



ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΤΗΛΕΜΑΤΙΚΗΣ

# Ανδρέας Ρουσούνελος Βασίλειος Κοκκινογένης Θεοχάρης Παυλίδης

Δομές Δεδομένων Εργασία Κατακερματισμού

## Περιεχόμενα

Εργασία	3
Επεξήγηση κώδικα	3
Ενδεικτικές εκτελέσεις (screenshots):	5
Ομαλή εκτέλεση τριών unit test	5
Screenshots	5
Ομαλή εκτέλεση προγράμματος μεγάλου κειμένου (txt αρχείου)	6
Screenshots	6
Ομαλή εκτέλεση προγράμματος μικρού κειμένου (txt αρχείου)	8
Screenshots	8
Ομαλή εκτέλεση προγράμματος με πολλαπλά αρχεία κειμένων	9
Screenshots	9
Ανεπιτυχής εκτέλεση προγράμματος χωρίς ορίσματα / με λάθος ορίσματα / με όρισμα μη υπάρχον αρχείο	11
Screenshots	11
Εν μέρει επιτυχής εκτέλεση προγράμματος με ένα σωστό όρισμα και με ένα λάθος όρισμα	12
Screenshots	12
Εν μέρει επιτυχής εκτέλεση προγράμματος με ένα υπάρχον αρχείο και με ένα μη υπάρχον αρχείο	13
Screenshots	13

# Εργασία

## Επεξήγηση κώδικα

### Dictionary:

Το interface Dictionary περιέχει χρήσιμες μεθόδους για την επεξεργασία λεξικών και παρόμοιων δομών δεδομένων όπως εισαγωγή, διαγραφή και αναζήτηση στοιχείων.

### HashMap:

Η κλάση HashMap μετράει τις συχνότητες των λέξεων σε αρχεία κειμένου. Δημιουργεί έναν πίνακα που κρατάει τα txt files, πραγματοποιεί έλεγχο εισόδου του χρήστη, δημιουργεί έναν πίνακα κατακερματισμού και χωρίζει τις γραμμές του αρχείου κειμένου σε λέξεις. Αν μια λέξη στο αρχείο κειμένου δεν υπάρχει στον πίνακα κατακερματισμού τότε ορίζει την συχνότητά της σε 1, αλλιώς αυξάνει την συχνότητά της κατά 1. Τέλος εκτυπώνει την επικεφαλίδα, τις λέξεις που έχει προσθέσει στον πίνακα κατακερματισμού και τις συχνότητές τους.

### OpenAddressHashTable:

Η κλάση `OpenAddressHashTable` δημιουργεί έναν πίνακα κατακερματισμού με τη χρήση ανοικτής διευθυνσιοδότησης και γραμμικής διερεύνησης. Ο constructor της δημιουργεί τον πίνακα κατακερματισμού, αρχικοποιεί τον αριθμό των στοιχείων του σε 0 και παράγει μια νέα συνάρτηση κατακερματισμού με την βοηθητική μέθοδο `generateHashFunction`. Η εισαγωγή - μέθοδος `put` - υλοποιείται με την βοήθεια δύο βοηθητικών μεθόδων `insert` και `rehashIfNeeded`. Η `insert` προσθέτει τα στοιχεία ή ανανεώνει το `value` τους αν υπάρχουν ήδη. Η `rehashIfNeeded` εξετάζει αν το μέγεθος του πίνακα είναι υπερβολικά μεγαλύτερο ή μικρότερο σε σύγκριση με τα στοιχεία που περιέχει (αν ο αριθμός των στοιχείων ξεπεράσει το 80% του μεγέθους του πίνακα, αυτός υποδιπλασιάζεται, διαφορετικά αν τα στοιχεία καλύπτουν λιγότερο από το 25%, ο πίνακας υποδιπλασιάζεται). Η διαγραφή - μέθοδος `remove` - συνδυάζει την `rehashIfNeeded` και τον αλγόριθμο διαγραφής από τις διαφάνειες. Χρησιμοποιεί μια ακόμα βοηθητική μέθοδο, την `getElement`, για να πάρει την θέση του πίνακα που ορίζει η συνάρτηση κατακερματισμού για το συγκεκριμένο στοιχείο. Η αναζήτηση υλοποιείται από τις μεθόδους `get` και `contains`. Οι δύο μέθοδοι λειτουργούν παρόμοια, με την διαφορά ότι η `get` επιστρέφει το στοιχείο που αναζητείται ενώ η `contains` επιστρέφει λογικό `ναι` αν το στοιχείο υπάρχει. Η `isEmpty` μηδενίζει τον αριθμό των στοιχείων του πίνακα, η `size` επιστρέφει τον αριθμό τους και η `clear` θέτει όλα τα στοιχεία του πίνακα σε `null`. Η μέθοδος `insert` δημιουργεί ένα νέο `HashIterator` που χρησιμεύει στην διάσχιση του πίνακα κατακερματισμού με έξυπνες προτάσεις `for`. Τέλος, κάθε στοιχείο αντιπροσωπεύεται από ένα `EntryImpl` το οποίο είναι απλά μια κλάση που υλοποιεί τις μεθόδους `getKey` και `getValue` της διασύνδεσης `Dictionary.Entry`.

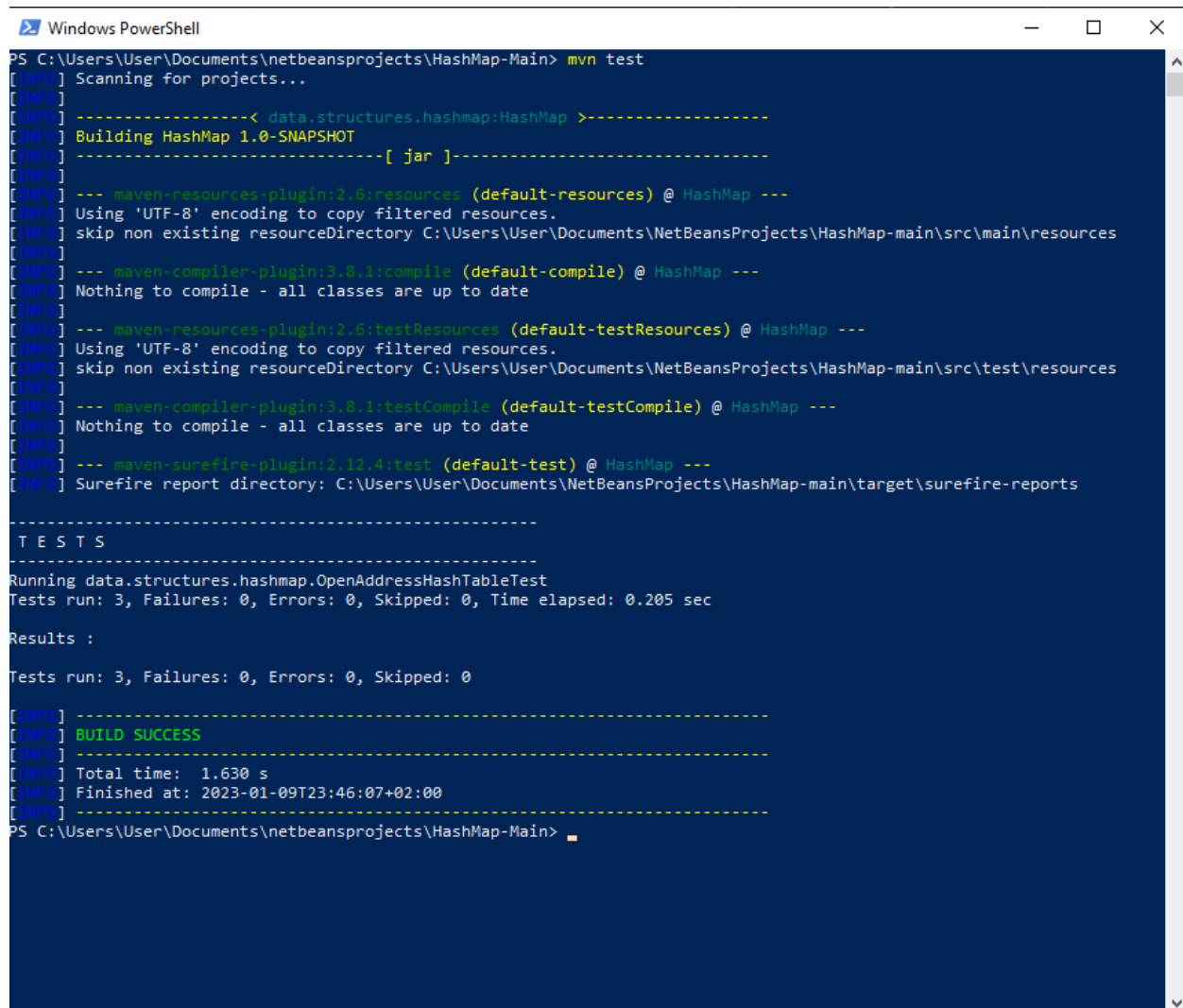
### OpenAddressHashTableTest:

Η κλάση `OpenAddressHashTableTest` δοκιμάζει τα όρια του πίνακα κατακερματισμού με 3 `unit tests`. Το πρώτο τεστ που ορίζει η μέθοδος `testPutAndGet` δοκιμάζει τις λειτουργίες εισαγωγής και αναζήτησης του πίνακα κατακερματισμού. Το δεύτερο τεστ `testRemoveAndContains` εξετάζει την διαγραφή και την μέθοδο `contains` του πίνακα κατακερματισμού. Τέλος το τρίτο τεστ `testClearAndIsEmpty` εξετάζει τις μεθόδους `clear` και `isEmpty` του πίνακα. Για περισσότερες πληροφορίες για το πως λειτουργεί η κάθε μέθοδος ανατρέξτε στα σχόλια του πηγαίου κώδικα.

## Ενδεικτικές εκτελέσεις (screenshots):

*Ομαλή εκτέλεση τριών unit test*

### Screenshots



```
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> mvn test
[INFO] Scanning for projects...
[INFO]
[INFO] -----< data.structures.hashmap:HashMap >-----
[INFO] Building HashMap 1.0-SNAPSHOT
[INFO] -----[ jar ]-----
[INFO]
[INFO] --- maven-resources-plugin:2.6:resources (default-resources) @ HashMap ---
[INFO] Using 'UTF-8' encoding to copy filtered resources.
[INFO] skip non existing resourceDirectory C:\Users\User\Documents\NetBeansProjects\HashMap-main\src\main\resources
[INFO]
[INFO] --- maven-compiler-plugin:3.8.1:compile (default-compile) @ HashMap ---
[INFO] Nothing to compile - all classes are up to date
[INFO]
[INFO] --- maven-resources-plugin:2.6:testResources (default-testResources) @ HashMap ---
[INFO] Using 'UTF-8' encoding to copy filtered resources.
[INFO] skip non existing resourceDirectory C:\Users\User\Documents\NetBeansProjects\HashMap-main\src\test\resources
[INFO]
[INFO] --- maven-compiler-plugin:3.8.1:testCompile (default-testCompile) @ HashMap ---
[INFO] Nothing to compile - all classes are up to date
[INFO]
[INFO] --- maven-surefire-plugin:2.12.4:test (default-test) @ HashMap ---
[INFO] Surefire report directory: C:\Users\User\Documents\NetBeansProjects\HashMap-main\target\surefire-reports

-----
T E S T S
-----
Running data.structures.hashmap.OpenAddressHashTableTest
Tests run: 3, Failures: 0, Errors: 0, Skipped: 0, Time elapsed: 0.205 sec

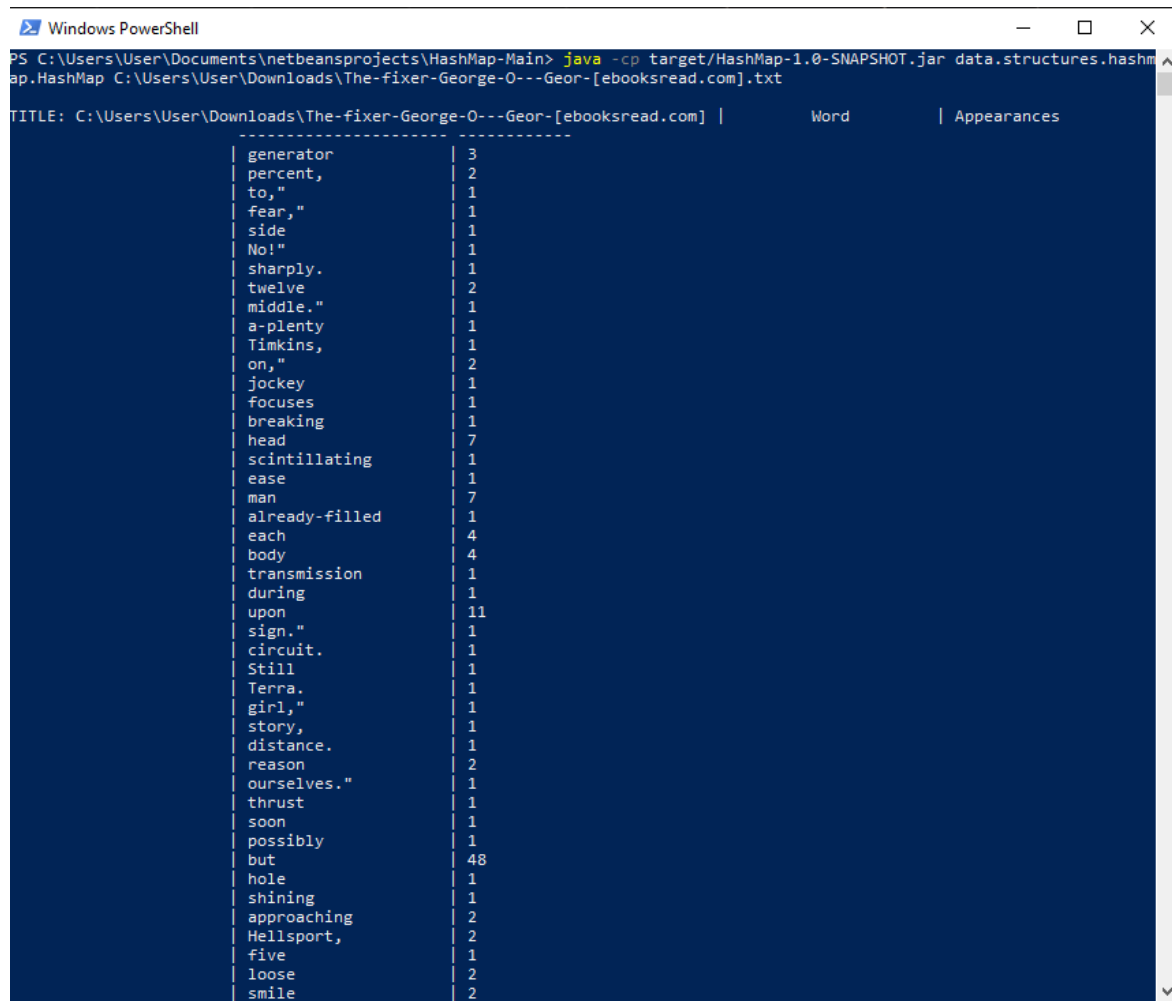
Results :

Tests run: 3, Failures: 0, Errors: 0, Skipped: 0

[INFO] -----
[INFO] BUILD SUCCESS
[INFO] -----
[INFO] Total time: 1.630 s
[INFO] Finished at: 2023-01-09T23:46:07+02:00
[INFO] -----
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main>
```

Ομαλή εκτέλεση προγράμματος μεγάλου κειμένου (txt αρχείου)

## Screenshots



```
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashmap.HashMap C:\Users\User\Downloads\The-fixer-George-O---Geor-[ebooksread.com].txt
```

Word	Appearances
generator	3
percent,	2
to,"	1
fear,"	1
side	1
No!"	1
sharply.	1
twelve	2
middle."	1
a-plenty	1
Timkins,	1
on,"	2
jockey	1
focuses	1
breaking	1
head	7
scintillating	1
ease	1
man	7
already-filled	1
each	4
body	4
transmission	1
during	1
upon	11
sign."	1
circuit.	1
Still	1
Terra.	1
girl,"	1
story,	1
distance.	1
reason	2
ourselves."	1
thrust	1
soon	1
possibly	1
but	48
hole	1
shining	1
approaching	2
Hellsport,	2
five	1
loose	2
smile	2

Αρκετά scroll πιο κάτω...

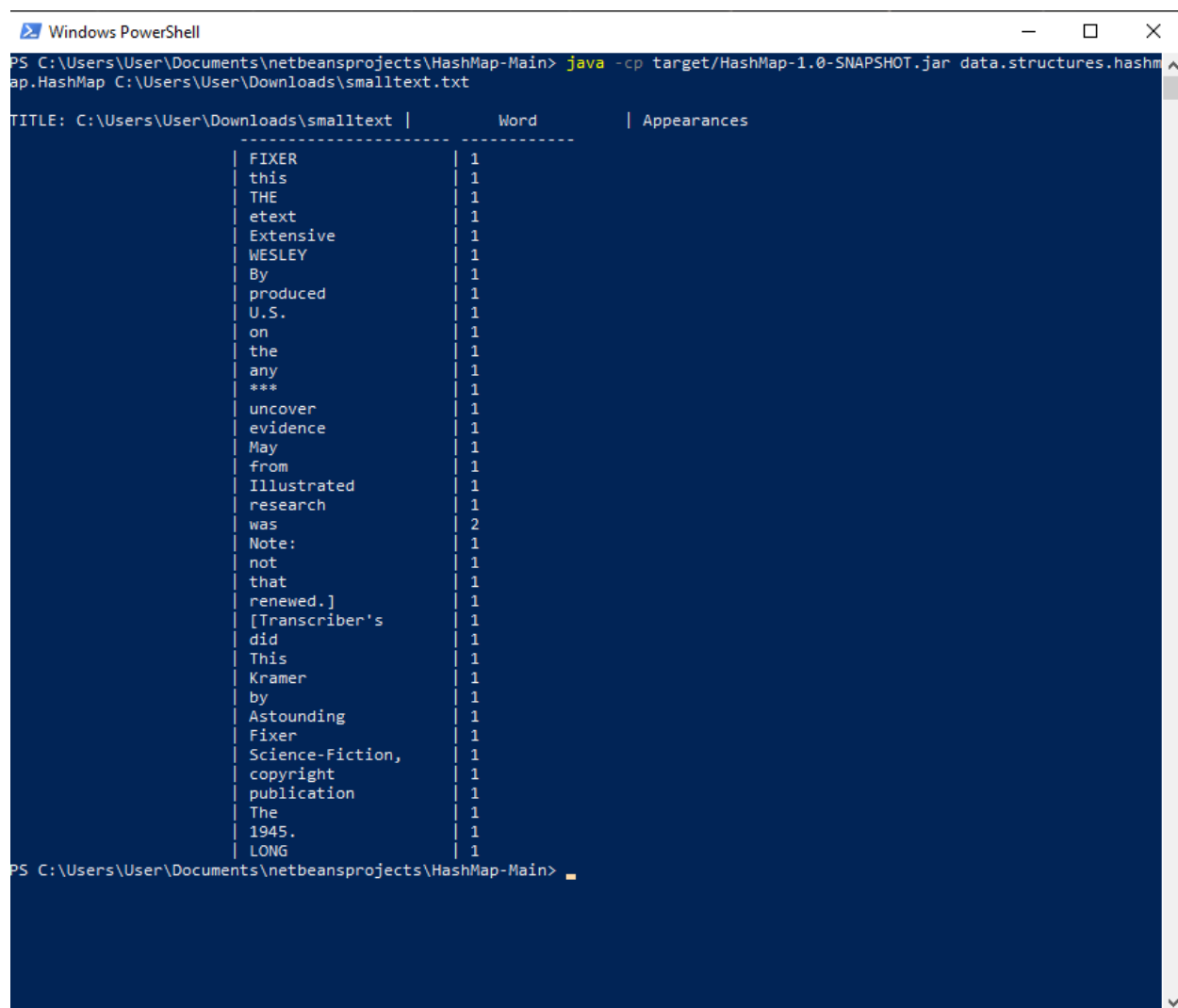
```
Windows PowerShell

Remember      1
try."         1
inexplicable  1
electro-combining 1
Darn          1
none.         1
screech.      1
fractional    2
spoiled       1
seeing        1
lens-element  1
andryorelitis 1
sleep,        1
follow,       1
matter,       1
was           136
yield?"       1
Theodi,       10
may,          1
gray-green    1
big           1
shell         1
relieve       1
We--and       1
"Obviously.   1
blooey        1
hey?          1
put           3
wearily.      1
Well,         4
aware         1
unseeing      1
fool!         1
met           1
hands.        1
figure.       1
proportions,  1
illustrated   1
fellow,"      1

PS C:\Users\User\Documents\netbeansprojects\HashMap-Main>
```

Ομαλή εκτέλεση προγράμματος μικρού κειμένου (txt αρχείου)

## Screenshots



```
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashmap.HashMap C:\Users\User\Downloads\smalltext.txt
```

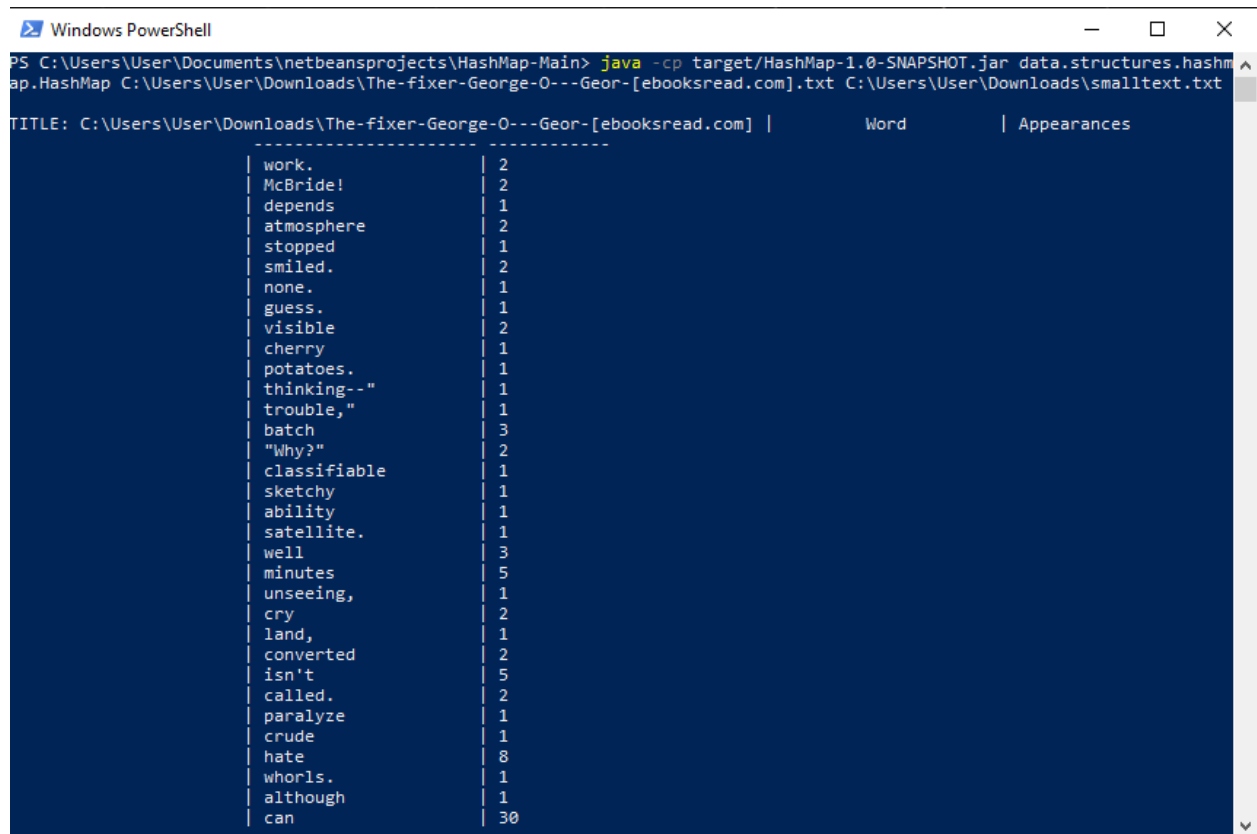
TITLE: C:\Users\User\Downloads\smalltext	Word	Appearances
	FIXER	1
	this	1
	THE	1
	etext	1
	Extensive	1
	WESLEY	1
	By	1
	produced	1
	U.S.	1
	on	1
	the	1
	any	1
	***	1
	uncover	1
	evidence	1
	May	1
	from	1
	Illustrated	1
	research	1
	was	2
	Note:	1
	not	1
	that	1
	renewed.]	1
	[Transcriber's	1
	did	1
	This	1
	Kramer	1
	by	1
	Astounding	1
	Fixer	1
	Science-Fiction,	1
	copyright	1
	publication	1
	The	1
	1945.	1
	LONG	1

```
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main>
```



Ομαλή εκτέλεση προγράμματος με πολλαπλά αρχεία κειμένου

## Screenshots



```
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashmap.HashMap C:\Users\User\Downloads\The-fixer-George-O---Geor-[ebooksread.com].txt C:\Users\User\Downloads\smalltext.txt

TITLE: C:\Users\User\Downloads\The-fixer-George-O---Geor-[ebooksread.com] | Word | Appearances
-----|-----|-----
| work. | 2 |
| McBride! | 2 |
| depends | 1 |
| atmosphere | 2 |
| stopped | 1 |
| smiled. | 2 |
| none. | 1 |
| guess. | 1 |
| visible | 2 |
| cherry | 1 |
| potatoes. | 1 |
| thinking--" | 1 |
| trouble," | 1 |
| batch | 3 |
| "Why?" | 2 |
| classifiable | 1 |
| sketchy | 1 |
| ability | 1 |
| satellite. | 1 |
| well | 3 |
| minutes | 5 |
| unseeing, | 1 |
| cry | 2 |
| land, | 1 |
| converted | 2 |
| isn't | 5 |
| called. | 2 |
| paralyze | 1 |
| crude | 1 |
| hate | 8 |
| whorls. | 1 |
| although | 1 |
| can | 30
```

Αρκετά scroll πιο κάτω...

```
Windows PowerShell

| "Sandrake!" | 1
TITLE: C:\Users\User\Downloads\smalltext | Word | Appearances
-----
| FIXER | 1
| that | 1
| The | 1
| Extensive | 1
| the | 1
| publication | 1
| was | 2
| on | 1
| research | 1
| etext | 1
| copyright | 1
| evidence | 1
| by | 1
| U.S. | 1
| did | 1
| WESLEY | 1
| any | 1
| Illustrated | 1
| not | 1
| produced | 1
| By | 1
| Astounding | 1
| [Transcriber's | 1
| Fixer | 1
| *** | 1
| May | 1
| 1945. | 1
| LONG | 1
| Science-Fiction, | 1
| uncover | 1
| this | 1
| Note: | 1
| This | 1
| from | 1
| Kramer | 1
| THE | 1
| renewed.] | 1
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main>
```

χωρίς ορίσματα / με λάθος ορίσματα / με όρισμα μη υπάρχον αρχείο

```
Windows PowerShell
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashmap.HashMap
>>> Not enough arguments
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashmap.HashMap wrongargument
>>> No text files
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashmap.HashMap missingfile.txt
>>> Error: file missingfile.txt not found
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main>
```

Εν μέρει επιτυχής εκτέλεση προγράμματος

με ένα σωστό όρισμα και με ένα λάθος όρισμα

```
Windows PowerShell
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashm
ap.HashMap C:\Users\User\Downloads\smalltext.txt wrong.nottxt
>>> IGNORED: wrong.nottxt

TITLE: C:\Users\User\Downloads\smalltext | Word | Appearances
-----|-----|-----
| THE | 1 |
| May | 1 |
| evidence | 1 |
| [Transcriber's | 1 |
| Fixer | 1 |
| that | 1 |
| Kramer | 1 |
| by | 1 |
| the | 1 |
| The | 1 |
| *** | 1 |
| uncover | 1 |
| FIXER | 1 |
| renewed.] | 1 |
| from | 1 |
| Astounding | 1 |
| By | 1 |
| was | 2 |
| U.S. | 1 |
| research | 1 |
| this | 1 |
| etext | 1 |
| did | 1 |
| copyright | 1 |
| on | 1 |
| Science-Fiction, | 1 |
| WESLEY | 1 |
| any | 1 |
| produced | 1 |
| Extensive | 1 |
| Illustrated | 1 |
| Note: | 1 |
| publication | 1 |
| 1945. | 1 |
| This | 1 |
| LONG | 1 |
| not | 1 |
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main>
```

Εν μέρει επιτυχής εκτέλεση προγράμματος

με ένα υπάρχον αρχείο και με ένα μη υπάρχον αρχείο

```
Windows PowerShell
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main> java -cp target/HashMap-1.0-SNAPSHOT.jar data.structures.hashmap.HashMap C:\Users\User\Downloads\smalltext.txt missingfile.txt
TITLE: C:\Users\User\Downloads\smalltext | Word | Appearances
-----
| renewed.] | 1
| etext | 1
| by | 1
| 1945. | 1
| WESLEY | 1
| [Transcriber's | 1
| Kramer | 1
| did | 1
| produced | 1
| Extensive | 1
| U.S. | 1
| This | 1
| that | 1
| not | 1
| on | 1
| *** | 1
| any | 1
| was | 2
| uncover | 1
| Science-Fiction, | 1
| May | 1
| Illustrated | 1
| By | 1
| evidence | 1
| this | 1
| Fixer | 1
| copyright | 1
| Astounding | 1
| LONG | 1
| FIXER | 1
| Note: | 1
| publication | 1
| THE | 1
| from | 1
| The | 1
| research | 1
| the | 1
>>> Error: file missingfile.txt not found
PS C:\Users\User\Documents\netbeansprojects\HashMap-Main>
```