**What are business metrics?**

What is the highest and best use of a business analyst or a business data analyst?

Answer. To find out the right question to ask. Then, find the best answer you possibly can in the time, and with the resources available. Ensuring that your answer has practical impact by translating it in to a specific call to action. A specific recommended action, communicated to the relevant decision makers. Using visual metaphors and non-technical language.

What is always the right question?

Answer. What change, in our business processes can and should we make right now? To increase revenues, maximize profitability, or reduce our risk.

What do we mean by, right now?

Well, there's several answers. The ideal answer is something like, our company has engineered real time, computerized systems. That optimize for certain dynamic metrics. By adjusting our response to customer and market input. With a delay for processing or latency, of no more than a fraction of a second.

The next best answer is what we call Just in Time. A computerized response is not possible. So we respond to individual customers with a human being. Where that human being has access to the full historical record, of that customers interactions with the company. Whether through buying things, customer support calls, actions on our website, etc.

The third best answer, which is still good, is that the change that we are recommending, can't happen overnight. But it can happen as soon as it has been empirically tested and shown to work. We do testing by making one small change at a time to our website. Creating two different versions, A and B, and comparing which performs better. This is called AB testing. In any case, an AB test can begin within 30 days of a recommendation. And the recommended change can be made as soon as enough data is gathered. To determine that one choice is clearly better than the other.

So metrics are special numbers that help us to ask and answer the right business question.

What decision should we make? What process should we change right now? Metrics are numbers that we can impact when we change our business processes.

So, for example, for a local bricks and mortar retail clothing store rooted in the state of North Carolina. Where we have a 4.75% sales tax, the sales tax is not a metric. There's nothing that we can do to change it. It's basically a fact of nature for us like the weather. On the other hand the percentage of people who viewed our ad on a particular website. Then used it as a link to c lick through to our homepage. Can change every time the ad contents and appearance change. Or when the time of day changes, where we run the ad, on what websites it appears etc., etc.

**Distinguishing Revenue, Profitability, and Risk Metrics**

All business metrics can be classified into three broad categories. Revenue metrics, profitability metrics, and risk metrics.

One way to help distinguish for a particular metric, which of these three categories it falls into, is to think about what people in the company depend on this information and will ask for it. Revenue metrics relate to sales and marketing. Profitability metrics to efficiency and logistics, production, and operations. And risk metrics to risk management, and are widely used by a company's creditors, and outside investors.

So, revenue metrics are outward facing. They tell us something about how well or badly the company is marketing and selling its products. The company sales force, typically led by a vice president of sales, will want to know how many units of each product were sold over a given time interval and how this compares to the same time interval last year and the year before. They will want to look at sales by region, by product and by new versus repeat customers. And they will want to know about the sales funnel, the potential future customers who have been identified, and where they are in the step by step process of moving towards making a purchase. Much more on the sales funnel later. Meanwhile, the marketing team, typically led by a VP of marketing, will want to know how effective any marketing campaigns may be. How many people have seen a particular advertisement or email marketing piece or mail offer? What percentage have responded, etc.? Everything that relates directly or indirectly to selling is a revenue metric.

Profitability metrics have to do with the efficiency of the processes by which the company creates and delivers its products and services to customers. These are operational metrics, sought after by those people in the company responsible for production. Typically led in a large company by the chief operating officer. Anything that relates to how much cash is tied up in the form of unsold inventory, how much production is unsaleable due to spoilage or wastage, we didn't sell those mangoes in time, and now they're rotten, like that. Or at the other extreme, how often the company is unable to meet urgent customer requests and loses sales because of insufficient production or inventory. What portion of products off a production line are rejected as defective, how much is spent on variable costs, raw materials and labor, per unit product, etc., etc., etc. These are all efficiency metrics. Note that even a company with large and rapidly increasing revenues will fail to be profitable if it cannot deliver its offerings efficiently. Large established companies with relatively little room to increase revenues can often achieve significant increases in profitability by focusing on improving operational efficiencies.

Finally, risk metrics have to do with tracking and where possible reducing the many potential dangers a company faces. For example, if a company is spending a large portion of its net cash flow every month on interest on its debts, then even a small drop in revenues caused by some external shock, like a recession, could cause the company to become insolvent and collapse. Secured creditors have the right to seize a company's assets if they're not paid on time and that would close down the business. Net cash out is always the most important metric to track. How many months can the company survive at the present burn rate? Another example of risk metric is churn, a company with a subscription based revenue model that has a very high churn rate, the rate at which new subscribers drop off within a year, runs the risk that over time, there are fewer targets who have never been customers and it becomes impossible to maintain revenue growth or even hold steady. The greater the reliance on long term recurring revenue customers, the less dependent a company is on constantly successfully converting new prospects into clients and that's a lot less risky. Other examples of risk metrics are specific to the financial industry. Banks that issue credit cards are in the business of tracking how much exposure they have to potential customer defaults at any time and what percentage of their customers are expected to default in the next six months or are in default now. Money managers use volatility of returns and something called the maximum historical drawn down from high water mark, which we'll explain later, as a proxy for their portfolio's risk exposure. I've noticed in general that most risk metrics are related in one way or another to leverage. Anyone whose survival depends on their ability to pay back a large amount of borrowed money faces a magnified risk from any misfortune. Maybe this will help you to remember the categories.

**Distinguishing Traditional and Dynamic Metrics**

Traditional business metrics include standard financial and managerial accounting categories, such as quarterly statements of net cash flow, profits and losses, and changes to balance sheet items such as shareholder's equity.

Traditional business metrics have their origins primarily in paper and pencil after the fact reporting. Some of these metrics were once innovative too, like 500 years ago, when bankers in Florence, Italy invented double entry book keeping. Business decisions, that might be made based on these metrics, will happen usually after long deliberation, sometimes on a scale of months to years. They remain extremely important and worth of our study, but we just don't have time for that.

We want Dynamic Business Metrics that are defined and can be communicated in a manner that conveys urgency. Metrics that address the right question, what change in our business processes can we make right now to increase revenues, maximize profitability, or reduce risk?

Two attributes make a business metric dynamic. First, will the metric change significantly over intervals of a month or less? If not, it's not very dynamic. For example, the monthly rent a standalone retail store in a mall pays on its three-year real estate lease is of course a business metric related to its efficiency and profitability, but it is not a dynamic business metric. There's no point in tracking it because it won't change anytime soon. On the other hand, if a national retail chain with 1,000 mall-based stores is individually negotiating and signing an average of about seven new three-year leases each week, it can and should track average monthly rent per square foot on new real estate leases as an important dynamic metric, against which to set goals and track progress.

Second, are their specific actions the company can take that can visibly or significantly impact the metric in the short term? If not, then the metric doesn’t lend itself to dynamic tracking. Whether the metric is dynamic, may also depend on the business context. For example, if the 1,000 store retail chain we just mentioned is neither adding nor closing stores and is simply renewing leases on current space, where the old leases had pre-negotiated terms to be extended, It will be difficult to make much impact on the average monthly rent per square foot. On the other hand, if that retail chain publicly announces that it is going to close 25% of its US retail stores, as the GAP chain did in June of 2015, it may be able to go back to landlords with the proposition, either we leave at the end of the current lease and you'll need to find a new tenant, if you can, who may not pay as much per square foot as we do. Or let's renegotiate our lease terms now. In this case, the average monthly rent per square foot for newly renegotiated leases would be a dynamic metric for the GAP. It's a place they could save money. This is why announcing all your bad news at once is often a good business strategy.

How much impact a business change can have on the metric, is another important thing that we need to observe. If the metric is noisy, then lots of things are affecting it. If it’s twitchy, that means it's very specifically offended by what we're doing. Traditional metrics, like quarterly revenues, are impacted by dozens of different factors, many of which are completely outside the control of our business. The great new advertising campaign that we launched may not even show up in quarterly revenue metrics if, in fact, most of our customers are government agencies on long-term contracts with a very long sales cycle.

Total revenues is always an aggregate number, and as they say on the Kiss metrics website, aggregate data is kind of worthless. Dynamic metrics are twitchy. Small changes in process, in our process, can lead to big impact. For example the percentage of people who will fill an online shopping cart and take their shopping cart all the way to purchase is extremely sensitive to average page load times. Studies have shown that pages that load in three or more seconds are much less likely to lead to sales than pages that load in less than three seconds. 40% of web users will abandon completely, a web page that does not load in three seconds. This is why there's an entire industry devoted to what's called edge caching. Content delivery networks, like Akamai that retail stores pay to store copies of their websites locally, physically near their customers around the world, saving a precious second or two in load times.

By the way, through a recent project my students at Duke did for a website performance monitoring company, we learned that on mobile devices many retails companies' web pages are taking 20 seconds or more to load. It appears that many global retail companies do not yet test their mobile based loading times from remote locations around the world. If your company does not already do load time performance testing for mobile devices in all global markets where it is active, it should start immediately.

**Egger’s Roast Coffee Case Study Part 1 – Definitions**

Next, we will be going through a case study to illustrate the critical difference between profitability metrics and cash flow metrics. Part one of this case study will talk about some financial definitions. Business analytics is not financial accounting, but it does draw upon some accounting concepts for its metrics. We will keep this brief. So in about the next ten minutes, I'm going to orient you to all the financial accounting terms and concepts you will need to grasp essential business metrics. I will explain these very few most important terms through the example of a start-up venture that I'm organizing, Egger’s Roast Coffee.

Egger's Roast Coffee plans to buy raw coffee beans in bulk, roast and package them and sell them wholesale at higher prices, obviously, to a grocery store chain. In the next video, we will explore what happens with Egger’s Roast Coffee when you have different types of growth scenarios and we will show how a rapid growth scenario can actually be very dangerous in terms of cash flow even if it's fine in terms of profits. Cash flow and profits are concepts that are so distinct that corporate financial reports discuss them separately. Profits are presented in a profit and loss or a P and L statement and cash flow is in a statement of cash flows.

Unprofitable companies can survive and thrive for years or decades. Take Amazon, for example, but unprofitable companies that run out of cash disappear without trace, like eToys or pets.com or Webvan. Even companies that are massively profitable can collapse if they can't meet their short term cash obligations quickly enough, as happened to the world's largest and most profitable insurance company AIG over a few days in the fall of 2008. Profitable companies actually run out of cash and go out of business quite often. One of the most common triggers for this kind of disaster is uncontrolled or unplanned sales growth. Too much success, too quickly can kill you. I am illustrating this danger through two different revenue scenarios for Egger's Roast Coffee.

Scenario one, which involves safe, steady revenues. And scenario two, rapidly increasing, excitingly great. But ultimately, fatal sales growth. If I, as your teacher did not make sure before we do anything else that your company will never run out of cash due to your inability to track the right business metrics, I would not be doing my job.

All of the Egger's Roast Coffee business model assumptions, the financial terms and concepts discussed in this video and the next and the two scenarios in the case study are spelled out in a detailed Excel spreadsheet that you can download. I recommend reviewing carefully after you watch the video and before you take the quiz for this lesson the materials in the spreadsheet.

So as I said, I’m starting a coffee roasting business. My business model is simple and I think it is guaranteed to work. I have a friend who works as a buyer for a giant retail grocery chain and he has promised to buy from me for $6 per pound, that's the wholesale price as much Egger’s Roast coffee as I can produce. I will sign a supply contract with the grocery store chain, agreeing to deliver as much coffee as they need. With financial penalties, if I failed to deliver the full amount on time, but at a price that guarantees me a profit for every pound.

My friend thinks that his giant grocery store chain can sell an unlimited amount. We're going to market this coffee as artisanal and authentic, roasted with love by me personally in Durham, North Carolina. The grocery chain pays all of its suppliers, including me on average 60 days after it receives the merchandise. This gives it time to sell the merchandise itself. These are common terms known as Net 60.

I expect that the average time it will take me to roast, package and deliver beans from the time I receive the raw green beans is 30 days. I have located a source of high quality raw green coffee beans in bulk for $2 per pound. Their terms are cash on delivery on COD. I pay for the beans when they arrive at the door.

Now variable costs are things that cost me more the more beans I roast. The obvious example that I've just mentioned are the raw green beans themselves. I buy each pound for $2 and if I buy one more pound, I pay another $2. That is a variable costs. Other variable costs of production for me are fuel used in the roasting process. The paper packaging the beans are wrapped in. Hourly wages for workers who help load the raw beans into the roaster and roasted package beans onto the delivery trucks. Transportation cost to deliver the beans to my customer stores. I expect all of these additional variable costs to be about $2 per pound, so I expect the total variable cost to be about $4 per pound.

Next, I need to consider my capital investment. I will need to make one major capital investment. I need to buy a machine that can roast and package a large number of beans quickly and consistently. I've looked into it and a roaster than can handle the volume of beans I'm expecting costs $540,000. It should last three years before it will need to be replaced. I plan to pay cash for it, I’m planning on putting $800,000 total cash into this business as we get organized and that will leave me with a nice cash cushion of $260,000 in the bank when we begin operations. I also can borrow up to an additional $800,000 from friends and family, if I need to in an emergency.

From the point of view of financial accounting in order to determine if I'm selling coffee at a profitable price or not, I need to find a way to include in the cost of roasting beans, the $540,000 that I already spent on the machine. I could say that the first pound of coffee cost me $540,004, but that would not be fair. The later pounds would get to use the roaster for free.

The traditional solution, which I use here is to take the $540,000 amounts and allocate it or parcel it out over time, so that some fraction of that money is assigned to each bag of coffee that is roasted by the machine. Based on the idea that the coffee roasting machine's useful life is three years, I’m going to allocate the $540,000 purchase price over three years or 36 months. One-thirty sixths or $15,000 each month. If I produced 25,000 pounds of beans each month, allocating that cost to each pound equally results in reducing my profits on each pound of beans sold by 15,000 divided by 25,000 or $0.60 per pound.

Allocation of capital expenditures over time, because stuff wears out is called depreciation. I know that machines don't really wear out exactly steadily, but we make this type of assumption in accounting all the time. You can picture a straight line that slopes down to the right, so the machine assumed to be worth exactly $15,000 less each month. This approach, which is again, the most common type of depreciation is called straight line depreciation. For simplicity here, we will consider depreciation to be a fixed cost. Fixed, because making one more pound of coffee or one less does not impact it. The machine is not considered to wear out faster or slower, if we run more or fewer beans through it.

Certain other of my costs are also the same each month, whether I roast one pound of coffee or 10,000 pounds. Rent for the building, utilities, insurance, required business licenses. Salaries for an office manager and bookkeeper, these are all the so called general and administrative or GA expenses. I also will assign a fraction of these GA expenses to each pound of coffee, just as I did with the capital expenditure depreciation. GA expenses of $10,000 per month allocated across 25,000 pounds of production should be 10,000 divided by 25,000 or $0.40 per pound that first year.

My cost per pound will drop, if my production increases. So in the model that I've presented at a selling price of $6 per pound and if one year volume of 300,000 pounds, it looks like my variable cost of $4 per pound plus fix costs of $1 per pound should result in a profit or net earnings of $1 per pound or $300,000 the first year. I will record or book this revenue when I receive an order. And so, I'm already showing a profit of $25,000 my very first month of operations. So what could possibly go wrong?

Well, if examined from the point of view of profitability, the business looks great. But if examined from the point of view of how cash moves into and out of the company's bank account, the picture is a little less reassuring. Cash flow analysis is all about timing. You may recall that we pay for new beans COD, cash on delivery and then we need 30 more days to process the beans before they can arrive at the customer site. The customer then pays us net 60. This means that on average when we receive an order, we spend cash on that same day, but we get paid cash by the customer 90 days later. This situation known as negative float is extremely common in all different kinds of businesses.

I will point out that certain businesses, like insurance and banking get paid cash now to deliver a service later. This is a delightful situation known as positive float. However, monies our customers owe us for products already delivered, but which we have not yet received payment on are called accounts receivable. So we deliver something and then we have an account receivable on that amount of money that we’re owed until it's paid. The money has been booked for purposes of calculating our profits and losses, but it does not show up on our cash flow statement, because we haven't actually gotten it yet. We track how long it has been that money we have not yet received is owed us through a metric known as aged account receivables or aged receivables for short.

That's all the financial accounting that you will need.

**Egger’s Roast Coffee Case Study Part 2 – How a Profitable, Growing Company can go Bankrupt**

Now we're going to look at how a profitable growing company can go bankrupt.

In scenario one, we have steady sales. Let's take a look at scenario one in the spreadsheet, a handout I provided. Edgar's Roast Coffee is selling 25,000 pounds of coffee every month at $6 per pound at an average total cost $5 per pound for a net profit of $1 per pound. So it all sounds good. The P&L statement matches sales that are booked in January. $150,000 for 25,000 pounds at $6 per pound. To the variable costs of January production which are $100,000 plus January's share of the fixed costs. $25,000 total. $10,000 GA expense. And $15,000 for one months' straight-line depreciation of capital equipment resulting in a new profit of 150 minus 125 or $25,000.

The cash flow statement represents the same events quite differently. Variable expenses of $100,000 and fixed expenses of $10,000 for buying and roasting 25,000 pounds of beans need to be paid for in cash in January. But the account receivable of $150,000 created in January is not paid in cash until April. In fact, no money at all comes into our business until April. But we still have the $110,000 cash expense triggered by February's order that must be paid in February. And another $110,000 cash expense triggered by March's order that must be paid in March. Altogether we need $330,000 more cash before the first payment comes in in April. Already my $260,000 reserve of cash is not enough. So I rush around and borrow money from friends and family and in April, cash flow turns positive and the situation starts to improve. My original reserve of $260,000was at least $70,000 too small. Assuming that I can get the extra money together in time to buy beans in March and stay on schedule, cash flow does turn positive in April with a net of $40,000 coming in. But I can't pay back that extra emergency $70,000 that I needed until May. The positive cash flow position develops slowly. By December, when I already have $300,000 in net profits on paper, I still have $510,000 less cash on hand than I had before I started the business. Although cash flow turned positive in April, it still takes me until January of year 3, 25 months after launch of my company to recover my original $800,000 cash investment.

Now we'll consider scenario two when I have rapid sales growth. It turns out Edgars Roast coffee is a huge regional and then nationwide success. The orders we receive are literally doubling every quarter. 25,000 pounds per month for January through March, then 50,000 pounds for April through May, 100,000 pounds for June through September, 200,000 pounds for October through December, and so on. So let's look at the scenario two. Profit and loss statement first. Under scenario 2, when our sales increase from 25,000 to 50,000 pounds in April, profits more than double. From $25,000, or $1 per pound on a 25,000 pound order to $75,000, or $1.50 per pound on the much larger 50,000 pound order. The reason why our marginal profits, profits per units of production, increase is that when the number of beans we process doubles, only the variable costs vary. They increase by $4 for every pound of additional production. But our fixed costs, the general and administrative expense and depreciation, are fixed. They stay the same when we produce more so the accounting allocation of $25,000 in fixed costs to total pounds of beans drops from $1 per pound at 25,000 pounds to $0.50 per pound at 50,000 pounds production. Rising marginal profits on higher sales and production volumes are typical in most businesses and are known as economies of scale. The lower a product's variable cost, relative to its fixed cost, the greater the economies of scale. This is why internet businesses, where variable costs are often little more than the cost of electricity, have such dramatic economies of scale. According to our projected profit and lost statement, this pattern continues by the time we are selling 200,000 pounds of beans while our variable costs remain $4 per pound, our fixed costs are now only 12 and a half cents per pound for a total cost of $4.12 and a half cents per pound and a total profit of $1.87 and a half cents per pound. Under this rapid growth scenario by the end of year one, we have sold 1,125,000 pounds of Edgars Roast Coffee and made a theoretical net profit of $1.95 million.

But now let's take a look at how scenario two plays out in our cash flow statement. Unfortunately our cash flow statement. Which reflects the true timing of cash moving in and out of our company paints a very different picture than our PNL. Under the scenario one cash flow, $110,000 in cash went out the door in April to process 25,000 pounds of beans, but we also received $150,000 cash in April. Finally getting payment for the 25,000 pounds of beans that we had processed and delivered back in January. This cash influx saved us from needing to borrow any more money, but now in scenario two. Because the customer has put in a new, higher order for April of 50,000 pounds, we need to spend $210,000 cash instead of $110,000 cash. So cash flow stays negative, and our emergency borrowing is now up to $130,000. May and June are going to be just as bad for cash flow. We're going to have $60,000 negative cash flow each month until July when we finally start receiving cash for the larger orders that started in April. Got to hang on until July. Unfortunately in July we receive an order not for 50,000 more, but for 100,000 more pounds of beans. Instead of the $90,000 positive cash flow we expected, $300,000 cash in as payment for the 50,000pounds of beans we produced in April. And $210,000 cash out to produce another 50,000 pounds of beans. We face a further $110,000 negative cash flow. $300,000 cash in as payment, as we mentioned before, but $410,000 cash out to produce another 100,000 pounds of beans. Now I start to panic. My emergency borrowing is up to $360,000 with no end in sight. Maybe I can hang on until October when we will receive the first of the larger $600,000 cash payments. By the end of September I owe my friends $580,000 on top of $800,000 of my own money. I'm nearing the limit of what they can do to help me. Then catastrophe strikes. At the beginning of October we receive a new larger order for 200,000 pounds of beans. Each month for three months. Now I get my Excel spreadsheet model out. And I begin plugging in the expected orders and expected cash flows, and it becomes clear that to survive until the end of December, we will need a total of $2,010,000 in cash. My $800,000 plus an additional $1.21M of borrowing. That is more than I can raise from my friends and family. Beans keep arriving that we can't pay for. We face financial penalties from the grocery chain if we fail to deliver on the full amount that they have requested under the terms of our supply contract with them.

So if we fail to deliver 200,000 pounds of beans in October, they can withhold the $600,000 payment for the 100,000 pounds of beans that we delivered back in June. And so on. The gaping hole appears to be getting larger and larger. If their orders keep doubling every quarter, it's now become clear to me based on my spreadsheet analysis, that when we will need close to $5 million more cash than we have by June of next year. Perhaps, if I had understood this negative float problem better before launching the business, I could have arranged what is called accounts receivable financing or factoring, a short term loan against accounts receivable. But these things take time to arrange and we're unable to pay our bills right now. The fuel company isn't paid and stops delivering fuel to power the roasters. Workers stop coming to work because we cannot pay them. The company is bankrupt. And can't even be reorganized. Having broken our contract with our buyer we owe them large cash penalties. In our now severed relationship with them was the company's only distribution channel. The lights go out on Edgars Roast Coffee. Yet, from a financial accounting point of view, every transaction we did was profitable. How do we somehow been able to come up with the mountains of cash needed. We would have had a profit of $1.95M our very first year. Now you know at least one way that profitable companies can go bankrupt.

**Revenue Metrics – Traditional Enterprise Sales Funnel**

We'll talk now about the Traditional Enterprise Sales Funnel. As you recall, revenue metrics are outward facing. They have to do with a customer’s outreach to customers. They measure how well or badly a company is identifying potential customers. Communicating to them about its offerings and value proposition to them and ultimately, selling to them. After watching this video you will be able to identify, and track, the most important business metrics that define success in traditional enterprise sales.

In this course, we use the term enterprise sales a bit loosely, to refer to any sale that requires involvement by full time sales people in your company. Enterprise sales historically refer to big sales of complex products or services, an elaborate piece of capital equipment, like a wind turbine or something like a bidding process to win the right to act as a general contractor on construction of an office building. Enterprise sales, at a minimum, require talking to potential customers on the phone, getting to know them, and more importantly, giving them a chance to get to know you. So they almost always involves ending your people or yourself to travel out to the customer site and hold in person meetings, sometimes many meetings before a sale is made.

Enterprise sales are expensive to do and therefore the payoff for each enterprise sale must be large. Companies that attempt enterprise sales for low-priced to medium-priced items typically lose money on average on every sale and either change their sales model or go out of business. Different companies in different industries have different numbers for what would qualify as a minimum size enterprise sale, but I'll give you a general rule based on the cost of operating in the United States. I would say you would want to have a minimum of $250,000 for a one-time sale, or $100,000 per year on a recurring sale to justify an enterprise sales effort.

Imagine your product is an electric delivery van that's non-polluting and has low operating costs but needs to be operated within a 75 mile radius of its charger station. One of your vans costs $50,000. It would make no sense to target a small business that might buy one van with your enterprise sales campaign. Even if they buy it, you would end up spending more money than the marginal profit. Instead, you need to target companies that replace a minimum of two to three delivery trucks every year as the old ones wear out and ideally far more.

The sales funnel metrics are as follows. You start with a lead. A lead is a person whose name and contact information you have, who you know works at a company that owns delivery trucks, so they might be buying more sometime. The first step is to qualify the lead. While definitions differ, I would say that to be a qualified lead your sales team needs to establish two things. First, that the target company your lead works for wants to buy at least two or three delivery vans from someone in the next year. And second, that they could afford to buy your vans if they wanted to. In other words, they have a plan to buy, and a budget to buy, vans in your price range. Hey, as an aside, I can't tell you how many people are happy to talk to you on the phone for hours about how their company wants very much to buy the thing you're making. Or, if you're still doing pre-launch market research, or considering making, and you only find out months or years later that they have no money in their budget to buy any such thing. Such people, also, often like to invite you to travel out and demonstrate your product in person at your own expense without informing you that no one with any buying authority will be at the demo. Apparently many people are quite bored sitting at their desks, and are happy to talk at length with and hold follow up meetings with anyone friendly, even if they will never buy anything from them.

Note that you have established that the lead is qualified by reference to the target company's plans and budget. Now, you need to identify the person, within the target company organization, who has the budget, and authority to decide on the purchase. They need to be able to sign a purchase order. They are the correct decision maker. The person you originally contacted may be able to tell you who the correct decision maker in their organization is, but they may not choose to identify them. Because doing so may mean you'll stop calling them. Really, not kidding. Not surprisingly, often your original contract won't tell you that they have no power to decide anything, and that their opinions don't even matter, and certainly don't matter to the actual decision maker. When the correct decision maker tells you that they want to engage with you about the possibility of buying your product, yes we might buy vans from you, that is an expression of interest. The next big step will be to meet directly with the correct decision maker. This may require, particularly in large complex organizations, many meetings with other people first. Various gatekeepers and lower level functionary’s and lots of requests for documentation. You can literally spend forever holding seemingly necessary and productive meetings in the intermediate levels of any large company. The work product of many people is meetings apparently, so the more meetings they hold the better in their view. Great sales people seem to have a knack for getting directly to the top guy or gal and short circuiting the potentially limitless meeting process. If you have such a salesperson cherish them. Keep them happy.

Once the decision maker is persuaded, you’ve overcome all their objections and addressed all their concerns and you’ve negotiated terms and a price and they tell you directly that their company will buy x amount of vans from you for y dollars, you have what I call a soft circle sale. Why do we call it a soft circle? The number of things that can still go wrong before you actually close the deal with a legal contract, never mind getting paid, is incredible. Here are a few examples that I have come across in my business. After getting a definite and unambiguous yes from the correct decision maker, here's some things that can happen.

The decision maker never takes another call from you, stops answering your email, and pretends you don't exist. Or, you eventually learn from someone else that the decision maker has quit or been fired, or has transferred to a different office, and you will need to start the sales process all over again, with a new person, who has never heard of you or your product. Or, the target company is suddenly acquired, or it is in acquisition talks, and all pending contracts are suspended indefinitely. Or, the target company goes bankrupt. Or, your own company decides to stop making the product or offering the service that you just sold, but only tells you about that decision after you get a yes from the customer.

So, to review, the key enterprise sales metrics are new leads. New qualified leads. Expressions of interest. Meetings with the correct decision maker. Getting the decision making to say yes, that's a soft circle sale. And then the number of actual official binding contract sales. And that's when you book the revenue for financial accounting purposes.

**Revenue Metrics - Amazon.com as a Leading Example of Use of Dynamic Metrics - Part 1**

Enterprise sales that are dependent on personal one-on-one meetings between sales people and potential customers, they represent one extreme on a sales continuum, a place where traditional revenue metrics predominate. At the other extreme, online retailers entice potential customers to make purchases without interacting with a human being at all. Yet successful online retailers manage the customer experience through their web or mobile based interfaces so that the sales process does not feel impersonal or robotic. And they manage it empirically to optimize revenue from each customer visit. We will study in detail the customer experience and related metrics for one company that does a brilliant job of defining and exploiting dynamic revenue metrics, and that is Amazon.com.

If your current company is in retail, there is certainly some aspect of the Amazon approach that you can apply as current best practice. Even if you are not in retail, the methods Amazon uses to study its visitors' click stream data, the pattern of clicks, cursor movements, movements from page to page, are very important data analytic methods that will apply to any business with a website. The beauty of Amazon's underlying computer system design is evident in how seamlessly what the user sees combines a large amount of pre-processed data from various databases and indexes, with real-time responses to the user's query, their clicking activity, they’re all-important clickstream.

Real-time customization means that a company can customize each individual's user experience while that user’s session is going on in real time, based on that user's historical and clickstream data, informed by detailed records of what that visitor and similar visitors did in the past. This is our highest expression of the overall goal of making a business process change right now. I'd like to show you an example and take you behind the scenes of one of my own Amazon book searches to explain what work Amazon had to do in the past, and what it is engineered in real time to supports this current user experience.

I'll start with a text search in books, typing in the three words information, theory, learning. In under three seconds, Amazon presents me with a webpage with 12 books visible. Let me walk you through how this web page with this particular display list of 12 books that Amazon chose to show me is generated. Amazon maintains a very large database, we’ll call at the Book ID database, where every book itself is assigned separate record number and location specified by a unique Book ID. Amazon also maintains an up-to-date text index, so that each word in a text search can be used to retrieve exactly the book records in the database whose title contains that word. Sometimes also text within book might be indexed.

For my search, Amazon identifies more than 1,200 books in its database that have at least one of my three search words in its title or in other indexed text. Now is when it gets interesting. Amazon could pick 12out of 1,200 at random, in which case, the books would almost certainly not be books that I want. I would probably click down, look at another page of 12 or so, maybe two pages, and then give up without buying anything, after some browsing through largely irrelevant junk. Welcome to internet search circa 1995. Instead, Amazon begins to apply its own dynamic revenue metrics. Amazon ranks the 1200 possibly relevant items by their predicted relevance to me. Or put another way, they’re ranked by the probability that someone who typed in my search terms will buy the item on this visit. They show me just the top 12 of 1200that Amazon's prior data analysis predicts I am most likely to buy, how cool is that?

I infer, although I do not have direct inside knowledge of Amazon's system, that Amazon is ranking the books they choose to show me using a two-step process. First they study my text string of three words, information, theory, and learning, and match it to a predefined list of categories, high level subject area. This list of categories is what old fashioned library card catalogs called a subject index. Because a subject index contains a much smaller or more controlled number of words than all the words that people could run queries on, this type of index is called controlled vocabulary index.

Amazon does matching between my text string, the text I typed in on the one hand, and their controlled vocabulary index on the other hand. They do this by using a giant thesaurus that finds the best synonym or synonyms for the words that I’ve typed in in their subject area index. This is a very effective method to find for people what they actually want but didn’t know how to ask for, instead of giving them what they thought they wanted but will be disappointed by. Maybe they should invent a system that does the same thing for dating. Oh yeah that exists also.

My clue that this synonym matching is what Amazon is doing behind the scenes comes from the names of the subject areas that Amazon has decided to show me as most relevant. They appear in the upper left-hand corner of the first page, with the number of books in the Amazon catalog that are cross-listed in each subject area. Note that although all ten subject areas are in fact relevant to me, only three of the ten subject areas contain even one of the words I typed into my original query. Very cool. The high level subject index also expands to include subtopics. If I click on AI and Machine Learning, for example, I get AI and Machine learning, computer vision, pattern recognition, intelligence and semantics, neural networks, machine theory, and so on. We have a tree of categories and sub-categories. Amazon defines best sellers by how well they are selling relative only to those other books that fall into the same subject area subcategories.

I told you that I suspect Amazon has a two-step process for choosing what 12 books to display to me first. Their first step is using the thesaurus of synonyms to retrieve the most relevant categories in Amazon’s own subject index. The second step involves identifying the best-selling books within the subject subcategories my search terms fit best. It is these best-sellers, weighted by their topical relevance that Amazon displays to me. The reason why the intermediate step of identifying subject categories is necessary for Amazon, is that if they just ranked just all 1200 books by their rate of sales without regard to the subcategory, that would bury any specialized books that I and maybe only a few thousand other people are interested in, beneath the weight of the most general interests best sellers. So, we can infer that the dynamic topline metrics Amazon is using are, first what are the subject area categories in the controlled vocabulary index most relevant to this user’s exact typed query terms. Within the subject subcategories that most closely fit his query, what books are we selling the most of right now?

**Revenue Metrics - Amazon.com as a Leading Example of Use of Dynamic Metrics - Part 2**

What have we discovered about the Amazon database behind this individualized magic? We already know that for every unique book ID, Amazon must store a tree of subject area categories and subtopics, subcategories, that the book is located in, and a record of how each book ranks in terms of Amazon's own recent sales within each of the subject area categories and subcategories in which it is cross-listed. We are about to see what other dynamic revenue metrics Amazon tracks for each book, and how it stores and uses that information.

The top book on the list of 12, remember out of 1,200 possible that Amazon has retrieved and is showing me, is Information theory, Inference, and Learning Algorithms by David MacKay. Which happens to be one of my favorite books which I first read when my brother gave it to me many years ago, and which we now use as a text book in my Duke Master of Engineering Management Data Mining course. I can click on the book to get more information about it. Look carefully at the new page layout. First, we get Frequently Bought Together with a suggestion that we buy the Cover and Thomas Elements of Information Theory. Note the subtle selling suggestion, that learning information theory in relation to machine learning requires buying more than one book. Because people frequently, quote frequently unquote, bought the Cover and Thomas Elements of Information Theory textbook, which is also excellent together with the MacKay book.

To my surprise, Amazon is not offering discount for purchasing both together. Obviously because they have found that doing so does not increase sales enough to justify it. We don't actually know what percentage of buyers of MacKay also bought Cover and Thomas or vice versa. But Amazon is clearly tracking every purchase of more than one book at a time, and maintaining an up-to-date database for every book it sells of the other books sold during the same user session to the same customer ID. Amazon offers viewers the top 100 books in terms of frequency that people bought at the same time that they bought the MacKay textbook. This kind of data is called co-occurrence data. It can be represented in a co-occurrence matrix, much more on this later.

Co-occurring sales is obviously a very important dynamic metric to Amazon. A question now arises to which I do not know the answer. If you do know, please send me a note. Does Amazon list the top 100 co-occurring sales books in simple order from the highest co-occurrence to the lowest? If yes, then Cover and Thomas is not actually the book bought most frequently with MacKay. That would be Pattern Recognition and Machine Learning by Christopher Bishop. Of course, Amazon did not say that it was bought together with MacKay most frequently, only frequently. It is likely that the frequently bought together book offered is chosen through A/B testing experiments by Amazon that suggests that showing Cover together with MacKay increased sales or profits by more than showing Bishop or maybe any of the other 98co-occurring books on the list. In this case, Amazon has created best upsell revenue potential in the frequently bought together marketing slot.

On the other hand, if they are presenting the 100 books out of ranked order, maybe Cover and Thomas really is the book most frequently bought with MacKay. But, Amazon knows we've already been given information about it and we didn't bite on their offer to buy both together. So Amazon makes a small change in the true ordering in the list of 100 to put the second-most popular book to buy together with MacKay first. Because it is a better use of the top slot to display a new title to us. It is a mystery to me. But in either case, the combined placement in the Frequently Bought Together display and the top 100 frequently bought together display, is certainly optimized through A/B testing to maximize sales.

In addition to tracking what books people buy together at the same time, Amazon maintains a co-occurrence database of every book that a web user clicked on while browsing. Wow, that is a lot of detail. Clickstream data about everything I ever looked at but did not buy. How do I know that Amazon does this? Because, at the bottom of the same page is a separate co-occurrence list of the books people most commonly bought during the same session when they viewed the MacKay textbook without buying it.

So we know that Amazon is maintaining co-occurrence matrix that lists for every visit that results in the sale of at least one book to other possibilities. First, when book one is bought, it has a list of all other books bought. And second, when book one is bought, it has a list of all other books viewed. By using this second set of records, Amazon can identify what books were ultimately purchased by people who viewed but did not buy any given book.

So, to summarize, Amazon's first line of promotion is the profit maximizing frequently bought together metric. Their second metric is a kind of recommendation engine. What's likely to be also be bought based on the assumption that we do ultimately purchase the first book. The third metric is a recommendation based on the purchases people made who looked at but did not buy the first book.

**Profitability/Efficiency Metrics: Inventory Management**

We've already identified some of the most important revenue business metrics, focusing on enterprise sales. Now we'll look at examples of profitability or efficiency metrics, focusing for now on metrics that are widely used to evaluate the efficiency of a company's inventory management. And we'll let you in on the secret of how to pay the lowest possible price for a hotel room.

Why the focus on inventory management? Inventory management is one of the primary ways that operating companies can reduce their costs and maximize profitability without lowering the quality of what they sell in any way. Let's define inventory. Inventory is typically a finished product waiting to be bought, typically sitting on a shelf in a retail store or somewhere between the factory and the customer in a logistics warehouse. The more time a product just lies around between being made and being sold, the less efficient for the manufacturer. Time spent in inventory erodes profits.

Here are the four main reasons that inventory time should be minimized. First, negative float. You should be familiar with this problem from the Edgars Roast Coffee case study. A manufacturer supplying a retailer typically will not get paid until the item is sold at retail. But the money it spent to make that product is already gone. This negative float requires working capital which also entails an interest expense. The longer the negative float in terms of time, the more interest you pay. If a company's credit is not good, and the negative float increases unexpectedly, it could run out of cash waiting for its inventory to sell. For example, retail clothing stores are especially vulnerable to being forced into bankruptcy due to cash flow problems caused by unsold inventory in their stores. In the US, three large retail chains, Wet Seal, Cache, and Simply Fashion, all declared bankruptcy just in the first six months of 2015.

The second reason the fixed cost of storage. It costs money every day to heat, light, air condition, pay rent on, guard and other ways maintain a warehouse or retail store. The longer a product sits on the shelf, the more the costs of storage that need to be allocated to it. For example, assume that a furniture warehouse and showroom has fixed costs of $100,000 per year, and it has room to display 200 fully setup luxury dining room tables, and these tables sell for an average price of $1,200. If the average inventory turnover were one time per year, that would mean that we would be selling 200 tables per year and we would need to allocate the $100,000 across those 200 units, or $500 in storage costs to every table. If the average inventory turnout were five times per year, then we would only need to allocate $100 per table, which seems more reasonable. The third is wastage. On average, inventory items lose value the longer that they are in inventory. Some inventory items are subject to complete wastage. Their value goes to zero if they are not sold immediately.

Consider fresh-baked bread or restaurant food that must be sold on the day it is made. Those are extreme examples where every unit of extra inventory not sold is a total loss. Other examples of inventory wastage include any product with an expiration date. In the US and other countries many grocery items have a regulated sell-by date after which they cannot legally be sold. A special case of complete wastage that we will consider separately in a few minutes are hotel rooms and airline seats. If they are not sold by their “expiration date”, their value as inventory goes to zero.

The fourth reason is obsolescence. Items that stay in inventory for six months or more may in reality be items that no one want to buy anymore ever. Maybe production of that item exceeded worldwide demand, or the product has become obsolete and been replaced by something better, or fashions have changed. The longer the item is in inventory, the greater the risk it will ultimately need to be disposed of without a sale for a complete loss. For these reasons, average number of days inventory is held, abbreviated as days inventory, is a very useful business metric to track.

Now not many companies are willing to make their days inventory data public. However, I will share with you a handy trick for estimating days inventory for publicly traded companies, based on numbers that they are required to include in their annual reports. Public company annual reports will include the value of inventory on hand at the end of the year as well as the total annual cost of goods sold. Inventory on hand is valued using the same accounting method that is used to value cost of goods sold, so this is an apples-to-apples comparison. The company's inventory on hand at the end of the year divided by the total annual cost of goods sold and then multiplied by 365 for days of the year is a very good estimate of average days inventory. So we'll work an example for Walmart, courtesy of investopedia.com, which is a great website with excellent glossary of business metrics. For the fiscal year ended January 2014, the Walmart annual report gave yearend inventory of 44.9 billion and annual cost of goods sold 358.1 billion. Walmart's estimated days inventory turnover for the year is simply 44.9 divided by 358.1, or one-eighth, times 365, or approximately 46 days. Walmart sells all its inventory on average within a 46-day period, which is quite impressive when you consider the number of slow-selling items every Walmart needs to keep on hand. A Walmart store typically has over 100,000 distinct product SKUs on hand. So the days inventory metric is great, but it does not tell us anything about which of our thousands or tens of thousands of items are not selling as quickly as expected and why. It could be that particular products are seasonal, are no longer in fashion, or getting returned by customers as defective. These are all dynamic metrics that we want to know immediately.

All companies that have embraced big data culture, and this includes the major US retail grocery chains such as Walmart, 7-Eleven, Costco, etc., all track inventory turnaround and days inventory at the individual product SKU level. They further subdivide that analysis by geographic region of the country and even by individual store. This is done by tracking the date and quantity when any item is reordered by local management, and often also when the item is delivered and stocked on the shelf. And that's compared against cash register sales data that identifies the exact time and SKU for all sales. So for example, if a store in North Carolina sells 50 snow shovels per year, and all those sales are between December and March, it does not make sense for them to hold snow shovel inventory between April and November. In a best practices system, each products has its own expected days inventory, adjusted for region and seasonal factors using a mathematical model that helps determine how much shelf space should be allocated to it. Any time the actual days inventory can be found to be different than what is expected, that is a warning sign that needs to be looked at. Another issue is when you have too few of an item on the shelf. We can't measure directly how many times potential customer went away unhappy and without buying anything because what they wanted was out of stock. But a good proxy for the empty-handed lost customer metric is a number of times inventory of any particular SKU reached zero. Having at least one of everything that you might be expected to sell on hand is highly desirable in terms of brand reputation. As Bloomberg reporter Megan McArdle stated in a 2013 article, Walmart has tried to cut the number of SKUs it carries, but ended up having to put them back because it cost them in complaints and sales.

The higher the volume of unit sales of a product, the larger the potential lost revenue every time we see a zero level inventory. It would be an excellent practice any time a potential customer is observed to walk out of a big box retailer empty-handed to send a friendly marketing researcher after them with a $5 gift certificate in their hand and the question, was there a particular product that you were looking for that we didn't have on the shelf? In fact, I noticed yesterday that at my local grocery store, which is part of a very large national chain, the staff are now trained to ask at checkout, was there anything you couldn't find today? I'm tempted to answer, yes, emu eggs, every single time, in hopes that some data analyst will decide that they should stock emu eggs. Try it yourself.

**Profitability/Efficiency Metrics: Hotel Room Occupancy Optimization**

Not all wasting assets in inventory are products in the traditional sense of an object sitting on a shelf. Two important examples are airline tickets and hotel rooms. Here is why these can be thought of as a completely wasting asset. An airline has a regularly scheduled flight, let's say from Raleigh Durham Airport to San Francisco. The minimal variable cost to the airline if it fills that seat is just the extra fuel to lift perhaps an additional 100 kilograms of a person and their baggage and maybe a bottle of water and some peanuts. Almost all the cost associated with that seat are fixed, or sunk costs. The airline can't spend less on its airplane lease or on maintenance, or pay the pilot and crew less just because there was an empty seat on that flight. The money spent to create the seat is completely wasted if the seat is not sold in time.

Similarly, on any given night a hotel room that goes unrented represents a lost opportunity for the hotel owner. The variable cost of renting a room is very low the supply, the cost of supplying clean linens, and maybe a tiny bottle of shampoo. Almost all the cost of making that hotel room available is unrecoverable fixed cost, it is sunk cost.

Visiting Coursera for a workshop this year I stayed in The Grand Hotel in Sunnyvale from a Saturday through a Friday night. I was quoted a different room price for the same exact room for each night of the week. This is variable pricing. How does a hotel come up with its variable pricing? Business analytics of course. We will look at a simplified model of applying business analytics to the Occupancy rate metric.

Assume that a hotel has been charging single fixed rate of $150 for all its basic rooms, every night, for the past two years. Looking at the data from the last two years, the hotel has a slightly above average occupancy rate of 66.4%. I'm making these numbers up. The average hotel occupancy rate in the US was 62.3% and 64.4% in 2013 and 14. So far, nothing useful. However, if we recalculate average occupancy rates sorted by day of the week, we see the following pattern. There are clearly many more people traveling to the hotel during the week and Wednesday is a peak day, with far fewer travelers on Saturday and Sunday. Business travelers want to be home for weekends and holidays. This tells us this is mostly a business hotel.

We can infer from these weekly data, that we might be able to increase overall occupancy rates, if we charged a somewhat lower rate for Mondays, Tuesdays, and Thursdays, and a much lower rate for rooms on Fridays, Saturdays, and Sundays. Wednesday is a special case. If we dig a little deeper into the occupancy data for Wednesday, it is clear that there are seasonal patterns to occupancy as well as weekly patterns. So the Wednesday before the Thanksgiving holiday, let's say it's 35% occupancy. Wednesday during the week after Christmas, only 40%. The Wednesdays in August, 75%. Our average Wednesday we'll say is 95%. And our Wednesdays in September, October March and April were 100%, meaning the hotel was completely sold out on Wednesdays in those months. So we are turning away an unknown number of guests. We have more demand than supply in September, October, March, and April on Wednesdays. It's not clear how much demand exceeds supply, but it does appear that we are charging less than we could for those nights.

So our strategy for Wednesday would be to charge maybe three different prices depending on time of year. Less than $150 for holidays and August, the same $150 for most Wednesdays and more than $150 for Wednesdays in the four peak business travel months. By how much should we raise or lower prices? Here, we get into the arena of complex models that weigh price against demand. The best basis for these models is empirical. Try different prices and see whether the rooms sell too slowly or too quickly.

One way to do this is to list lots of room on online travel agency sites, or OTA sites. OTAs are not ideal for hotels because they charge commissions of 10 to 25% of the room rate, so hotels would rather sell direct to consumers. However, a typical strategy would be to assign a block of rooms on a particular night to an OTA at a pre agreed price for that block. If the OTA fails to sell all of them the unsold rooms are released back to the hotel to try to sell directly. OTA's provide a mechanism for hotels to test different prices on different dates to see whether or not that block of rooms sells out at that price.

A challenge hotels face when experimenting with their published or rack price, tat's the undiscounted room rate, is that it becomes a new maximum for even the price insensitive, last minute business traveler. Ideally, our hotel would like to be able to offer some rooms at a lower rate, even all the way down to their theoretical breakeven price, what's called the floor rate, to reach vacation travelers while at the same time maintaining higher prices for our business travelers. There is one way that this is possible. The very lowest prices for a given room are typically found in these so called opaque, inventory market.

Over opaque inventory pricing, OTA's, and at various times this has been available through Expedia, Hotwire, Priceline, Qunar, and Travelocity, they offer rooms in a given city by price only, maybe also providing some basic rating and category information such as the number of stars for the hotel. But they do not disclose the hotel name or any other information that would allow you to identify what hotel they’re talking about. Buyers in an opaque inventory market only learn what hotel they'll be staying at after their purchase is complete and all purchases are non-changeable and nonrefundable. This type of uncertainty and unchanged ability does not appeal to business travelers. So this market makes it possible for hotels to maintain at least two separate prices for the same room on the same night. Although exact statistics on the size of the opaque inventory market are scarce, I suspect that this market is about 2 to 4% of hotel room sales in the United States and Canada. And it may not be growing. Fancy hotels do not want to publicize that they are renting some rooms at 30% to 60% below their usual rates for identical rooms on the same night. Also, opaque pricing deals tend to be barely profitable if at all for hotels.

Prices tend to be near what's called the floor rate, which is an accounting calculation of the hotel's break even per room, based on allocating fixed costs per room plus variable costs to one room. Instead of opaque pricing, hotels can and do use various models to sell rooms at prices that fall between the rack rate and the floor rate. For example, when a hotel realizes that some rooms may go unsold they, can offer discounts off of the rack rate to a mailing list or a mobile text list of customers who have shown some loyalty to that hotel and that brand in the past. People who have some brand connection or brand loyalty will be willing to pay more than the opaque inventory floor price.

Question. When should a hotel offer a room at 80%of its rack price to a loyal customer? Answer, where there is a less than 80% chance that the room will be rented at the list price before it expires. In conclusion, a hotel room typically has at least three potential prices for each night. A rack, or listed rate, the floor rate and an intermediate promotional rate the hotel can offer on a last minute basis to loyal customers. Only the floor rate is constant. The listed and promotional rates vary by day of the week and by season of the year. So in theory, a hotel could rent the same level of room in one year for over 1,000 different prices.

**Risk Metrics: Leverage and Reputational Risk**

The most common sources of business risk seem to stem from excessive leverage. When a company owes more money than itis worth, it is unlikely to survive. If you're in a business that depends upon making loans, and you are also leveraged, the risk to you if your customers default is magnified many times.

In the United States, before the financial crisis of 2008, banks were allowed to borrow $33 for every $1 that they held in equity. So imagine a bank that could borrow money at 2% and lend it out at 3%. If they had $10 million in equity they could borrow $330 million in debt, lend out that money at 3% and make a gross profit before expenses of 3.3 million per year. Or a 33% return on their 10 million in equity. However, only a very small portion of their loans needed to become unrecoverable or in default for the bank to become insolvent. 10 million out of 333 million is only 3%.

Another type of risk is reputational, damage to your brand and ability to sell in the future. One area where any restaurant chain, manufacturer, grocery store, or retailer is extremely vulnerable to bad publicity and damage to reputation is when they have sold a food product that is tainted or contaminated and makes customers sick.

One successful example of risk management using big data in the product recall space is Costco’s rapid response to product recalls. The metric for Costco is how long it takes them to contact every potentially at-risk customer, when what's called a level one product recall occurs. Because Costco requires all of its customers to purchase an annual membership, and their membership database contains address information and often email and mobile phone numbers as well, it's relatively easy for Costco to track down customers. This is easy because Costco stores a record of every purchase made by every customer at the individual product ID, or SKU level. So that data is indexed and when a product is recalled by the manufacturer, as happened twice in the spring of 2014, with peanut butter and again with peaches. Within 24 hours, Costco was able to identify all of the more than 50,000 customers, who had bought the product and used the fastest available means to contact them, in some cases email and phone calls, but in any case a letter to a mailing address, and they were able to prevent many, many people from getting sick.

An even more interesting example of using Costco's database to track an epidemic to its source occurred in January, February of 2010, when 272 people in 44 different states became ill with a single genetically identical strain of salmonella. United States Centers for Disease Control, or CDC, sought to track down the source of contamination, but were initially unsuccessful. Until someone had the very bright idea of analyzing Costco purchase records. For those in the group of 272 who might have shopped at Costco. Since Costco membership reaches about one-quarter of US families, it would be reasonable to assume that this was about 70 people. It was then possible to identify what food items those 70 had purchased in common. The only common product was pepper salami, manufactured by a small sausage company, Danielle International, based in Pascoag, Rhode Island, and the contaminant turned out to be one single 25-pound box of crushed red pepper provided by wholesaler, wholesome spice, based in Brooklyn, New York. Not so wholesome after all.