



Proyecto Integrador I

2021-1 Taller Experimentos

Nombre: Sebastian Villa Ávila Jhon Sebastian Ijaji Ortiz Alejandro Arce Rendón Alexander Samacá Burbano	Código: A00361589 A00362423 A00358633 A00362108
Profesor: Juan Manuel Reyes	

Diseño e Implementación del Experimento

Planeación y Realización

Se ha solicitado desarrollar un experimento que evalúe el tiempo de ejecución de un programa encargado del ordenamiento de arreglos, entendiendo que un experimento es “ un cambio en las condiciones de operación de un sistema o proceso, que se hace con el objetivo de medir el efecto del cambio en una o varias propiedades del producto o resultado” (2008, Gutierrez, p. 7) Partiendo del anterior concepto se plantea que la unidad experimental evaluada será un arreglo de datos primitivos de tamaño finito. Además, se identifican los siguientes factores y se clasifican en controlables y no controlables, obteniendo la siguiente lista:

Factores Controlables

- Algoritmo de Ordenamiento (Burbuja, Selection, QuickSort, HeapSort, etc)
- Tamaño del arreglo (10^1 , 10^2 , 10^3 , 10^4 , etc)
- Estado de los valores en el arreglo (en orden aleatorio, ordenado ascendente, ordenado descendente)
- Lenguaje de programación (Java, C#, Python, etc)
- Cantidad de procesos que se están ejecutando en el computador mientras se ejecuta el algoritmo

Factores no Controlables

- RAM del computador donde se ejecuta el algoritmo (2GB, 4GB, 8GB, 16GB, etc)
- Procesador del computador donde se ejecuta el algoritmo
- Sistema Operativo
- Nivel de fragmentación del disco duro del computador donde se ejecuta el algoritmo
- Tamaño del registro del procesador (8bits, 16bits, 32bits, 64bits)

Por otro lado, se identificaron los factores de estudio, los cuales son Algoritmo de ordenamiento, el tamaño de arreglo, el estado de los valores y Cantidad de procesos que se están ejecutando . A continuación, se identifican los niveles y tratamientos en la siguiente tabla:

Nivel de Ordenamiento	Nivel de Estado de los datos	Nivel de procesos ejecutados	Nivel de tamaño de arreglo	Tratamiento
1	1	1	1	1
1	1	1	2	2
1	1	1	3	3
2	1	1	1	4
2	1	1	2	5
2	1	1	3	6
1	2	1	1	7
1	2	1	2	8
1	2	1	3	9
2	2	1	1	10
2	2	1	2	11
2	2	1	3	12
1	1	2	1	13
1	1	2	2	14
1	1	2	3	15
2	1	2	1	16
2	1	2	2	17
2	1	2	3	18
1	2	2	1	19
1	2	2	2	20
1	2	2	3	21
2	2	2	1	22
2	2	2	2	23
2	2	2	3	24

Por último, se define que la variable de respuesta va a ser una variable cuantitativa denominada tiempo de ejecución. Asimismo, con el objetivo claro y los factores identificados se procederá a implementar el experimento para obtener las variables de respuesta y llegar a conclusiones pertinentes sobre el tema.

Análisis

Para realizar la fase de análisis se debe tener en cuenta que se cuenta con un total de 24 tratamientos, el programa diseñado realiza el experimento para 12 tratamientos posibles ya que el nivel de procesos en ejecución se realiza manualmente. Es decir, en un primer momento se ejecuta el programa cuando el equipo esté corriendo únicamente el proceso de compilación del programa y se recolectan los datos y las variables de respuesta, seguidamente se ejecutan otros procesos y vuelve y se corre el programa obteniendo nuevamente los datos y variables de respuesta.

Una vez obtenidos los 2400 datos se realiza una clasificación, se obtiene la media aritmética de cada tratamiento y se realiza un análisis de ANOVA para llegar a la interpretaciones y conclusiones realizadas en las siguientes etapas.

Este análisis se encuentra adjunto en un archivo de Excel en la carpeta docs

Interpretación

Para realizar el análisis de Anova se dividieron los tratamientos en tres grupos diferentes con base al nivel de tamaño del arreglo (10^1 , 10^2 , 10^3) obteniendo la siguiente división:

grupo 1:

Tratamiento	Media
Tratamiento #1	0,012335
Tratamiento #4	0,015629
Tratamiento #7	0,028731
Tratamiento #10	0,013331
Tratamiento #13	0,013638
Tratamiento #16	0,013308
Tratamiento #19	0,020911
Tratamiento #22	0,01196
Menor	0,01196
Mayor	0,028731

grupo 2:

Tratamiento	Media
Tratamiento #2	0,059515
Tratamiento #5	0,060734
Tratamiento #8	0,077107
Tratamiento #11	0,057826
Tratamiento #14	0,082469
Tratamiento #17	0,061447
Tratamiento #20	0,072939
Tratamiento #23	0,057742
Menor	0,057742
Mayor	0,082469

grupo 3:

Tratamiento	Media
Tratamiento #3	0,590891
Tratamiento #6	3,531977
Tratamiento #9	0,701925
Tratamiento #12	2,528497
Tratamiento #15	0,61303
Tratamiento #18	3,595354
Tratamiento #21	0,717256
Tratamiento #24	2,423142
Menor	0,590891
Mayor	3,595354

Cada una de las tabla muestra la media de los tiempos que se tardaron en cada uno de los tratamientos y a la vez en cada grupo cuál fue el menor tiempo y el mayor tiempo.

Una vez realizado el análisis se llegaron a las siguientes conclusiones de cada grupo

Grupo 1:

- Dado que 1,15 es menor al valor f crítico 2,02 aceptamos H_0 por lo que decimos que no hay diferencias entre los tratamientos que utilizan un tamaño de arreglo de 10^1 , haciendo uso un nivel de significancia del 5%

Grupo 2:

- Dado que 10,07 es mayor al valor f crítico 1,94 rechazamos H_0 por lo que decimos que existe algún par de medias de tratamientos que son distintos que utilizan un tamaño de arreglo de 10^2 , haciendo uso un nivel de significancia del 5%

Grupo 3:

- Dado que 722,49 es mayor al valor f crítico 2,02 rechazamos H_0 por lo que decimos que existe algún par de medias de tratamientos que son distintos que utilizan un tamaño de arreglo de 10^3 , haciendo uso un nivel de significancia del 5%

Control y conclusiones finales

El diseño de experimentos es una técnica que puede ayudar a conocer un proceso. Permite averiguar la forma como diversos factores presentes en él influyen sobre la respuesta y ajustarlos en los niveles que optimicen los resultados.

El objetivo de este experimento es determinar cómo diversos factores influyen en el tiempo de ejecución de un programa encargado de ordenar arreglos. Por tanto, tras realizar el proceso de ejecución, recolección y análisis de datos se llega a las siguientes conclusiones.

En primer lugar, para ordenar un arreglo de tamaño diez la mejor opción es utilizar el tratamiento 22 que utiliza el insertion sort como método para ordenar datos, en un inicio los datos se encuentran ordenados ascendentemente y se están ejecutando 2 procesos. En segundo lugar para un arreglo de tamaño 100 la mejor opción es el tratamiento 23 que utiliza el insertion sort como método para ordenar datos, en un inicio los datos se encuentran ordenados ascendentemente y se están ejecutando 2 procesos. Por último, para un arreglo de tamaño 1000 la mejor opción es el tratamiento 3 que utiliza el merge sort como método para ordenar datos, en un inicio los datos se encuentran ordenados aleatoriamente y se está ejecutando 1 proceso.

En conclusión, es evidente cómo cada uno de los factores de estudio influye significativamente en las variables de respuesta. Es decir, cada factor que se tomó influyó en el tiempo que se tarda en ejecutar el programa, se probaron distintas combinaciones entre cada factor y se determinaron cuál era el más efectivo para cada caso.