Since I had some students ask about it, allow me to show you how all this Calculus relates to some cool things in Number Theory. I do have to stress that NONE of this is required reading for Math 408D!

In everything that follows, you should remember that the infinite "harmonic" series $\sum(1/n)$ diverges; more precisely the partial sum $\sum_{n=1}^{n=T}(1/n)$ is approximately equal to $\ln(T)$.

Now, I invite you to expand

$$\left(1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}\right)\left(1+\frac{1}{3}+\frac{1}{9}\right)\left(1+\frac{1}{5}\right)$$

using the distributive law (that is, to "FOIL" it). What you will get will be 24 of the terms in the harmonic series. Specifically you will get terms of the form $1/n$ for which n has a prime factorization of the form $n = 2^i 3^j 5^k$ with $0 \le i \le 3$, $0 \le j \le 2$, and $0 \le k \le 1$. If you include more, similar, terms inside these three big parentheses, you simply get more summands $1/n$ of the same form but with larger i, j, k. If you multiply by similar expressions involving powers of $1/7$ or $1/11$ etc. then the expansion will include more of the summands in the harmonic series — those whose reciprocals involve these new primes.

In a formal way, then, we deduce that the whole of the harmonic series can be represented as a product of multiple infinite series:

$$\sum_{n\ge1}\frac{1}{n} = \left(1+\frac{1}{2}+\frac{1}{4}+\dots\right)\left(1+\frac{1}{3}+\frac{1}{9}+\dots\right)\left(1+\frac{1}{5}+\frac{1}{25}+\dots\right)\dots$$

including one long factor for each prime p, namely

$$\left(1+\frac{1}{p}+\frac{1}{p^2}+\frac{1}{p^3}+\dots\right)$$

You should recognize that this is a geometric series: its first term is 1 and the ratio between any two consecutive terms is always $1/p$. Since that ratio is less than 1, the geometric series converges, and you know the sum is $1/(1-(1/p))$. In symbols, that gives us a powerful connection to prime numbers now:

$$\sum_{n\ge1}\frac{1}{n} = \prod_{\text{all primes } p}\left(1-\frac{1}{p}\right)^{-1}$$

(The Π is a capital Greek letter Pi and it indicates a product in the same way that a Σ indicates a sum.)

What does this equation tell us? For starters, note that that geometric series for each prime separately has a finite sum — for example it's 2 when $p = 2$, and $3/2$ for $p = 3$ and so on. If, for some reason, you thought there were only finitely many primes in the universe, then the right side of this last equation would just be some number; but we already know the left side is a *divergent* infinite series, so this would be a contradiction. The divergence of the harmonic series thus proves that the list of primes goes on forever!

Mathematicians hate to let simple ideas go to waste so they try to re-use them. In this case, if we put a fixed exponent on every single fraction, we can reuse the prime-factorization idea above. The conclusion is that for every number s we have

$$\sum_{n\geq 1}\frac{1}{n^s} = \prod_{\text{all primes } p}\left(1-\frac{1}{p^s}\right)^{-1}$$

For any real number $s>1$ the series on the left will converge now, which tells us something about the factors on the right.

Indeed, we can turn this into a statement about infinite series by taking logarithms: using properties of logarithms we rewrite the last equation as

$$\log\left(\sum_{n\geq 1}\frac{1}{n^s}\right) = \sum_{\text{all primes } p} -\log\left(1-\frac{1}{p^s}\right)$$

Here I want to stress that for any $s>1$ the left side is a *number*, so we are saying that the infinite series on the right side converges! For example, when $s=2$, I happen to know not only that the series $\sum(1/n^2)$ converges but I know its sum — it's $\pi^2/6$. (That's not obvious!) So I now have a convergent series:

$$\log(\pi^2/6) = (-\log(3/4)) \ + \ (-\log(8/9)) \ + \ (-\log(24/25)) \ + \ (-\log(48/49)) \ + \ \ldots$$

The fact that this series is convergent tells us that the individual terms must converge to zero, and in fact must be converging to zero fairly quickly. (We always keep in mind the harmonic series, whose terms converge to zero but "not fast enough" to make the harmonic series converge.) Numerically this is evident: the series above is

$$0.4977003032 = 0.2876820722 + 0.1177830357 + 0.0408219948 + 0.0206192868 + \ldots$$

To make a little more headway, we can use the theory of Taylor Series, which we will get to at the end of Chapter 11. Among many other examples, it will tell us that for small numbers x we may compute logarithms using infinite series:

$$-\log(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots$$

Or, if we just want an estimate, we will show $-\log(1-x)$ will always lie between x and $x+x^2$ when $0<x<1/2$. Either way, a previous equation now takes on this form:

$$\log\left(\sum_{n\geq 1}\frac{1}{n^s}\right) = \sum_{\text{all primes } p}\frac{1}{p^s} + \text{something small}$$

(Here the "small thing" can be estimated as a function of s but stays bounded for all $s > 0$.) It's a little tricky handling the limits involved in partial sums as s approaches 1 but very roughly speaking you know that when $s = 1$ the partial sum of the first N terms on the left is about $\log(N)$ so you expect the same on the right: the sum of the reciprocals of the primes less than N comes out to around $\log(\log(N))$. For this to be true for all values of N sort of forces the nth prime to be approximately $n \log(n)$.

So you see, we have learned not only that the primes go on forever but they don't even get all that far apart. We can predict approximately where the primes are. This is known as "The Prime Number Theorem", which was mostly guessed-at in the 19th century and finally proved in 1896: if $a_n = n/\ln(n)$ and b_n =the number of primes less than n, then the sequence $c_n = a_n/b_n$ converges to 1. That is: b_n is pretty close to $n/\ln(n)$.

This theorem is about the *limit* of the sequence c_n, but the actual values of c_n get close to 1 even when n is small. There are 25 primes less than 100, so c_{100} works out to be 0.87. You can also interpret the theorem probabilistically: if you pick a number near n at random, you will hit a prime with probability $1/\ln(n)$. So for example about 4% of all Social Security numbers are prime.

Here's a mind-bender for you. First, allow me to repeat what I said above: the sum of the reciprocals of all the primes less than N comes out to around $\log(\log(N))$. In particular, the series $1/2 + 1/3 + 1/5 + 1/7 + 1/11 + \ldots$ diverges. The partial sums get larger and larger without any upper bound. Go far enough out, they will exceed 10, or 100, or 1000. And yet: if you add up the sum of the reciprocals of ALL the primes that have been written down so far, that partial sum comes out to less than 5.0 ! Even more amazing: based on the amount of time and space in the physical universe, the sum of the reciprocals of all the primes that anyone will EVER write down will never exceed 5.0 ! This is a DIVERGING sum, but we will never manage to "witness" the divergence. It's all because the function $\log(\log(x))$ grows sooooooo slowly.

For comparison, the "twin primes" (those that differ by 2, such as 17 and 19) obviously include only some of the primes and so the sum of all of *their* reciprocals will have to be smaller than the sum running over all primes. What's not obvious is that in fact the twin primes constitute a MUCH sparser set of numbers than the set of all primes — so much so that the sum of the reciprocals of all the twin primes is actually a convergent series. (This was proved just 100 years ago by Viggo Brun.) That series converges to less than 2.0 . If there were only finitely many twin primes, that would of course explain why the sum converges, and so you might conjecture that there is, in fact, a last twin prime. As of this moment, no one can disprove that statement! Most mathematicians expect that even though the twin primes get sparser and sparser as you venture into larger and larger numbers, the list of them never comes to an end; but until someone can actually prove this, it's just a conjecture, and the convergence of Brun's series should make you just a little skeptical of this conjecture. (By the way did you ever hear of the "Pentium Bug"? It was a computer hardware glitch that was discovered while people were trying to find the sum of this series!)

Incidentally, the value of the series

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s}$$

obviously depends on the value of s, that is, ζ is a function of s. With some trickery we can define $\zeta(s)$ in such a way that it makes sense for most complex numbers s, too. The *Riemann Hypothesis* is the claim that that this function is never equal to zero except at a couple of well-understood types of values of s. As hinted at above, this would have implications to the field of Number Theory. As of now, the Riemann Hypothesis is still just a conjecture — we have no proof yet that it's true. It is a "Millennium Problem" now: someone has offered a one-million-dollar prize for a proof of this conjecture.

Here's one other sequence factoid related to the harmonic series (but not to primes). We have commented that the partial sums of the harmonic series are roughly $\ln(N)$. Well, how close is that estimate? In other words, let $a_n = 1 + (1/2) + (1/3) + \ldots + (1/n)$, and let $b_n = \ln(n)$. What can be said of the sequence $c_n = a_n - b_n$? Well, $c_1 = 1$ and after that it's not hard to show that c_n is a decreasing sequence. It's a little harder to show that each c_n is positive, but once you do, you can use a theorem in the book to know that the sequence c_n converges. The limit of the sequence is called the Euler-Mascheroni constant γ. We know half a trillion digits of this number but we don't know for sure that it's irrational, nor do we have a handy formula for it in terms of π or e or square roots or logs or anything. But if you ever need a really good estimate for one of the partial sums of the harmonic series, you can use the statement that

$$\sum_{n=1}^{n=T} \left(\frac{1}{n} \right) \approx \log(T) + \gamma$$

in the sense that the difference between the left and right sides is (a) positive and (b) increasingly close to 0 as T gets larger.

In a similar way the Meissel-Mertens constant is the limit of the sequence $C_n = A_n - B_n$ where A_n is the sum of the reciprocals of the *primes* less than n, and $B_n = \log(\log(n))$; this sequence also converges (to approximately 0.26). Google had a little fun with the digits of these two constants a few years ago. I invite you to read all about this (and all the other topics in this essay) in Wikipedia.

So, see? What you're learning about sequences and series is part of what mathematicians do every day!