Navid Hariri

Professor Barnes

CS 182

September 29, 2018

Report

The linear simulation simply uses constant values for mean, median, and investment and simply compounds the investment 3,582 times. There is one equation for the mean and another for the median. However, both equations are identical, except for the part that one equation uses the mean that was provided and the other uses the median that was provided. After the investment is compounded 3,582 times, the results for each are then printed out. This simulation uses constants, thus there is no need to run this simulation more than once because every single trial would give the exact same answer. Thus there is only one trial that is done.

The distributional simulation is very similar to linear, however, instead of using constants for the mean it uses a random Gaussian to randomly provide a number random number based on a Gaussian or bell curved distribution. Thus most of the numbers given will be within one standard deviation apart. Based on the empirical rule, approximately 68% of the data would fall within this range, and that is where the random number generated by the Gaussian method mostly comes from. The standard deviation of this curve would dictate how close or far each data point is from each other and adding the mean or median would dictate where the curve would be located. The mean is the best point of estimation, however, if the graph is skewed in any direction, the median then becomes the best point of estimation because the mean would then be leaning towards the tail. Thus when the random Gaussian number is generated, it is multiplied by the standard deviation that was given to influence how far apart the numbers are and then the mean or median is added to dictate where the curve should be located. There are two different simulation runs in this method, one based on mean and the other based on median. However, the equations are almost identical except one adds the mean given and the other adds the median given. This is done for 3,582 weeks, which would be one be equivalent to one trial. However, unlike the linear simulation, distributional simulation is taking different numbers every time. Thus every trial would give approximately different numbers. In statistics, the sample size would be determined by population and confidence level, however, it is also largely agreed upon that generally 1,000 samples gives a relatively accurate answer for most cases. Thus, the calculation is done 3,582 times for one sample, and then that cycle is repeated 1,000 times to get a large sample to get the min, max, mean, median, and standard deviation.

For the Monte Carlo simulation it is also very similar to the distributional simulation. It also uses a random number and needs 1,000 samples. However, unlike the distributional, the random numbers are picked from the text file given. The text file, contains the dates and the values for each week since the 1950's. Each time a random value is picked from the 3,582 values

and is then multiplied and added by the investment. Since the values picked are random we will again need 1,000 trails. After 1,000 trials, the data is then stored in an ArrayList, and the statistics of it such as the min, max, mean, median, and standard deviation is taken from those stored data points and printed out.

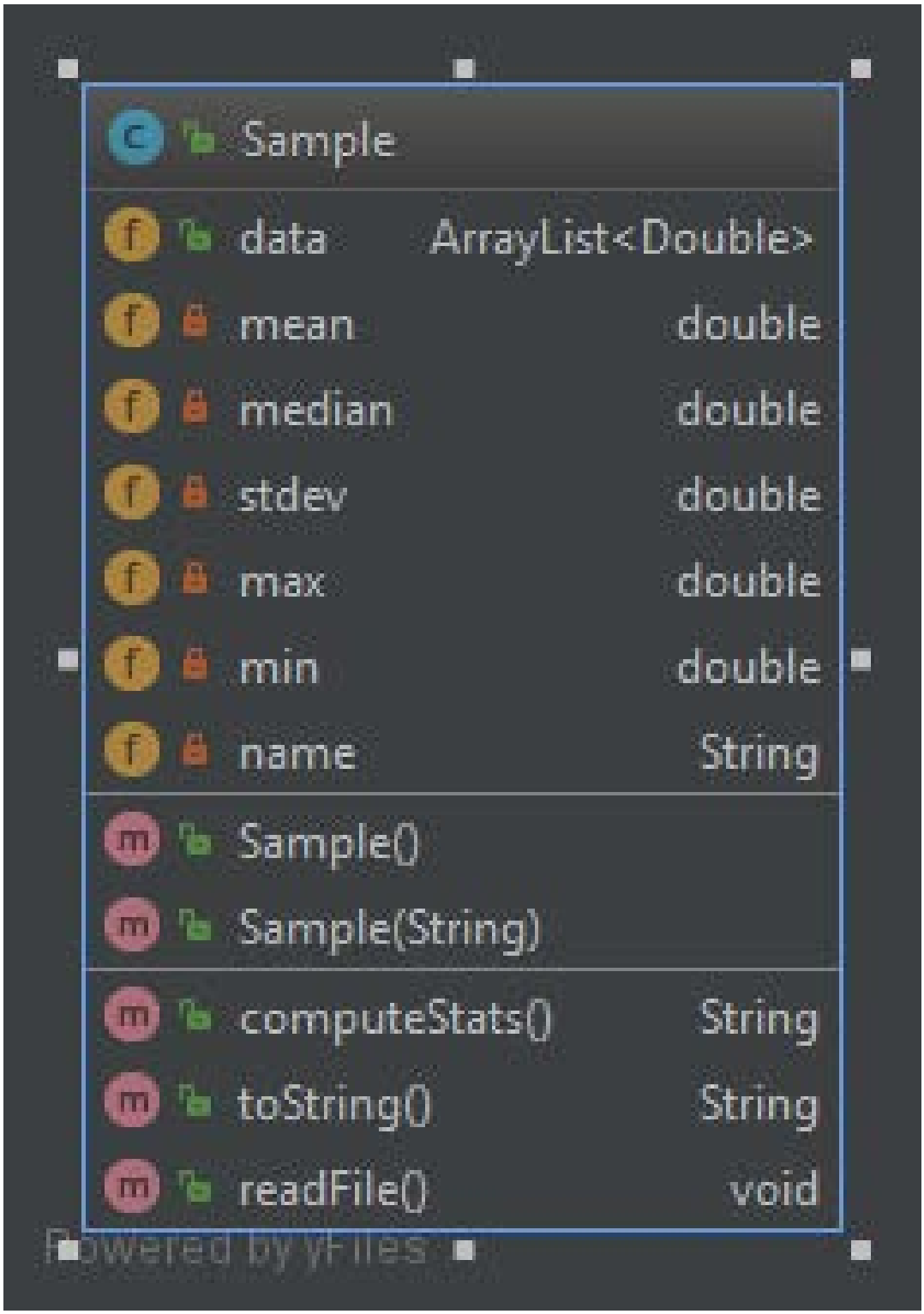
Results:

Linear Simulation for mean is: 30693.006635943697

Linear Simulation for median is: 3197682.649065718

Name	Size	Mean	Median	Standard Deviation
Distributional with Mean	1,000	32624.65	14884.18	55908.96
Distributional Simulation with median	1,000	3361298.94	1493400.28	7406670.34
Monte Carlo simulation	1,000	39271.27	17569.17	76460.22

UML Diagram for Sample class



UML Diagram for SimulateMarket class