



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Probabilidad y Estadística

*Comparación entre dos CPU de distinta
generación*



Alejandra Traveria Martí

Enric Teixidó Álvarez

Pol Beltrán Ancín

DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

1R CUATRIMESTRE DE 2022

RESUMEN

- Objetivo: Comparar dos procesadores del mismo proveedor, en este caso AMD Ryzen 5 5600X 3.7GHz y AMD Ryzen 7 5800X 3.8GHz, para ver si realmente hay una diferencia en el rendimiento de estos.
- Métodos: Hemos seleccionado tres *benchmarks* diferentes (*Cinebench R23*, *Geekbench 5 64bits*, *CPU-Z Benchmark 17*), con los cuales hemos realizado 27 test de rendimiento por cada procesador con cada uno de ellos obteniendo así 81 puntuaciones por procesador. Debido a que la puntuación varía drásticamente según el *benchmark* usado decidimos usar una puntuación basada en la media de los *scores* de los tres *benchmarks*, obteniendo así una Distribución Normal. Con la distribución obtenida hemos comprobado las premisas convenientes para poder ajustar los métodos estadísticos aprendidos en clase para comprobar si realmente existen diferencias en el rendimiento de los CPU.
- Resultados: El *score* promedio del AMD Ryzen 5 5600X 3.7GHz ha sido de 8005 y el del AMD Ryzen 7 5800X 3.8GHz 11165. Así entonces negamos H_0 (que afirma que no existe diferencia en los rendimientos) y, por lo tanto, sí que observamos una diferencia significativa en los *scores* de los procesadores.
- Conclusión: Basándonos en los resultados de nuestro estudio hemos observado que el AMD Ryzen 7 5800X 3.8GHz es un 39.47% más eficiente que el AMD Ryzen 5 5600X 3.7GHz. Este porcentaje se asemeja bastante a los datos proporcionados por la compañía AMD la cual aseguró que el Ryzen 7 era un 37% más eficaz que el Ryzen 5.

INTRODUCCIÓN

En el ámbito de la informática la importancia de un buen procesador es esencial, es uno de los componentes más importantes a la hora de determinar el rendimiento de un PC. El procesador es el cerebro y corazón de los ordenadores, por eso sus características son imprescindibles, ya que consideran la velocidad, la capacidad de memoria, seguridad y compatibilidad con los demás componentes. Por todo ello es vital a la hora de escoger un procesador saber si realmente es tan eficiente como los proveedores aseguran.

Es por ese motivo que nos sentimos motivados a realizar este estudio en el cual pondremos a prueba estos dos procesadores para valorar cuanta diferencia real existe entre dos CPU de distinta generación.

Como hemos mencionado brevemente en el resumen hemos empleado tres *benchmarks* diferentes para poner a prueba los dos procesadores. Un *benchmark* es una técnica utilizada para medir el rendimiento de un sistema o uno de sus componentes, es el resultado de la ejecución de un programa informático o un conjunto de programas en una máquina con el objetivo de estimar el rendimiento de un elemento en concreto y poder comparar el resultado con máquinas similares. Más adelante en la recogida de datos y observación de los resultados indagaremos un poco más acerca de los *benchmarks* que hemos seleccionado.

Partiendo de los resultados de los *benchmarks* podremos afirmar o desmentir si efectivamente el Ryzen 7 es o no mejor que el Ryzen 5.

RECOGIDA Y TRATAMIENTO DE LOS DATOS

Para estudiar la diferencia de rendimiento real de los dos procesadores hemos recabado distintas puntuaciones obtenidas mediante *benchmarks*, todos ellos pueden probar los procesadores tanto con pruebas *single-core* como *multi-core*, pero para este estudio solo hemos probado el *multi-core* ya que suele ser más exigente a la hora de poner a prueba el rendimiento de los CPU ya que todos ellos involucran todos los núcleos de la CPU y hacen uso del *hyperthreading* (es un proceso por el cual la CPU divide sus cores en cores virtuales también conocidos como *threads* y estos son tratados como cores físicos por el sistema operativo).

El *Cinebench R23* es un software utilizado para generar formas 3D. El *Geekbench 5 64bits* es un *benchmark* multi-plataforma que utiliza intensivamente la memoria del sistema, una memoria rápida mejora mucho el resultado. El *CPU-*

Z Benchmark 17 mide el tiempo que tiempo necesario para que el sistema complete todos los cálculos del *benchmark*. Para cada procesador ejecutamos cada *benchmark* 27 veces, es decir obtuvimos 81 scores por cada procesador.

Cuando ya tuvimos una tabla clara con los datos recogidos utilizamos el software estadístico R, *RStudio 2022.07.1+576*, para poder generar los diversos gráficos y cálculos.

Durante el tratamiento de los datos en R nos percatamos de la diferencia abismal de las puntuaciones dependiendo del *benchmark* utilizado y eso nos suponía un problema ya que no se cumplía la normalidad de la muestra, con el propósito de solucionar este problema decidimos hacer una media de las puntuaciones de los tres *benchmarks* para obtener un score representativo de los tres, arreglando así la inconveniencia y poder estudiar los datos partiendo de una Distribución Normal. En un principio habíamos distribuido los datos en subtablas que tenían en cuenta tanto el *benchmark* como el tipo de procesador, pero no era posible normalizar los datos debido a la disparidad de puntuaciones. Se puede ver más claro en el anexo 1 donde se puede ver el trabajo de los datos con R.

Todas las variables están medidas en puntuaciones de *benchmark*.

I. Diferencia de scores

Para poder ver si el rendimiento de los dos procesadores es distinto hemos usado la diferencia entre los scores de los *benchmarks* de cada uno de los procesadores. Por esta razón hemos creado una variable $d1 = (\text{Ryzen7\$Score} - \text{Ryzen5\$Score})$.

II. Premisas adecuadas

1. Normalidad de los datos
2. Aleatoriedad de las muestras, es decir, pruebas generadas al azar. Como en nuestro caso las pruebas eran realizadas aleatoriamente por los *benchmarks* podemos asegurar esta premisa.

III. Planteamiento de la prueba bilateral

$$H_0: \mu_{d1} = 0$$

$$H_1: \mu_{d1} \neq 0$$

IV. Cálculo del estadístico

Como nuestros datos no son independientes realizaremos un test para muestras pareadas. En un principio empezamos con un estadístico de tipo Z pero al encontrarnos con el problema de la normalidad decidimos utilizar el estadístico T.

$$T = \frac{(\bar{y} - \mu)}{s / \sqrt{n}} = \frac{(\bar{y} - \mu)}{se}$$

V. Estudio del p-value ($\alpha = 0.05$)

Si $p\text{-value} < \alpha$, entonces podremos rechazar la hipótesis nula con una confianza del $(1-\alpha)\%$.

VI. Intervalo de confianza de la diferencia

Con esta fórmula podremos dar un intervalo que contendrá la mediana de $d1$ con una confianza del $(1-\alpha)\%$.

$$IC(d1, 1 - \alpha) = \left[\bar{y} \pm t_{n-1, 1-\frac{\alpha}{2}} \cdot se \right]$$

OBSERVACIÓN DE LOS RESULTADOS

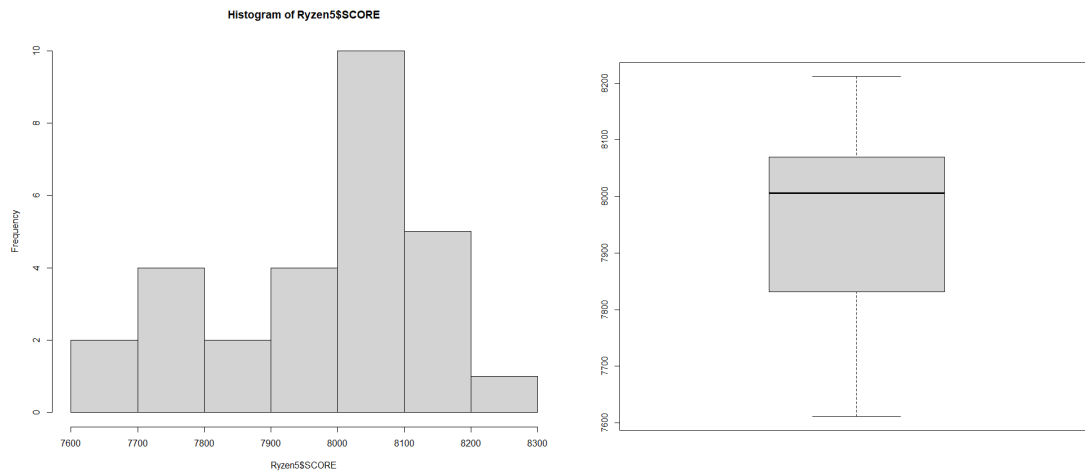
Descriptiva

La siguiente tabla contiene las medias i desviaciones típicas de los scores de cada CPU y de la variable $d1$.

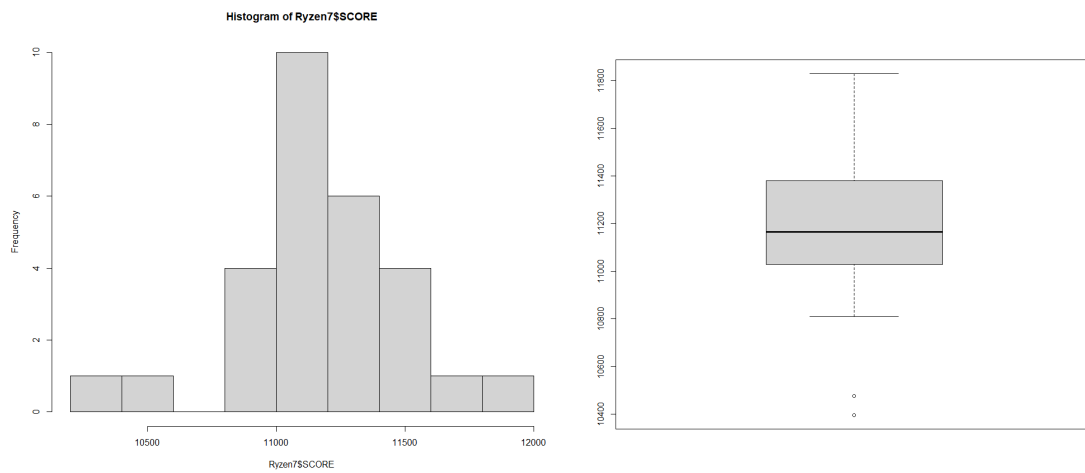
Procesador	Mediana	Desviación Típica
AMD Ryzen 5 5600X 3.7GHz	8005	157.4491
AMD Ryzen 7 5800X 3.8GHz	11165	323.3518
$d1$	3283	387.124

Los datos obtenidos están representados en las siguientes gráficas:

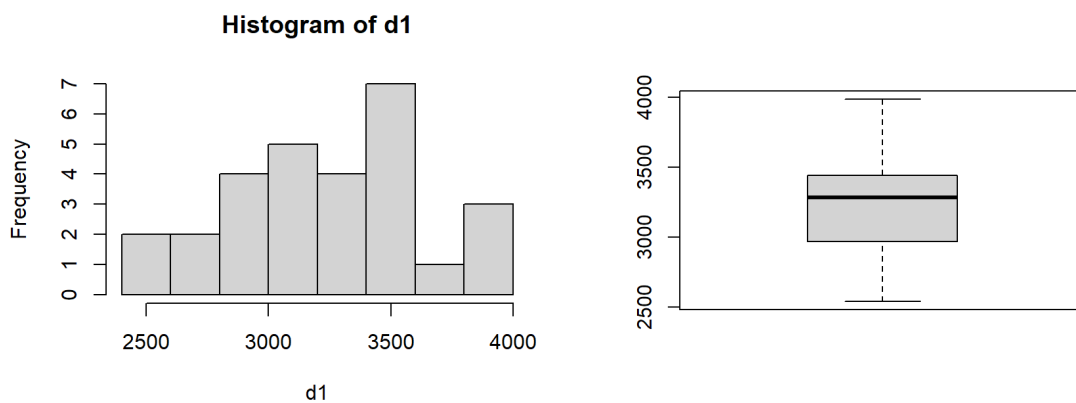
Ryzen 5:



Ryzen 7:



d1:



Podemos observar que las gráficas parecen seguir una distribución normal, obteniendo los picos de los scores según su frecuencia. Solo observando las gráficas de los procesadores podemos llegar a intuir que el Ryzen 7 muestra puntuaciones más elevadas que las den Ryzen 5.

INFERENCIA

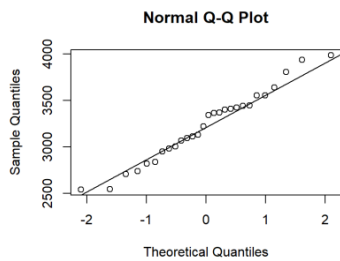


Ilustración 1. Gráfica QQnorm d1

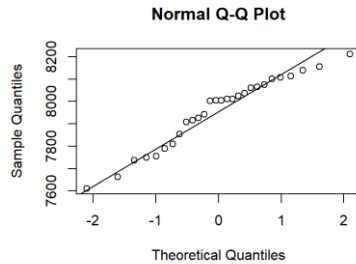


Ilustración 2. Gráfica QQnorm R5

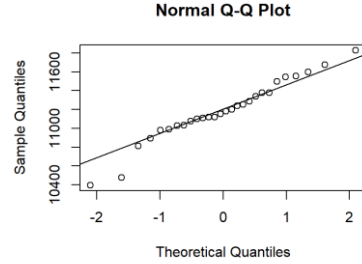


Ilustración 3. Gráfica QQnorm R7

Podemos observar en las gráficas Q-QPlot que los valores se ajustan bastante a una Distribución Normal ya que los cuantiles de las muestras se asemejan bastante a los cuantiles teóricos. Aunque también se pueden observar algunos valores que se alejan de la distribución que podrían tratarse de errores o información importante sobre el estudio.

Como hemos comentado, hemos trabajado con la variable diferencia d1 para poder trabajar con la diferencia de los scores de los diferentes procesadores. A raíz de esta variable hemos encontrado el estadístico T de la Distribución Normal con un valor de 44.156 que nos permite poner a prueba la hipótesis nula. Calculamos el p-value del estadístico con R y obtuvimos un valor muy próximo a 0 (siendo exactos $p\text{-value} < 2.2e - 16$). Podemos concluir entonces, que podemos negar la hipótesis nula con mucha certeza y, así mismo, afirmar que hay una clara diferencia de rendimiento entre los dos procesadores.

Seguidamente hemos calculado el intervalo de confianza para ajustar el valor de la mediana de d1. Como hemos observado que el Ryzen 7 obtiene mejores puntuaciones que el Ryzen 5 y como sabemos también que la diferencia no es 0, podemos afirmar con una confianza del 95% que, en mediana, el Ryzen 7 obtiene una mejor puntuación que el Ryzen 5 con 3230.417 de diferencia, con un IC = [3080.306, 3380.528].

Si calculamos el porcentaje de mejora relativa en el rendimiento, vemos que el Ryzen 7 es un 39.47% más eficaz que el Ryzen 5.

CONCLUSIÓN

Tras realizar el estudio estadístico hemos podido observar que el procesador AMD Ryzen 7 5800X 3.8GHz demuestra un mejor rendimiento que el AMD Ryzen 5 5600X 3.7GHz. Si es cierto que estos resultados pueden estar afectados por situaciones varias, como los componentes del ordenador con el que hemos realizado las pruebas, el sistema operativo del ordenador, etc. Desgraciadamente no podemos controlar estos factores en nuestro estudio estadístico, y es por eso que hay que tener estos hechos en mente a la hora de concluir el estudio.

ANEXO 1

Script R

```
# BLOQUE TRANSVERSAL

#####

# Leemos el archivo excel con los datos

library(readxl)

path <- 'C:/Users/eteix/Desktop/'

data <- data.frame(read_excel(paste0(path, 'DadesR.xlsx'))))

# Visualizamos la tabla

View(data)

# Creamos subtablas separadas por procesador

Ryzen5 <- subset(data, Procesador == "AMD Ryzen 5 5600X 3.7GHz")
Ryzen7 <- subset(data, Procesador == "AMD Ryzen 7 5800X 3.8GHz")
```

Descriptiva variables continuas

En las siguientes gráficas se puede ver con más claridad lo comentado en el apartado de recogida y tratamiento de datos. Podemos observar que dependiendo de los *benchmarks* las puntuaciones varían drásticamente, si nos fijamos, por ejemplo, en las puntuaciones del *CineBench* para el Ryzen 5 podemos ver que sus puntuaciones tienen una media de 10862, sin embargo, las puntuaciones del *CPU-Z Bench* para el Ryzen 5 tienen una media de 4780, y lo mismo sucede con las puntuaciones del Ryzen 7. Esto afectaba directamente en la normalidad de los datos y es por ello que al final nos decidimos por hacer la media de las tres puntuaciones para obtener un score más homogéneo que se ajustara más a una Distribución Normal que era lo que veníamos buscando.

Ryzen 5

```
summary(Ryzen5$Score)
```

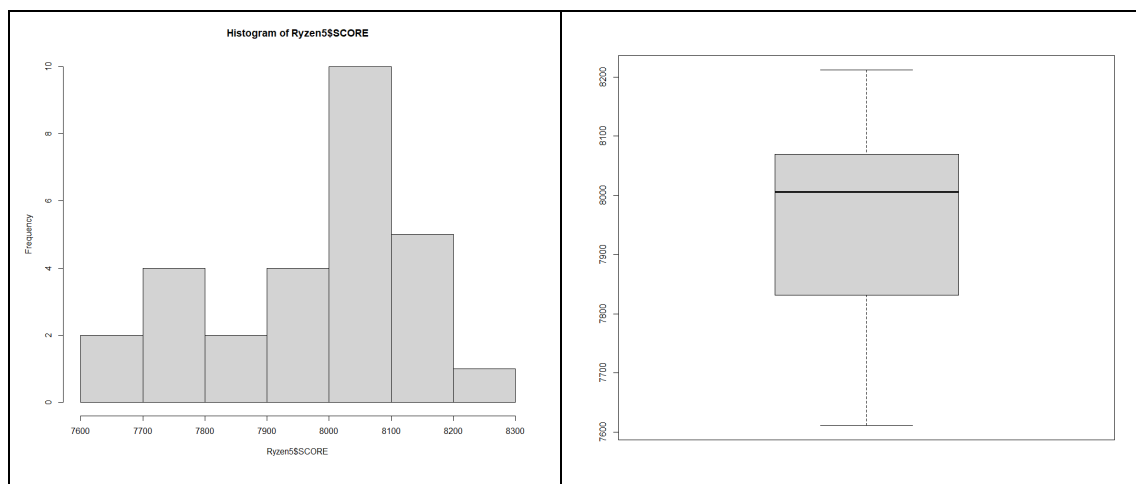
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
7611	7842	8005	7957	8064	8212

```
var(Ryzen5$Score) 24790.23
```

```
sd(Ryzen5$Score) 157.4491
```

```
hist(Ryzen5$Score)
```

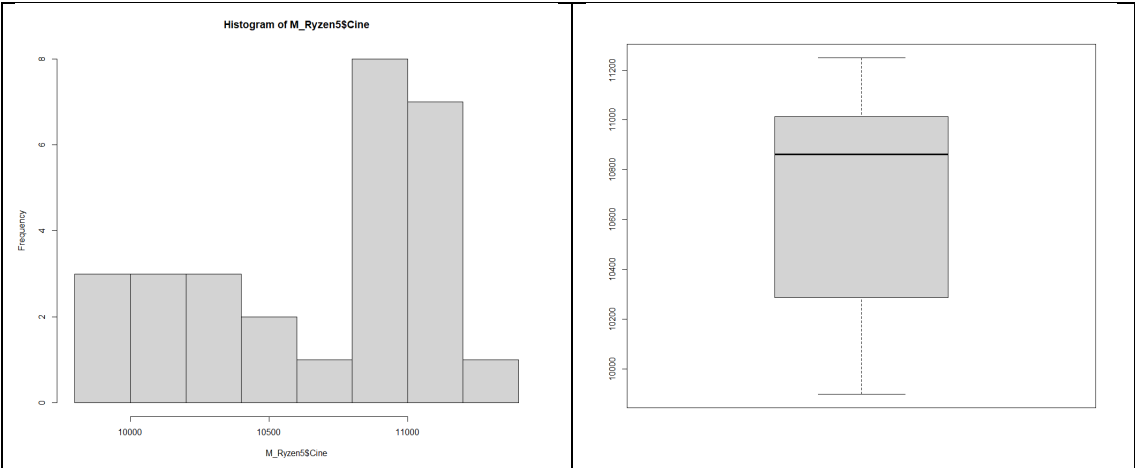
```
boxplot(Ryzen5$Score)
```



```
# Ryzen 5/Benchmark  
  
#Cine Bench  
  
summary(Ryzen5$Cine)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
9899	10308	10862	10680	11009	11250

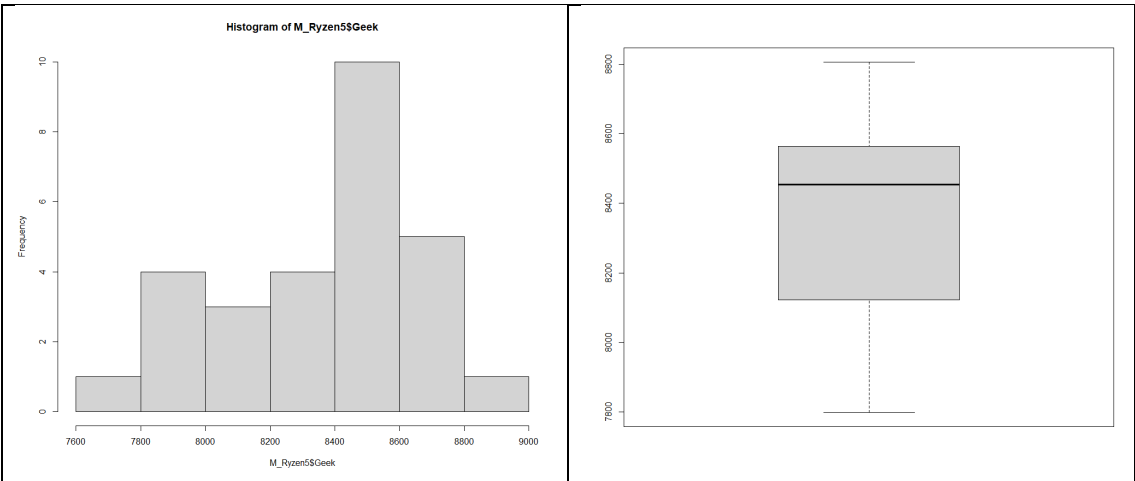
```
var(Ryzen5$Cine) 183091.1  
sd(Ryzen5$Cine) 427.8914  
hist(Ryzen5$Cine)  
boxplot(Ryzen5$Cine)
```



```
#Geek Bench  
  
summary(Ryzen5$Geek)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
7799	8126	8454	8355	8548	8806

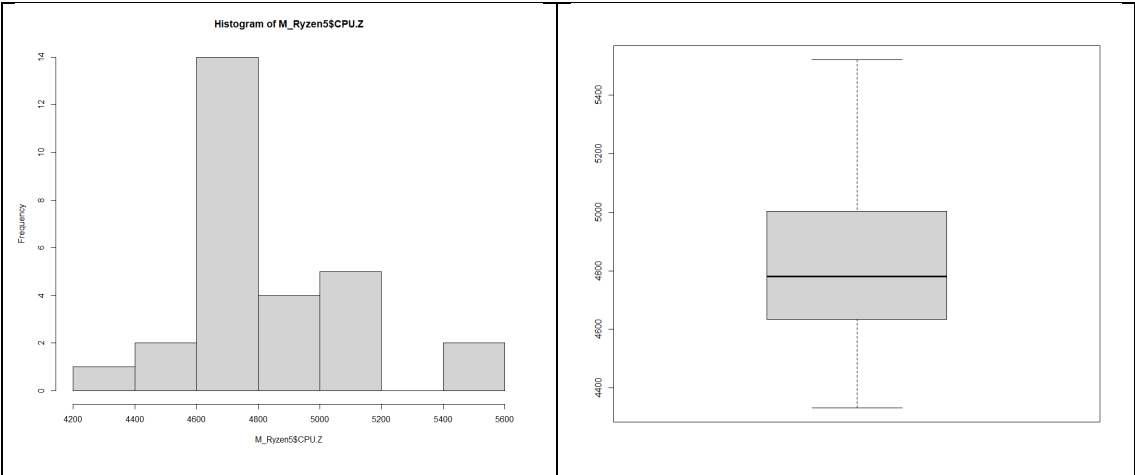
```
var(Ryzen5$Geek) 84628.69  
sd(Ryzen5$Geek) 290.9101  
hist(Ryzen5$Geek)  
boxplot(Ryzen5$Geek)
```



```
#CPU.Z Bench  
  
summary(Ryzen5$CPU.Z)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
4332	4643	4780	4836	4998	5521

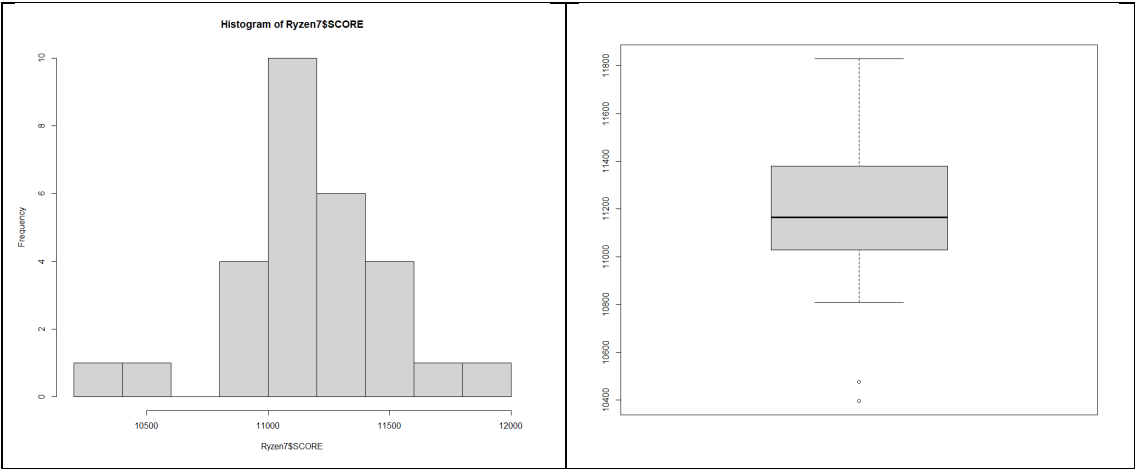
```
var(Ryzen5$CPU.Z) 69829.96  
  
sd(Ryzen5$CPU.Z) 264.2536  
  
hist(Ryzen5$CPU.Z)  
  
boxplot(Ryzen5$CPU.Z)
```



```
#Ryzen 7  
  
summary(Ryzen7$Score)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
10396	11031	11165	11187	11380	11830

```
var(Ryzen7$Score) 104556.4  
  
sd(Ryzen7$Score) 323.3518  
  
hist(Ryzen7$Score)  
  
boxplot(Ryzen$Score)
```



#Ryzen 7/Benchmark

#Cine Bench

summary(Ryzen7\$Cine)

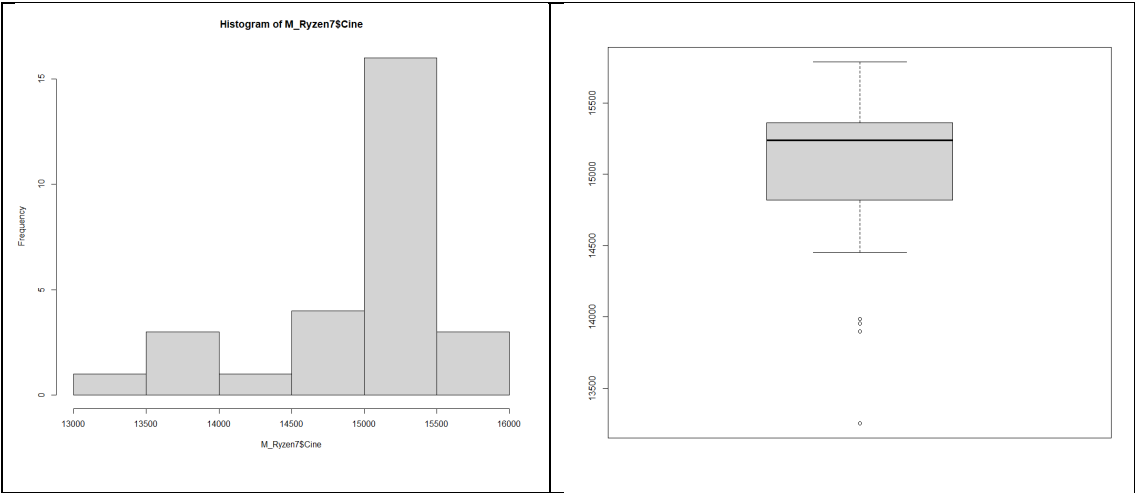
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
13254	14846	15240	15020	15347	15789

var(Ryzen7\$Cine) 370791.8

sd(Ryzen7\$Cine) 608.9268

hist(Ryzen7\$Cine)

boxplot(Ryzen7\$Cine)



#Geek Bench

summary(Ryzen7\$Geek)

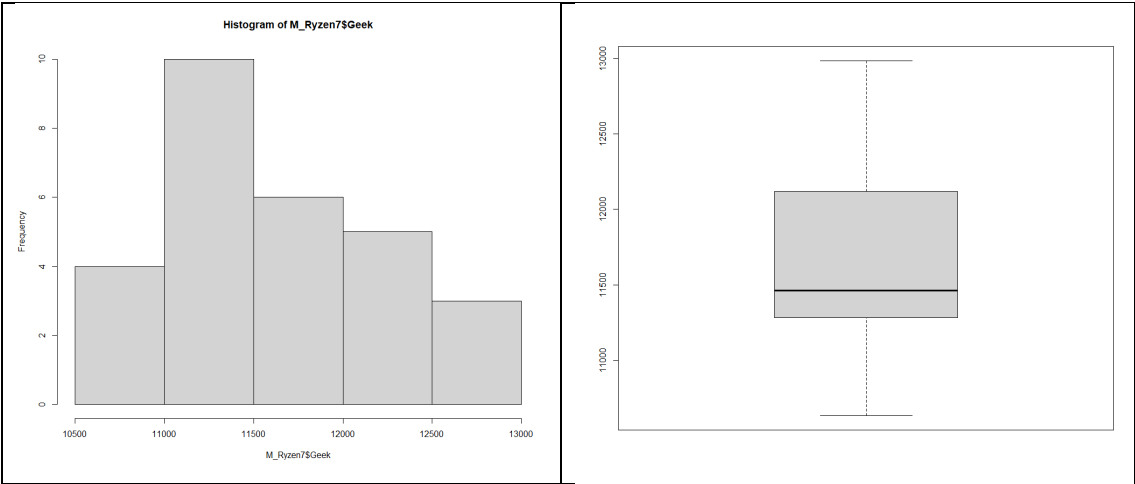
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
10633	11288	11461	11658	12115	12986

var(Ryzen7\$Geek) 388843

sd(Ryzen7\$Geek) 623.5728

hist(Ryzen7\$Geek)

boxplot(Ryzen7\$Geek)



```
#CPU.Z Bench
```

```
summary(Ryzen7$CPU.Z)
```

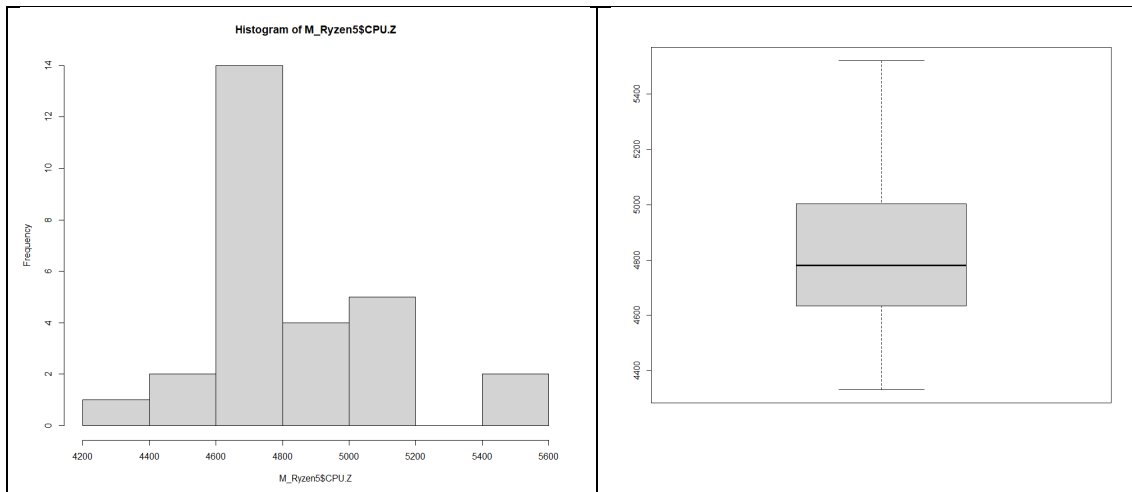
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
6154	6510	6663	6884	7132	8354

```
var(Ryzen7$CPU.Z) 307002.4
```

```
sd(Ryzen7$CPU.Z) 554.0779
```

```
hist(Ryzen7$CPU.Z)
```

```
boxplot(Ryzen7$CPU.Z)
```



#Comprovacio de la "Normalitat" de les mostres

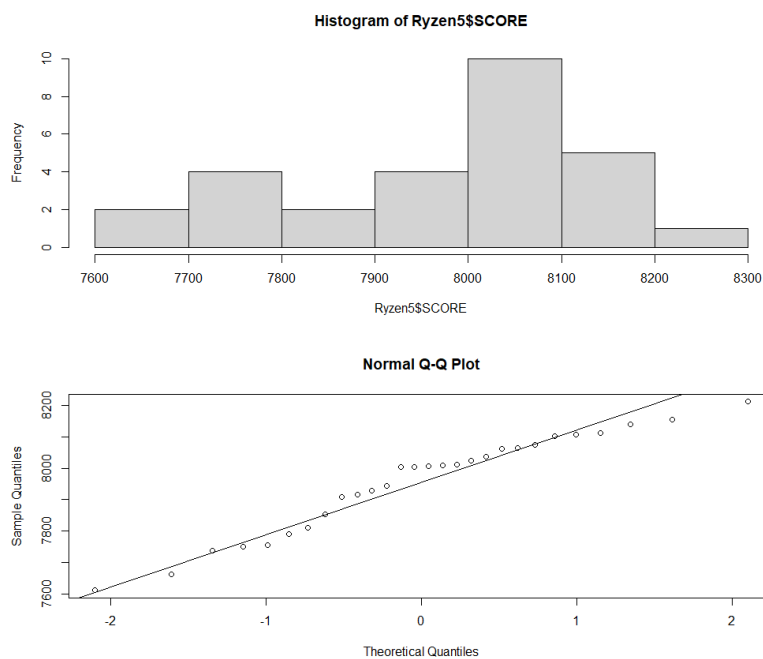
```
#Ryzen 5
```

```
par(mfrow=c(,))
```

```
hist(Ryzen5$Score)
```

```
qqnorm(Ryzen5$Score)
```

```
qqline(Ryzen5$Score)
```



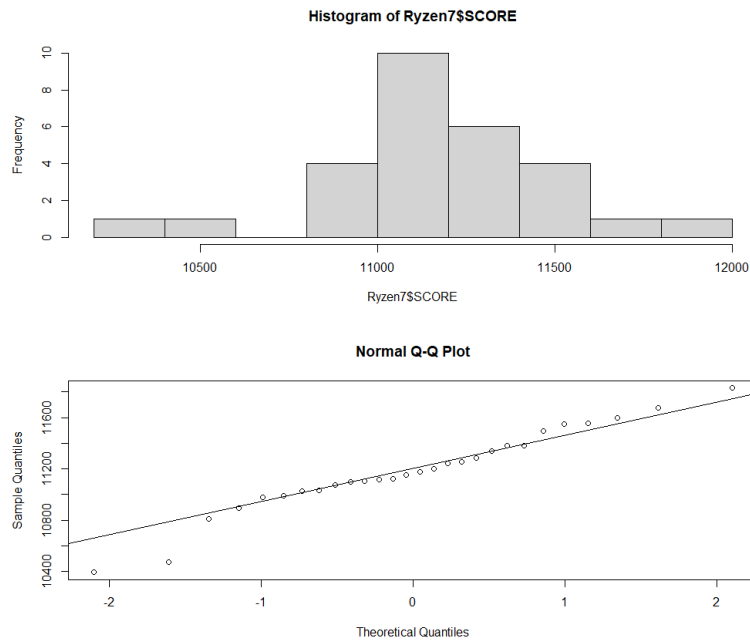
```
#Ryzen 7

par(mfrow=c(2,1))

hist(Ryzen7$Score)

qqnorm(Ryzen7$Score)

qqline(Ryzen7$Score)
```



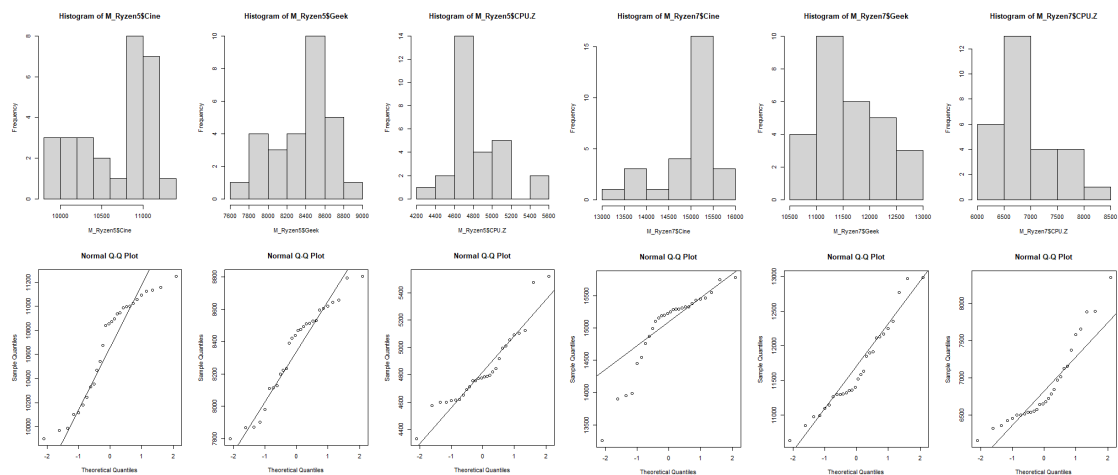
#Test "Normalidad"

```
>Shapiro.test(Ryzen5$Score) w = 0.94864, p-value = 0.1831
```

```
>Shapiro.test(Ryzen7$Score) w = 0.96546, p-value = 0.4655
```

#Ryzen 5/Benchmarks

#Ryzen 7/Benchmarks

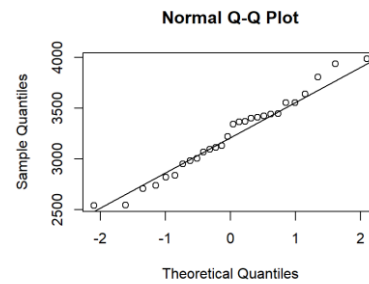
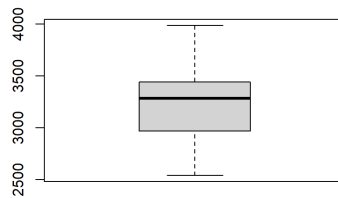
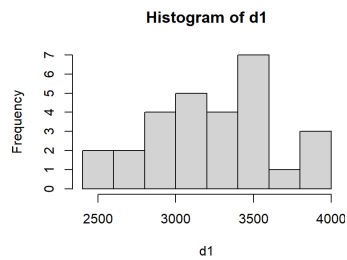


En estos gráficos Q-QPlot podemos ya acabar de ver el problema comentado con los *benchmarks*, y la diferencia que supone en la normalidad de los datos.

#OBJECTIU 1

#Ryzen 7 mejor que Ryzen 5?

```
d1 = Ryzen7$Score - Ryzen5$Score
T = mean(d1) / (sd(d1) / sqrt(length(d1)))
Pvalue = 2 * (1 - pnorm(t, 0, 1))
t.test(Ryzen7$Score, Ryzen5$Score, paired = "TRUE")
```



```
> t.test(d1)

One Sample t-test

data: d1
t = 44.156, df = 27, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 3080.306 3380.528
sample estimates:
mean of x
 3230.417

> shapiro.test(d1)

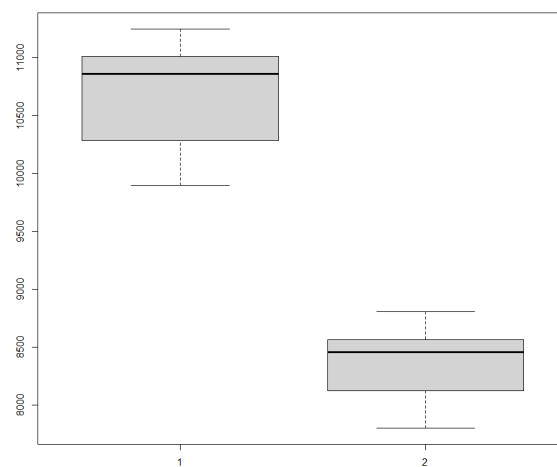
Shapiro-wilk normality test

data: d1
W = 0.97525, p-value = 0.7255
```

#OBJECTIU 2

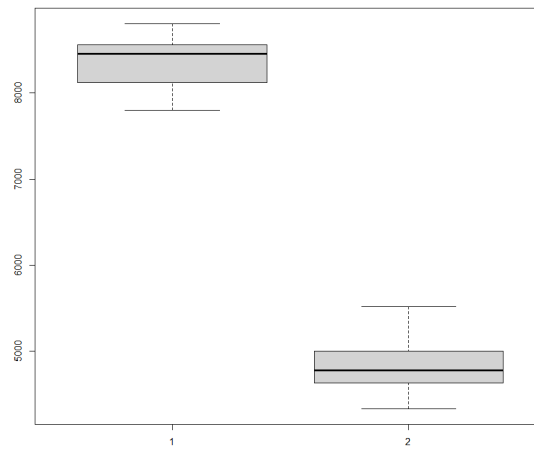
#El benchmark afecta en las puntuaciones obtenidas?

```
par(mfrow = c(1,1))
boxplot(Ryzen5$Cine, Ryzen5$Geek)
cor(Ryzen5$Cine, Ryzen5$Geek) -0.3095971
```



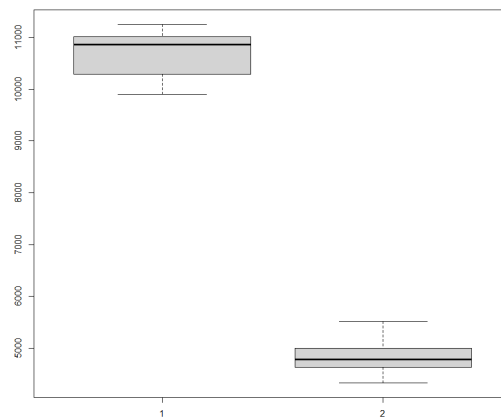
```
boxplot(Ryzen5$Geek, Ryzen5$CPU.Z)
```

```
cor(Ryzen5$Geek, Ryzen5$CPU.Z) -0.4711345
```



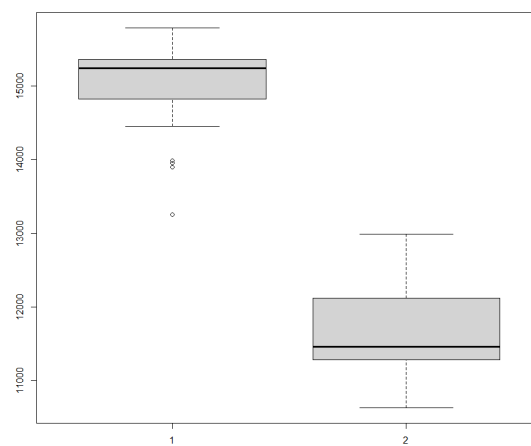
```
boxplot(Ryzen5$Cine, Ryzen5$CPU.Z)
```

```
cor(Ryzen5$Cine, Ryzen5$CPU.Z) 0.1550976
```



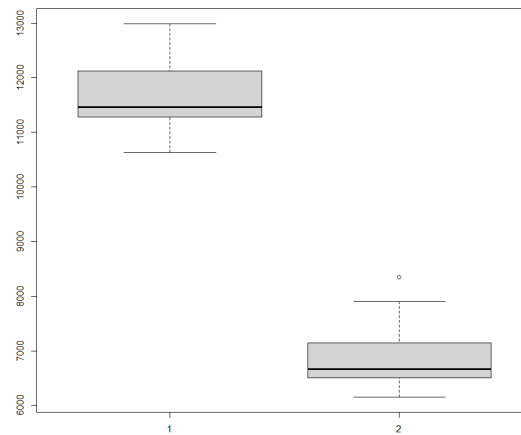
```
boxplot(Ryzen7$Cine, Ryzen7$Geek)
```

```
cor(Ryzen7$Cine, Ryzen7$Geek) -0.3095971
```



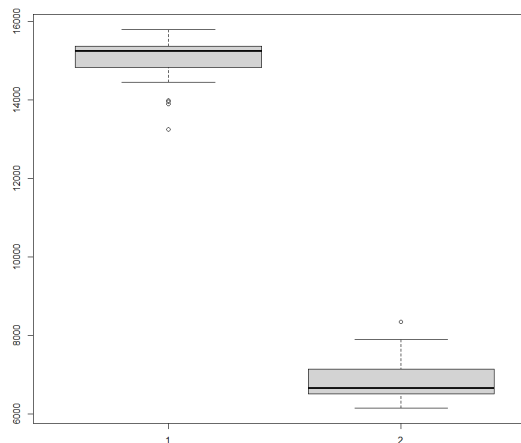
```
boxplot(Ryzen7$Geek, Ryzen7$CPU.Z)
```

```
cor(Ryzen7$Geek, Ryzen7$CPU.Z) -0.4711345
```



```
boxplot(Ryzen5$Cine, Ryzen5$CPU.Z)
```

```
cor(Ryzen5$Cine, Ryzen5$CPU.Z) 0.1550976
```



El objetivo principal del estudio era demostrar la diferencia de rendimiento de dos procesadores de distinta generación, pero como surgió el problema de las puntuaciones con los *benchmarks* decidimos además demostrar que las puntuaciones obtenidas vienen directamente relacionadas con el *benchmark* empleado en cada prueba, y esto tiene toda la lógica ya que cada *benchmark* pone a prueba una habilidad diferente del procesador y es adecuado suponer que no todos los procesadores presentan el mismo rendimiento en todos los aspectos. Como futuro estudio podría ser interesante estudiar como afectan las distintas pruebas al rendimiento.

ANEXO 2

Recogida de datos

Procesador	Cine	Geek	CPU-Z	SCORE
AMD Ryzen 5 5600X 3.7GHz	10102	8513	4597	7737,333
AMD Ryzen 5 5600X 3.7GHz	11003	8233	4772	8002,667
AMD Ryzen 5 5600X 3.7GHz	10467	8657	4601	7908,333
AMD Ryzen 5 5600X 3.7GHz	10243	7901	5104	7749,333
AMD Ryzen 5 5600X 3.7GHz	9987	8806	4576	7789,667
AMD Ryzen 5 5600X 3.7GHz	10115	8794	4920	7943
AMD Ryzen 5 5600X 3.7GHz	10541	8111	4776	7809,333
AMD Ryzen 5 5600X 3.7GHz	10872	8439	5012	8107,667
AMD Ryzen 5 5600X 3.7GHz	10353	8201	5521	8025
AMD Ryzen 5 5600X 3.7GHz	9970	8532	4332	7611,333
AMD Ryzen 5 5600X 3.7GHz	11124	7867	5475	8155,333
AMD Ryzen 5 5600X 3.7GHz	10180	8608	4994	7927,333
AMD Ryzen 5 5600X 3.7GHz	10329	8115	4823	7755,667
AMD Ryzen 5 5600X 3.7GHz	10675	7981	5091	7915,667
AMD Ryzen 5 5600X 3.7GHz	10988	8421	4784	8064,333
AMD Ryzen 5 5600X 3.7GHz	11132	8595	4612	8113
AMD Ryzen 5 5600X 3.7GHz	10852	8622	4712	8062
AMD Ryzen 5 5600X 3.7GHz	10945	8225	4845	8005
AMD Ryzen 5 5600X 3.7GHz	11250	8129	4651	8010
AMD Ryzen 5 5600X 3.7GHz	11054	8494	4755	8101
AMD Ryzen 5 5600X 3.7GHz	11095	7799	5123	8005,667
AMD Ryzen 5 5600X 3.7GHz	10897	8529	4797	8074,333
AMD Ryzen 5 5600X 3.7GHz	9899	8469	4619	7662,333
AMD Ryzen 5 5600X 3.7GHz	10996	7872	4691	7853
AMD Ryzen 5 5600X 3.7GHz	10841	8512	4756	8036,333
AMD Ryzen 5 5600X 3.7GHz	10938	8644	5054	8212
AMD Ryzen 5 5600X 3.7GHz	11156	8476	4788	8140
AMD Ryzen 5 5600X 3.7GHz	11026	8392	4615	8011
AMD Ryzen 7 5800X 3.8GHz	15300	12356	7372	11676
AMD Ryzen 7 5800X 3.8GHz	14545	12965	7158	11556
AMD Ryzen 7 5800X 3.8GHz	15155	11586	7899	11546,67
AMD Ryzen 7 5800X 3.8GHz	13254	11846	8354	11151,33
AMD Ryzen 7 5800X 3.8GHz	15453	11632	6512	11199
AMD Ryzen 7 5800X 3.8GHz	15789	11915	6785	11496,33
AMD Ryzen 7 5800X 3.8GHz	14452	12986	6318	11252
AMD Ryzen 7 5800X 3.8GHz	13954	12112	7655	11240,33
AMD Ryzen 7 5800X 3.8GHz	15753	12765	6971	11829,67
AMD Ryzen 7 5800X 3.8GHz	15386	11519	7891	11598,67
AMD Ryzen 7 5800X 3.8GHz	14996	10846	6843	10895
AMD Ryzen 7 5800X 3.8GHz	13899	10974	6554	10475,67
AMD Ryzen 7 5800X 3.8GHz	15108	11899	6354	11120,33
AMD Ryzen 7 5800X 3.8GHz	15198	12171	6645	11338

AMD Ryzen 7 5800X 3.8GHz	15228	11298	7012	11179,33
AMD Ryzen 7 5800X 3.8GHz	15312	11365	6681	11119,33
AMD Ryzen 7 5800X 3.8GHz	15195	12125	6535	11285
AMD Ryzen 7 5800X 3.8GHz	15554	11324	6426	11101,33
AMD Ryzen 7 5800X 3.8GHz	15465	11095	7581	11380,33
AMD Ryzen 7 5800X 3.8GHz	14873	11403	6154	10810
AMD Ryzen 7 5800X 3.8GHz	15252	10994	6724	10990
AMD Ryzen 7 5800X 3.8GHz	15334	11295	6452	11027
AMD Ryzen 7 5800X 3.8GHz	15432	11354	6537	11107,67
AMD Ryzen 7 5800X 3.8GHz	13985	10633	6571	10396,33
AMD Ryzen 7 5800X 3.8GHz	14765	12250	7123	11379,33
AMD Ryzen 7 5800X 3.8GHz	15294	11305	6498	11032,33
AMD Ryzen 7 5800X 3.8GHz	15289	11145	6502	10978,67
AMD Ryzen 7 5800X 3.8GHz	15328	11268	6637	11077,67