

Algoritmos de ordenamiento, un análisis estadístico comparativo

Sorting algorithms, a comparative statistical analysis

Salgado Gallegos, Mireya^{1*}; Pérez Merlos, Juan Carlos²; Albarrán Trujillo, Silvia Edith¹

¹Licenciatura en Ingeniería en Computación, Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca-Estado de México.

²Licenciatura en Ingeniería en Electrónica, Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca-Estado de México.

*msalgadog@uaemex.mx

DOI: <https://doi.org/10.53766/CEI/2022.43.03.05>

Resumen

El impacto que las computadoras y la informática han tenido en todos los aspectos de la sociedad, la capacidad de desarrollar, analizar e implementar algoritmos está ganando más atención (Nasar 2019). En este contexto, el presente artículo tiene por objetivo realizar un análisis estadístico comparativo del tiempo de ordenación entre los algoritmos de ordenamiento de inserción, selección, burbuja simple y recursivo y quicksort. Para ordenar 1,000, 10,000 y 100,000 datos se realizaron 50 pruebas aplicando los algoritmos de ordenamiento, se aplicó el análisis ANOVA con diferentes pruebas y los valores de p-valor obtenidos fueron menores al valor de significancia de $\alpha=0.05$, el resultado permite concluir que estadísticamente sí existen diferencias significativas entre los tiempos de ordenación de los algoritmos ya mencionados siendo el algoritmo quicksort el más rápido en todas las pruebas.

Palabras clave: Algoritmos de ordenamiento, análisis ANOVA, análisis comparativo entre los algoritmos de ordenamiento

Abstract

The impact that computers and informatics have had on all aspects of society, the ability to develop, analyze and implement algorithms is gaining more attention (Nasar 2019). In this context, this article aims to carry out a comparative statistical analysis of the sorting time between the insertion, selection, simple and recursive bubble and quicksort algorithms. To order 1,000, 10,000 and 100,000 data, 50 tests were carried out applying the ordering algorithms, the ANOVA analysis was applied with different tests and the p-values obtained were less than the significance value of $\alpha=0.05$, it is possible to conclude that statistically, there are significant differences between the sorting times of the algorithms mentioned before, with the quicksort algorithm being the fastest in all tests.

Keywords: Sorting algorithms, ANOVA analysis, comparative analysis between sorting algorithms

1 Introducción

Está bastante claro que vivimos en la era del Big Data, y ahora, el gran problema es manejar esa gran cantidad de datos en la computadora y clasificarlos es una tarea esencial (Prajapati y col; 2017), de tal manera que más allá de la recopilación de todos estos, nos hemos vuelto dependientes de los algoritmos para ordenamiento de elementos con la finalidad de llegar a algo significativo (Nasar 2019), tratando de mejorar la eficiencia de búsqueda de los datos en la computadora (Prajapati y col; 2017).

El ordenamiento de elementos es un proceso que se encuentra muy a menudo en la vida cotidiana (Nasar 2019), su objetivo es hacer que el elemento sea más fácil de buscar, insertar y eliminar (Htwe Htwe 2019; Yang Yu y Gan 2011). La ordenación es un proceso de reorganización de una lista

de elementos en el orden correcto (Adhikari 2007; Buradagunta y col; 2020), de aquí que los algoritmos de ordenación son considerados como una parte importante de la gestión de datos (Adhikari 2007), en la informática y programación (Htwe Htwe 2019; Prajapati y col; 2017; Yang y col; 2011; Zhao y col; 2016) y en las ciencias computacionales (Buradagunta y col; 2020; Nasar 2019).

Un algoritmo de ordenación es un proceso computacional que facilita a los usuarios la clasificación de datos de forma rápida y automática considerando la velocidad en la clasificación de datos además de la eficiencia de la memoria (Iskandar y col; 2020).

La mayoría de los algoritmos de ordenación funcionan comparando los datos que se ordenan (Adhikari 2007) y cada algoritmo de ordenamiento es mejor en alguna situación, por ejemplo, el algoritmo de ordenamiento por inserción es preferible al de ordenación rápida cuando se tienen pocos

elementos (Chowdhury 2016), además la implementación del algoritmo (matriz, vector, bases de datos, etc.) también será muy diferente, lo cual puede hacer que el mismo algoritmo que es rápido en un caso sea lento en el otro (Chowdhury 2016).

Generalmente los tipos de ordenamiento se clasifican en dos: 1) interno y 2) externo. En el tipo interno, los datos se almacenan en la memoria durante la ordenación, mientras que en el externo los datos se almacenan fuera (por ejemplo, en el disco duro y se cargan en la memoria cuando es necesario). En el tipo 1) se tienen diferentes algoritmos de ordenación como pueden ser: selección, burbuja, inserción, fusión, rápida (quicksort), montículo, radix, shell, entre otros (Buradagunta y cols.). Este trabajo consideró los ordenamientos internos de inserción, selección, burbuja (simple y recursivo) y quicksort.

Con base en lo anterior, un gran número de informáticos han trabajado mucho en los algoritmos de ordenación usándolos con frecuencia en diferentes procesos (Htwe Htwe 2019) y realizando diferentes investigaciones con la implementación de estos.

Yang y cols. (2011), a través de la descripción de cinco algoritmos de clasificación: burbuja, selección, inserción, fusión y quicksort resumieron la complejidad de tiempo y espacio de cada uno de estos

En 2016, Zhao y col., realizaron un análisis comparativo de 6 algoritmos de clasificación a partir de la complejidad y estabilidad del tiempo algorítmico en entorno Java (Zhao y col; 2016). En el mismo año, Chowdhury (2016) en su estudio, determinó la eficiencia de los diversos algoritmos de ordenamiento basada en el tiempo y número de intercambios utilizando ensayos aleatorios, implementados en el lenguaje Java.

Prajapati y col. (2017), discutieron el rendimiento de diferentes algoritmos de clasificación con sus ventajas y desventajas basados en una comparación de su desempeño bajo diferentes parámetros (Prajapati y col; 2017). De la misma manera, Manaseer y Al Hwaitat (2018) estudiaron el rendimiento de dos algoritmos de clasificación paralelos (bubble y bucket) y los evaluaron en rendimiento considerando la velocidad y la eficiencia.

Khreisat (2018) proporciona un estudio empírico de tres variaciones del algoritmo de quicksort, considerando, de cada uno, la cantidad de comparaciones realizadas y los tiempos de ejecución cuando se utilizan para ordenar matrices de enteros generados aleatoriamente.

Recientemente en el 2019, Htwe realizó una comparación de diferentes algoritmos de ordenamiento en función de la eficiencia del tiempo de ejecución (Htwe Htwe 2019). Así también en el mismo año, Nasar (2019) proporciona una introducción a la complejidad y eficiencia computacional a través de un análisis matemático de algoritmos de clasificación generados por estudiantes. García y Velázquez (2019) analizan el rendimiento en tiempo, de los algoritmos de ordenamiento: burbuja, par impar, rango,

conteo, radix, fusión, bitonic y clasificación rápida; todos paralelizados y ejecutados en una instancia de la Máquina Virtual de Java (JVM).

Asimismo, Alif y col. (2019), implementaron tres algoritmos de ordenación: burbuja, selección e inserción utilizando el lenguaje Verilog HDL para determinar el algoritmo con mejor rendimiento de hardware y pueda ser usado como un bloque en cualquier sistema con paralelismo.

Kumar (2020) realizó un estudio empírico sobre las diferentes técnicas que implementan los algoritmos de ordenamiento como: dividir y vencer, disminuir y vencer, transformar y vencer, junto con la comparación de varios parámetros que deben tenerse en cuenta al desarrollar un algoritmo.

Buradagunta y col. (2020), realizaron un estudio comparativo entre números aleatorios positivos y negativos como experimentación aplicando diferentes algoritmos como UNH, selección, burbuja, inserción, mezcla, rápido.

Iskandar y col. (2020), realizaron una nueva prueba de dos métodos de ordenamiento: burbuja e inserción basada en la comparación de dos lenguajes de programación, Java y Visual Basic 2010 con el objetivo de descubrir qué algoritmo tiene un menor consumo de memoria en el proceso de clasificación utilizando Java o Visual Basic 2010.

En general, después de una ardua investigación documental, las diferentes aportaciones concluyen que entre los algoritmos existe una gran diferencia en el tiempo de ordenamiento de los elementos independientemente de otros factores.

Con base en lo anteriormente expuesto, este trabajo basa su interés en un análisis comparativo en tiempos de ejecución de los algoritmos de ordenamiento desde una perspectiva estadística ya que como se aprecia en párrafos anteriores, existen muchos estudios comparativos (sólo se mencionaron algunos) entre diferentes algoritmos de ordenamiento en los cuales concluyen que sí hay diferencias entre estos en los tiempos de ordenamiento y eficiencia pero ninguno realizado estadísticamente, objetivo en el que se basa esta investigación definiendo la hipótesis: Existen diferencias estadísticamente significativas en los tiempos de ejecución entre los algoritmos de ordenamiento.

2 Metodología

Para el desarrollo de este trabajo se siguió la siguiente metodología:

- Investigación documental
- Experimentación de los algoritmos de ordenamiento
- Análisis de los tiempos de ejecución del ordenamiento de elementos de los cinco algoritmos: inserción, selección, burbuja (simple y recursivo) y quicksort
- Realización de las pruebas de normalidad

- Realización de las pruebas de homogeneidad de varianzas
- Análisis ANOVA
- Interpretación de resultados
- Conclusiones

2.1 Desarrollo

Algunos autores consideran que, para medir el rendimiento de cada algoritmo, el factor más importante es el tiempo de ejecución que utiliza para ordenar un dato (Chowdhury 2016), razón por la cual esta investigación se basa en el factor de tiempo de ejecución del ordenamiento para realizar el análisis estadístico comparativo entre los cinco algoritmos antes mencionados.

Para la experimentación en la obtención de los tiempos de ejecución de cada uno de los algoritmos se consideró:

- Archivos con 1,000, 10,000 y 100,000 elementos a ordenar.
- Un grupo de 50 alumnos que ejecutaron los algoritmos de ordenamiento de inserción, selección, burbuja (simple y recursivo) y quicksort.
- Cada alumno ejecutó cada algoritmo con los 1,000, 10,000 y 100,000 datos desordenados. Cabe mencionar que con el algoritmo de burbuja recursivo no se realizó la prueba de los 100,000 elementos ya que el programa se ciclaba con la recursividad, lo cual no sucedía con los archivos de los 1,000 y 10,000 datos.
- Los algoritmos fueron todos codificados en Java y con funciones para determinar el tiempo de ejecución.
- Los tiempos de ordenamiento de los algoritmos fueron concentrados por cada participante en un documento compartido (Figura 1).

Fig. 1. Documento concentrador de tiempos de ordenamiento.

- Las características de las computadoras donde se ejecutaron los algoritmos fueron las siguientes:
 - Intel(R) Core (TM) i7-4770
 - CPU a 3.40GHz

- 16 GB Memoria RAM
- Sistema operativo de 64 bits
- Procesador x64

3 Resultados y Discusión

A partir del documento concentrador de tiempos de ordenamiento (Figura 1), se procedió a realizar el análisis estadístico de estos utilizando el software estadístico IBM SPSS Statistics® Ver 20.

Para el análisis estadístico, se definieron dos variables: Ordenamiento y Tiempo, variable independiente y dependiente respectivamente para cada uno de los conjuntos de datos (1,000, 10,000 y 100,000), donde Ordenamiento refiere a cada uno de los algoritmos de ordenamiento y Tiempo al tiempo de ejecución que se llevó el algoritmo en ordenar cada conjunto de datos (ver Figura 2).

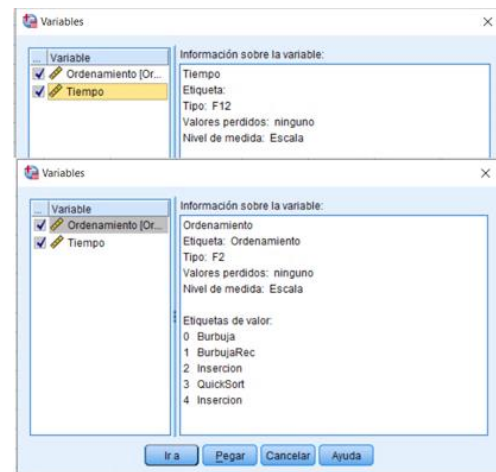


Fig. 2. Variables para cada conjunto de datos.

El análisis de los datos se inició con la prueba de normalidad para cada algoritmo y para cada conjunto de datos (1,000, 10,000 y 100,000), ésta se llevó a cabo a través de la prueba de Kolmogorov-Smirnov debido a que el número de datos es mayor a 30.

Los niveles de significancia obtenidos se presentan las Tablas 1-3.

Tabla 1. Prueba de normalidad para 1,000 elementos.

		Pruebas de normalidad					
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
Ordenamiento	Tiempo	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Burbuja		.084	50	.200 [*]	.959	50	.077
BurbujaRec		.092	50	.200 [*]	.957	50	.065
Insercion		.071	50	.200 [*]	.981	50	.596
QuickSort		.105	50	.200 [*]	.954	50	.050
Insercion		.075	50	.200 [*]	.979	50	.524

Tabla 2. Prueba de normalidad para 10,000 elementos.

		Pruebas de normalidad					
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
Ordenamiento	Tiempo	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Burbuja		.063	50	.200 [*]	.987	50	.860
BurbujaRec		.092	50	.200 [*]	.959	50	.083
Insercion		.089	50	.200 [*]	.959	50	.081
QuickSort		.102	50	.200 [*]	.956	50	.062
Seleccion		.102	50	.200 [*]	.972	50	.267

Tabla 3. Prueba de normalidad para 100,000 elementos.

Ordenamiento	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Tempo						
Burbuja	.107	50	.200 [*]	.953	50	.043
Insercion	.091	50	.200 [*]	.979	50	.509
QuickSort	.086	50	.200 [*]	.981	50	.590
Seleccion	.087	50	.200 [*]	.968	50	.188

Considerando que las hipótesis de la prueba de normalidad son:

H_0 : Los datos provienen de una distribución normal.

H_1 : Los datos no provienen de una distribución normal.

donde los valores de P_valor :

$P_valor \geq \alpha$ se acepta H_0

$P_valor < \alpha$ se acepta H_1

con un valor de significancia de $\alpha=0.05$ y con base en los resultados de las Tablas 1-3, se puede ver que los P_valor (.2000) de cada una de las pruebas están por encima del 0.05 por lo que se acepta la hipótesis nula H_0 afirmando que los datos provienen de una distribución normal.

Una vez que se comprobó que los datos tienen una distribución normal, se prosiguió a realizar la prueba de homogeneidad de varianza o varianzas iguales (llamada también homocedasticidad), ésta se realizó con la prueba de Levene (>30 elementos). Para hacer esta prueba, se consideraron cinco grupos, los cuales corresponden a los cinco algoritmos de ordenamiento.

Para las hipótesis de la prueba de varianzas se tienen:

H_0 : Los datos tiene varianzas iguales.

H_1 : Los datos no tienen varianzas iguales.

$P_valor \geq \alpha$ se acepta H_0

$P_valor < \alpha$ se acepta H_1

con un valor de significancia de $\alpha=0.05$

Los resultados obtenidos con la prueba de Levene para cada conjunto de datos son los presentados en las Tablas 4-6

Tabla 4. Prueba de homocedasticidad para 1,000 elementos.**Prueba de homogeneidad de varianzas**

Tiempo			
Estadístico de Levene	gl1	gl2	Sig.
97.260	4	245	.000

Tabla 5. Prueba de homocedasticidad para 10,000 elementos.**Prueba de homogeneidad de varianzas**

Tiempo			
Estadístico de Levene	gl1	gl2	Sig.
72.092	4	245	.000

Tabla 6. Prueba de homocedasticidad para 100,000 elementos.**Prueba de homogeneidad de varianzas**

Tiempo			
Estadístico de Levene	gl1	gl2	Sig.
59.134	3	196	.000

A partir de los datos obtenidos en las Tablas 4-6 se rechaza la hipótesis nula H_0 , ya que los P_valor (.000) están por debajo de 0.05, entonces se afirma que los datos no tienen varianzas iguales.

Para la comparación de grupos independientes la hipótesis nula debe ser aceptada al ser uno de los supuestos para realizar una prueba estadística de comparación de grupos entonces podría pensarse que no procedería la aplicación de esta prueba. Sin embargo, al respecto Anderson y cols. (2008), en su libro establecen que: "Si los tamaños de las muestras son iguales, el análisis de varianza no es sensible a desviaciones..." (Anderson y col; 2008), lo cual apoya a Hildebrand y Ott (1998) quienes afirmaron que este supuesto es importante si los tamaños de las muestra son sustancialmente distintos y mencionan que cuando todos los grupos son iguales en tamaño, el efecto de las varianzas enormemente desiguales es mínimo (Hildebrand y Ott 1998). En relación con esto, el número de elementos que se están analizando es de 50 para cada uno de algoritmos de ordenamiento, lo cual permite proceder a la aplicación de la prueba estadística de comparación entre grupos independientes mediante el análisis ANOVA.

La hipótesis que se propone a prueba en el ANOVA es que las medias poblacionales son iguales, si éstas son iguales, significa que los grupos no difieren en la variable dependiente (tiempo de ordenamiento). Con base en esto, las hipótesis planteadas para el análisis ANOVA son:

H_0 : Las medias poblacionales son iguales.

H_1 : Al menos dos medias poblacionales son distintas.

$P_valor \geq \alpha$ se acepta H_0

$P_valor < \alpha$ se acepta H_1

con un valor de significancia de $\alpha=0.05$

Con base en lo anterior, se comprueba si los diferentes grupos (algoritmos de ordenamientos) definidos por la variable ordenamiento difieren en la variable *Tiempo*.

En las Tablas 7-9 se presentan los análisis ANOVA para cada uno de los datos trabajados.

Tabla 7. ANOVA de un factor para 1,000 elementos.

Tiempo					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	.021	4	.005	233.554	.000
Intra-grupos	.005	245	.000		
Total	.026	249			

Tabla 8. ANOVA de un factor para 10,000 elementos.

Tempo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	1.779	4	.445	4485.832	.000
Intra-grupos	.024	245	.000		
Total	1.803	249			

Tabla 9. ANOVA de un factor para 100,000 elementos.

Tempo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	14053.791	3	4684.597	3606663.831	.000
Intra-grupos	.255	196	.001		
Total	14054.046	199			

Con los resultados de las Tablas 7-9 se aprecia que los niveles de significación (.000) son menores que 0.05, se rechaza la hipótesis H_0 de igualdad de medias, es decir, existen diferencias significativas entre los grupos, lo cual también permite afirmar que sí hay diferencia significativa entre los tiempos de ejecución en los algoritmos de ordenamiento.

Específicamente, con las pruebas post hoc (Tablas 10-12) se pueden detectar en qué grupos se presentan esas diferencias significativas y haciendo referencia a los resultados de estas Tablas 10-12, se puede ver que en la prueba post hoc de los 1,000 elementos (Tabla 10) es en la única en la que no se encuentran diferencias significativas como lo son:

- HSD de Tukey: Burbuja-Inserción (0.067), Burbuja-Selección (.385) e Inserción-Selección (.913)
- Scheffé: Burbuja-Inserción (0.142), Burbuja-Selección (.529) e Inserción-Selección (.948)

En las Tablas 11 y 12 es realmente significativa la diferencia entre los algoritmos de ordenamiento en cuanto al tiempo de ordenación se refiere.

Tabla 10. Pruebas post hoc para 1,000 elementos.

Variable dependiente: Tempo		Comparaciones múltiples		Intervalo de confianza al 95%	
(i) Ordenamiento	(j) Ordenamiento	Diferencia de medias (i-j)	Error típico	Límite inferior	Límite superior
HSD de Tukey	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
BurbujaRec	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
Insercion	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
QuickSort	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
Scheffé	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000

* La diferencia de medias es significativa al nivel .05.

Tabla 11. Pruebas post hoc para 10,000 elementos.

Variable dependiente: Tempo		Comparaciones múltiples		Intervalo de confianza al 95%	
(i) Ordenamiento	(j) Ordenamiento	Diferencia de medias (i-j)	Error típico	Límite inferior	Límite superior
HSD de Tukey	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
BurbujaRec	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
Insercion	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
QuickSort	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
Scheffé	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000

* La diferencia de medias es significativa al nivel .05.

Tabla 12. Pruebas post hoc para 100,000 elementos.

Variable dependiente: Tempo		Comparaciones múltiples		Intervalo de confianza al 95%	
(i) Ordenamiento	(j) Ordenamiento	Diferencia de medias (i-j)	Error típico	Límite inferior	Límite superior
HSD de Tukey	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
BurbujaRec	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
Insercion	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
QuickSort	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000
Scheffé	Burbuja	Inserción	.000	-.00000000	.00000000
	Burbuja	Selección	.000	-.00000000	.00000000
	Burbuja	QuickSort	.000	-.00000000	.00000000
	Inserción	Burbuja	.000	.00000000	.00000000
	Inserción	Selección	.000	.00000000	.00000000
	Inserción	QuickSort	.000	.00000000	.00000000

* La diferencia de medias es significativa al nivel .05.

Bakieva y col. (2010), trabajaron con un modelo de contraste de medias que se presenta en la Figura 3 (Bakieva y col; 2010). Siguiendo este proceso y al encontrar que las variancias obtenidas en las Tablas 4-6 no son iguales se procede a obtener las pruebas post hoc aplicando T2 de Tamhane, T3 de Dunnett y Games-Howell.

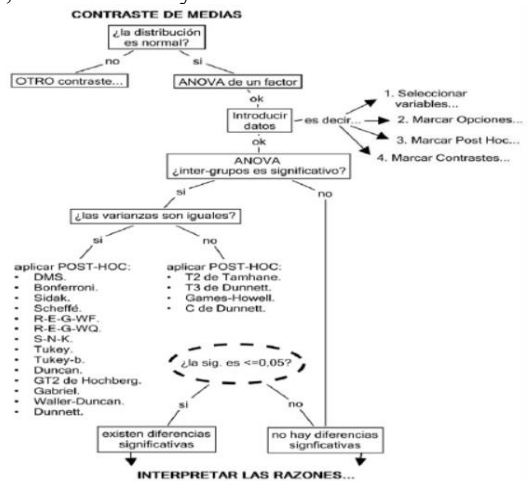


Fig. 3. Proceso para el contraste de medias. Fuente: (Bakieva y cols; 2010).

Se procedió a aplicar estas pruebas y los resultados obtenidos se muestran en las Tablas 13-15.

Tabla 13. Pruebas post hoc aplicando T2 de Tamhane y T3 de Dunnett y Games-Howell para 1,000 elementos.

Table with columns: Variable dependiente: Tiempo, (i) Ordenamiento, (j) Ordenamiento, Diferencia de medias (i, j), Error t-Student, Sig., Intervalo de confianza al 95%, Límite inferior, Límite superior. Rows include Tamhane, Dunnett, and Games-Howell tests for various sorting algorithms like Burbuja, QuickSort, and Selección.

* La diferencia de medias es significativa al nivel .05.

Tabla 14. Pruebas post hoc aplicando T2 de Tamhane y T3 de Dunnett y Games-Howell para 10,000 elementos.

Table with columns: Variable dependiente: Tiempo, (i) Ordenamiento, (j) Ordenamiento, Diferencia de medias (i, j), Error t-Student, Sig., Intervalo de confianza al 95%, Límite inferior, Límite superior. Rows include Tamhane, Dunnett, and Games-Howell tests for various sorting algorithms like Burbuja, QuickSort, and Selección.

* La diferencia de medias es significativa al nivel .05.

Tabla 15. Pruebas post hoc aplicando T2 de Tamhane y T3 de Dunnett y Games-Howell para 100,000 elementos.

Table with columns: Variable dependiente: Tiempo, (i) Ordenamiento, (j) Ordenamiento, Diferencia de medias (i, j), Error t-Student, Sig., Intervalo de confianza al 95%, Límite inferior, Límite superior. Rows include Tamhane, Dunnett, and Games-Howell tests for various sorting algorithms like Burbuja, QuickSort, and Selección.

* La diferencia de medias es significativa al nivel .05.

Donde se aprecia que el nivel de significancia (.000) de cada uno de los algoritmos y de todas las pruebas es menor al 0.05, con lo cual también se puede afirmar que sí existe diferencia significativa en el tiempo de ordenación entre los algoritmos de ordenamiento de inserción, selección, burbuja (simple y recursivo) y quicksort.

4 Conclusiones

Con los resultados obtenidos se puede concluir que:

Se logró realizar un análisis estadístico comparativo entre los algoritmos de ordenación considerando los tiempos de ejecución de 1,000, 10,000 y 100,000 elementos.

Se demostró que sí existe diferencia significativa de los tiempos de ejecución entre los algoritmos de ordenación.

Se corroboran los resultados obtenidos de muchas investigaciones referentes a las diferencias que existen entre los algoritmos de ordenamientos realizados con base en distintos factores: tiempos de ejecución, eficiencia, etc.

Se confirma que el algoritmo quicksort sigue siendo uno de los más rápidos.

Referencias

Adhikari, P. (2007). Review On Sorting Algorithms A comparative study on two sorting algorithms. A Term Paper Submitted to the Faculty of Dr. Gene Boggess, Mississippi State University, In the Department of Computer Science & Engineering, Mississippi State, Mississippi.
Alif, A. F., Islam, S. M. R., & Deb, P. (2019). Design and implementation of sorting algorithms based on FPGA. Paper presented at the 2019 International Conference on Computer, Communication,


- Chemical, Materials and Electronic Engineering (IC4ME2).
- Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2008). *Estadística para administración y economía*. México: Thomson/Southwestern.
- Bakieva, M., González Such, J., & Jornet, J. (2010). *SPSS: ANOVA de un factor*. Grupo de Innovación Educativa. Universitat de Valencia.
- Buradagunta, S., Bodapati, J. D., Mundukur, N. B., & Salma, S. (2020). Performance Comparison of Sorting Algorithms with Random Numbers as Inputs *Ingénierie des Systèmes d'Information*, 25(1), 113-117.
- Chowdhury, N. (2016). *A Java Based Tool to Monitor Execution Time of Different Sorting Algorithms*. East West University.
- García, D. G., & Velázquez, J. L. A. (2019). Comparación entre algoritmos de ordenamiento paralelizados en Java. *Pistas Educativas*, 41(133).
- Hildebrand, D. K., & Ott, L. (1998). *Estadística aplicada a la administración ya la economía*: Addison Wesley Logman de MÚxico, SA.
- Htwe Htwe, A. (2019). Analysis and Comparative of Sorting Algorithms. *International Journal of Trend in Scientific Research and Development*, 3(5), 1049-1053.
- Iskandar, I. D., Amirulloh, I., Pertiwi, M. W., Kusmira, M., Hikmah, A. B., & Supriadi, D. (2020). Analysis of bubble sort and insertion sort algorithm on memory efficiency using data mining approach. *Jurnal Pilar Nusa Mandiri*, 16(1), 89-96.
- Khreisat, L. (2018). A Survey of Adaptive QuickSort Algorithms. *International Journal of Computer Science and Security (IJCSS)*, 12(1), 1.
- Kumar, S. (2020). An Empirical Study on Algorithms-Conquer Sorting Techniques. Available at SSRN 3568396.
- Manaseer, S., & Al Hwaitat, A. K. (2018). Measuring Parallel Performance of Sorting Algorithms Bubble Sort and Bucket Sort on IMAN. *Modern Applied Science*, 12(10).
- Nasar, A. (2019). A Mathematical Analysis of student-generated sorting algorithms. *The Mathematics Enthusiast*, 16(1), 315-330.
- Prajapati, P., Bhatt, N., & Bhatt, N. (2017). Performance Comparison of Different Sorting Algorithms. *International Journal of Latest Technology in Engineering, Management & Applied Sciences*, 6.
- Yang, Y., Yu, P., & Gan, Y. (2011). *Experimental study on the five sort algorithms*. Paper presented at the 2011 Second International Conference on Mechanic Automation and Control Engineering.
- Zhao, L., Liu, X., & Shao, X. (2016). *Comparative Analysis of Numerical Sorting Algorithms in Java Language*. Paper presented at the 2016 6th

International Conference on Advanced Design and Manufacturing Engineering (ICADME 2016).


Recibido: 23 de enero de 2022

Aceptado: 26 de mayo de 2022

Salgado Gallegos, Mireya: Ingeniera en Computación en la Universidad Autónoma del Estado de México en 1994. Obtuvo el grado de Maestría en Ingeniería en Informática en 2007 en la misma Institución y es Doctora en Ingeniería Industrial en Tecnologías de Información en el 2018 por la Universidad Anáhuac Norte de México.

 <http://orcid.org/0000-0003-2675-9456>

Pérez Merlos, Juan Carlos: Ingeniero en Electrónica en Instrumentación egresado del Instituto Tecnológico de Ciudad Guzmán Jalisco en 1986. Obtuvo el grado de Maestro en Ingeniería en Informática en 1998 en la Universidad Autónoma del Estado de México y es Doctor en Ingeniería Industrial en Tecnologías de Información en el 2018 por la Universidad Anáhuac Norte de México. Correo electrónico: jcjcjc63@yahoo.com

 <https://orcid.org/0000-0001-6189-5125>

Albarrán Trujillo, Silvia Edith: Ingeniera en Computación en la Universidad Autónoma del Estado de México en 1994. Obtuvo el grado de Maestra en Administración en 1999 en la misma Institución y es Doctora en Ingeniería Industrial en Tecnologías de Información en el 2019 por la Universidad Anáhuac Norte de México). Correo electrónico: seat@uaemex.mx

 <https://orcid.org/0000-0003-2620-5277>

