alias norminette=flake8

Para este módulo, utilizaremos datos de MATERIALES ESCOLARES GRATUITOS DE GAP-MINDER.ORG, LICENCIA CC-BY.

Te animamos a que eches un vistazo a los datos disponibles si quieres formarte en la manipulación de datos o realizar visión de datos.

Capítulo III

Ejercicio 00

	Ejercicio 00
Ejercicio 00: Cargar mi conjunto de datos	
Directorio de entregas: ex00/	
Archivos a entregar: cargar_csv.py	
Funciones permitidas: Pandas o cualquier biblioteca para la manipulación de conjuntos de datos.	

Crea una función que tome una ruta como argumento, escriba las dimensiones del conjunto de datos y lo devuelva. Debes manejar los casos de error y devolver None si la ruta es incorrecta, el formato es incorrecto...

```
def carga(ruta:cadena)->Conjunto de datos:(Tienes que adaptar el tipo de devolver(según tu biblioteca)
    # tu código aquí
```

Su probador de scripts:

```
desde load_csv importar carga
imprimir(cargar("años_de_esperanza_de_vida.csv"))
```


```
> probador de python.py
Cargando conjunto de datos de dimensiones (195, 302)
País 1800 1801 1802 1803 ... 2096 2097 2098 2099 2100 Afganistán 28,2 28,2
28,2 28,2 ... 76,2 76,4 76,5 76,6 76,8
...
$>
```



Puede mostrar el conjunto de datos en cualquier formato que desee, el formato proporcionado no es restrictivo.

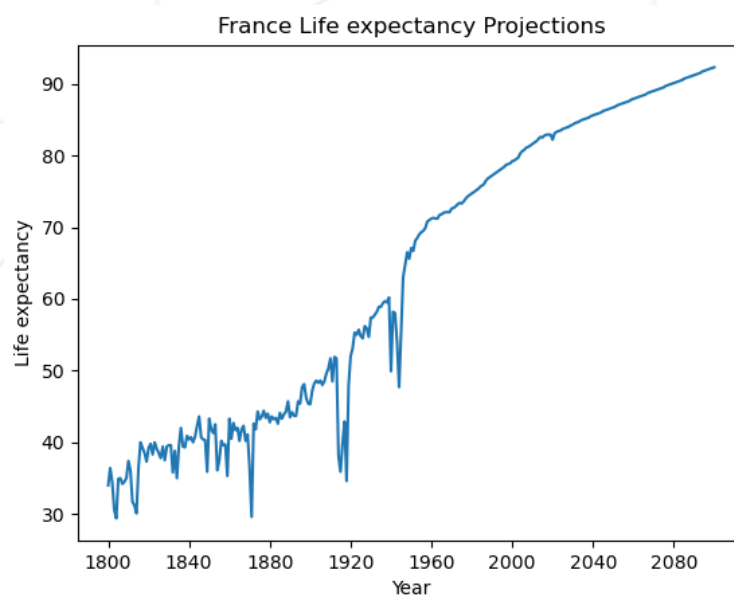
Capítulo IV

Ejercicio 01

	Ejercicio 01
Ejercicio 01: dibuja mi país	
Directorio de entregas: <i>ex01/</i>	
Archivos a entregar: <i>cargar_csv.py</i> , <i>aff_life.py</i>	
Funciones permitidas: <i>matplotlib</i> , <i>seaborn</i> o cualquier biblioteca para visualización de datos	


Crea un programa que llame a la función de carga del ejercicio anterior, cargue el archivo `life_expectancy_years.csv` y muestre la información del país de tu campus. Tu gráfico debe tener un título y una leyenda para cada eje.

Para ejemplo, Para los 42 campus de Francia tendremos este resultado.



Capítulo V

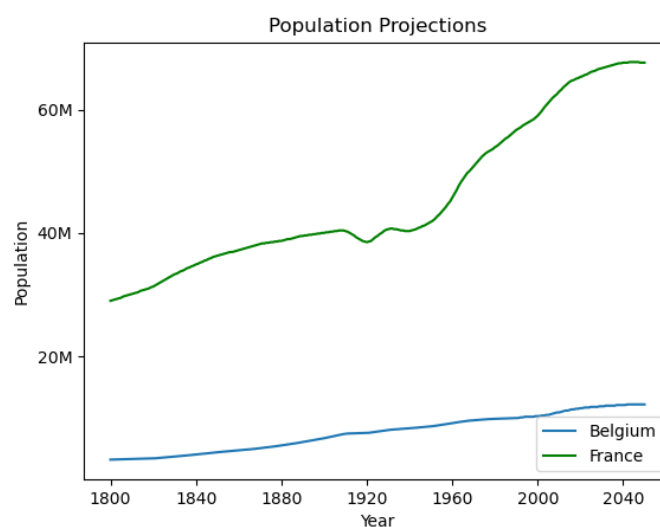
Ejercicio 02

	Ejercicio 02
Ejercicio 02: compara mi país	
Directorio de entregas: <i>ex02/</i>	
Archivos a entregar: <i>cargar_csv.py</i> , <i>aff_pop.py</i>	
Funciones permitidas: <i>matplotlib</i> , <i>seaborn</i> o cualquier biblioteca para visualización de datos	

Cree un programa que llame a la función de carga del primer ejercicio, cargue el archivo `population_total.csv` y muestre la información del país de su campus en comparación con otros países de Tu elección. Su gráfico debe tener un título, una leyenda para cada eje y una leyenda para cada gráfico.


Debes mostrar los años del 1800 al 2050.

Por ejemplo, Para los 42 campus de Francia tendremos este resultado.



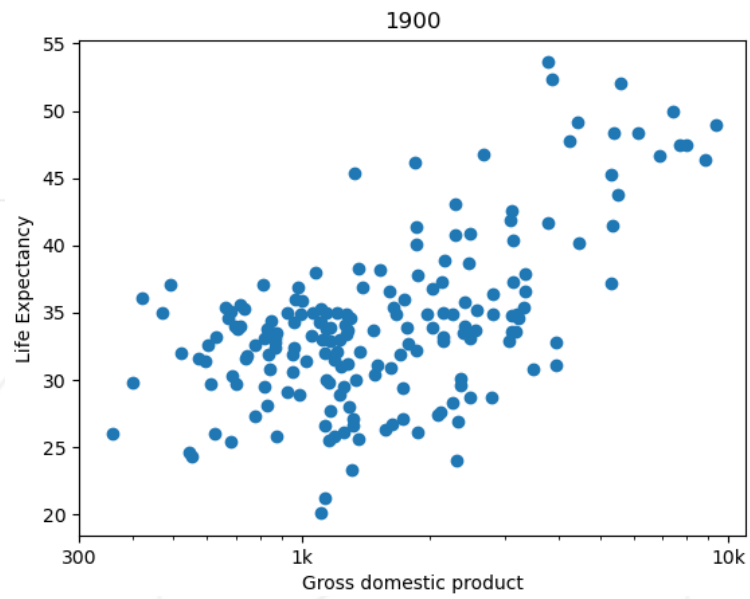
Capítulo VI

Ejercicio 03

	Ejercicio 03
Ejercicio 03: dibuja mi año	
Directorio de entregas: <i>ex03/</i>	
Archivos a entregar: <i>load_csv.py</i> , <i>proyección_life.py</i>	
Funciones permitidas: <i>matplotlib</i> , <i>seaborn</i> o cualquier biblioteca para visualización de datos y su biblioteca de <i>ex00</i>	

Cree un programa que llame a la función de carga del primer ejercicio, cargue los archivos "income_per_person_gdppercapita_ppp_inflation_adjusted.csv" y "life_expectancy_years.csv", y muestre la proyección de la esperanza de vida en relación al producto nacional bruto del año 1900 para cada país.

Tu gráfico debe tener un título, una leyenda para cada eje y una leyenda para cada gráfico. Debes mostrar el año 1900.



¿Ves una correlación entre la esperanza de vida y el producto interno bruto?

Capítulo VII

Presentación y evaluación por pares

Entregue su tarea en suGitRepositorio como de costumbre. Solo el trabajo dentro de su repositorio será evaluado durante la defensa. No dude en volver a verificar los nombres de sus carpetas y archivos para asegurarse de que sean correctos.



El proceso de evaluación se realizará en el computador del grupo evaluado.