Fractional Greedy Knapsack Growth Rate Investigation
CS 4310 – Design & and Analysis of Algorithms

Fall 2016

Maximilian Ellsberry

12/8/16

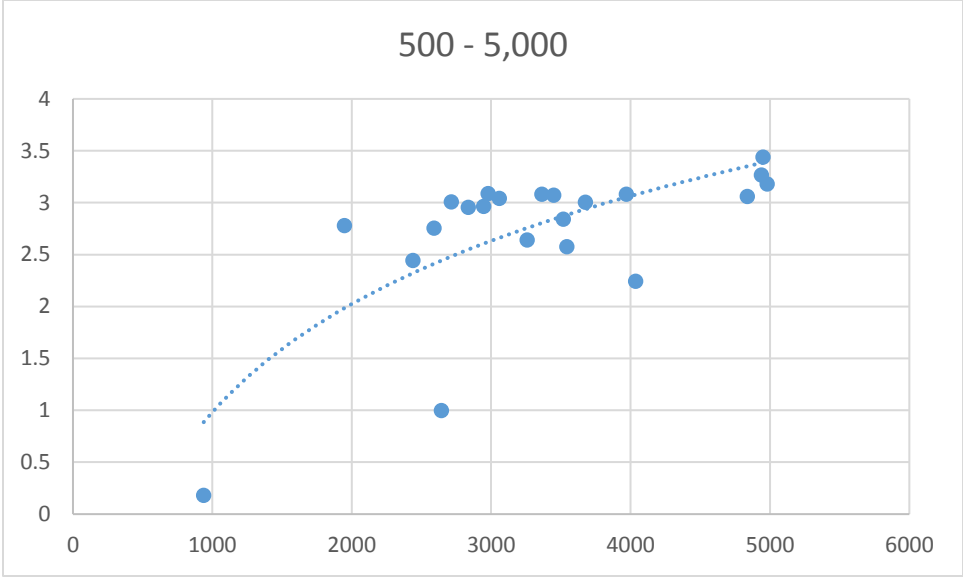
Hypothesis:

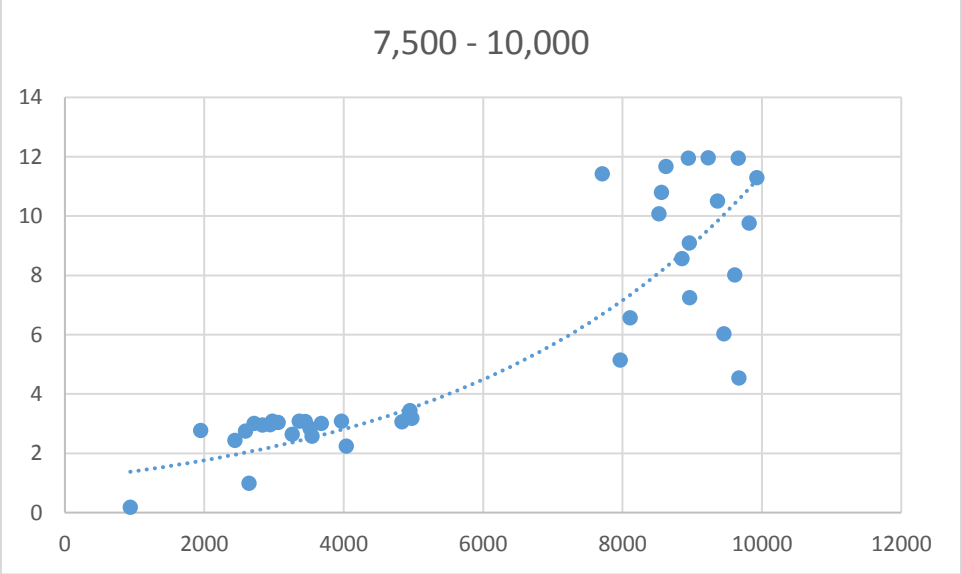
The Fractional Greedy Knapsack has a time complexity of O(n log (n))

Test Design:

1) Implement the Fractional Greedy Knapsack algorithm in python
2) Run code for different total Knapsack sizes
    a) Collected 40 data points for analysis
    b) Test three ranges of sizes for knapsack
        I. Test One: 500 – 5,000
        II. Test Two: 7,500 – 10,000
    c) Analyze the data points to make sure that growth rate will match that of our hypothesis
    d) Each data point is the average of a trial ran 10 times to minimize the outliers
3) At this point we can just look at the graph to see if our data is consisted if not visually then we must re work out experiment
4) After determining that the rate of change is similar use mathematical analysis to verify hypothesis

Data Evaluation:



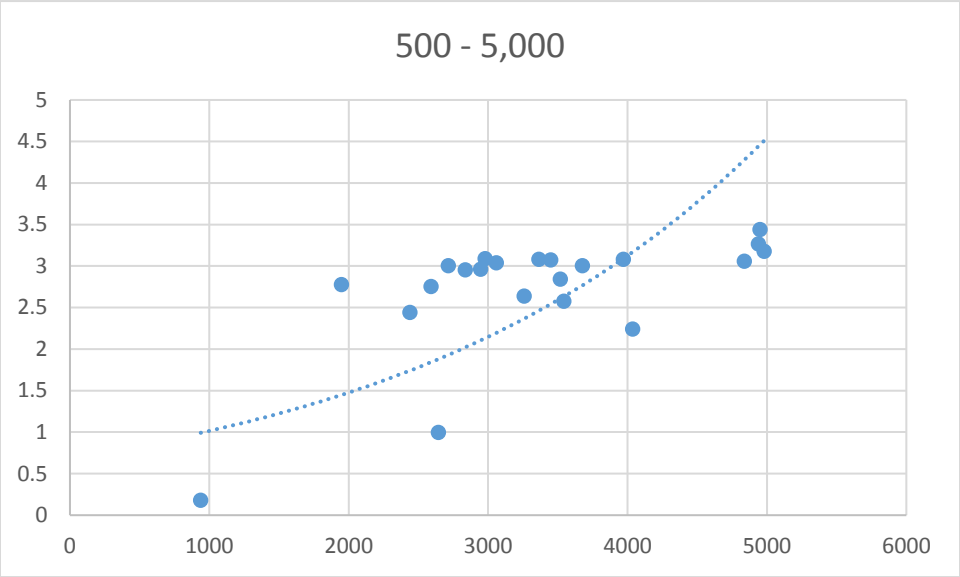


The two charts above are from the two ranges. The first range being sets between the size of 500 and 5,000. The second chart consists of both the first set and the second set with sizes from 7,500 and 10,000. One set seems to grow logarithmic and the other exhibits exponential. I will proceed with the mathematical analysis.

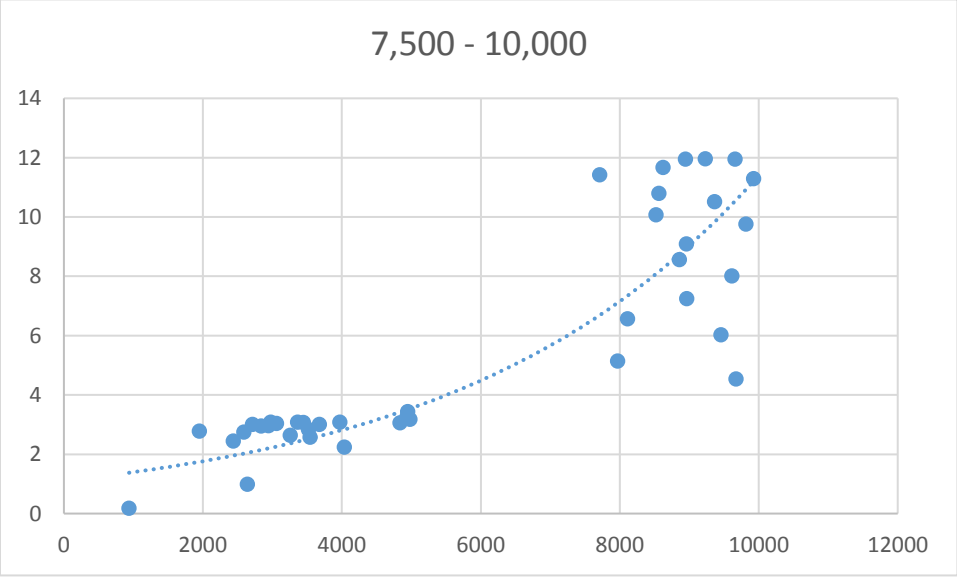
As mentioned in the example I used the Microsoft Excel to form a simple trend line analysis. For the sets of data I used all of the possible trend lines. I will go through the different trend lines and choose whether I will reject or accept them as the best fit for the data sets. I actually got very different trend lines for the 500 – 5,000 and the full set. So I will have to show both results for both sets.

Exponential Trendline:

The 500 – 5,000 set can be ruled out because we can visually see the trend line doesn't resemble the data.

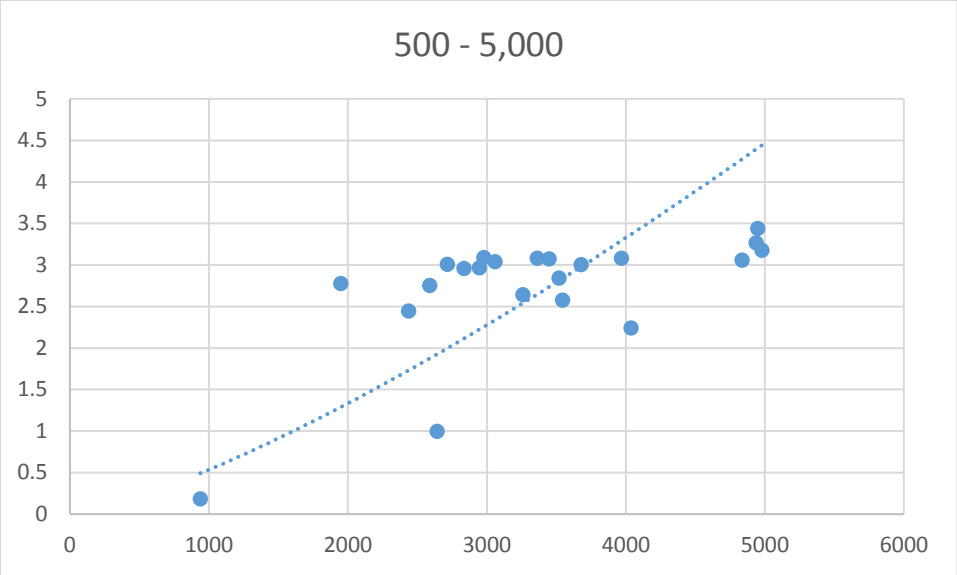


The 7,500 – 10,000 set visually matches up with the exponential visually and we can proceed to further investigate and will be explained in the conclusion.

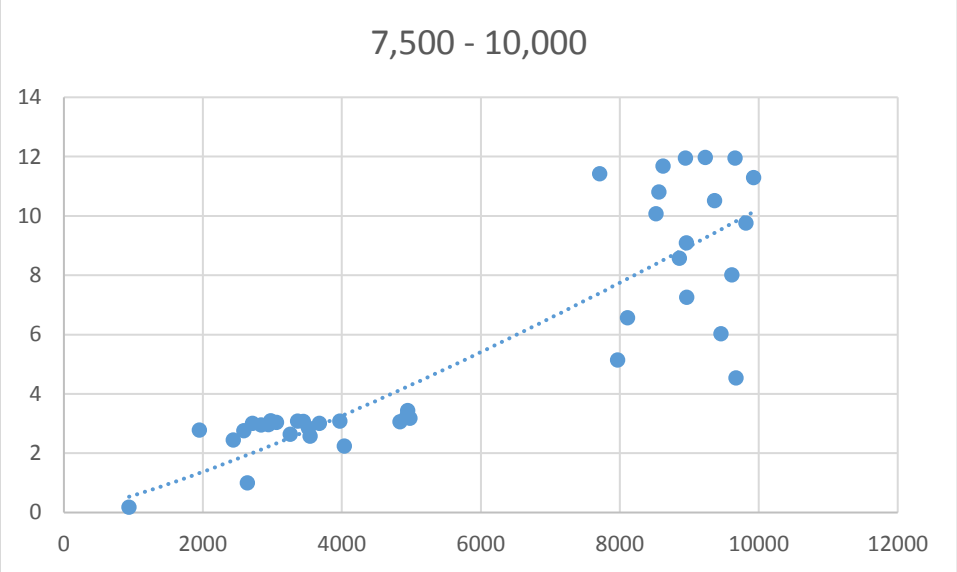


Power Trendline:

This data can be rejected because the trendline does not visually match up.

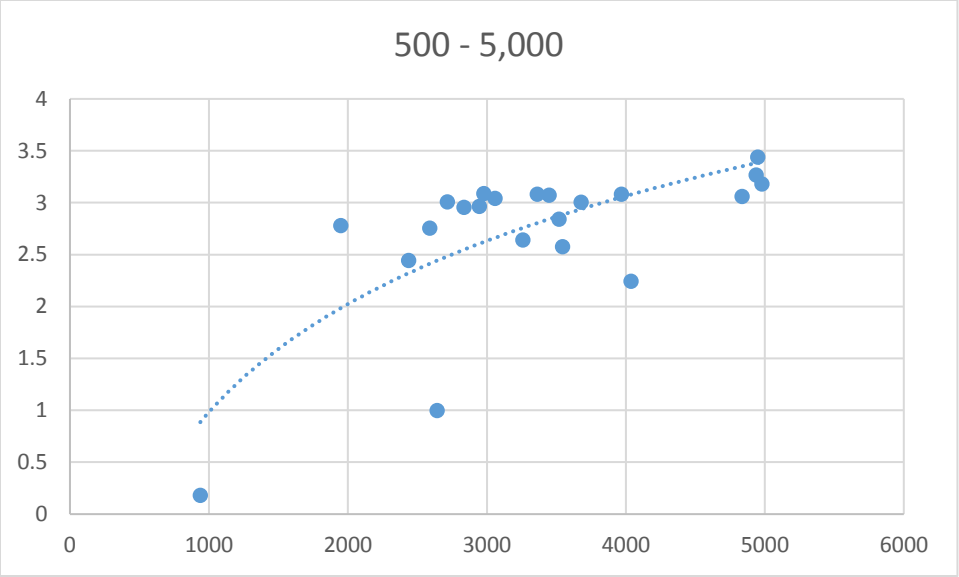


This data set can reject this trend line visually. I believe there is too many outliers for this data set to match up with the trend line.

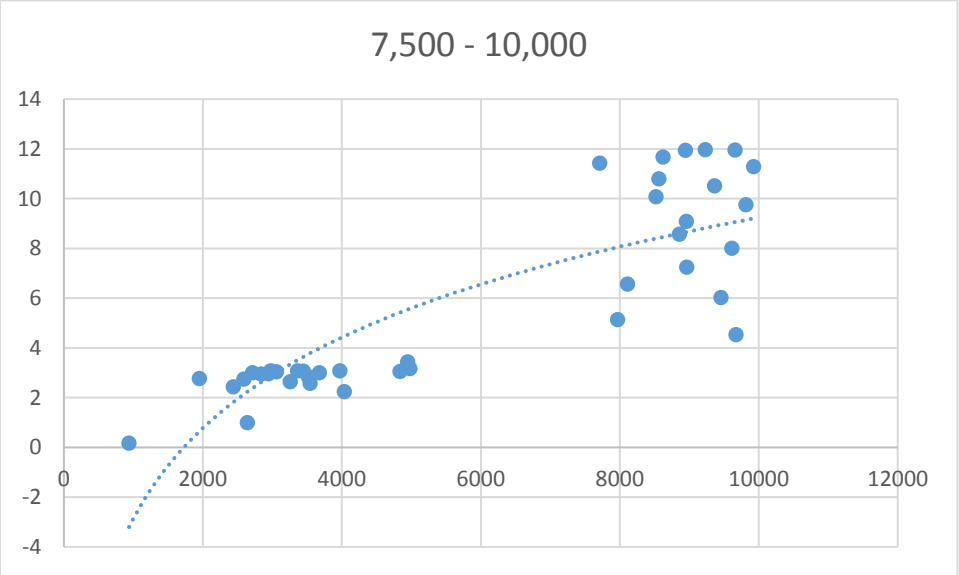


Logarithmic Trendline:

The first data set closely resembles the logarithmic trend with a few outliers. I think it is possible to accept this trendline as a possibility.

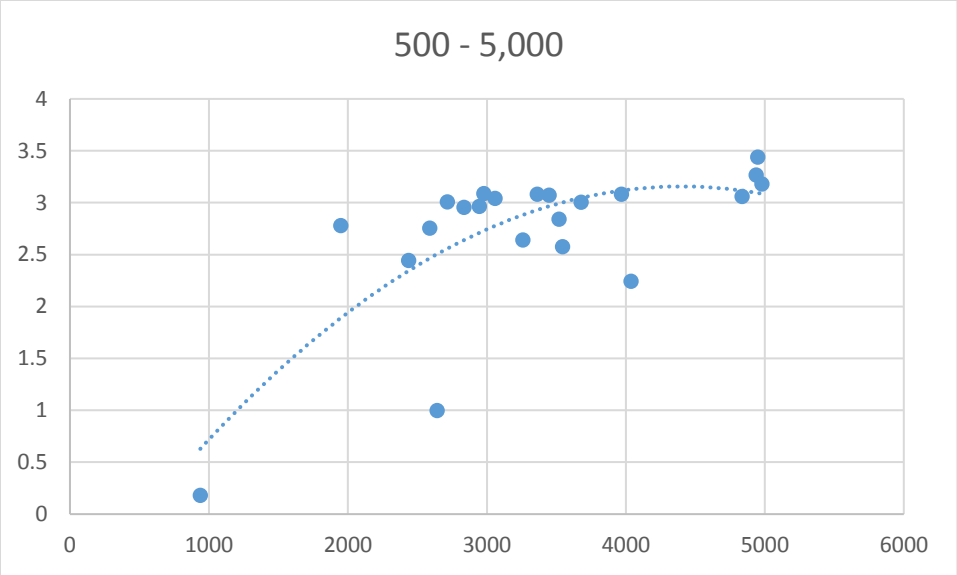


This set can be rejected through visually inspection. The data clearly does not follow the trend line.

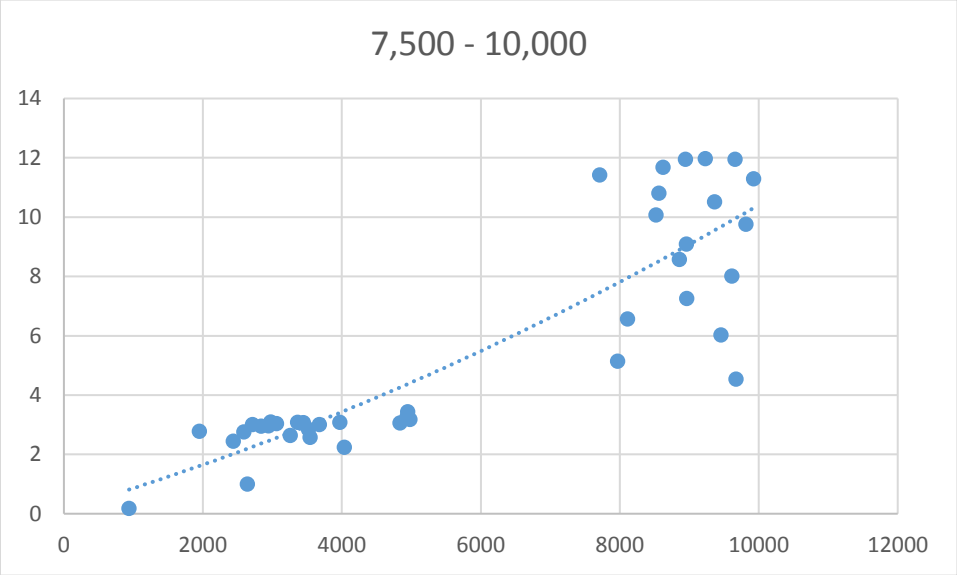


Polynomial Trendline:

The first data set resembles the polynomial trendline with the exception of a few outliers. It means that I will accept this as a possible trendline for the first set of data.

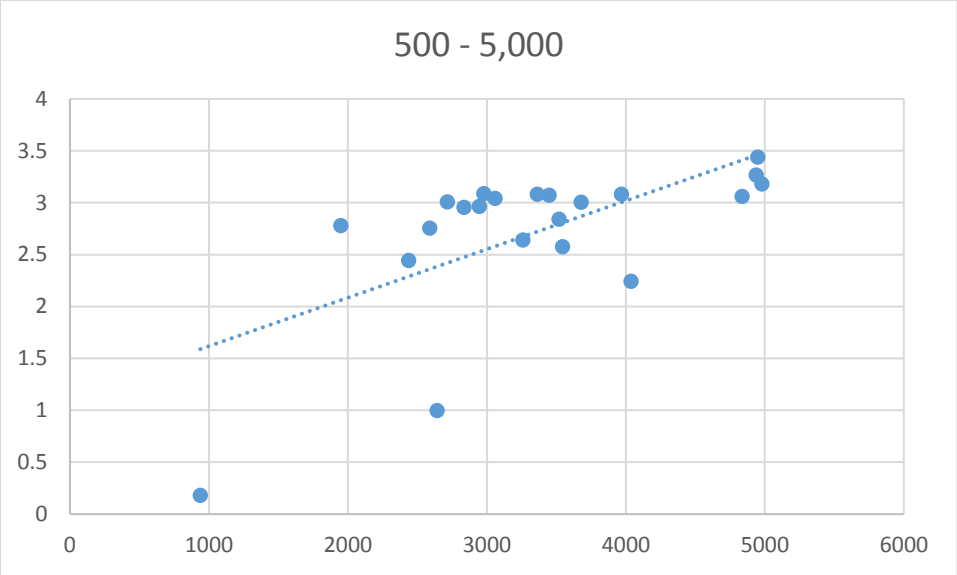


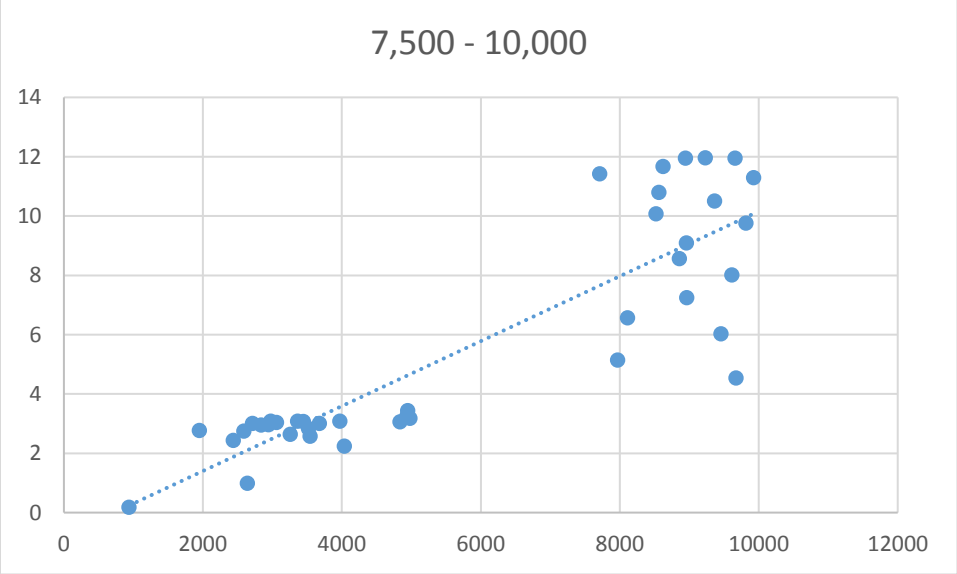
Through visual inspection we can reject this trendline as a possibility for the second data set.



Linear Trendline:

Both data sets can reject this trendline through visual inspection. The data trendline does not resemble the data at all.





Conclusion:

After careful analysis I determined that there are two possible possibilities that support the first set of data. The possibilities are Polynomial and Logarithmic trendlines are possible. This data analysis does show that the knapsack is indeed of time complexity O (n log n).