

Linked List Implementation and Statistics Package

David Alejandro Vásquez Carreño

August 2020

1 Abstract

In this paper, a java implementation of a Linked List will be showed, using java collection API to standardize the functions this new implementation has to offer, as well as to offer the same utilities as other implementations, like ArrayList or List. This new data structure will be used to store sample

data for statistics matters, like mean or standard deviation calculation. The use of those packages is shown in the lecture of files, containing data used in sample experiments, and be able to make an approximation to population variables.

2 Introduction

Massive data storage is a important matter when taking sample data from population, and it is useful to make forecast, approximations and estimates. In a pandemic situation, high level languages help us to manipulate big amount of data with advanced data structures. Governments make full

use of statistics methodologies, making estimates and measuring population variable, like number of unemployed's, born per day or the acceptance of certain policy. At the moment of take all that data, it is very important to make full use of computational systems to process all that data, and generate approximations as quick as possible. At the time of create a proper

implementation of a data structure, java provide us with three main tools to do that, encapsulation, extend a existing one or create a totally new one. Looking at all the tools and APIs that java has to offer, it is very useful to follow all previous functions that collections has to the users, like add or remove objects. Our approach allow us to keep it simple and standard with all java data structures, while adding new capabilities for statistics matters.

3 Problem

New data is generated every minute, and decisions need to be made in base to that data, considering pandemic data that need to be processed as quick as possible, to be presented and converted into information. DBMS allow us to maintain those data protected and maintained for a long time, but is not data structures, even if volatile, give us more flexibility at the moment of making calculations with DB items, and because not everything need to be made into a database, then great and efficient data structures have come this far to high level languages.

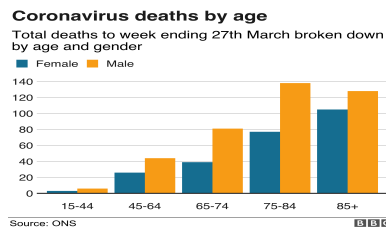


Figure 1: Corona virus deaths by age

The disposition of this kind of data as quick as possible is very important in hard moments, because it allow us to make wise decisions, so efficient implementations of data management must be studied and implemented.

4 Solution

Our implementation is a double linked list between nodes, and reference to the head, where new elements are added to the end of the list. The real value of the list are stored into the nodes.

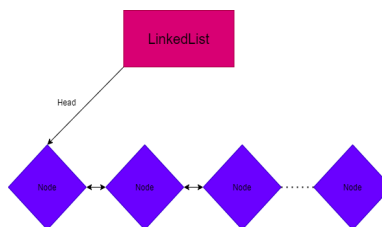


Figure 2: Linked List references

Figure 2 shows the reference between objects, where the only connection between the nodes and the list is the head, and nodes are connected in pairs.

The main goal of the linked list is to store values, so it used in a Statistics package that take all values from the list to make calculations of population variables.

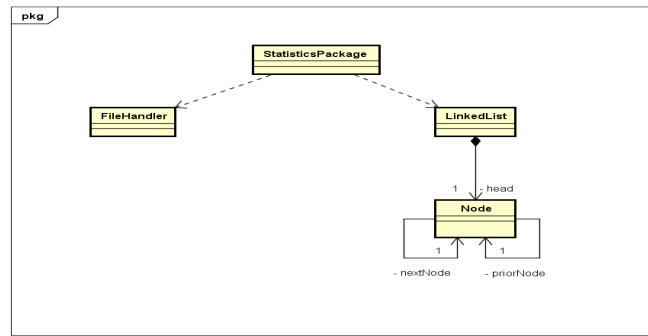


Figure 3: Class Diagram

Statistics package makes full use of the linked list and file input for receiving data for calculations. Proper test were made to the two main classes, Statistics Package and Linked List, with the focus in adding and removing, in the Linked List, and into the Statistics mean and standard deviation were tested.

```

T E S T S
-----
Running edu.escuelaing.arep.mavengit.LinkedListTest
Tests run: 5, Failures: 0, Errors: 0, Skipped: 0, Time elapsed: 0.423 sec
Running edu.escuelaing.arep.mavengit.StatisticsPackageTest
Tests run: 2, Failures: 0, Errors: 0, Skipped: 0, Time elapsed: 0.019 sec
Results :
Tests run: 7, Failures: 0, Errors: 0, Skipped: 0

[INFO] -----
[INFO] BUILD SUCCESS
[INFO] -----
[INFO] Total time: 11.986 s
[INFO] Finished at: 2020-08-13T03:02:02-05:00
[INFO] Final Memory: 14M/47M
[INFO] -----
  
```

Figure 4: Test successful

5 Conclusions

Statistics needs must be satisfied with high performance data structured that allow developers to easily manipulate data and make prognostics with sample data. A linked list is able to satisfy those needs, but, is this the best

data structure?. Best data structures should be used with high amounts of data, that give us better complexities and better performance.

References

- [1] how2examples. *The Java Collections API*. URL: <http://how2examples.com/java/collections>. (accessed: 13.08.2020).
- [2] Victor S.Adamchik. *Linked Lists*. URL: <https://www.cs.cmu.edu/~adamchik/15-121/lectures/Linked%20Lists/linked%20lists.html>. (accessed: 13.08.2020).