Analysis of Algorithms I

BLG 335E

Assignment – I

Hasan Fatih Durkaya

150200074

Part-2

A-) How2Run:

I used classes to implement what was expected. From the csv file I got all the values and set them as attributes of ‘book’ class .I wrote a class definition with getters and setters to make this more appropriate and readable. I wrote class definition in the different files book.h and book.cpp . To run implementation I direct my terminal to HW1 folder and run
“g++ -g blg335_hw1.cpp book.cpp -o hw1” to compile my code and then run “./hw1” to run my code.

B-)

```
14917,Desire and Duty: A Sequel to Jane Austen's Pride and Prejudice,Ted Bader/Marilyn Bader,2.86,965429903,9.78E+12,en
35819,Lair of the White Worm,Bram Stoker,2.86,646418424,9.78E+12,eng,120,2276,244,10.01.2002,Deodand Publishing
6644,Le Divorce,Diane Johnson,2.88,452284481,9.78E+12,eng,320,3602,295,7.01.2003,Plume
28227,Checkpoint,Nicholson Baker,2.88,1400079853,9.78E+12,eng,128,654,60,4.12.2005,Vintage
3918,Divorcio a la Francesa: Le Divorce,Diane Johnson/Carlos Milla Soler/Roberto Fernandez Sastre,2.88,8420466557,9.79E
33032,Queen of the Underworld,Gail Godwin,2.89,345483197,9.78E+12,eng,368,541,100,1/30/2007,Ballantine Books
8881,Planet of the Apes: The Fall,William T. Quick,2.92,60086203,9.78E+12,eng,288,53,2,6.04.2002,HarperEntertainment
41073,Children of the Thunder / The Tides of Time / The Crucible of Time,John Brunner,2.92,051712310X,9.78E+12,eng,800,
41154,Secret Identity (Lost  #2),Catherine Hapka,2.93,786890916,9.78E+12,eng,176,195,11,1.01.2006,Voice
32181,Zelda's Cut,Philippa Gregory,2.94,6511775,9.78E+12,eng,432,65,12,2.01.2001,HarperCollins Publishers
19608,The Doctor's House,Ann Beattie,2.96,743235010,9.78E+12,eng,288,164,23,2.11.2003,Scribner
28585,Judge Dredd (Audio Cassette),Neal Barrett Jr./Martha Banta,2.96,553476793,9.78E+12,eng,0,6,0,7.01.1995,Random Hou
4221,The Guide to Dan Brown's the Solomon Key,Greg  Taylor,2.98,875168167,9.78E+12,en-US,183,34,2,9/30/2005,DeVorss & C
35760,Mistaken Identity,Nayantara Sahgal,2.98,8172235224,9.79E+12,eng,324,0,0,12/30/2016,Harper Collins
159,Dinner with Anna Karenina,Gloria Goldreich,2.99,778322270,9.78E+12,eng,360,411,65,1/28/2006,Mira Books
30786,Verserade tankar,Lennart Hellsing/Fibben Hald,3,9177129334,9.79E+12,swe,128,1,0,9.01.1999,Alfabeta Bokförlag
33581,The Work of Work Servitude  Slavery and Labor in Medieval England,Allen J. Frantzen/Ruth Mazo Karras/Madonna J. H
19084,Aristotle's Ethics (SparkNotes Literature Guides),SparkNotes,3,158663822X,9.78E+12,eng,80,2,0,6/13/2003,SparkNote
2452,Theocritus: Select Poems: Select Poems,Theocritus/Kenneth James Dover,3,862921473,9.78E+12,grc,395,1,1,6.01.1991,B
2756,New Media Language,Jean Aitchison/Diana M. Lewis,3,415283043,9.78E+12,eng,209,2,0,5/22/2003,Routledge
26800,C++ Programmer's Notebook,Jim Keogh/John Shapley Gray,3,130887013,9.78E+12,eng,528,3,0,8/16/2001,Prentice Hall PT
3167,Phaedrus/Apology/Crito/Symposium,Plato/Benjamin Jowett,3,1420926845,9.78E+12,en-US,144,19,0,1.01.2006,Digireads.co
16726,Cypress Gardens (Images of America: Florida),Mary M. Flekke/Sarah E. MacDonald/Randall M. MacDonald,3,073854339X,
32189,Sukhoi Su-27 Flanker (WarbirdTech  #42),Yefim Gordon/Peter Davison,3,1580070914,9.78E+12,eng,104,6,0,4.01.2006,Sp
28617,A Free Enquiry Into the Vulgarly Received Notion of Nature,Robert Boyle/Michael Hunter/Edward B. Davis,3,52156796
36266,Magical Creatures (Easy To Read! Easy To Draw!),Joan Holub/Dana Regan,3,843104368,9.78E+12,eng,48,1,0,10/27/2003,
36461,Proof of Concept,Larry Young/Damian Couceiro/Kieron Dwyer,3,1932051295,9.78E+12,eng,134,6,1,1/18/2005,AIT Planet
37902,Turn of the Cards,Georgina Grey,3,449239691,9.78E+12,eng,0,3,1,5.12.1979,Fawcett Coventry
24594,Selections from the Oakland Tribune Archives (Images of America: California),Annalee Allen,3,073854678X,9.78E+12,
```

```
36461 --> 3
37902 --> 3
24594 --> 3
40597 --> 3
8889 --> 3
16710 --> 3
41378 --> 3
42410 --> 3
10713 --> 3.02
42926 --> 3.02
15375 --> 3.03
33075 --> 3.04
41261 --> 3.04
31004 --> 3.06
15384 --> 3.07
38570 --> 3.07
18412 --> 3.07
15379 --> 3.07
29320 --> 3.08
2152 --> 3.08
21615 --> 3.09
18837 --> 3.09
37491 --> 3.09
681 --> 3.09
2126 --> 3.1
36361 --> 3.11
15610 --> 3.11
32941 --> 3.11
20518 --> 3.11
```

C-)

Yes, there has been significant changes comparing the full data.

Execution time, number of partitions and number of swaps with fulll data:

```
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$ g++ -g blg335_hw1.cpp book.cpp
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$ ./a.out
Running time of quicksort1 with full data: 0.011614 seconds
Number of partitions: 10944
Number of swaps: 99512
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$
```

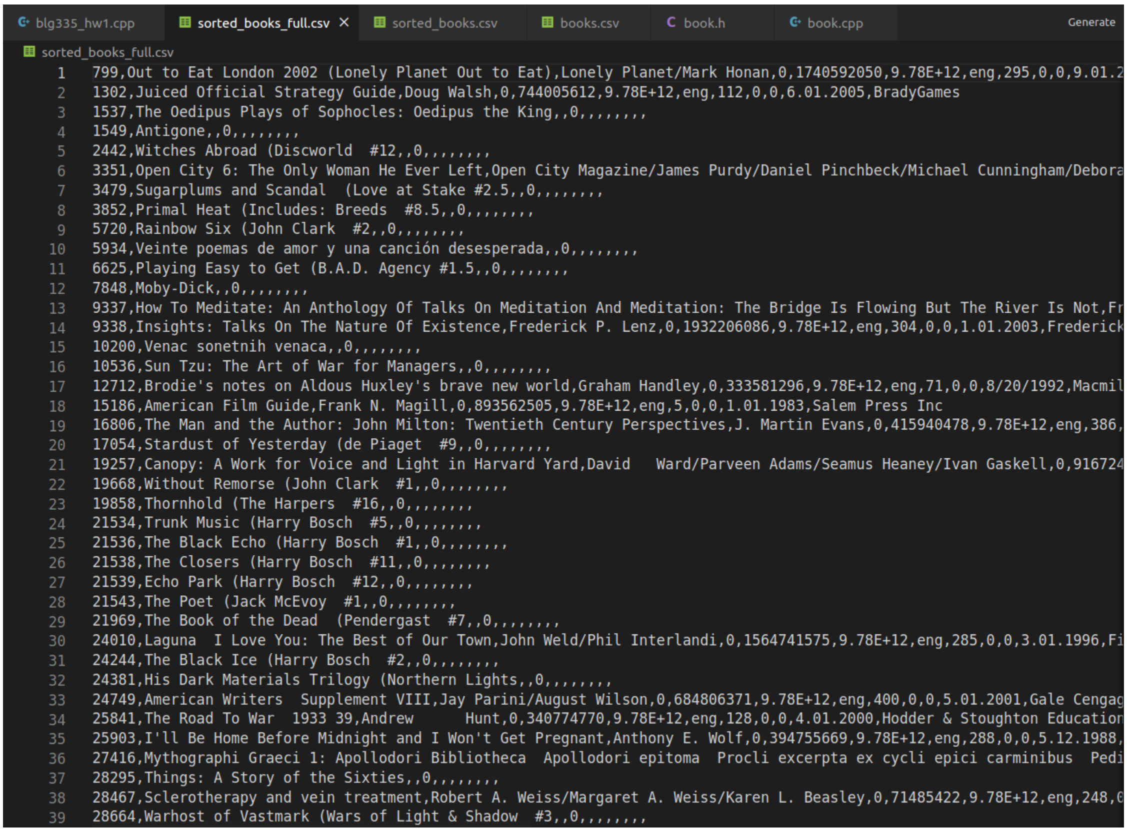
Execution time, number of partitions and number of swaps with half of data:

```
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$ g++ -g blg335_hw1.cpp book.cpp
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$ ./a.out
Running time of quicksort1 with half of data: 0.003577 seconds
Number of partitions: 5401
Number of swaps: 53295
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$
```

Execution time, number of partitions and number of swaps with quarter of data:

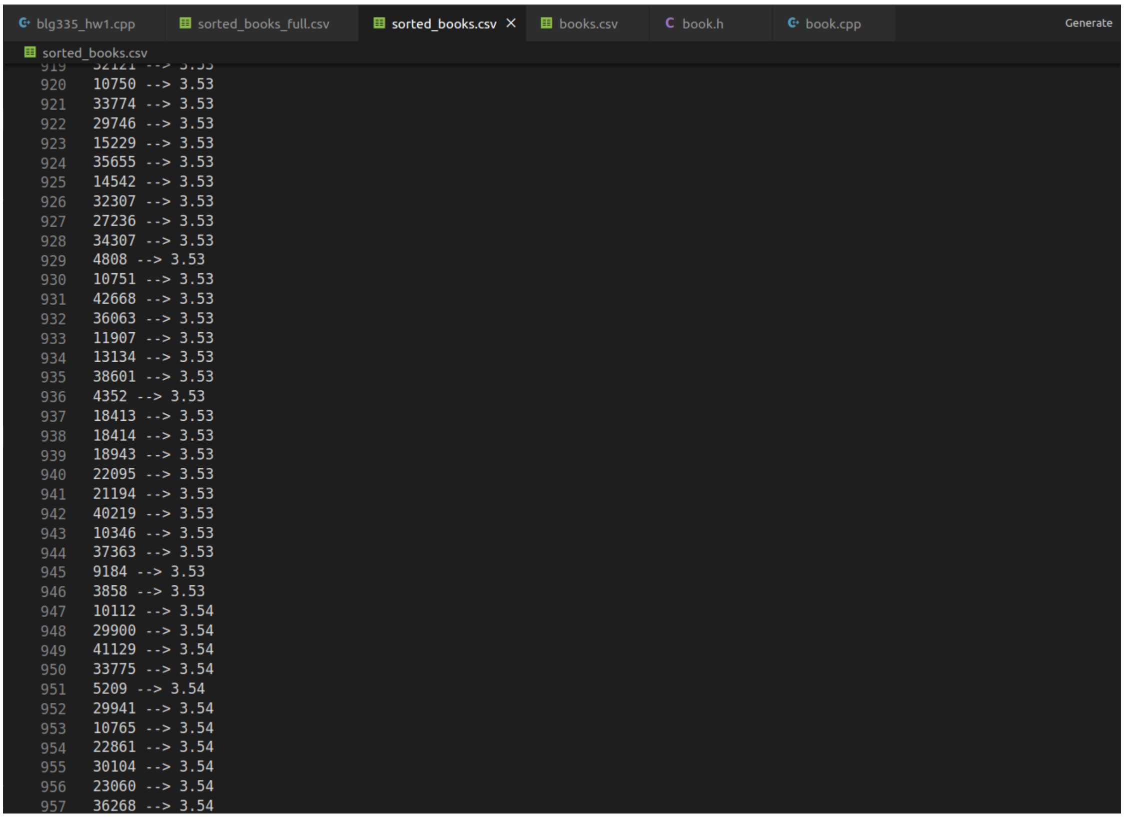
```
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$ g++ -g blg335_hw1.cpp book.cpp
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$ ./a.out
Running time of quicksort1 with quarter of data: 0.001413 seconds
Number of partitions: 2637
Number of swaps: 24681
hfd@hfd-ABRA-A5-V15-8:~/blg_335/assignments/HW1$
```

Sorted books list with all attributes from the start.



```
blg335_hw1.cpp    sorted_books_full.csv ×    sorted_books.csv    books.csv    book.h    book.cpp    Generate
sorted_books_full.csv
799,Out to Eat London 2002 (Lonely Planet Out to Eat),Lonely Planet/Mark Honan,0,1740592050,9.78E+12,eng,295,0,0,9.01.2
1302,Juiced Official Strategy Guide,Doug Walsh,0,744005612,9.78E+12,eng,112,0,0,6.01.2005,BradyGames
1537,The Oedipus Plays of Sophocles: Oedipus the King,,0,,,,,,,,
1549,Antigone,,0,,,,,,,,
2442,Witches Abroad (Discworld  #12,,0,,,,,,,,
3351,Open City 6: The Only Woman He Ever Left,Open City Magazine/James Purdy/Daniel Pinchbeck/Michael Cunningham/Debora
3479,Sugarplums and Scandal  (Love at Stake #2.5,,0,,,,,,,,
3852,Primal Heat (Includes: Breeds  #8.5,,0,,,,,,,,
5720,Rainbow Six (John Clark  #2,,0,,,,,,,,
5934,Veinte poemas de amor y una canción desesperada,,0,,,,,,,,
6625,Playing Easy to Get (B.A.D. Agency #1.5,,0,,,,,,,,
7848,Moby-Dick,,0,,,,,,,,
9337,How To Meditate: An Anthology Of Talks On Meditation And Meditation: The Bridge Is Flowing But The River Is Not,Fr
9338,Insights: Talks On The Nature Of Existence,Frederick P. Lenz,0,1932206086,9.78E+12,eng,304,0,0,1.01.2003,Frederick
10200,Venac sonetnih venaca,,0,,,,,,,,
10536,Sun Tzu: The Art of War for Managers,,0,,,,,,,,
12712,Brodie's notes on Aldous Huxley's brave new world,Graham Handley,0,333581296,9.78E+12,eng,71,0,0,8/20/1992,Macmil
15186,American Film Guide,Frank N. Magill,0,893562505,9.78E+12,eng,5,0,0,1.01.1983,Salem Press Inc
16806,The Man and the Author: John Milton: Twentieth Century Perspectives,J. Martin Evans,0,415940478,9.78E+12,eng,386,
17054,Stardust of Yesterday (de Piaget  #9,,0,,,,,,,,
19257,Canopy: A Work for Voice and Light in Harvard Yard,David   Ward/Parveen Adams/Seamus Heaney/Ivan Gaskell,0,916724
19668,Without Remorse (John Clark  #1,,0,,,,,,,,
19858,Thornhold (The Harpers  #16,,0,,,,,,,,
21534,Trunk Music (Harry Bosch  #5,,0,,,,,,,,
21536,The Black Echo (Harry Bosch  #1,,0,,,,,,,,
21538,The Closers (Harry Bosch  #11,,0,,,,,,,,
21539,Echo Park (Harry Bosch  #12,,0,,,,,,,,
21543,The Poet (Jack McEvoy  #1,,0,,,,,,,,
21969,The Book of the Dead  (Pendergast  #7,,0,,,,,,,,
24010,Laguna  I Love You: The Best of Our Town,John Weld/Phil Interlandi,0,1564741575,9.78E+12,eng,285,0,0,3.01.1996,Fi
24244,The Black Ice (Harry Bosch  #2,,0,,,,,,,,
24381,His Dark Materials Trilogy (Northern Lights,,0,,,,,,,,
24749,American Writers  Supplement VIII,Jay Parini/August Wilson,0,684806371,9.78E+12,eng,400,0,0,5.01.2001,Gale Cengag
25841,The Road To War  1933 39,Andrew      Hunt,0,340774770,9.78E+12,eng,128,0,0,4.01.2000,Hodder & Stoughton Education
25903,I'll Be Home Before Midnight and I Won't Get Pregnant,Anthony E. Wolf,0,394755669,9.78E+12,eng,288,0,0,5.12.1988,
27416,Mythographi Graeci 1: Apollodori Bibliotheca  Apollodori epitoma  Procli excerpta ex cycli epici carminibus  Pedi
28295,Things: A Story of the Sixties,,0,,,,,,,,
28467,Sclerotherapy and vein treatment,Robert A. Weiss/Margaret A. Weiss/Karen L. Beasley,0,71485422,9.78E+12,eng,248,0
28664,Warhost of Vastmark (Wars of Light & Shadow  #3,,0,,,,,,,,
```

Sorted books list with only bookID and rating.



```
blg335_hw1.cpp    sorted_books_full.csv    sorted_books.csv ×    books.csv    book.h    book.cpp    Generate
sorted_books.csv
10750 --> 3.53
33774 --> 3.53
29746 --> 3.53
15229 --> 3.53
35655 --> 3.53
14542 --> 3.53
32307 --> 3.53
27236 --> 3.53
34307 --> 3.53
4808 --> 3.53
10751 --> 3.53
42668 --> 3.53
36063 --> 3.53
11907 --> 3.53
13134 --> 3.53
38601 --> 3.53
4352 --> 3.53
18413 --> 3.53
18414 --> 3.53
18943 --> 3.53
22095 --> 3.53
21194 --> 3.53
40219 --> 3.53
10346 --> 3.53
37363 --> 3.53
9184 --> 3.53
3858 --> 3.53
10112 --> 3.54
29900 --> 3.54
41129 --> 3.54
33775 --> 3.54
5209 --> 3.54
29941 --> 3.54
10765 --> 3.54
22861 --> 3.54
30104 --> 3.54
23060 --> 3.54
36268 --> 3.54
```

D-)

Changes that we make in the normal form of quicksort will effect some constant times but for asymptotic upper bounds we mainly focus on the highest order terms .

For the best case , that means after the each partition operation is done , each subarray contains half of the previous step's data. This means that pivotsa re always the middle element.

$$T(n) <= 2T(n/2) + \Theta(n)$$

We can solve this recurrence equation with master method

$$T(n) = aT\left(\frac{n}{b}\right) + n^c \quad a \geq 1, b \geq 1, c > 0$$
$$\Downarrow$$
$$T(n) = \begin{cases} \Theta(n^{\log_b a}) & a > b^c \\ \Theta(n^c \log_b n) & a = b^c \\ \Theta(n^c) & a < b^c \end{cases}$$

and when we solve this equation we get -->

$$T(n) = O(nlogn)$$

Fort he average case, we can suppose that we always split the data 9 to 1.

Recurrence equation will be

$$T(n) <= T(9n/10) + T(n/10) + \Theta(n)$$
$$= T(9n/10) + T(n/10) + cn$$

We can solve this equation with recursion tree

$$T(n) = T(9n/10) + T(n/10) + n$$

$<= (9cn/10) \log(9n/10) + (cn/10)\log(n/10) + n$

$<= (9cn/10) \log(n) + (9cn/10)\log(9/10) + (cn/10)\log(n) + (cn/10)\log(1/10) + n$

$<= cn\log n + 9cn/10 \log(9) – (9cn/10) \log 10 – (cn/10)\log 10 + n$

$<= cn\log(n) – n(c\log 10 – (9c/10)\log 9 - 1)$

$T(n) <= cn\log n$ if $c\log 10 – (9c/10)\log 9 – 1 > 0$ which is always true if

$c > 10/\log 10$

therefore average case running time is :

$$T(n) = \Theta(n\log n)$$

For the worst case we must always selected pivot must always stay in the bottom or top of the array/subarray.(highest or lowest)

Cost of partitionn will be $\Theta(n)$

Recurrence equation for quicksort:

$$T(n) = T(n - 1) + T(0) + \Theta(n)$$
$$= T(n - 1) + \Theta(n)$$

We can solve this recurrence by iteration.

$$T(n) = \Theta(n) + T(n - 1)$$
$$= \Theta(n) + \Theta(n - 1) + \Theta(n - 2) + \ldots + \Theta(1)$$
$$= \ (k = 1 \text{ to } n)\sum \Theta(k)$$
$$= \Theta([k = 1 \text{ to } n]\sum k)$$
$$= \Theta(n^2)$$

E-)

For the worst case scenarios to appear pivot must always be in the beginning or end of the array or subarray. When this situation occurs we always divide array to (n-1) and 1 length parts. This means we always select the highest or lowest element.This happens when we select the first or last element and array is sorted. To overcome this problem Pivot randomization will be a good solution but even in randomized quick sort algorithms there is a small chance to select always extreme terms.

For instance our array(data) is [0,1,2,3,4,5,6] N = 6

[0,1,2,3,4,5,6] → N

[0] [1,2,3,4,5,6] → 1+ (N – 1)

[1] [2,3,4,5,6] → 1+ (N – 2)

[2] [3,4,5,6] → 1+ (N – 3)

[3] [4,5,6] → 1+ (N – 4)

[4] [5,6] → 1+ (N – 5)

[5] [6] → 2

$$O(N^2)$$

İf we have used randomization pivot will be selected differently and there will be larger partitions this would increase the speed of algorithms.