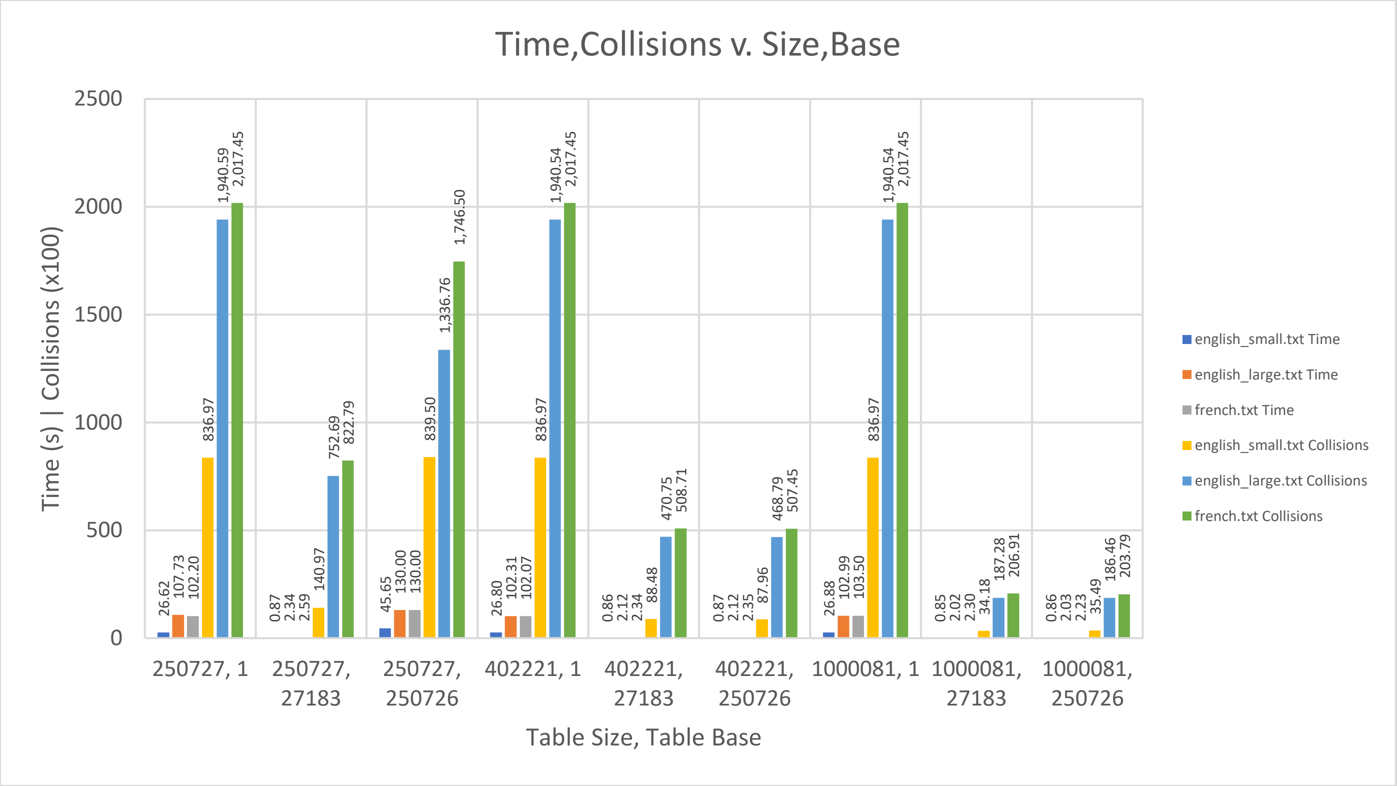
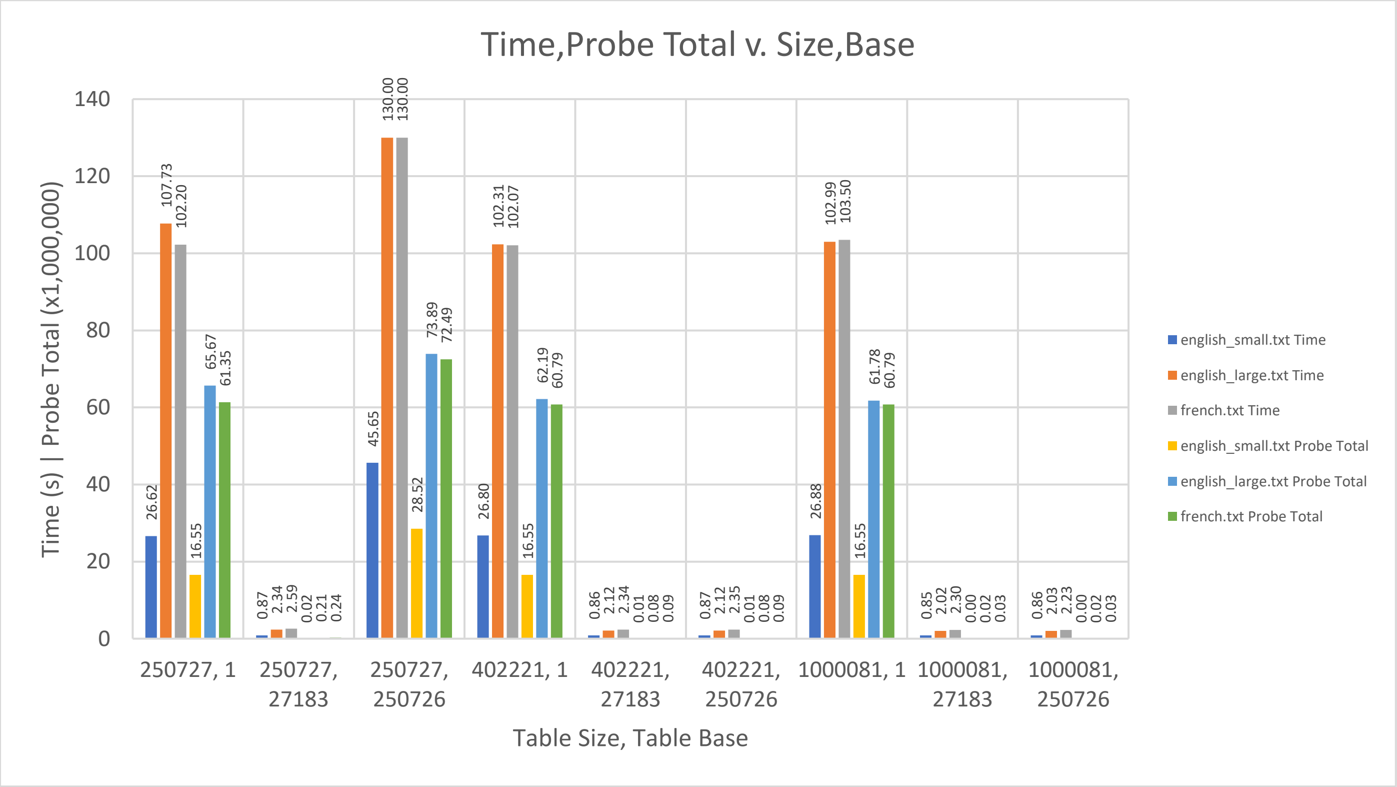
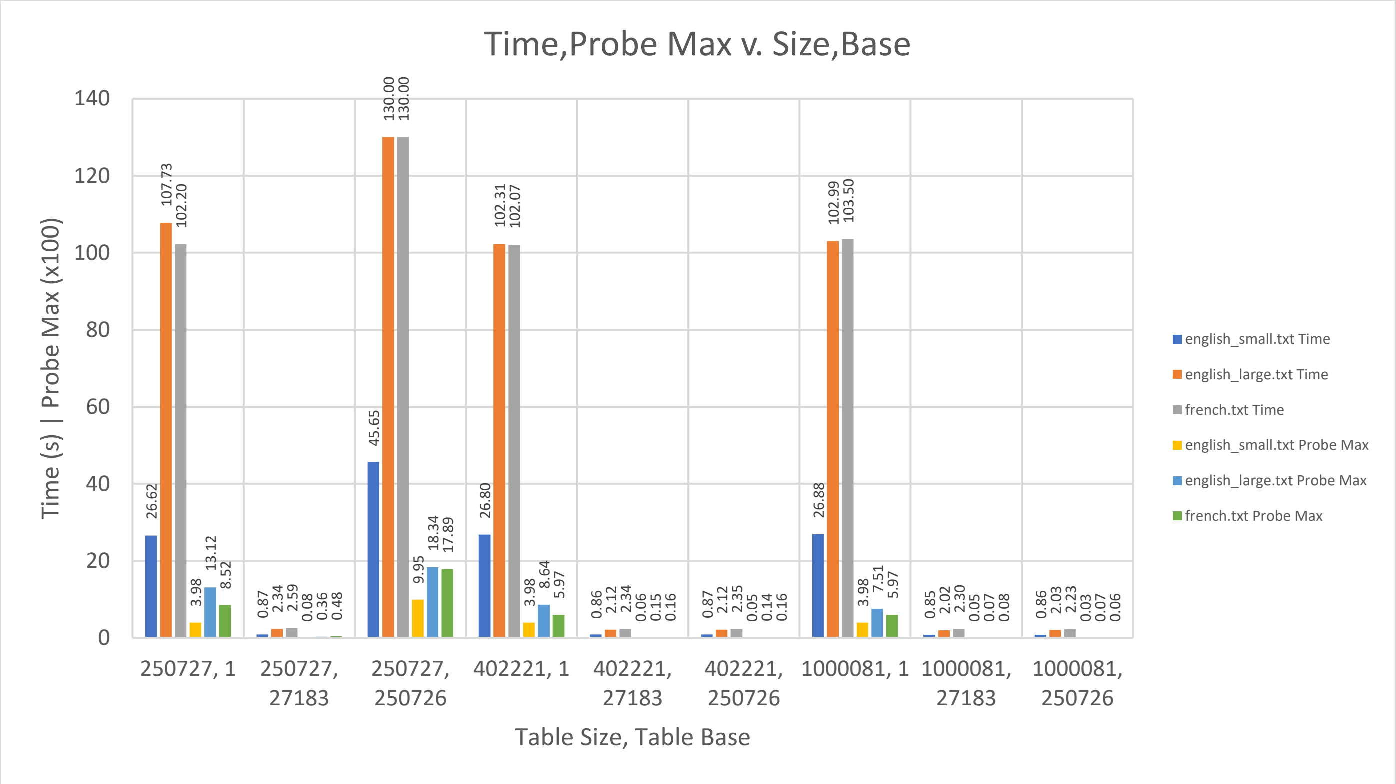
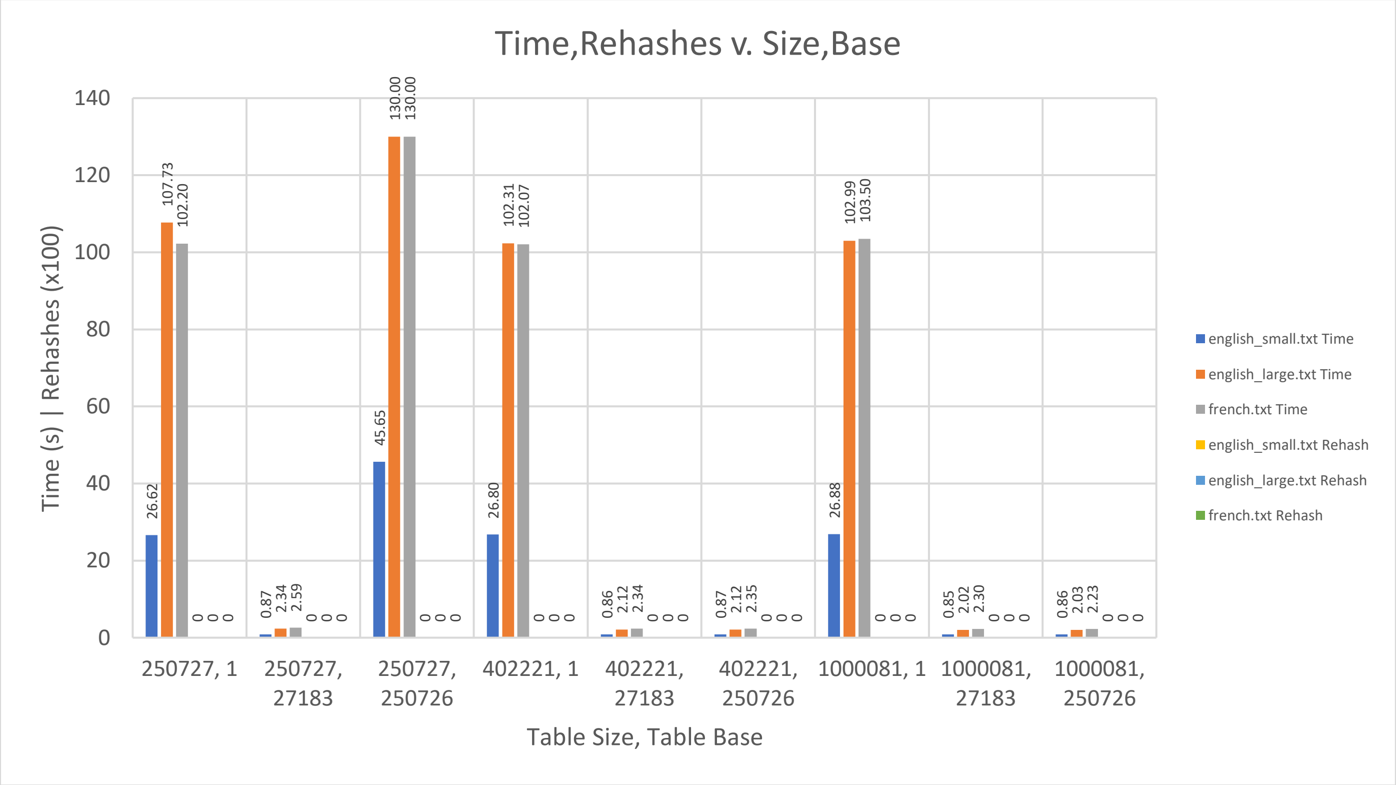
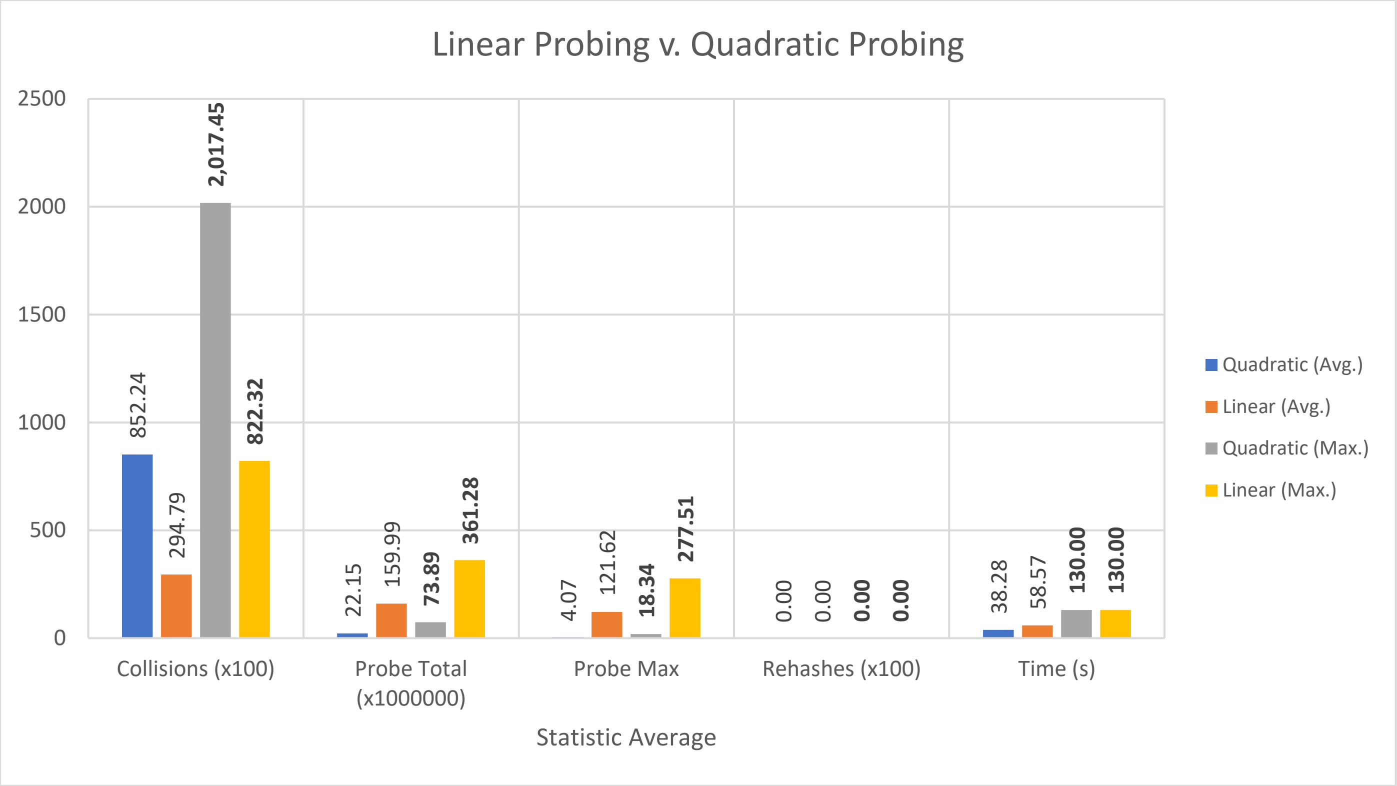

Time,Collisions v. Size,Base
Time (s) | Collisions (x100)
2500
2000
1500
1000
500
0
26.62
107.73
102.20
836.97
1,940.59
2,017.45
0.87
2.34
2.59
140.97
752.69
822.79
45.65
130.00
130.00
839.50
1,336.76
1,746.50
26.80
102.31
102.07
836.97
1,940.54
2,017.45
0.86
2.12
2.34
88.48
470.75
508.71
0.87
2.12
2.35
87.96
468.79
507.45
26.88
102.99
103.50
836.97
1,940.54
2,017.45
0.85
2.02
2.30
34.18
187.28
206.91
0.86
2.03
2.23
35.49
186.46
203.79
250727, 1
250727, 27183
250727, 250726
402221, 1
402221, 27183
402221, 250726
1000081, 1
1000081, 27183
1000081, 250726
Table Size, Table Base
english_small.txt Time
english_large.txt Time
french.txt Time
english_small.txt Collisions
english_large.txt Collisions
french.txt Collisions


Time,Probe Total v. Size,Base
Time (s) | Probe Total (x1,000,000)
140
120
100
80
60
40
20
0
26.62
107.73
102.20
16.55
65.67
61.35
0.87
2.34
2.59
0.02
0.21
0.24
45.65
130.00
130.00
28.52
73.89
72.49
26.80
102.31
102.07
16.55
62.19
60.79
0.86
2.12
2.34
0.01
0.08
0.09
0.87
2.12
2.35
0.01
0.08
0.09
26.88
102.99
103.50
16.55
61.78
60.79
0.85
2.02
2.30
0.00
0.02
0.03
0.86
2.03
2.23
0.00
0.02
0.03
250727, 1
250727, 27183
250727, 250726
402221, 1
402221, 27183
402221, 250726
1000081, 1
1000081, 27183
1000081, 250726
Table Size, Table Base
english_small.txt Time
english_large.txt Time
french.txt Time
english_small.txt Probe Total
english_large.txt Probe Total
french.txt Probe Total


Time,Probe Max v. Size,Base
Time (s) | Probe Max (x100)
140
120
100
80
60
40
20
0
26.62
107.73
102.20
3.98
13.12
8.52
0.87
2.34
2.59
0.08
0.36
0.48
45.65
130.00
130.00
9.95
18.34
17.89
26.80
102.31
102.07
3.98
8.64
5.97
0.86
2.12
2.34
0.06
0.15
0.16
0.87
2.12
2.35
0.05
0.14
0.16
26.88
102.99
103.50
3.98
7.51
5.97
0.85
2.02
2.30
0.05
0.07
0.08
0.86
2.03
2.23
0.03
0.07
0.06
250727, 1
250727, 27183
250727, 250726
402221, 1
402221, 27183
402221, 250726
1000081, 1
1000081, 27183
1000081, 250726
Table Size, Table Base
english_small.txt Time
english_large.txt Time
french.txt Time
english_small.txt Probe Max
english_large.txt Probe Max
french.txt Probe Max


Time,Rehashes v. Size,Base
Time (s) | Rehashes (x100)
140
120
100
80
60
40
20
0
26.62
107.73
102.20
0
0
0
0.87
2.34
2.59
0
0
0
45.65
130.00
130.00
0
0
0
26.80
102.31
102.07
0
0
0
0.86
2.12
2.34
0
0
0
0.87
2.12
2.35
0
0
0
26.88
102.99
103.50
0
0
0
0.85
2.02
2.30
0
0
0
0.86
2.03
2.23
0
0
0
250727, 1
250727, 27183
250727, 250726
402221, 1
402221, 27183
402221, 250726
1000081, 1
1000081, 27183
1000081, 250726
Table Size, Table Base
english_small.txt Time
english_large.txt Time
french.txt Time
english_small.txt Rehash
english_large.txt Rehash
french.txt Rehash



- max_time = 120

- As already stated in the explanation of task 3, high collision counts do not directly affect the performance of the hash table but simply indicate that poor performance may be observed.

- Again, the better indicators of performance are the Probe Total and Probe Max statistics. Especially Probe Total, which keeps track of total length probed, this is the most performance taxing aspect of the hash table. Thus, reducing the Probe Total and Probe Max statistics would lead to better timed performance.

- When comparing linear and quadratic probing it is clear that quadratic probing is a faster implementation with lower averaged times and much lower average as well as maximum Probe Total and Probe Max values. This means that less time is spent probing the hash table.

- It is important to note that linear probing resulted in 12 time-outs whilst quadratic probing only resulted in 2 time-outs. This is reflected in the average and maximum collision statistics. Quadratic probing completed 10 more combinations – and would've progressed further in the remaining two – than linear probing, explaining why it observed more collisions on average.

- All combinations executed here did not require a rehash as even the smallest hash table (size 250727) was larger than the maximum unique words observed (length 202358 – french.txt). With quadratic probing, the hash table will resize when the quadratic probe path cannot find an empty slot for the new key-value pair. This did not occur in any instances and thus no rehashes took place.

Overall, probe max and probe total remain the best indicators for performance. The effectively communicate the performance benefits – in the form of smaller clusters and less clustering – of quadratic probing over linear probing as well.