

Instrucciones: Escriba el código en Python 3 que funcione como describe el enunciado. Entregue su solución en un cuaderno de examen u hojas grapadas. Firme al entregar el examen. No se permite el uso de dispositivos electrónico. Tiene permitido utilizar material escrito o impreso de temas respectivos al curso. Si realiza el examen en lápiz, pierde el derecho a reclamos.

Puntaje: 100 puntos

Nota: A menos que esté escrito explícitamente, no será necesario realizar validación de datos de entrada. Suponga que los datos que se ingresan siempre son válidos.

Enunciado

Una empresa productora de fármacos está en proceso de mejorar un medicamento que disminuye la duración del resfriado común. Ya han realizado varios ensayos de laboratorio, pero ahora desean realizar una prueba con personas. Por esto, han conducido un experimento con varios pacientes voluntarios, y necesitan ayuda para analizar los datos de este estudio empírico.

El experimento fue llevado a cabo en pacientes con síntomas del resfriado. Cada voluntario fue asignado al azar a uno de dos grupos: el grupo Control o el grupo Tratamiento. Los miembros del grupo Tratamiento deberían tomar el nuevo medicamento; mientras que los del grupo Control deberían tomar un medicamento disponible en el mercado. Para cada participante se midió la cantidad de días que estuvo resfriado.

Los resultados de este experimento se guardaron en un archivo llamado `farmaco.csv`. Cada fila de este archivo corresponde a un paciente voluntario que participó en el experimento. Para cada paciente, se guardó información correspondiente a cuál de los 2 grupos fue asignado (Tratamiento o Control), su género (F o M), y la cantidad de días que reportó estar resfriado. El siguiente es un ejemplo de cómo se vería este archivo:

grupo,genero,dias		
tratamiento,M,4.885	control,F,6.662	tratamiento,M,7.097
control,F,6.583	control,M,6.322	tratamiento,M,7.385
tratamiento,F,9.454	tratamiento,M,7.927	tratamiento,M,4.728
tratamiento,M,6.646	control,F,6.415	control,F,6.843
control,F,7.83	tratamiento,F,5.985	tratamiento,M,7.841
tratamiento,M,6.738	tratamiento,M,5.332	control,M,5.879
control,F,7.431	tratamiento,F,8.09	tratamiento,F,6.043
tratamiento,M,5.499	tratamiento,F,4.544	control,M,6.598
tratamiento,F,6.134	control,M,6.123	tratamiento,M,8.559
control,F,7.922	control,F,7.443	control,M,6.597
tratamiento,F,5.925	control,M,7.253	tratamiento,F,7.008
control,F,7.057	tratamiento,F,8.925	control,F,8.59
control,F,5.437	control,F,6.595	tratamiento,M,7.955
control,M,6.613	control,F,8.483	
	tratamiento,M,5.44	

Para determinar si el medicamento tiene un efecto que influye la cantidad de días del resfriado, se debe comparar las estadísticas de ambos grupos. Las estadísticas a comparar son la media y la desviación estándar. La media

(μ) de un grupo es simplemente el promedio de los días de malestar de los pacientes que pertenecen a ese grupo. Este promedio se calcula como la suma de todos los elementos, dividido entre la cantidad de miembros del grupo. La desviación estándar (σ) simplemente es una medida de separación de la media. Esta se puede calcular con la siguiente fórmula:

$$\sigma = \frac{1}{N-1} \sqrt{\sum_{i=1}^N (x_i - \mu)^2}$$

donde x_i es la cantidad de días de cada uno de los N pacientes que pertenecen al grupo.

Por último, para determinar si el efecto del medicamento produce una «diferencia significativa» entre los 2 grupos de pacientes, se compara la diferencia entre las medias de los 2 grupos. Para eso, se calcula un valor de la diferencia verdadera entre ambos grupos, llamado Z. Sea grupo 1 el grupo Tratamiento y el grupo 2 el Control, el valor Z se calcula con la siguiente fórmula:

$$Z = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

Además de mostrar el valor Z, se deberá definir si esta diferencia es suficiente para que el medicamento tenga efecto. Si se cumple que Z es mayor a 1.96, o menor a -1.96, significa que el medicamento es lo suficientemente diferente al disponible actualmente. Si por lo contrario, el valor Z está entre -1.96 y 1.96, el medicamento tiene el mismo efecto que el actualmente disponible.

La empresa desea saber entonces, para cada uno de los grupos, su media, desviación estándar, y cantidad de voluntarios. Además, deberá mostrar el valor Z de los grupos, y si la este valor causa que las medias de los 2 grupos sea «significativa» o «no significativa». El siguiente es un ejemplo de cómo debería ser la salida del programa:

Grupo	N	Prom	Desv

Control	20	6.934	0.189
Tratamiento	22	6.734	0.307

Valor Z: 2.570

Estadísticamente hay diferencia

Finalmente, el programa debe estar programado en el paradigma orientado a objetos. Se utilizará un diseño de 2 clases: *Grupo* y *Experimento*. *Grupo* debe encargarse de almacenar la información correspondiente de la cantidad de días de enfermedad para miembro del grupo. *Experimento* está encargada de almacenar las 2 instancias de Grupo, leer la información del archivo, y mostrar los resultados.

Evaluación

1. **(45%)** Implementa la clase `Grupo`
 - a. **(10%)** El grupo tiene atributos que le permiten almacenar los días de enfermedad de cada paciente.
 - b. **(35%)** El grupo implementa los siguientes métodos:
 - i. **(5%)** Un constructor para inicializar los atributos de la clase.
 - ii. **(3%)** `agregar(d)` que agrega la nueva observación de día d al grupo
 - iii. **(2%)** `cantidad()` que retorna la cantidad de participantes de cada grupo.
 - iv. **(10%)** `media()` que calcula y retorna la media de días del grupo.
 - v. **(10%)** `desviacion()` que calcula y retorna la desviación de días del grupo.
 - vi. **(5%)** `mostrar()` que muestra el nombre, la cantidad de participantes, media y desviación del grupo.
2. **(55%)** Implementa la clase `Experimento`
 - a. **(5%)** El experimento tiene un atributo que le permite representar los 2 grupos.
 - b. **(50%)** El experimento implementa los siguientes métodos:
 - i. **(5%)** Un constructor para inicializar los atributos de la clase.
 - ii. **(30%)** `leer_datos()` que abre el archivo `farmaco.csv`, lee la información de este, y agrega la cantidad de días de cada paciente al grupo correspondiente.
 - iii. **(15%)** `mostrar()` que muestra, para las estadísticas de cada grupo, el valor Z , y si existe o no una diferencia significativa entre los grupos.

Nota: En cada punto se evaluará la correctitud, completitud, y buenas prácticas de programación. Estas incluyen, pero no se limitan a identificadores significativos para variables e indentación adecuada.

Nota: El diseño mostrado es una sugerencia. Puede utilizar otro diseño de clases, métodos y atributos si lo considera adecuado; siempre y cuando implemente la funcionalidad descrita.