

***Datos experimento compresión y
descompresión Libro “Quijote de la
Mancha” grupo 5.***

Integrantes: Felipe Barrientos.

Roberto Verdugo.

Felipe Parra.

Martin Pizarro.

Oroi Teave.

Tabla de datos

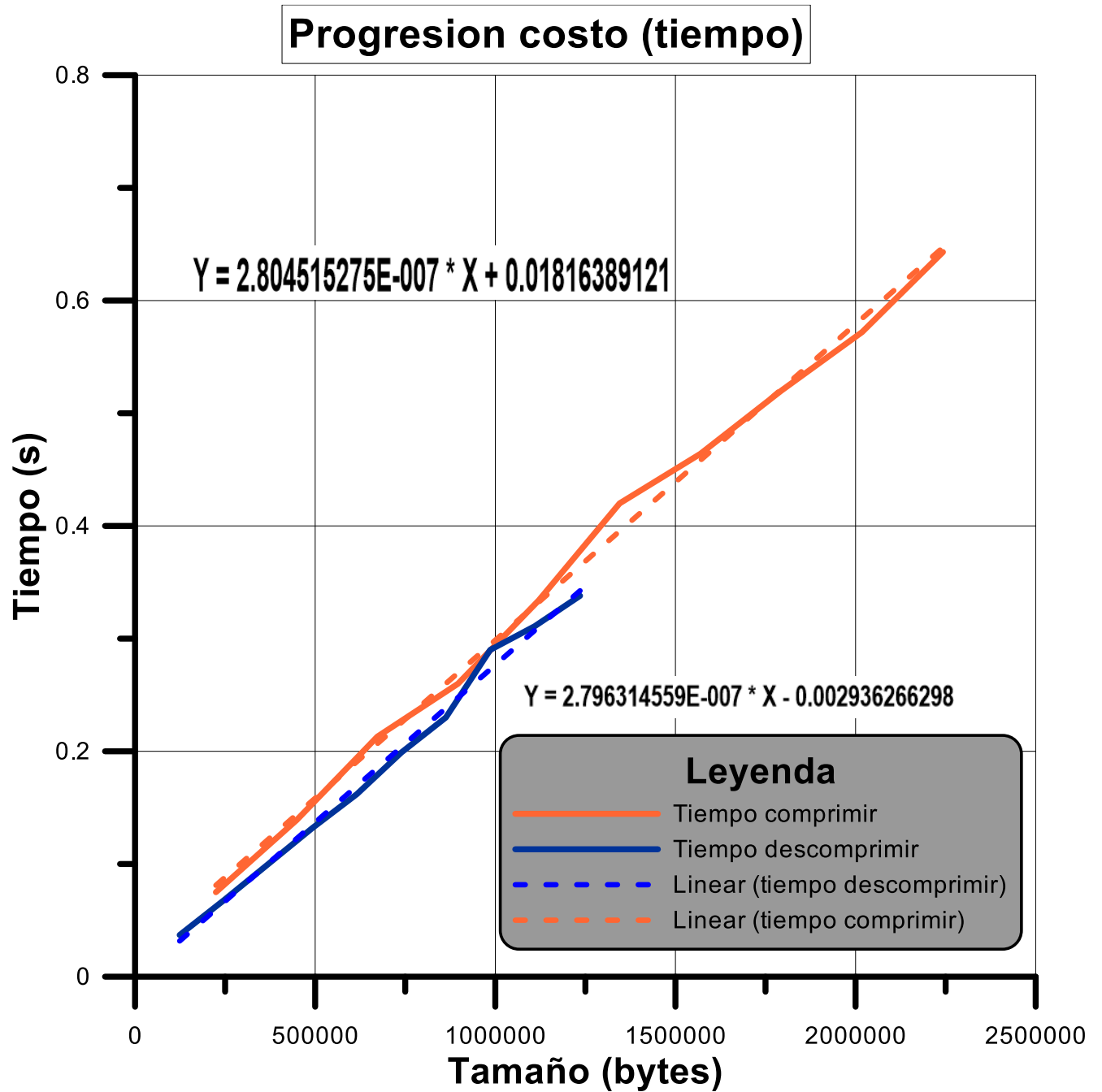
Tabla utilitario creado

Parte	Tamaño (bytes)	Tiempo en comprimir (s)	Tamaño comprimido (bytes)	Tiempo en descomprimir (s)	Tasa de compresión
1	224449	0.075	123869	0.037	%55.18
2	448606	0.139	247118	0.068	%55.08
3	672603	0.213	369908	0.100	%54.99
4	896489	0.260	492560	0.132	%54.94
5	1120652	0.334	615680	0.162	%54.93
6	1345012	0.420	739412	0.199	%54.97
7	1569331	0.464	862622	0.230	%54.96
8	1793664	0.520	985903	0.290	%54.96
9	2018074	0.572	1109388	0.311	%54.97
10	2242524	0.643	1235190	0.338	%55.08

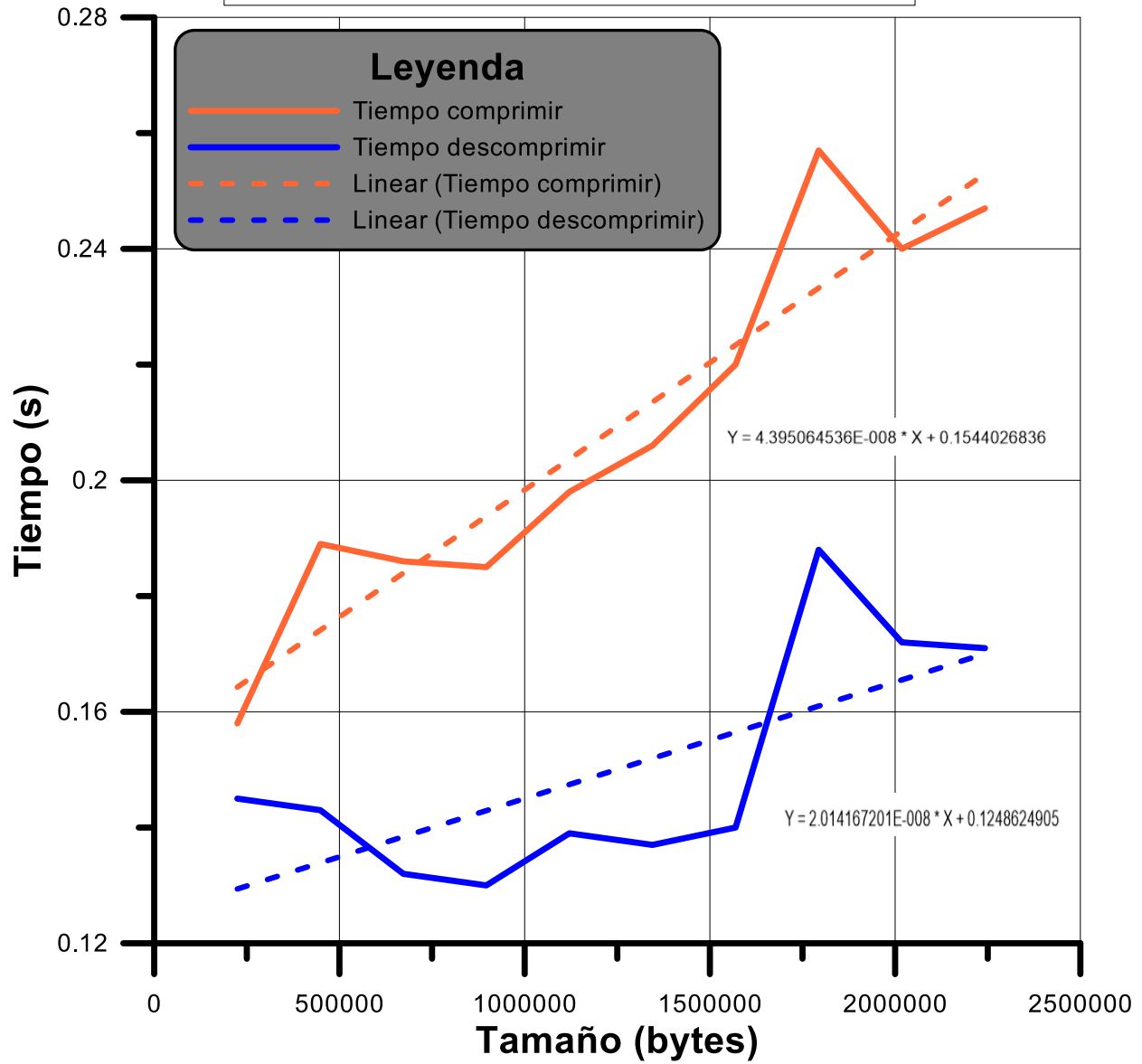
Tabla programa Winrar

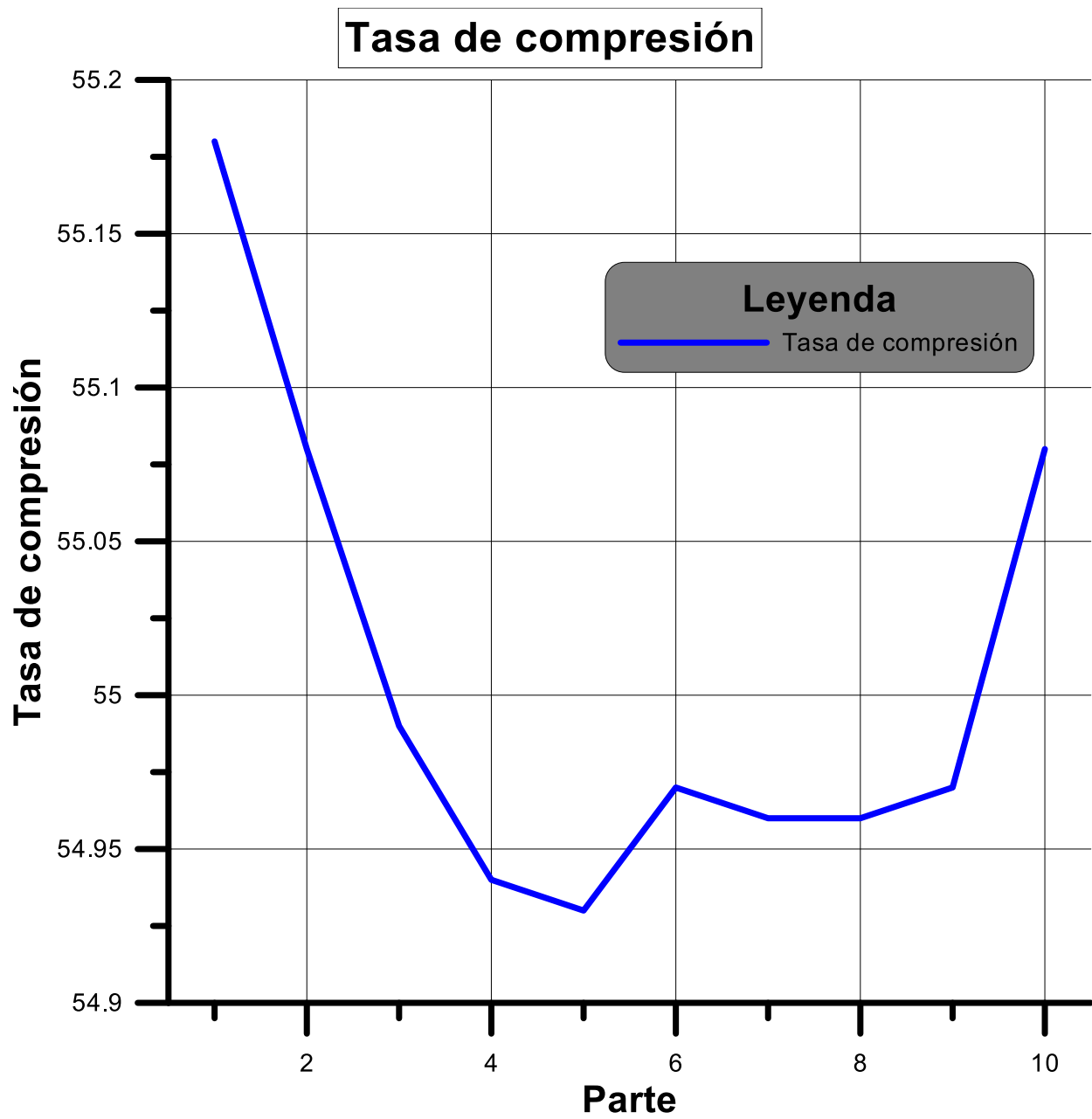
Parte	Tamaño (bytes)	Tiempo en comprimir Winrar (s)	Tamaño comprimido (bytes)	Tiempo en descomprimir Winrar (s)	Tasa de compresión
1	224449	0.158	77843	0.145	%34.00
2	448606	0.189	147779	0.143	%32.00
3	672603	0.186	215841	0.132	%32.00
4	896489	0.185	283350	0.130	%31.00
5	1120652	0.198	350778	0.139	%31.00
6	1345012	0.206	417890	0.137	%31.00
7	1569331	0.220	484806	0.140	%30.00
8	1793664	0.257	551720	0.188	%30.00
9	2018074	0.240	617564	0.172	%30.00
10	2242524	0.247	684308	0.171	%30.00

Gráficos

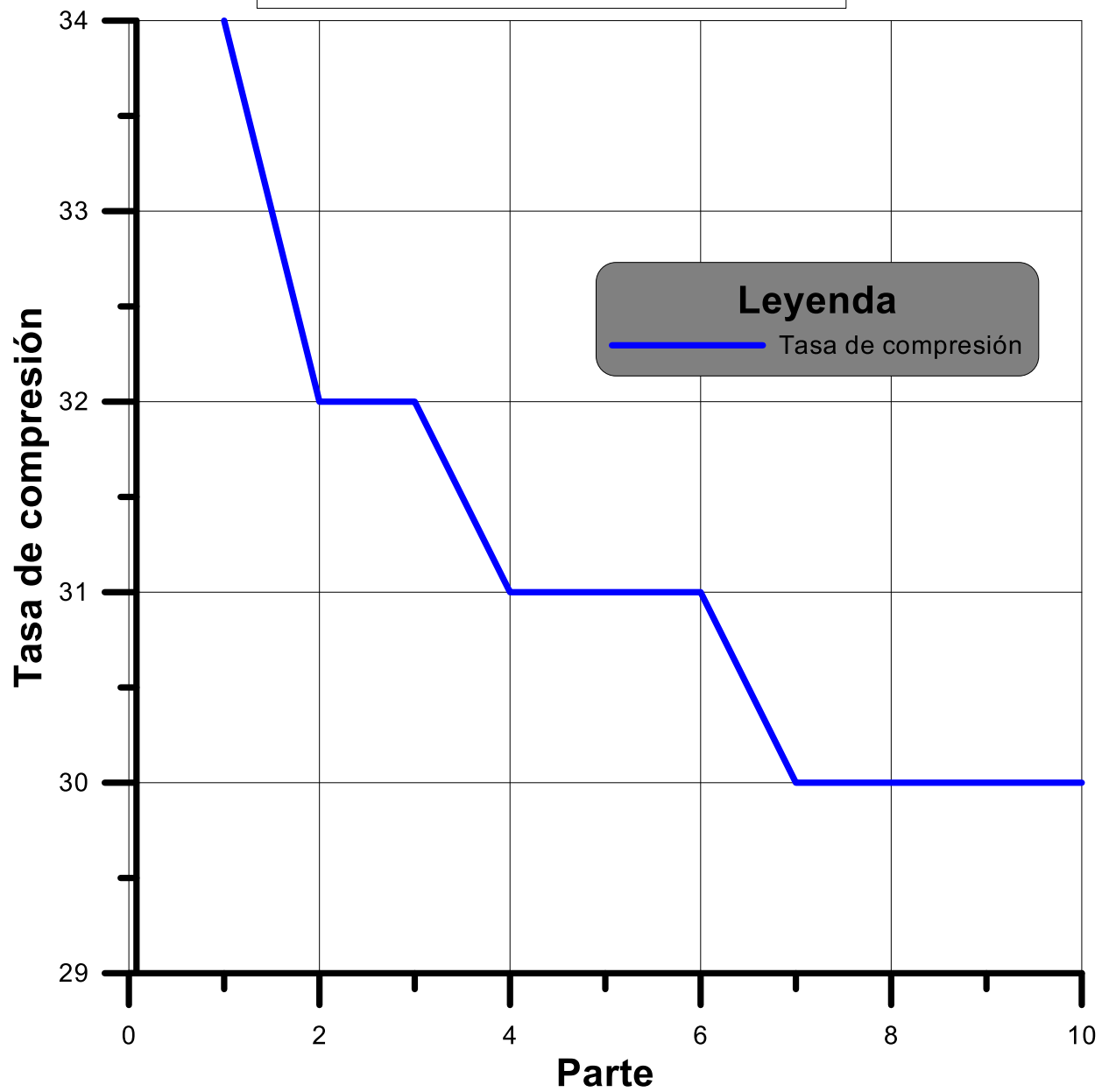


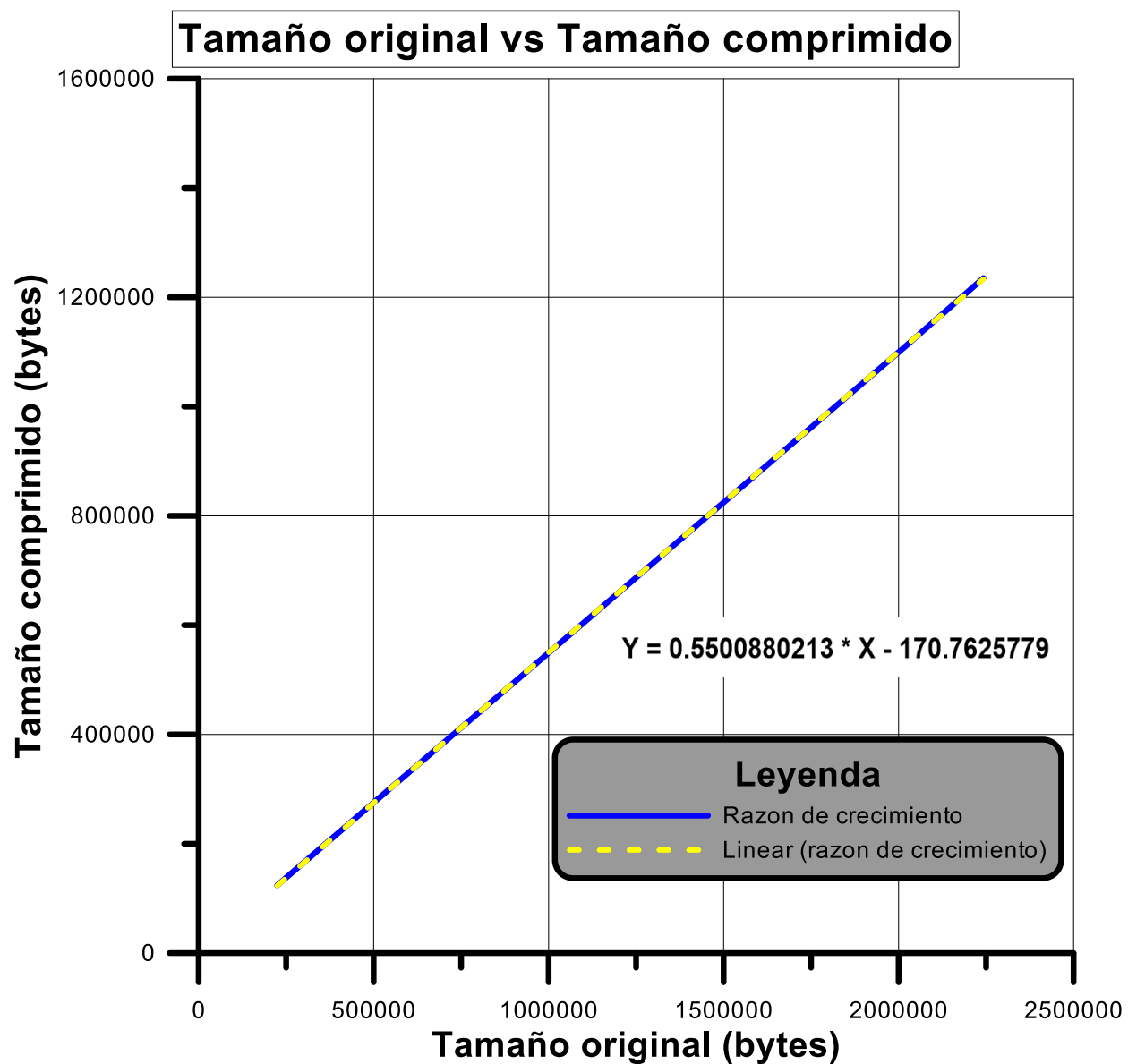
Progresion costo (tiempo) Winrar

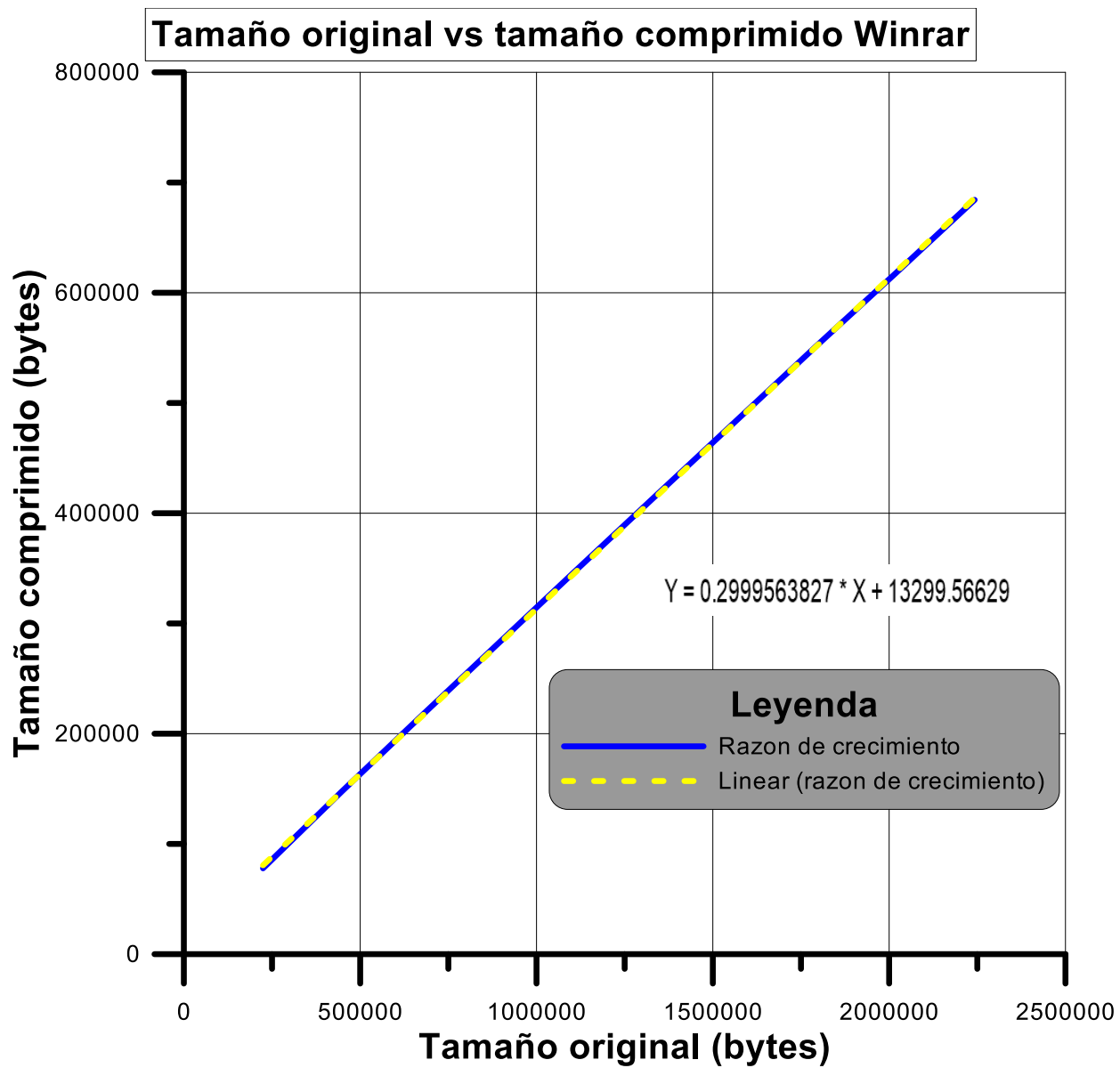




Tasa de compresión Winrar







Análisis de gráficos.

Grafico progresión costo (utilitario).

Analizando el algoritmo y los tiempos creemos que la diferencia de tiempo entre comprimir y descomprimir se debe a que en la función comprimir debemos realizar la lectura entera del archivo .txt dos veces, la primera para crear la tabla y la segunda al momento de reescribir el texto en binario con el código de huffman.

Grafico tasa de compresión (utilitario).

Analizando el grafico y conociendo que la eficacia del algoritmo de huffman reside la cardinalidad de caracteres y la variación de la frecuencia de ellas, se puede especular que desde la parte 1 a 5 se iba encontrando caracteres nuevos y que la diferencia entre las frecuencias era relativamente baja, pero desde la parte 5 a 10 las frecuencias de los caracteres se comenzaron a desigualar para algunos casos.

Grafico tamaño original vs tamaño comprimido (utilitario).

Analizando los datos entregados por el grafico se tiene que la pendiente de la gráfica está determinada por la tasa de crecimiento que el utilitario genera, también se puede decir que al ser una recta bastante definida (sin picos) se puede decir que la cantidad de bytes está bien distribuida, cumpliendo así, la distribución de los décimos pedida.