

Assignment 6

Question 1:

For the K80:

```
nvcc -arch=compute_37 -code=sm_37 -Xcompiler -fopenmp -lcuda question2_p1_skj64.cu -o  
question2_p1_skj64  
nvcc -arch=compute_37 -code=sm_37 -Xcompiler -fopenmp -lcuda question2_p2_skj64.cu -o  
question2_p2_skj64
```

For the P100:

```
nvcc -arch=compute_60 -code=sm_60 -Xcompiler -fopenmp -lcuda question2_p1_skj64.cu -o  
question2_p1_skj64  
nvcc -arch=compute_60 -code=sm_60 -Xcompiler -fopenmp -lcuda question2_p1_skj64.cu -o  
question2_p2_skj64
```

For the V100:

```
nvcc -arch=compute_70 -code=sm_70 -Xcompiler -fopenmp -lcuda question2_p1_skj64.cu -o  
question2_p1_skj64  
nvcc -arch=compute_70 -code=sm_70 -Xcompiler -fopenmp -lcuda question2_p2_skj64.cu -o  
question2_p2_skj64
```

*So the highlighted lines are the ones I used to compile my program.

Question 2:

Part 1 Outputs:Brute Force

These are the 3 outputs for epsilon = 5 and N=100 for a P100 GPU:

```
Size of dataset (MiB): 0.001526  
Total time to transfer from CPU to GPU(s): 0.128923  
total blocks: 1  
Total time for GPU Kernel(s): 0.000065  
Total number of points within epsilon (GPU): 100  
Total time (s): 0.129017  
Total transfer time (s): 0.128942
```

```
Size of dataset (MiB): 0.001526  
Total time to transfer from CPU to GPU(s): 0.128750  
total blocks: 1  
Total time for GPU Kernel(s): 0.000064  
Total number of points within epsilon (GPU): 100  
Total time (s): 0.128842  
Total transfer time (s): 0.128769
```

```
Size of dataset (MiB): 0.001526
Total time to transfer from CPU to GPU(s): 0.126877
total blocks: 1
Total time for GPU Kernel(s): 0.000064
Total number of points within epsilon (GPU): 100
Total time (s): 0.126969
Total transfer time (s): 0.126895
```

The average kernel time is: 0.0000643

The average total time is: 0.128276

These are the 3 outputs for epsilon = 5 and N=1000 for a P100 GPU:

```
Size of dataset (MiB): 0.015259
Total time to transfer from CPU to GPU(s): 0.124229
total blocks: 1
Total time for GPU Kernel(s): 0.000549
Total number of points within epsilon (GPU): 1102
Total time (s): 0.124807
Total transfer time (s): 0.124248
```

```
Size of dataset (MiB): 0.015259
Total time to transfer from CPU to GPU(s): 0.123923
total blocks: 1
Total time for GPU Kernel(s): 0.000546
Total number of points within epsilon (GPU): 1102
Total time (s): 0.124497
Total transfer time (s): 0.123942
```

```
Size of dataset (MiB): 0.015259
Total time to transfer from CPU to GPU(s): 0.124359
total blocks: 1
Total time for GPU Kernel(s): 0.000544
Total number of points within epsilon (GPU): 1102
Total time (s): 0.124932
Total transfer time (s): 0.124378
```

The average kernel time is: 0.0005463

The average total time is: 0.1247453

These are the 3 outputs for epsilon = 5 and N=10000 for a P100 GPU:

```
Size of dataset (MiB): 0.152588
Total time to transfer from CPU to GPU(s): 0.126614
total blocks: 10
Total time for GPU Kernel(s): 0.005094
Total number of points within epsilon (GPU): 17644
Total time (s): 0.131737
Total transfer time (s): 0.126633
```

```
Size of dataset (MiB): 0.152588
Total time to transfer from CPU to GPU(s): 0.128548
total blocks: 10
Total time for GPU Kernel(s): 0.005098
Total number of points within epsilon (GPU): 17644
Total time (s): 0.133674
Total transfer time (s): 0.128566
```

```
Size of dataset (MiB): 0.152588
Total time to transfer from CPU to GPU(s): 0.123765
total blocks: 10
Total time for GPU Kernel(s): 0.005097
Total number of points within epsilon (GPU): 17644
Total time (s): 0.128890
Total transfer time (s): 0.123784
```

The average kernel time is: 0.0050963

The average total time is: 0.1314336

These are the 3 outputs for epsilon = 5 and N=100000 for a P100 GPU:

```
Size of dataset (MiB): 1.525879
Total time to transfer from CPU to GPU(s): 0.124197
total blocks: 98
Total time for GPU Kernel(s): 0.099994
Total number of points within epsilon (GPU): 879688
Total time (s): 0.224223
Total transfer time (s): 0.124218
```

```
Size of dataset (MiB): 1.525879
Total time to transfer from CPU to GPU(s): 0.124277
total blocks: 98
Total time for GPU Kernel(s): 0.100009
Total number of points within epsilon (GPU): 879688
Total time (s): 0.224316
Total transfer time (s): 0.124298
```

```
Size of dataset (MiB): 1.525879
Total time to transfer from CPU to GPU(s): 0.125945
total blocks: 98
Total time for GPU Kernel(s): 0.100064
Total number of points within epsilon (GPU): 879688
Total time (s): 0.226040
Total transfer time (s): 0.125967
```

The average kernel time is: 0.1000223

The average total time is: 0.2248596

These are three outputs for epsilon = 5 and N=100 for a CPU with 4 cores with O3:

```
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question1resultsforn100 5

Size of dataset (MiB): 0.001526
Total Count of Distance less than Eps: 100
Total time (s): 0.000214
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question1resultsforn100 5

Size of dataset (MiB): 0.001526
Total Count of Distance less than Eps: 100
Total time (s): 0.000508
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question1resultsforn100 5

Size of dataset (MiB): 0.001526
Total Count of Distance less than Eps: 100
Total time (s): 0.001657
```

Total average:0.000793

These are three outputs for epsilon = 5 and N=1000 for a CPU with 4 cores with O3:

```
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question1resultsforn1000 5

Size of dataset (MiB): 0.015259
Total Count of Distance less than Eps: 1102
Total time (s): 0.006556
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question1resultsforn1000 5

Size of dataset (MiB): 0.015259
Total Count of Distance less than Eps: 1102
Total time (s): 0.005135
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question1resultsforn1000 5

Size of dataset (MiB): 0.015259
Total Count of Distance less than Eps: 1102
Total time (s): 0.004960
```

Total Average:0.0055503

These are three outputs for epsilon = 5 and N=10000 for a CPU with 4 cores with O3:

```
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelF
64$ ./question1resultsforn10000 5

Size of dataset (MiB): 0.152588
Total Count of Distance less than Eps: 17644
Total time (s): 0.275809
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelF
64$ ./question1resultsforn10000 5

Size of dataset (MiB): 0.152588
Total Count of Distance less than Eps: 17644
Total time (s): 0.270894
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelF
64$ ./question1resultsforn10000 5

Size of dataset (MiB): 0.152588
Total Count of Distance less than Eps: 17644
Total time (s): 0.280864
```

Total Average:0.2758557

These are three outputs for epsilon = 5 and N=100000 for a CPU with 4 cores with O3:

```
64$ ./question1resultsforn100000 5

Size of dataset (MiB): 1.525879
Total Count of Distance less than Eps: 879688
Total time (s): 26.998895
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelF
64$ ./question1resultsforn100000 5

Size of dataset (MiB): 1.525879
Total Count of Distance less than Eps: 879688
Total time (s): 26.976587
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelF
64$ ./question1resultsforn100000 5

Size of dataset (MiB): 1.525879
Total Count of Distance less than Eps: 879688
Total time (s): 26.829187
```

Total Average: 26.9348896667

Answer for Question 1 Part 1:

So over all we can see that the time it took to do the actual calculations in the kernel for the GPU is a lot faster than the average total time for our CPU calculations from the data above. Although some of the total times were similar to the CPU's for a small N value. So they were similar when calculating the distances between less amounts of points. Although when we get a large amount of points such as N=100000 the GPU is much faster than the CPU as you can see from the data above. The total time average for N=100000 and epsilon = 5 for the GPU was 0.2248596 secs while the CPU average total time was 26.9348896667 secs. I found this

interesting but I could kind of understand why the GPU is much faster with a large amount of points. My understanding is based on the idea that in order for the graphics to be presented on the screens we see today they need to make lots of calculations with large amounts of data to be presented to us and very quickly.

Part 2 Outputs: Improved Solution

These are the 3 outputs for epsilon = 5 and N=100 for a P100 GPU:

```
Size of dataset (MiB): 0.001526
Total time to transfer from CPU to GPU(s): 0.123079
total blocks: 1
Total time for GPU Kernel(s): 0.000061
Total number of points within epsilon (GPU): 100
Total time (s): 0.123168
Total transfer time (s): 0.123098
```

```
Size of dataset (MiB): 0.001526
Total time to transfer from CPU to GPU(s): 0.128972
total blocks: 1
Total time for GPU Kernel(s): 0.000061
Total number of points within epsilon (GPU): 100
Total time (s): 0.129062
Total transfer time (s): 0.128991
```

```
Size of dataset (MiB): 0.001526
Total time to transfer from CPU to GPU(s): 0.123681
total blocks: 1
Total time for GPU Kernel(s): 0.000060
Total number of points within epsilon (GPU): 100
Total time (s): 0.123771
Total transfer time (s): 0.123700
```

The average kernel time is: 0.0000606

The average total time is: 0.1253336

These are the 3 outputs for epsilon = 5 and N=1000 for a P100 GPU:

```
Size of dataset (MiB): 0.015259
Total time to transfer from CPU to GPU(s): 0.122816
total blocks: 1
Total time for GPU Kernel(s): 0.000898
Total number of points within epsilon (GPU): 1102
Total time (s): 0.123742
Total transfer time (s): 0.122834
```

```
Size of dataset (MiB): 0.015259
Total time to transfer from CPU to GPU(s): 0.123394
total blocks: 1
Total time for GPU Kernel(s): 0.000903
Total number of points within epsilon (GPU): 1102
Total time (s): 0.124326
Total transfer time (s): 0.123413
```

```
Size of dataset (MiB): 0.015259
Total time to transfer from CPU to GPU(s): 0.141833
total blocks: 1
Total time for GPU Kernel(s): 0.000899
Total number of points within epsilon (GPU): 1102
Total time (s): 0.142761
Total transfer time (s): 0.141852
```

The average kernel time is:0.0009

The average total time is: 0.1302763

These are the 3 outputs for epsilon = 5 and N=10000 for a P100 GPU:

```
Size of dataset (MiB): 0.152588
Total time to transfer from CPU to GPU(s): 0.123517
total blocks: 10
Total time for GPU Kernel(s): 0.016552
Total number of points within epsilon (GPU): 17644
Total time (s): 0.140099
Total transfer time (s): 0.123537
```

```
Size of dataset (MiB): 0.152588
Total time to transfer from CPU to GPU(s): 0.123171
total blocks: 10
Total time for GPU Kernel(s): 0.016564
Total number of points within epsilon (GPU): 17644
Total time (s): 0.139766
Total transfer time (s): 0.123192
```

```
Size of dataset (MiB): 0.152588
Total time to transfer from CPU to GPU(s): 0.126949
total blocks: 10
Total time for GPU Kernel(s): 0.016529
Total number of points within epsilon (GPU): 17644
Total time (s): 0.143509
Total transfer time (s): 0.126969
```

The average kernel time is:0.0165483

The average total time is: 0.1411246

These are the 3 outputs for epsilon = 5 and N=100000 for a P100 GPU:

```
Size of dataset (MiB): 1.525879
Total time to transfer from CPU to GPU(s): 0.127837
total blocks: 98
Total time for GPU Kernel(s): 0.245257
Total number of points within epsilon (GPU): 879688
Total time (s): 0.373125
Total transfer time (s): 0.127858
```

```
Size of dataset (MiB): 1.525879
Total time to transfer from CPU to GPU(s): 0.123908
total blocks: 98
Total time for GPU Kernel(s): 0.245225
Total number of points within epsilon (GPU): 879688
Total time (s): 0.369164
Total transfer time (s): 0.123929
```

```
Size of dataset (MiB): 1.525879
Total time to transfer from CPU to GPU(s): 0.124695
total blocks: 98
Total time for GPU Kernel(s): 0.245218
Total number of points within epsilon (GPU): 879688
Total time (s): 0.369945
Total transfer time (s): 0.124717
```

The average kernel time is:0.24523

The average total time is: 0.3707447

These are three outputs for epsilon = 5 and N=100 for a CPU with 4 cores with O3:

```
64$ ./question2resultsforn100 5

Size of dataset (MiB): 0.001526
Total Count of Distance less than Eps: 100
Total time (s): 0.005062
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question2resultsforn100 5

Size of dataset (MiB): 0.001526
Total Count of Distance less than Eps: 100
Total time (s): 0.000175
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question2resultsforn100 5

Size of dataset (MiB): 0.001526
Total Count of Distance less than Eps: 100
Total time (s): 0.006781
```

The Total average is: 0.004006

These are three outputs for epsilon = 5 and N=1000 for a CPU with 4 cores with O3:

```
64$ ./question2resultsforn1000 5
Size of dataset (MiB): 0.015259
Total Count of Distance less than Eps: 1102
Total time (s): 0.002918
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question2resultsforn1000 5
Size of dataset (MiB): 0.015259
Total Count of Distance less than Eps: 1102
Total time (s): 0.002601
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question2resultsforn1000 5
Size of dataset (MiB): 0.015259
Total Count of Distance less than Eps: 1102
Total time (s): 0.001242
```

The Total Average is:0.0022537

These are three outputs for epsilon = 5 and N=10000 for a CPU with 4 cores with O3:

```
64$ ./question2resultsforn10000 5
Size of dataset (MiB): 0.152588
Total Count of Distance less than Eps: 17644
Total time (s): 0.058489
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question2resultsforn10000 5
Size of dataset (MiB): 0.152588
Total Count of Distance less than Eps: 17644
Total time (s): 0.058369
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450Parallel
64$ ./question2resultsforn10000 5
Size of dataset (MiB): 0.152588
Total Count of Distance less than Eps: 17644
Total time (s): 0.057025
```

The Total Average is:0.057961

These are three outputs for epsilon = 5 and N=100000 for a CPU with 4 cores with O3:

```
64$ ./question2resultsforn100000 5
Size of dataset (MiB): 1.525879
Total Count of Distance less than Eps: 879688
Total time (s): 6.669768
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelP
64$ ./question2resultsforn100000 5
Size of dataset (MiB): 1.525879
Total Count of Distance less than Eps: 879688
Total time (s): 5.345110
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelP
64$ ./question2resultsforn100000 5
Size of dataset (MiB): 1.525879
Total Count of Distance less than Eps: 879688
Total time (s): 5.614908
rjena@rjena-Lenovo-Yoga-2-Pro:~/CS450ParallelP
```

The Total Average is:5.8765953

Answer for Question 2 :

So to get it to go faster I decided to implement a similar solution to what I had in assignment 5. So basically I treated the TID like it was the outerloop in the cpu version. So then I set the inner loops index to TID+1. So this gets rid of half of the distance calculations. Therefore I end up incrementing a counter by 2. Also I implemented an if statement to make sure it doesn't calculate the distance if it is 0. Then I started the counter off at N because that's how many distance calculations would be 0. So overall this just got rid of a lot of unnecessary calculations and made the algorithm $(N^2)/2$. Still if you look at the data above and compare it to the last part the times are only faster by very little. Although when I tested larger values for N I ended up getting a recognizable time difference. This could have been improved a lot more using shared data for the data points we are passing through the GPU function.