There are two environment variables in my Lab7

The first is sys.argv[1] which is the input filename(without .txt)

The second is sys.argv[2] which is the hash function name

You can select hash\_func or hash\_funcOne or hash\_funcTwo as the argv[2]

Same factor with different hash functions factor is equal to 0.7

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| data  hash\_function | the maximum  appeared times | which word appeared the maximum times | the number of probes | the number of collisions |
| hash\_func | 4988 | the | 108215587 | 26570 |
| hash\_funcOne | 4988 | the | 19550570 | 28917 |
| hash\_funcTwo | 4988 | the | 175405356 | 30208 |

Same hash function with different factors using hash\_func

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| data  factor | the maximum  appeared times | which word appeared the maximum times | the number of probes | the number of collisions |
| 0.6 | 4988 | the | 121148424 | 29774 |
| 0.7 | 4988 | the | 108215587 | 26570 |
| 0.8 | 4988 | the | 110255665 | 27200 |

Same hash function with different factors using hash\_funcOne

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| data  factor | the maximum  appeared times | which word appeared the maximum times | the number of probes | the number of collisions |
| 0.6 | 4988 | the | 20004469 | 31813 |
| 0.7 | 4988 | the | 19550570 | 28917 |
| 0.8 | 4988 | the | 20803854 | 29579 |

Same hash function with different factors using hash\_funcTwo

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| data  factor | the maximum  appeared times | which word appeared the maximum times | the number of probes | the number of collisions |
| 0.6 | 4988 | the | 172961870 | 33418 |
| 0.7 | 4988 | the | 175405356 | 30208 |
| 0.8 | 4988 | the | 200914820 | 30838 |