



Examen Extraordinario (Práctico) Febrero 2022

INSTRUCCIONES

- Utiliza tu computadora con Spyder para resolver los ejercicios.
- Puedes utilizar todo el material de consulta que necesites, apuntes, libros, hojas, códigos de ejercicios anteriores, etc.
- No está permitido utilizar la conexión a internet con ningún dispositivo electrónico (reloj, laptop, celular, computadora, etc.) para consultas.
- Solo puedes utilizar los módulos Pandas, Pyplot y Math. No puedes utilizar ningún otro módulo (**import**)
- Cada ejercicio vale 10 puntos.
- Guarda el código de cada ejercicio como **ejercicio_1.py, ejercicio_2.py, ejercicio_3.py, ejercicio_4.py, ejercicio_5.py**
- Cuando termines de resolver los problemas, comprime los cinco archivos de los ejercicios en una carpeta **.zip** o **.rar** y nombra la carpeta comprimida de la siguiente manera: **apellido1_apellido2_nombre1_nombre2_nombre3.zip**, todo en **minúsculas y sin acentos**, por ejemplo: **gomez_carranza_juan_carlos.zip**
- Envía el archivo comprimido al correo jc.gomez@ugto.mx

EJERCICIOS

- Escribe un programa en Python que lea un archivo que contenga una lista de países y de ciudades que están en esos países. Después, que pida una ciudad al usuario y devuelva en qué país se encuentra esa ciudad, en caso de que la ciudad no se encuentre, que devuelva 'No sé'. En el archivo, una línea representa un país y sus ciudades, la primera palabra de la línea es el país y el resto de las palabras son las ciudades, todas separadas por espacios. Crear el archivo para probar en formato .txt. Transformar los nombres a minúsculas para comparar.

Ejemplo:

Archivo: US Boston Pittsburgh Washington Seattle
UK London Edinburgh Cardiff Belfast
México Guanajuato Querétaro CDMX Toluca Torreón
Brasil Río Brasilia SaoPaulo Recife PortoAlegre

Dado: `ciudad = 'Guanajuato'`

Salida: Mexico

Explicación: Hay cuatro países: US, UK, Mexico y Brasil, cada uno con una lista de ciudades que se encuentran en él. Guanajuato es una ciudad que se encuentra en Mexico.



2. Escribe un programa en Python que lea un archivo de texto, lo separe en palabras, encuentre aquellas que terminen en 'ando' o 'iendo' y elimine esos segmentos de las palabras, pero solo cuando la primera letra de la palabra esté en mayúscula. La salida se debe guardar en otro archivo, guardando tanto las palabras no modificadas como las sí modificadas. Utiliza la expresión regular `\w+` para capturar las palabras. Solo se deben modificar las palabras que contengan exclusivamente letras (no números o caracteres especiales).

Ejemplo:

Archivo entrada: hola estoy corriendo Saltando y Cantando Corr1iendo

Archivo salida: hola estoy corriendo Salt y Cantando corr1iendo

Explicación: Hay cuatro palabras que terminan en 'ando' o 'iendo' (corriendo, Saltando, Cantando, Corr1iendo), de éstas solo 3 empiezan con mayúsculas (Saltando, Cantando, Corr1iendo), finalmente, solo 2 están compuestas por solo letras (Saltando, Cantando) que son las que se modifican

3. Escribe un programa en Python que dado un número entero $n > 3$, cree una lista de listas de $n \times n$, llena con números enteros al azar entre 1 y 100, y que después imprima todos los bloques continuos de 3×3 junto con el promedio de cada bloque.

Ejemplo:

Dado: $n = 4$

Se genera al azar la siguiente lista de listas (como si fuera un arreglo de 4×4):

```
m = [[1, 2, 3, 4],
      [5, 6, 7, 8],
      [9, 10, 11, 12],
      [13, 14, 15, 16]]
```

Salida: `[[1,2,3],[5,6,7],[9,10,11]], 6`
`[[2,3,4],[6,7,8],[10,11,12]], 7`
`[[5,6,7],[9,10,11],[13,14,15]], 10`
`[[6,7,8],[10,11,12],[14,15,16]], 11`

Explicación: A continuación, se muestran los bloques de 3×3 que se generan:

<code>[[1, 2, 3, 4],</code>	<code>[[1, 2, 3, 4],</code>	<code>[[1, 2, 3, 4],</code>	<code>[[1, 2, 3, 4],</code>
<code>[5, 6, 7, 8],</code>	<code>[5, 6, 7, 8],</code>	<code>[5, 6, 7, 8],</code>	<code>[5, 6, 7, 8],</code>
<code>[9, 10, 11, 12],</code>	<code>[9, 10, 11, 12],</code>	<code>[9, 10, 11, 12],</code>	<code>[9, 10, 11, 12],</code>
<code>[13, 14, 15, 16]]</code>	<code>[13, 14, 15, 16]]</code>	<code>[13, 14, 15, 16]]</code>	<code>[13, 14, 15, 16]]</code>

4. El archivo adjunto `titanic.csv` contine las siguientes variables de interés (PassengerId, Survived, Pclass, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin, Embarked). Escribe un programa en Python que lea el archivo con Pandas y calcule lo siguiente:
- El resumen de los cinco números, la media, la desviación estándar y dibuje el diagrama de caja para la variable Age (edad) de forma independiente para cada



-
- valor de la variable Sex (sexo: female, male). Ignorar los valores que sean 0 en la variable Age.
- b. La correlación entre las variables Age (edad) y Fare (tarifa) de forma independiente para cada sexo. Debes separar los datos de Age y Fare por sexo y para cada sexo calcular la correlación entre esas dos variables.
5. Utilizando el mismo archivo titanic.csv, escribe un programa en Python que calcule la probabilidad condicional de morir dado que el pasajero viajó en 1a, 2a o 3a clase, utilizando las variables Pclass (1, 2, 3) and Survived (0, 1). La probabilidad condicional se calcula como $P(A|B) = \frac{P(A \cap B)}{P(B)}$. En donde $P(A|B)$ es la probabilidad que A suceda (morir) dado B (ser hombre/mujer), $P(A \cap B)$ es la probabilidad que A (morir) y B (ser hombre/mujer) ocurran juntos, y $P(B)$ es la probabilidad de que B (ser hombre/mujer) ocurra. Las probabilidades se calculan dividiendo la frecuencia de un evento entre el total de eventos.

Ejemplo:

Dado: muere, clase 2
muere, clase 2
muere, clase 2
sobrevive, clase 2
muere, clase 1
sobrevive, clase 1
sobrevive, clase 1

B = clase 2

Salida (para morir dado que es clase 2): $P(A \cap B) = 3/7$, $P(B) = 4/7$

$$P(A|B) = \frac{3/7}{4/7} = \frac{21}{28} = 0.75$$

Explicación: Hay 7 pares de datos en total. Hay 3 casos en que ser clase 2 y muere aparecen juntos por lo que $P(A \cap B) = 3/7$. Finalmente, hay 4 casos en que aparece 2 en la clase por lo que $P(B) = 4/7$