Geometric mean essay

Geometric mean From Wikipedia, the free encyclopedia Jump to: navigation, search The geometric mean, in mathematics, is a type of mean or average, which indicates the central tendency or typical value of a set of numbers. It is similar to the arithmetic mean, which is what most people think of with the word " average," except that instead of adding the set of numbers and then dividing the sum by the count of numbers in the set, n, the numbers are multiplied and then the nth root of the resulting product is taken. For instance, the geometric mean of two numbers, say 2 and 8, is just the square root (i. e.

the second root) of their product, 16, which is 4. As another example, the geometric mean of 1, ? , and ? is the cube root (i. e. , the third root) of their product (0. 125), which is ?.

The geometric mean can be understood in terms of geometry. The geometric mean of two numbers, a and b, is simply the side length of the square whose area is equal to that of a rectangle with side lengths a and b. That is, what is n such that n? = a ? b? Similarly, the geometric mean of three numbers, a, b, and c, is the side length of a cube whose volume is the same as that of a rectangular prism with side lengths equal to the three given numbers. The geometric mean only applies to positive numbers. [1] It is also often used for a set of numbers whose values are meant to be multiplied together or are exponential in nature, such as data on the growth of the human population or interest rates of a financial investment.

The geometric mean is also one of the three classic Pythagorean means, together with the aforementioned arithmetic mean and the harmonic mean.

Contents [hide] •1 Calculation o1. 1 Relationship with arithmetic mean of logarithms •2 Notes and references •3 See also •4 External links [edit] CalculationThe geometric mean of a data set [a1, a2, …

, an] is given by . The geometric mean of a data set is less than or equal to the data set's arithmetic mean (the two means are equal if and only if all members of the data set are equal). This allows the definition of the arithmetic-geometric mean, a mixture of the two which always lies in between. The geometric mean is also the arithmetic-harmonic mean in the sense that if two sequences (an) and (hn) are defined: and then an and hn will converge to the geometric mean of x and y. [edit] Relationship with arithmetic mean of logarithmsBy using logarithmic identities to transform the formula, we can express the multiplications as a sum and the power as a multiplication.

This is sometimes called the log-average. It is simply computing the arithmetic mean of the logarithm transformed values of ai (i. e. the arithmetic mean on the log scale) and then using the exponentiation to return the computation to the original scale. I.

e. , it is the generalised f-mean with f(x) = ln x. Therefore the geometric mean is related to the log-normal distribution. The log-normal distribution is a distribution which is normal for the logarithm transformed values.

We see that the geometric mean is the exponentiated value of the arithmetic mean of the log transformed values, i. e. emean(ln(X)). Calculating Geometric Means Geometric Mean Calculations by Dr. Joe Costa Definition of Geometric Mean Mathematical definition: The n-th root of the product of n numbers. Practical definition: The average of the logarithmic values of a data set, converted back to a base 10 number.

Geometric Means for Water Quality Standards Many wastewater dischargers, as well as regulators who monitor swimming beaches and shellfish areas, must test for and report fecal coliform bacteria concentrations. Often, the data must be summarized as a “ geometric mean” (a type of average) of all the test results obtained during a reporting period. Typically, public health regulations identify a precise geometric mean concentration at which shellfish beds or swimming beaches must be closed. A geometric mean, unlike an arithmetic mean, tends to dampen the effect of very high or low values, which might bias the mean if a straight average (arithmetic mean) were calculated. This is helpful when analyzing bacteria concentrations, because levels may vary anywhere from 10 to 10, 000 fold over a given period.

As explained below, geometric mean is really a log-transformation of data to enable meaningful statistical evaluations. Other Uses of Geometric Means Besides being used by scientists and biologists, geometric means are also used in many other fields, most notably financial reporting. This is because when evaluating investment returns and fluctuating interest rates, it is the geometric mean, not the arithmetic mean, that tells you what the average financial rate of return would have had to have been over the entire investment period to achieve the end result. Financial Return CalculationFor financial investment return calculations, the geometric mean is calculated on the decimal multiplier equivalent values, not percent values (i. e.

, a 6% increase becomes 1. 06; a 3% decline is transformed to 0. 97. Just follow the steps outlined in the section below titled Calculating Geometric Means with Negative Values).

The equation is also flipped around when calculating the financial rate of return if you know the starting value, end value, and the time period. This equation is used in these cases when the average rate of return is needed (or population growth rate): Note: If you subtract 1 from the equation above, this is your compound interest rate. To use this equation, if years= 5, this is the “ fifth root”, which is the same as raising to the power of 1/5 or 0. 2). Problem submitted by a student: “ A recent article suggested that if you earn $25, 000 a year today and the inflation rate continues at 3 percent per year, you’ll need to make $33, 598 in 10 years to have the same buying power. .

.. Confirm that this statement is accurate by finding the geometric mean rate of increase” Solution using a formula in Excel: = Power(33598/25000,. 1)= 1. 03When to Use or Not Use Geometric Mean Geometric mean is often used to evaluate data covering several orders of magnitude, and sometimes for evaluating ratios, percentages, or other data sets bounded by zero. If your data covers a narrow range (I have seen it stated that the largest value must be at least 3x the smallest value), or if the data is normally distributed around high values (i.

e. skew to the left), geometric means and log transformations may not be appropriate. Do not use geometric mean on data that is already log transformed such as pH or decibels (dB). Geometric Mean CalculationHow do you calculate a geometric mean? The easiest way to think of the geometric mean is that it is the average of the logarithmic values, converted back to a base 10 number.

However, the actual formula and definition of the geometric mean is that it is the n-th root of the product of n numbers, or: Geometric Mean = n-th root of (X1)(X2)… (Xn) Where X1, X2, etc.

represent the individual data points, and n is the total number of data points used in the calculation. If this is the definition of geometric mean, why is my first statement true, that geometric mean is really the average of the log values? Consider this example. Suppose you wanted to calculate the geometric mean of the numbers 2 and 32. This simple example can be done in your head. First, take the product; 2 times 32 is 64. Because there are only two numbers, the n-th root is the square root, and the square root of 64 is 8.

Therefore the geometric mean of 2 and 32 is 8. Now, let’s solve the problems using logs. In this case, we will convert to base-2 logs so that we can solve the problem in our head (in fact, any base could be used). Converting our numbers, we have: 2= 21 32= 25 21 x 25 = 26 (= 64) he square root of 26 is 23 (= 8) Of course, the short cut to solve the problem is to take the average of the two exponents (1 and 5) which is 3, and 23 is 8. Problem: Can you calculate the geometric mean of these 5 numbers, in your head? 23, 25, 28, 23, 21 (These values of course equal 8, 32, 256, 8, and 2) (Hint: The 5 exponents add up to 20.) Click for the answer.

From the discussion above, you can see that the calculation of the Geometric Mean can be performed by either of two procedures on a calculator, depending upon which functions are available. Computer-based spreadsheet programs like Excel have built geometric mean functions, and in general you should use these (see below) to save time if a computer with the appropriate software is available. Calculation Procedure 1: Multiply all of the data points, and take the n-th root of this product. Example: Suppose you have this beach monitoring data from different dates: (data are Enterococci bacteria per 100 milliliters of sample) 6 ent. /100 ml 50 ent. /100 ml 9 ent.

/100 ml 1200 ent. /100 ml Geometric Mean = 4th root of (6)(50)(9)(1200) 4th root of 3, 240, 000 Geometric Mean = 42. 4 ent. /100 ml On a good scientific calculator, you would multiply the numbers together, press equal, then the root key, then the number 4 to get the forth root (or enter 0. 25 with the exponent key on the last part). Calculation Procedure 2: Take the average of the logs, then convert to a base 10 number Of course, many calculators do not have a root key that allow the calculation of any root, so you must use the logarithm function, which is typically more widely available on calculators.

To use this calculation procedure, you must have a calculator which will give logarithms (log or ln) and anti-logarithms (exp or e). The first step in calculating the Geometric Mean using this method is to determine the logarithm of each data point using your calculator. Next, add all of the data point logarithms together and divide this sum by the number of data points (n). In other words, take the average of the logs.

Next, convert this log average back to a base 10 number using the antilogarithm function key on the calculator. Example (using previous data):

log 6= 0. 7815 log 50= 1. 69897 log 9= 0. 95424 log 1200= 3. 07918 Sum= 6.

51054 The logarithm of the Geometric Mean is 6. 51054/4 = 1. 62764 (the average of the logs) From your calculator, determine the number whose logarithm is 1. 62764 (use the antilogarithm key), and you will find that the Geometric Mean = 42.

4 ent. /100 ml This process works whether or not you use natural logs (“ ln” key) or base 10 logs (“ log” key). That is, on your calculator you could do ln(x1), ln(x2), etc. then use the ‘ ex’ key on the average of the logs, or you would do log(x1), log(x2), etc. hen use the ’10x’ key on the average of the logs.

(key names may vary among calculators). Incidentally, for this example data set, the arithmetic mean (average) of the four data points is: Arithmetic Mean = (6 + 50 + 9 + 1200)/4 = 1265/4 Arithmetic Mean = 316. 3 colonies/100 ml The geometric mean is always less than the arithmetic mean (except of course if all the data points have an identical value). On most scientific calculators your key sequences to calculate the geometric mean would be: enter a data point, press either the Log or ln function key, ecord the result or store it in memory, calculate the mean or average of these log values, calculate the antilog value of this mean (’10x’ key if you used ‘ Log’ key, ‘ ex’ key if you used ‘ ln’ key) Excel #Num! overflow error In Excel and Quattro an error may be obtained in the geometric mean function if you apply the function to a very long list of numbers. This occurs because of a numeric overflow error (the product of the numbers is so large the software cannot compute them the way the software is written). If this occurs, you can use an " array formula.

An array formula is one that repeats the same calculations over an array (list) of numbers. This " average of the logs" formula will work fine in such situations: {= EXP(AVERAGE(LN(A1: A200)))} Do not enter the curly brackets. Enter the formula "= EXP(A…

. ", then create the array formula by pressing Control+Shift+Enter simultaneously on your keyboard while your cursor is inside the formula cell. Change A1 and A2 to the actual locations of the first and last values of the data set. Calculating Geometric Means in SpreadsheetsRather than using a calculator, it is far easier to use spreadsheet functions.

For example, in Microsoft Excel™ the simple function " GeoMean" is provided to calculate the geometric mean of a series of data. For example, if you had 11 values in the range A1…

A10, you would simply write this formula in any empty cell: '= geomean(A1: A10)'. In Corel Quattro™ spreadsheets, the function is '@geomean(A1.. A10)'. In both programs, you can enter values directly inside the parentheses (x1, x2, x3) instead of referencing a range of cells. Calculating Geometric Means with Zero ValuesThe calculation of the Geometric Mean may appear impossible if one or more of the data points is zero (0).

In these cases, however, the convention used is that a value of either ' 1', one half the limit of detection, or some other substitution is allowed for each zero or " less than" value, so that the information contained in these data is not lost. For example, the US Food and Drug Administration in its shellfish sanitation program regulations requires the substitution of a value that is one significant digit less than the detection limit [i. e. " less than 2" becomes " 1.

"]. Because of how geometric mean is calculated, the precise substitution value generally does not appreciably affect the result of the calculation, and ensures that all the data remains usable. Here is an example with a non detect (and assuming the detection limit was 2 bacteria per 100 milliliters): 1100 0 (" less than 2") 30 13000 Geometric Mean = 4th root of 1100 X 1 X 30 X 13000 = 4th root of 429, 000, 000 Geometric Mean = 143. 9 Incidentally, substituting 1. 9 for the less than value results in a geometric mean of 169.

0, which is nearly statistically different (alpha= 0. 5) using a t-test using the substituted value 1. 0. See additional comments in the bacteria data section below.

Debate on the use of substitutions of below reporting limits and other censored data Many statisticians have criticized common procedures for providing substituted values for non-detects or below-reported-limits value data. Other alternatives, such as " delta log-normal models" have also received criticism and even legal challenges when applied to regulatory discharges permits. These problems and alternative analysis strategies are presented in Helsel (1990, 2005) and EPA (2002). These references also contain useful citations to other publications. Statistical tests on bacterial data All statistical tests used to evaluate variable bacterial data (i.

e. a range of values over orders of magnitude) should be employed using the means, variances, or standard deviations of the log-transformed data. However, a special problem is created when reporting standard deviations of log data. That is because plus or minus (+/-) a log constant creates unequal error bars when converting back to base 10 (see note below on plotting geometric means). To overcome other log transformation problems, values less than detection limits should be replaced with non-zero value to avoid log of zero errors. As noted above, certain regulatory programs, like the US FDA requires the substitution of a number one significant digit less than the detection limit [i.

e. " less than 2" becomes " 1. 9"], or one significant digit larger for exceedances [i. e.

" Greater than 1600" becomes " 1700"] under their shellfish sanitation program regulations. Other agencies have required models to predict the variance of these below-reported-limits data. Another special problem that exists with bacteria testing is that bacterial plates can be inundated with bacteria so that bacteria colony forming units are expressed as exceeding a certain number. These " greater than" values are similarly converted for geometric mean calculations (FDA requires conversion to the next significant digit ("> 1200" becomes " 1300"). Regulatory programs like these also have water quality standards that incorporate median values and 90th percentile values because of concerns about possible non-normal distributions of even the log-transformed data. The calculated means and variances of log-transformed data can be plugged into a t-Test to evaluate whether there is a statistical between two stations.

To answer the question whether there is a statistical difference among three or more stations, use an ANOVA test. When analyzing log-transformed data, you may be surprised to find that two sites with remarkably different arithmetic means may be not statistically different from one another. The substitution values for non-detects can sometimes affect the outcomes of statistical tests, especially in cases where a large percentage of the data are non-detect or zero. Helsel (1990, 2005) describes a variety of tests and approaches that are more robust and valid in evaluating this type of data. Plotting log transformed data It is relatively easy to plot log-transformed data in spreadsheet programs. When graphing standard deviations or standard errors around a mean, your error bars will be of equal size above and below the mean if you plot on log paper or apply a log scale in a spreadsheet program.

However, the error bars will be unequal if the y axis is not log transformed. Calculating Geometric Means with Negative ValuesLike zero, it is impossible to calculate Geometric Mean with negative numbers. However, there are several work-arounds for this problem, all of which require that the negative values be converted or transformed to a meaningful positive equivalent value. Most often this problem arises when it is desired to calculate the geometric mean of a percent change in a population or a financial return, which includes negative numbers. For example, to calculate the geometric mean of the values +12%, -8%, and +2%, instead calculate the geometric mean of their decimal multiplier equivalents of 1. 12, 0.

2, and 1. 02, to compute a geometric mean of 1. 0167. Subtracting 1 from this value gives the geometric mean of +1.

67% as a net rate of population growth (or financial return). Incidentally, if you do not have a negative percent value in a data set, you should still convert the percent values to the decimal equivalent multiplier. It is important to recognize that when dealing with percents, the geometric mean of percent values does not equal the geometric mean of the decimal multiplier equivalents. For example: Geometric mean of [12%, 4%, 2%] does not equal the Geometric mean of [1. 2, 1. 04, 1.

02]. 4. 6% does not equal 5. 9% Calculating Geometric Means with Both Large Negative and Positive Numbers Combined I have received a number of queries, particularly from those analyzing gene block microarray data sets, about how to calculate geometric means on data sets that include both very large and very positive numbers. The analysis of data to evaluate the similarity of gene blocks is a complex topic and the statistics of this field is evolving, and you should perform an internet search to find the latest thinking on this topic. However, in principal, comparing data sets consisting of very large negative and positive numbers together is an easy matter, and all that is required is to temporarily suspend the negative signs of the data.

Consider, for example, two sets sample data sets as follows: | A= {-5,-3,-2, 3} and B={-1, 0, 2, 4} The mean of data set A is -1. 75, and the mean of data set B is +1. 25. A simple Student’s t-test (assuming alpha= 0. 05 and equal sample variances) would suggest these samples are not statistically different from one another. This approach would be no different than if you were to calculate geometric mean in these two data sets: | A’={-100000, -1000, -100, 1000} and B’={-10, 1, 100, 10000} If you were to take off the negative signs, take the log, then add the negative sign back on, you could then compare the means of the A’ and B’ data sets.

In fact, you might have noticed that data sets A and B are really the log (base 10) transformed data sets A’ and B’. You might therefore conclude that A’ and B’ are not statistically different samples using the same t-test. Of course like any statistical analysis you have to make sure you have not violated the assumptions of the statistical test (in this case you must assume the log transformed data is normally distributed, and the sample variances were equal). Geometric Mean of Grouped Data A student recently posed this question: How do you calculate the geometric mean on grouped data? That is to say, when the data exists as a data range and frequency, what formula do you use? As per the discussion above, there are two ways to approach this problem: Method 1: (hardest for grouped data): Calculate the product of all the values in the data set (frequency of each mid-point value), then take the nth root of the product, with n being equal to the cumulative frequency. Method 2: (easiest for grouped data): Calculated the average weighted mean of the logarithm of each mid-point value, then convert this mean value back to a base 10 number. These two statements are best illustrated by the sample data set in the table below.

Arithmetic Mean CalculationGeometric Mean Calculation (Meth. 1) rangefrequencymidpointfreq x midln(midpt)freq x ln(midpt) 10 to ;= 20315452. 7088. 124 20 to ;= 309252253. 21928.

970 30 to ;= 405351753. 55517. 777 Total1744554. 871 arithmetic mean= 26. 176arithmetic mean of weighted ln= 3.

228 Using Method 1, you would take the 17th root of the product 15 x 15 x 15 x 25 x 25 x 25 x 25 x 25 x 25 x 25 x 25 x 25 x 35 x 35 x 35 x 35 x 35, which is also equal to 25. 221. In a spreadsheet, you would type this formula: =((15^3)*(25^9)*(35^5))^(1/17)As you might imagine, if you have large mid-point values or large frequencies, your calculator or spreadsheet program could not compute the formula because the intermediate numbers are impossibly large, and the result would be an error. To calculate geometric mean in these cases, you must use Method 2.

You might also consider the spreadsheet " array formula" method in the " Excel #Num! overflow error" callout box above. If your grouped data includes large negative numbers, you have no choice but think of a clever transformation to make the values positive and use Method 2. For Method 2, as shown in the table above, you would calculate the weighted mean of the natural logarithms of the mid-point values, which in this case is 3. 228. When the value is converted back to base 10, the geometric mean is 25. 221.

Interestingly, this problem is quite similar to one faced by the Buzzards Bay NEP, in evaluating the extent of oiling from an oil spill. In this case the data consisted of an average width and the length of the beach. For example, 1500 ft of beach may have had between 0 and 5 foot-wide band of oil, 10, 000 feet may have been documented to have a band of oil between 5 and 10 ft, etc. The length of beach oiled became the frequency for the interval. Whether geometric mean is an appropriate metric for evaluating this type of data, or any other data set, always needs to carefully considered.

Working Backwards This following problem was posed by a student: If Geomean(8, a)= 12, what is a? The question can be most easily be rephrased using the nth root definition of geometric mean. That is: square root of (8 x a)= 12 solve first by squaring both sides: (8xa)= 144 a= 144/8 = 18 Using logs, the mathematical solution is: First express the problem as the mean of logs: (ln(8)+ln(a))/2 = ln(12) Solving: n(8)+ln(a) = 2 x ln(12) ===; ln(a)=(2 x ln(12))-ln(8) ===; a = exp((2 x ln(12))-ln(8)) ===; a= exp(2. 8904) ===; a= 18 Answer The answer to the mental math problem above: The exponents add to 20, 20 divided by 5 is 4, so the geometric mean is 24 or 16. geometric mean Frames not supported Definition A measure of central tendency calculated by multiplying a series of numbers and taking the nth root of the product, where n is the number of items in the series.

The geometric mean is often used when finding an average for numbers presented as percentages. This content can be found on the following page: ttp://www. investorwords. com/cgi-bin/getword.

cgi? id= 5896= geometric%20mean Geometric Mean The average of a set of products, the calculation of which is commonly used to determine the performance results of an investment or portfolio. Technically defined as “ the ‘ n’th root product of ‘ n’ numbers”, the formula for calculating geometric mean is most easily written as: Where ‘ n’ represents the number of returns in the series. The geometric mean must be used when working with percentages (which are derived from values), whereas the standard arithmetic mean will work with the values themselves. The main benefit to using the geometric mean is that the actual amounts invested do not need to be known; the calculation focuses entirely on the return figures themselves and presents an “ apples-to-apples” comparison when looking at two investment options . Question Corner and Discussion Area ____ Applications of the Geometric Mean Asked by Senthil Manick on May 22, 1997: When would one use the geometric mean as opposed to arithmetic mean? What is the use of the geometric mean in general? The arithmetic mean is relevant any time several quantities add together to produce a total.

The arithmetic mean answers the question, “ if all the quantities had the same value, what would that value have to be in order to achieve the same total? ” In the same way, the geometric mean is relevant any time several quantities multiply together to produce a product. The geometric mean answers the question, “ if all the quantities had the same value, what would that value have to be in order to achieve the same product? ” For example, suppose you have an investment which earns 10% the first year, 50% the second year, and 30% the third year. What is its average rate of return? It is not the arithmetic mean, because what these numbers mean is that on the first year your investment was multiplied (not added to) by 1. 10, on the second year it was multiplied by 1.

60, and the third year it was multiplied by 1. 20. The relevant quantity is the geometric mean of these three numbers. The question about finding the average rate of return can be rephrased as: “ by what constant factor would your investment need to be multiplied by each year in order to achieve the same effect as multiplying by 1. 10 one year, 1.

60 the next, and 1. 20 the third? The answer is the geometric mean . If you calculate this geometric mean you get approximately 1. 283, so the average rate of return is about 28% (not 30% which is what the arithmetic mean of 10%, 60%, and 20% would give you).

Any time you have a number of factors contributing to a product, and you want to find the “ average” factor, the answer is the geometric mean. The example of interest rates is probably the application most used in everyday life. Here are some basic mathematical facts about the arithmetic and geometric mean: Suppose that we have two quantities, A and B. Taking their arithmetic mean we get the number (A+B)/2 which can be interpreted in a number of ways. One interpretation (probably the most common) is that this quantity is the midpoint of the two numbers viewed as points on a line.

Now suppose that we have a rectangle with sides of lengths A and B. The arithmetic mean can also be interpreted as the length of the sides of a square whose perimeter is the same as our rectangle. Similarly, the geometric mean is the length of the sides of a square which has the same area as our rectangle. It is known that the geometric mean is always less than or equal to the arithmetic mean (equality holding only when A= B).

The proof of this is quite short and follows from the fact that is always a non-negative number. This inequality can be surprisingly powerful though and comes up from time to time in the proofs of theorems in calculus.

____ Asked by G. Ellis, student, Southeast Bulloch High on January 16, 1997: Could you give the formula for the geometric mean for a series of numbers ifI am trying to get the compound annual growth rate for a series of number that include negative numbers? In general, you can only take the geometric mean of positive numbers.

The geometric mean of numbers is the nth root of the product . In your example, you are taking the mean of positive numbers. For example, if you're looking at an investment that increases by 10% one year and decreases by 20% the next, the simple rates of change are 10% and -20%, but that's not what you're taking the geometric mean of. At the end of the first year you have 1.

times what you started with (the original plus another tenth of it). At the end of the second year you have 0. 8 times what you started the second year with (the original minus one fifth of it). So, the numbers you are taking the geometric mean of are 1.

1 and 0. 8. This mean is approximately 0. 938. This means that, on average, your investment is being multiplied by 0.

938 (= 93. 8%) each year, a 6. 2% loss. So, the compound anual growth rate is (approximately) -6.

2%. ____ Asked by Paul van Esbroeck on October 5, 1997: (Question abridged from original posting)I am presently engaged in a dispute with a bank concerning a Stock Market Tracker GIC. I have found significantly different definitions for the terms " average percentage growth" and " average percentage growth rate", and in many instances I cannot distinguish between the use of the words " growth" vs. "

growth rate" in the literature. It was in investigating growth that I came to your question page about the geometric mean.

While I, like most people I asked, originally understood these GICs to be guaranteeing a rate of return equal to the rate of growth of the index, the bank has a different interpretation. Let's say a stock market index starts at 1000, and at the end of 1 month is 1010, at the end of 2 months is 1020, and so on, ending up at 1120 after one full year. The bank seems to be considering the growth by month end for each month (which is 10 for the first month, 20 for the second, and so on, with this growth being 120 by the end of the twelfth month), then averaging these growths, obtaining an " average growth" of 60 and an average percentage growth of 60/1000 or 6%. For other definitions there seems to be quite a bit of agreement that: •The real growth was 120/1000 = 12%.

The average percentage growth rate is 12% per year. •The annual compound growth rate is 12%. It seems to me that it is incorrect to average growth as done by the bank. Since an average should be of equal periods, in the banks formula we average growth for the first month with growth for the first 11 months.

Since the growth was 10 index units each month, should the average growth not be 10 units per month, but how is this different from the average growth rate? What if not growth, do you call the plot of (Index value minus Index value at some starting time) ? What if any, is the correct distinction between growth and growth rate ? Do you find the term " average percentage growth" as used by the bank problematic? First let's try to get the words sorted out. Then we can address the situation you describe. The word " growth" is often used quite loosely to mean any of the above concepts. The best and most correct definition of growth would be a quantity that is associated to a particular period in time, and describes the index value at the end of the period minus the index value at the beginning of the period. In your example, the growth for the period consisting of the first month was 10, and the growth for the period consisting of the first 12 months was 120. " Growth rate" describes growth per unit time.

It may vary with time. The " average growth rate" over a period is the growth divided by the length of the period. If the growth rate is constant over a period, then the average growth rate over the period will be the same as that constant value. In your example the growth rate was a constant 10 units per month.

The average growth rate over every period was also 10 units per month. For example, the average growth rate during the first month was 10 units per 1 month = 10 units per month. The average growth rate during the entire year was 120 units per 12 months = 10 units per month. These concepts are not relevant to typical investements, stock market indices, etc.

, because the growth rate tends to be proportional to the index value. After all, if a 10 dollar investment grew by 1 dollar over a year (an average growth rate of 1 dollar per year), you'd expect a 10, 000 dollar investment to grow by 1000 dollars per year: a very different growth rate! So instead, the relevant concepts are percentage growth and growth rates. The percentage growth over a period is the ratio of the growth to the starting value. In your example, the percentage growth during the first month was 1%. The percentage growth during the second month was about 0.

99% (10/1010). The percentage growth during the entire year was 12% (120/1000). The percentage growth rate is the ratio of the growth rate to the current value. Thus, if an index value is 1000 and it is growing at 10 units per month, it is experiencing a percentage growth rate of 1% per month. If its value is 1010 and it is growing at 10 units per month, its percentage growth rate is about 0. 99% per month.

And so on. In your exaple, by the end of the year it is still growing at 10 units per month and its value is 1120, so the percentage growth rate by the end of the year has dropped to about 0. 89%. The only tricky thing about percentage growth rate is in changing units: a percentage growth rate of 1% per month is not the same thing as 12% per year. If an index were growing at a constant percentage growth rate of 1% per month, after one month it would be 1.

1 times its starting value, after two months it would be times its starting value, and after 12 months it would be times its starting value. Therefore, a percentage growth rate of 1% per month is the same as a percentage growth rate of about 12. 7% per year. If the percentage growth is G over a period of T time units, then the average percentage growth rate over that period is To figure out the average percentage growth rate of your example over the entire year: if your units are years, then T= 1.

G = 12%, so the average percentage growth rate is per year. If your units are months, the average percentage growth rate is which is about 0. 49%

per month. These are each saying the same thing; an average percentage growth rate of 12% per year is the same as an average percentage growth rate of about 0. 949% per month. Often, in investment circles, when people refer to “ growth” or “ growth rate” they are meaning percentage growth and percentage growth rate.

Now, on to your particular situation. As you point out, it is not particularly meaningful to average the growths over time periods of wildly differing lengths. So, to average the growth during the first month with the growth during the first 12 months is not a reasonable thing to do. However, what is reasonable to do is to average the growths over overlapping time periods of similar lengths, to smooth out fluctuations in the index. For example, let’s suppose the end of the year happened to be a really bad day on the stock market, and instead of rising to 1120 the index fell back to 1000, jumping back up to 1120 the next day.

Would you really want to say that there was no growth at all during the year, just because the day you picked to evaluate the index happened to be a bad one? I think not. So, what is commonly done is to average the index value over a certain time period. For example, while what you describe is unreasonable, it could be perfectly reasonable to say that the GIC’s rate of return from 1996 to 1997 will be the percentage growth of the index’s average 1996 value to its average 1997 value. That way, fluctuations due to the particular day the index is measured will be smoothed out.

Now, if the index were a constant 1000 during all of 1996, and rose to 1120 in the manner you describe during 1997, you would unfortunately only be getting a 6% return from average 1996 value (1000) to average 1997 value (1060). But that's because some of the 1996 behaviour is being factored in as well. Another thing that could be reasonable is to average the percentage growths over different 12-month periods ending in the same year. For example, you could take the January 1996-January 1997 percentage growth, the February 1996-February 1997 percentage growth, and so on, and average them. This is a reasonable sort of average to take, since the percentage growths are being averaged over equal but overlapping periods.

These sorts of averages have the advantage of smoothing out wild fluctuations in the index. For example, if the index started 1996 at 1000 and rose at 10 units per month steadily, ending 1997 at almost 1240 but plummeting down to 1150 on the last day of 1997 due to a temporary correction in the markets, you would still get the benefit of the growths. Even though the December 31, 1996 to December 31, 1997 growth was from 1120 to 1150, just 2. 7% on the entire year, the November 30, 1996 to November 30, 1997 growth was from 1110 to 1230, about 11%.

So, by including these in the average, rather than just looking at the average percentage growth over a single year-long period, you get a much less volatile measure of performance. My suspicion is that the GIC you are concerned about probably employs some sort of legitimate averaging such as the above, and that the bank has miscommunicated the calculation method to you. Another possibility is that the bank may be measuring index value from start of investment to the average index value during the year preceding the end of investment. That means that during the first year of the investment you have the unreasonable calculation you described, but over the long term those effects are negligble. For instance, the 10-year return might be the average of the index return over the twelve periods all beginning on the date of invesment and ending on the twelve month-end dates during that 10th year. True, this is still averaging growth over periods of differing length, but the lengths range from 9 years to 10 years and this is much less of a difference than your example where they range from 1 month to 12 months! I will not venture to speculate further on what methods the bank may or may not be using for the GIC you have in mind.

However, I hope this clears up for you some of the mathematical ideas involved.