

Profesor: Dr. Oldemar Rodríguez Rojas
PF-1319 y PF-1320 Análisis de Datos II
Fecha de Entrega: Domingo 28 de agosto a las 12 media noche
Instrucciones:

- La solución a cada tarea se debe subir en el aula virtual, no pueden ser enviadas por correo.
- Las tareas son estrictamente individuales.
- Tareas idénticas se les asignará cero puntos.
- Todas las tareas tienen el mismo valor en la nota final del curso.
- Las tareas se pueden entregar tarde, pero cada día de atraso tendrá un rebajo de 20 puntos.

TAREA NÚMERO 1

Utilizando Python resuelva los siguientes ejercicios:

1. [5 puntos] ¿Cuál es el resultado?

a) $s = |(44^3 - 7!) \cdot (25/88)|$

b) $s = \sqrt{99} \cdot \pi^2$

c) $s = \log_2(38)$

d) $s = \sin(e)^{0.5}$

2. Dada una cadena de texto indique si la cantidad de caracteres en la cadena es un número par o impar. Realice una prueba para cada uno de los siguientes valores: “abracadabra”, “casa”, “sol”.

3. Dado un valor numérico indique a que día de la semana pertenece, siendo 1 = domingo y 7 = sábado. En caso que se digite un número fuera del rango imprima un mensaje indicando el error. Realice una prueba para cada uno de los siguientes valores: 4, 2, 9..

4. Suponga que tienen los valores a, b y c de una ecuación cuadrática $ax^2 + bx + c = 0$. Determine si la ecuación es degenerada ($a = 0$) o en caso contrario determine cuántas soluciones tiene (si el discriminante es mayor a cero tiene 2 soluciones, si es igual a cero tiene 1 solución, y si es menor a cero tiene 0 soluciones reales). Debe imprimir la cantidad de soluciones. Realice una prueba para cada uno de los siguientes casos: ($a = 9, b = 0, c = 2$), ($a = 0, b = -2, c = 3$), ($a = 2, b = -4, c = 2$). Recuerde que el discriminante se calcula como sigue: $\Delta = -b^2 + 4ac$ y por lo tanto las soluciones de la ecuación son: $x_1 = \frac{-b + \sqrt{-b^2 + 4ac}}{2a}$ y $x_2 = \frac{-b - \sqrt{-b^2 + 4ac}}{2a}$, o usando el discriminante: $x_1 = \frac{-b + \sqrt{\Delta}}{2a}$ y $x_2 = \frac{-b - \sqrt{\Delta}}{2a}$.

5. Dado $x = (15, 34, 72, 23, 91, 4, 201, 68, 56, 78)$ realice lo siguiente:

- Calcule la media, la varianza y la desviación estándar.
- Extraiga los primeros 3 valores.

- Indique el valor más pequeño del vector.
- Obtenga la sumatoria de todos los valores del vector.
- Obtenga la lista x invertida.

6. Realice la siguiente operación entre matrices:

$$A = \begin{pmatrix} -17 & 11 & 61 \\ 4 & 29 & 33 \end{pmatrix} - 12 \cdot \begin{pmatrix} -8 & 6 \\ -15 & 25 \\ 5 & -13 \end{pmatrix}^t$$

7. Dada la matriz cuadrada A que se presenta abajo, calcule, la suma de los elementos que conforman la diagonal. Es decir, $6 + (-8) + 4 = 2$.

$$A = \begin{pmatrix} 6 & 7 & -5 \\ 1 & -8 & -6 \\ 10 & 13 & 4 \end{pmatrix}$$

8. Genere, sin utilizar archivos, un **DataFrame** de la siguiente tabla de datos:

Nombre	Matematicas	Ciencias	Español	Historia	EdFisica	Genero
Lucia	7.0	6.5	9.2	8.6	8.0	F
Pedro	7.5	9.4	7.3	7.0	7.0	M
Ines	7.6	9.2	8.0	8.0	7.5	F
Luis	5.0	6.5	6.5	7.0	9.0	M
Andres	6.0	6.0	7.8	8.9	7.3	M
Ana	7.8	9.6	7.7	8.0	6.5	F

9. Utilizando la tabla creada en el punto anterior realice lo siguiente:

- Ejecute un `info()` de los datos.
- Muestre el resumen estadístico básico, uno incluyendo solo los tipos “numéricos” y otro excluyendo los tipos numéricos.
- Despliegue las primeras 3 columnas de la tabla de datos (Usando solamente `[]`).
- Despliegue las primeras 3 columnas de la tabla de datos (Usando solamente `iloc`).
- Despliegue las primeras 3 columnas de la tabla de datos (Usando solamente `loc`).
- Calcule la correlación entre Historia y Matematicas con la función `corrcoef()` de la biblioteca `numpy`.
- Construya un diccionario llamado `resumen` que tenga 4 campos `Media`, `Mediana`, `Máximo` y `Mínimo` que tienen la media, la mediana, el máximo y el mínimo respectivamente de la variable `Ciencias`.

10. Cargue la tabla de datos que está en el archivo `students.csv` haga lo siguiente:

- Calcule la dimensión de la Tabla de Datos.

- Calcule el resumen numérico de la tabla.
 - Calcule la suma de las columnas con variables cuantitativas (numéricas).
 - Calcule la moda de las columnas con variables cualitativas (categóricas).
11. Usando `for(...)` en **Python** muestre los números del 1 al 100 que terminan en 8, debe mostrarlos en orden inverso, es decir, de mayor a menor.
 12. Mediante un ciclo, calcule la sumatoria de los números enteros múltiplos de 17, comprendidos entre el 1 y el 300.
 13. Mediante un ciclo, guarde en una lista todos los números pares desde 10 hasta el 25.
 14. Programe una función que recibe tres objetos A , B , y C , si el tipo de dichos objetos es `str` retorna la unión de dichos caracteres separados por un espacio. En caso de que alguno de los tres objetos no sea `str` retorna `None`, use `isinstance` para determinar si el objeto es de tipo `str`.
 15. Programe una función que reciba un número y retorne `True` en caso de ser un número primo, de lo contrario retorna `False`. La definición de «número primo» dice que «Un número entero mayor que 1 se denomina número primo si y sólo si tiene como divisores positivos (factores) únicamente a sí mismo y a la unidad 1». Por ejemplo, son números primos: 2, 3, 5, 7, 11, 13, 17.
 16. Programe una función que recibe una lista numérica y retorna el número primo más pequeño que está en la lista. Puede utilizar la función creada en el punto anterior.
 17. Programe una función que reciba un objeto C tipo `str` (cadena de caracteres) y retorne la cantidad de vocales que posee la cadena de caracteres C . Por ejemplo: “Hola Mundo” \rightarrow 4.
 18. Programe una función que reciba un objeto C tipo `str` (cadena de caracteres) y una letra l y retorne la cantidad de veces que aparece la letra dentro de la cadena de caracteres C . Por ejemplo: “abracadabra” y “a” \rightarrow 5.
 19. Programe en **Python** una función que recibe tres valores A , B , y C y retorna el menor.
 20. Programe en **Python** una función que recibe un número n y retorna la sumatoria de los números enteros comprendidos entre el 1 y el n .
 21. Programe una función en **Python** que recibe un número n y realice la sumatoria de los números enteros múltiplos de 5, comprendidos entre el 1 y el n .
 22. Programe en **Python** una función que genera 50 números al azar entre 1 y 400. Luego de esos 50 números la función calcula y retorna qué porcentaje son pares.
 23. Programe en **Python** una función que genera 100 números al azar entre 1 y 500 y luego calcula cuántos están entre el 50 y 450, ambos inclusive.
 24. Desarrolle una función en **Python** que calcula el costo de una llamada telefónica que ha durado t minutos sabiendo que si $t < 1$ el costo es de 0.4 dólares, mientras que para duraciones superiores el costo es de $0.4 + (t - 1)/4$ dólares, la función debe recibir el valor de t .

25. Desarrolle una función en **Python** que reciba un vector de números reales y un número real x , tal que retorne el porcentaje de elementos menores o iguales a un valor x .
26. Desarrolle una función en **Python** que recibe 4 números (x_1, x_2, y_1, y_2) . La función debe calcular la distancia entre dos puntos y mostrar mediante texto el resultado, para ello use la fórmula:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$
27. Desarrolle una función que recibe la cantidad de entradas que una persona desea comprar para un espectáculo y el precio de las mismas (todas valen igual), luego debe calcular el pago a realizar por la(s) entrada(s) tomando en cuenta que se pueden comprar sólo hasta cuatro entradas, que al costo de dos entradas se les descuenta el 15 %, que al de tres entrada el 20 % y que a la compra de cuatro entradas se le descuenta el 25 %. Realice además una función de validación que evite que una persona compre más de cuatro entradas.
28. Desarrolle una función que reciba un número natural n (suponiendo que $n > 1$) y que construya y retorne un vector v de tamaño n tal que $v_k = v_{k-1}/3 + 0.5$ para $k = 1, \dots, n$ y siendo que $v_0 = 1$.
29. Desarrolle una función que construye y retorna una matriz A de tamaño $m \times n$ cuya entrada générica es $i^2 - j$, es decir $a_{ij} = i^2 - j$.
30. Desarrolle una función que recibe una matriz cuadrada A de tamaño $n \times n$ y calcula su traza (utilizando **for**), es decir, la suma de los elementos de la diagonal. Por ejemplo, la traza de la siguiente matriz:

$$\begin{pmatrix} 8 & 3 & 24 \\ 12 & 13 & -11 \\ 14 & 12 & -6 \end{pmatrix}$$

es 15.

31. Desarrolle una función que reciba dos números enteros a y b , tal que, en un diccionario retorne el Máximo Común Divisor (**MCD**) y el mínimo común múltiplo (**mcm**). Para calcular el **MCD** puede utilizar la función **gcd** del paquete **math**. La fórmula para calcular el **mcm** es la siguiente:

$$\text{mcm}(a, b) = \frac{a \cdot b}{\text{MCD}(a, b)}.$$

32. Desarrolle una función que recibe una matriz cuadrada A de tamaño $n \times n$ y retorna su transpuesta (utilizando **for**). Por ejemplo, la transpuesta de la matriz A :

$$A = \begin{pmatrix} 46 & 30 & 6 \\ 4 & 2 & 54 \\ -7 & -5 & -11 \end{pmatrix}$$

es la siguiente matriz A^t :

$$A^t = \begin{pmatrix} 46 & 4 & -7 \\ 30 & 2 & -5 \\ 6 & 54 & -11 \end{pmatrix}$$

33. Desarrolle una función en **Python** que recibe una lista de edades y retorna que porcentaje son mayores de 18 años y que porcentaje tienen 18 o menos años.
34. Desarrolle una función que recibe un **DataFrame** y un nombre de columna y que retorne un diccionario con los valores de la media, mínimo, máximo y varianza de dicha columna del dataframe. Verifique la correctitud de esta función usando la tabla `celulares.csv` y la columna `battery_power`.
35. Desarrolle una función que recibe un **DataFrame** y dos números de columna y que retorna la correlación entre esas dos variables. Verifique la correctitud de esta función usando la tabla `celulares.csv` y las columnas 5 y 6, que hacen referencia a las variables `four_g` y `int_memory`.

Entregables: Debe subir en el Aula Virtual el script de pruebas y un documento autoreproducible con la solución de la tarea.

