# Wk1

## What are Business Metrics?

1. Metrics help us ask the right questions

Ask the right questions—what process can and should we change right now? To increase revenues, maximize profitability, or reduce risk

And find the best answer in time and with the resources available

AB testing

Translate it into a specific recommended action

Communicate with decision makers

Using visual metaphors and non-technical language

What are metrics? Metrics are special numbers that help us to ask and answer the right question and that we can impact when we change our business processes.

Sales tax is not metric. Can’t change it. “Click-through” rates are a metric.

1. Distinguishing revenue, profitability and risk metrics

What people in the company depend on this information and will ask for it:

* 1. Revenue-sales and marketing: outward facing

examples:

how many units of each product were sold over a given time interval

how this compares to the same time interval last year and the year before

sales by region, by product and by new vs repeat customers

sales funnel (potential future customers have been identified, where they are in the step by step process of moving towards making a purchase)

marketing: how effective any marketing campaigns may be

how many people have seen a particular advertisement or email marketing or mail offer

what percentage have responded

**Revenue metrics are about selling!**

* 1. Profitability-logistics, production and operations: operationally oriented

efficiency by which company create and delivers its product and services to customers

examples:

how much case is tied up in the form of unsold inventory

how much production is unsaleable due to spoilage or wastage

how often the company is unable to meet urgent customer requests and loses sales

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

**Profitability is all about efficiency.** Large established companies with relatively little room to increase revenues can often achieve significant increases in profitability.

Risk-risk management, creditors and outside investors

* 1. Track and mitigate dangers

Net case out is always the most important metric to track.

Examples:

How many months can the company survive at the present burn rate

Churn rate: the rate at which new subscribers drop off within a year, fewer targets who have never been customers and impossible to maintain revenue growth or hold steady. The greater reliance on longer term recurring revenue customers, the less dependent a company is on constantly successfully converting new prospects into clients, less risky

Banks: how much exposure they have to potential customer defaults at any time (default: failure to meet the legal obligations) , what percentage of their customers are expected to default or are in default now

**Risk is closely tied to debt**

**Optimistic extroverts + fastidious perfectionists + informed skeptics**

1. Distinguishing traditional and dynamic metrics

Convey urgency: significant change over a month or less; specific actions that will significantly impact the metric (announcing all your bad news at once is often a good business strategy)

Traditional metrics like quarterly revenues are impacted by dozens of different factors, aggregate data is kind of worthless

Dynamic metrics are twitchy: small changes in process can lead to big impact

Eg: 40% customers will abandon a webpage that does not load in 3 seconds (edge caching, Akamai)

## Case Study for Financial Accounting Metrics

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 a 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.

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.

## Revenue, Profitability, and Risk Metrics

1. Revenue Metrics: Traditional Enterprise Sales Funnel

Full-time sales, complex equipment or services, expensive

Start from a lead=>qualify a lead (plan and budget to buy)=>identify the correct decision maker=>expression of interest=>negotiated terms and price🡺soft-circle sale

1. Revenue Metrics: online sales

Amazon.com as a leading example of use of dynamic metrics

Enterprise sales = personal vs online sales = impersonal

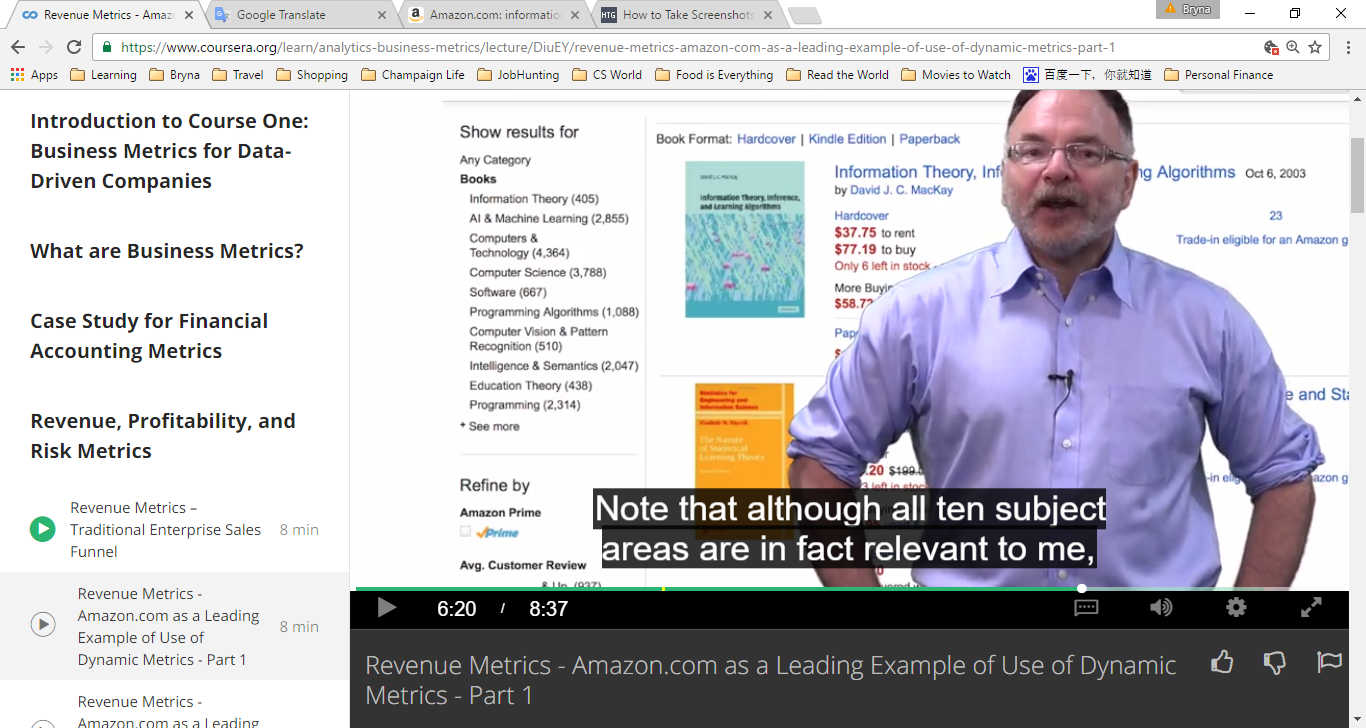
Real-time customization

Book ID Database: every book is assigned a separate record number and location specified by a unique book ID

Up-to-date text index

Amazon ranks the 1200 possibly relevant items by their predicted relevance to me and show me the top 12 of 1200 that Amazon’s prior data analysis predicts I am most likely to buy

1. Synonym matching: Matching with subject index or controlled vocabulary index => to find what people actually want but didn’t know how to ask for instead of giving them what they thought they wanted but will be disappointed by

Note that all 10 subject areas are in fact relevant to me, only 3 of the 10 contain even one of the words I typed in my original query ”information, theory, learning”

1. Identify the best-selling books within the subject subcategories my search terms fit best, top-sellers weighted by their topical relevance

Conclusion: 1. what are the subject area categories in the controlled vocabulary index most relevant to this user’s exact typed query terms? 2. Within the subcategories that most closely fit his query, what books are we selling the most right now?

**Co-occurrence data**： Amazon maintains 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

To notice that: 1). No discount for bundle, because doing so does not increase sales enough to justify it

2). Amazon seems not list the top 100 co-occurring sales book in simple order from high to low. A/B testing showing cover together with A increased sales or profits by more than showing with B=>create best upsell revenue potential

3). Amazon put the second-most popular book to buy together with my book first, we’ve already been given information about the most popular book and we didn’t bite on their offer(maybe?)

4). Amazon also tracks clickstream data: everything I ever looked at but did not buy

Conclusion: Amazon metric include 1). Frequently bought together metric

2). Recommendation engine, what’s likely to be also bought

3). Recommendation based on purchases made by people who looked at but did not buy the first book

1. Profitability/Efficiency Metrics: Inventory Management

3.1 Why to minimizing inventory days:

* negative float(Wet Seal, Cache, and Simply Fashion declared bankruptcy)
* fixed costs of storage
* wastage: fresh-based bread, value goes to zero; any product with an expiration date, hotel rooms or airline seats
* obsolescence: out-of-date

average number of days inventory is held (days inventory)

3.2 Days Inventory Estimate

public company annual reports include 1)i= inventory on hand the end of the year, 2) c=total annual cost of goods sold

estimated days inventory = i/c \* 365, Walmart = 46 days, 100,000 distinct product SKUs( stock keeping unit) on hand

tracking the date and quantity when any item is reordered, is delivered and stocked on the shelf

best practice:

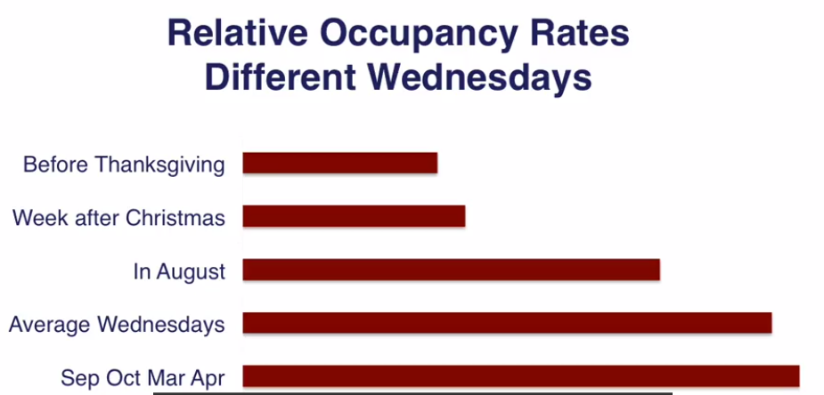
SKU reached zero=>unhappy customer, the higher the volume of unit sales of a product, the larger potential lost revenue every time we see a zero level inventory

1. Profitability/Efficiency Metrics: Hotel Room Occupancy Optimization

Airline tickets: almost all the cost associated with seat are fixed, sunk costs

The money spent to create the seat is completely wasted if the seat is not sold in time.

Variable pricing: business hotel



One way to try pricing strategy is to list rooms on OTA sites (cons: 10-25% commission), pros: if OTA fails to sell all of them, unsold rooms are released back and hotel can try to sell them directly.

Challenge: how to maintain two prices for two segments- business/vacation?

rack price (no discount)for the price-insensitive, last-minute business traveler;

floor price for vacation travelers

Opaque inventory market(2%-4%, prices are near floor rate), OTA offer rooms in a given city by price only with some basic rating/ category information, this uncertainty and un-changeability does not appeal to business travelers=> hotel can maintain two prices for the same room same night

Or offer some discount to people with brand connection or brand loyalty

At least three potential prices for each night: 1. rack/listed rate, 2. floor rate and 3. Intermediate promotional rate. 2 stays constant, 1 and 3 vary by day of the week and by season of the year

1. Risk Metrics: Leverage and Reputational Risk
   1. excessive leverage = low survival rate(when a company owes more money than it is worth, it’s unlikely to survive)

example: a small portion of unrecoverable loans will make bank insolvent(bankrupt)

---banks were allowed to borrow $33 for every $1 that they held in equity, a $10 million bank

can borrow $330 million in debt at 2% and lend it out at 3%=> make a 1% profit 3.3% million, which is 33% of their 10 million in equity. More than 3% ($10 million out of $330) loans in default will result in bankruptcy.

* 1. reputational risk: damage to your brand and ability to sell in the future

hotel chains, manufacturer, grocery story, retailers are extremely vulnerable to bad publicity

example: Costco’s rapid response to produce recalls, annual membership database contains address, email, phone information, easy to track down customers

track an epidemic to its source in 2010, 272 people in 44 states became ill, Costco membership reaches one-quarter of US families, able to identify what food items 70 people of the group had purchased in common.

# Wk2

## Introduction to Data-Centric Roles

1. different types of work that relate to data

Specialization will prepare learners for jobs as:

1. Business Analysts
2. Business Data Analysts

Started as Business (Data) Analysts + additional skill and experience:

Or master/PHD in statistics/CS

1. Data Scientists

Different training=>

1. Senior Software Engineer/ Project Manager
2. What role data plays in the success of various types of companies

* Strategic consulting: general business
* Traditional bricks-and-mortar companies:

Fully embraced the opportunities created by big data culture

Costco: good practice

* Strategic consulting: information technology
* Hardware and software companies
* Digital business companies(core value creation involves real-time analysis of data)

## The business Analyst

This title is used as a label for a vary range of jobs with quite different expectations

1. Local knowledge of the industry: ability to identify the most important and relevant business metrics, both industry-specific and general metrics
2. Ability to apply appropriate models to analyze those metrics(Excel): models are ways we represent a real world situation in simplified mathematical form
3. Ability to quantify the effectiveness of models used: choose the best measures available
4. Ability to listen, interview customers internal and external to define project requirements: translating demand into product features and services
5. Ability to identify patterns and trends in business data, make forecasts, organize information and display conclusions in charts and graphs: basic Excel skills

Ability to import and manage large data sets, develop and test different models, run optimizations using solver: intermediate Excel skills

1. Presentation skills: effective=>clear, concise, persuasive
2. Data visualization: make your conclusion/ recommendations intuitive to a non-technical audience

Business data analyst vs business analyst => analytical skill, think flexibly and explore opportunities

=> technical skills: SQL

=>higher position and better pay

## Data Scientist and Software Engineering Roles

Modeling, machine learning and forecasting

Capturing, storing, analyzing, sharing data

Communicating to decision-makers

In conclusion: grasp what is possible from big data and translate to engineering terms and translate needs of engineers into terms that the company’s non-technical management can understand

Bigger picture, how different tools and technologies can be integrated

Business data analyst skills + additional skills:

1. Intermediate to advanced knowledge of at least one of R, Matlab, or SAS
2. Advanced statistical methods: calculus-level probability and statistical inference course
3. Machine learning
4. Additional relational database skills: populate(put data into) and optimize(making pulling out data fast) SQL databases
5. Intermediate knowledge of the most important technologies used to communicate between the customer layer and database layer: Python, Java, JavaScript
6. Experience managing distributed and unstructured data including tools like Hadoop, MapReduce, Hive, Pig and Spark
7. Basic knowledge of tools and methods of natural-language processing for applications such as sentiment analysis
8. Experience with massively scalable could-based data hosting processing, with Amazon Web Services( AWS) , Microsoft Azure or Google Compute layer

Senior Software Engineer:

Lower-level language: C++

Highly scalable, low latency commercial relational Database: Oracle and Teradata

Experience of leading software development team

Experience with project management

Real-time and third-party data

## Companies as They Relate to Big Data

Ranked from most likely to least likely to be hiring for entry level business analyst positions

1. Strategic business consulting firms: paid for advice

Collect and analyze data to recommend changes

Deloitte consulting, McKinsey & Company, BCG, Bain & Company

Credit bureaus: Experian, TransUnion, Equifax, FICO(FICO scores, 90% US lending decisions)

comScore(track web clicks) and Argus(track credit card transactions & payments)

1. Brick-and-mortar companies: traditional businesses in all market sectors, hotel, financial, taxis, real estate and retail are facing dangerous competition from digital companies

Check-list:

* Offer a mobile version of its website: rapid loading time, use edge caching, sell directly from mobile apps for Android and IOS…
* Track your visitors’ full click stream: more than just count unique visitors or page views
* Define conversions of new visitors in two steps: register and become revenue-generating customers
* A/B testing of site features: optimize rates of conversion to revenue
* Know how to achieve and maintain a high Google Adrank
* Know the price per click-through for web advertising: both the conversion rate to revenue per click-through and the average lifetime value of a new customer
* Offer web based incentives: real-time online promotion to individual based on a preference analysis of specific history
* Allow customers to see what is on the shelf
* Allow customers to order ahead and pick-up
* Same day and faster delivery: collaborate with their-party shopping and delivery service
* Complete service as fast as digital
* Track all interactions with customers: located under a single customer record id in a database
* Develop membership programs
* Have point-of-sale customization: coupons for additional products aimed at past purchase interests when check out
* Rewards for high spend customers: the whales
* Track churn and outreach to quiet accounts
* Track SKU at the store level: to prevent excessive day’s inventory
* Track all zero-inventory items: to eliminate empty shelves
* Inventory optimization models: by store, region, season, even day of the week, balance minimizing days inventory and wastage against not running
* Offer effective price reduction and last-minute promotion programs to sell wasting inventory: hotel rooms, airline tickets

Cast study: Barnes & Noble

Nook, BN.com do not enhance any competitive differentiation

Its basic value proposition is as a place to gather with other people, sit, study, sell exclusive membership might be a solution

1. Consulting with software/IT systems integration focus: fortune 500 companies follow their advice

Accenture: hired by US government to manage healthcare.gov

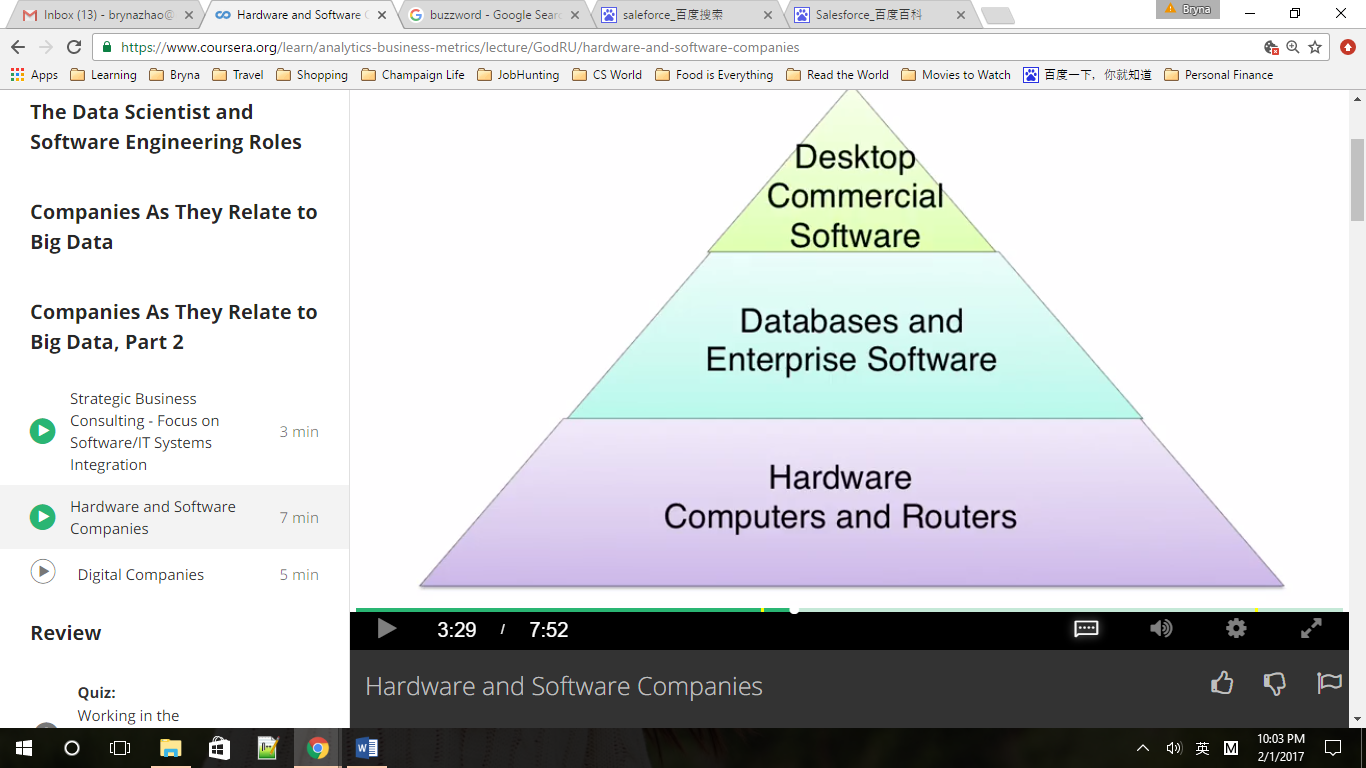
Palantir and Opera Solutions

1. Hardware and software companies: gather, store, manage, search, process, analyze, visualize and report world’s data

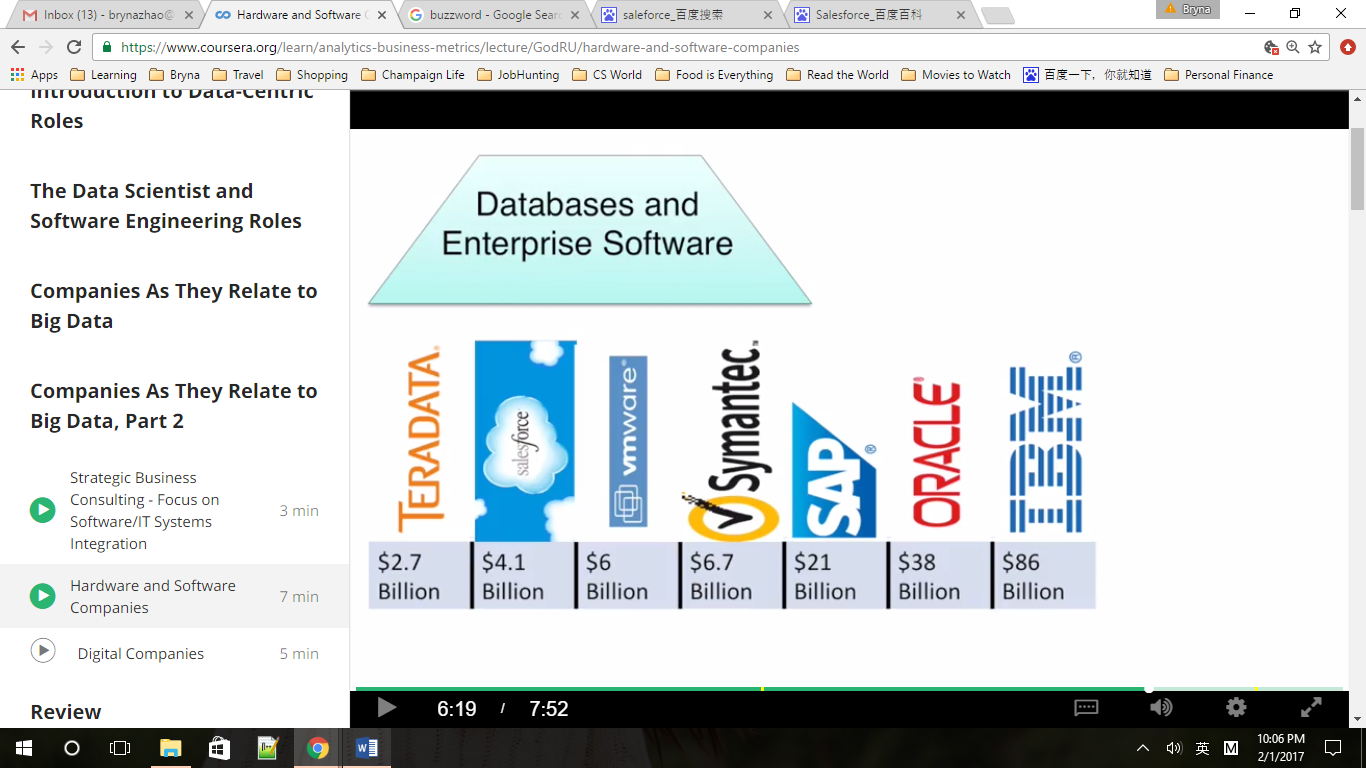
Unusual price dynamics, high up-front costs, very low variable costs

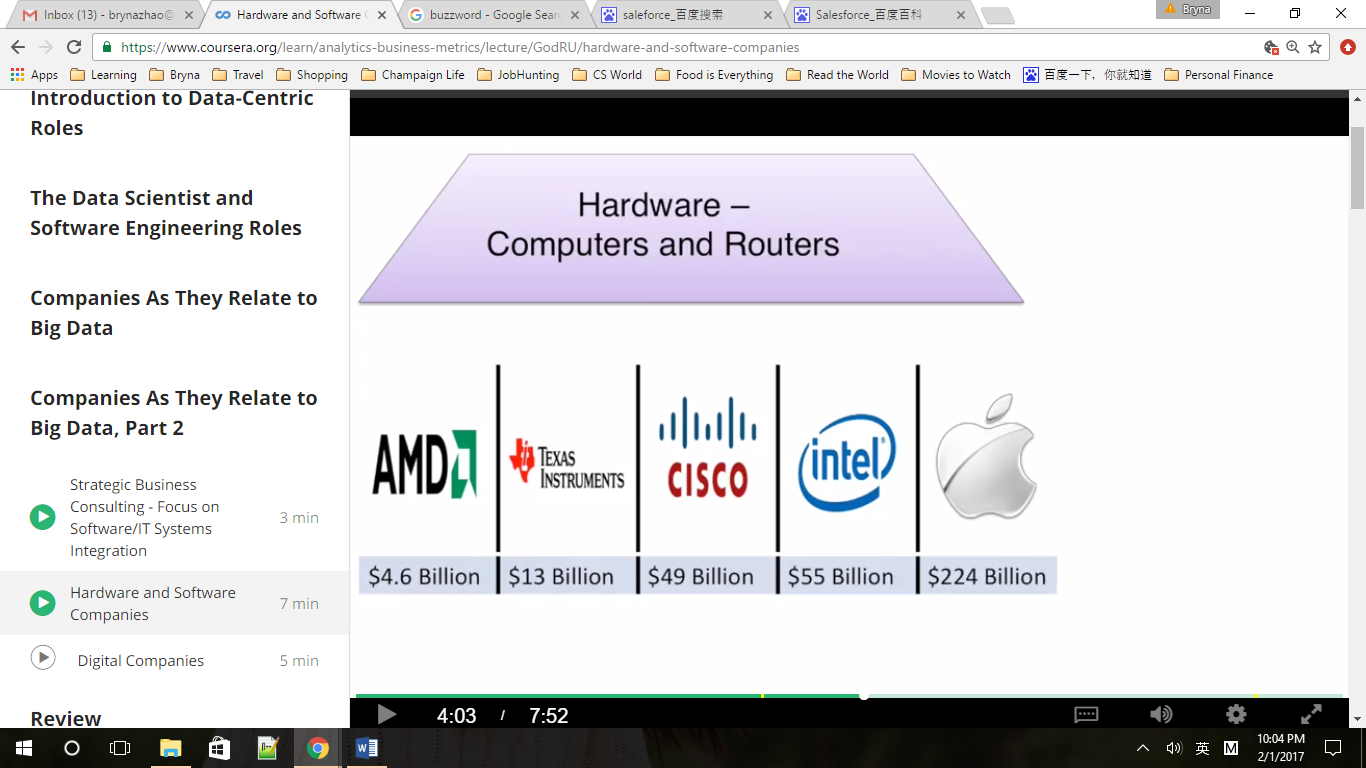
Big data culture is possible only because of today’s extremely low costs for storing and processing electronic information. To survive, hardware and software either need to figure out how to be the absolute lowest cost provider or they need to dominate a branded niche so effectively that when a lower cost alternative appears, the cost of switching away outweigh the benefits.

These companies provide core technologies for big data:









1. Digital companies: competing against established players in traditional markets, without the fixed costs. Create market sectors with no real equivalent before the internet. They rely on massive data processing, enhanced by machine learning systems to make real-time decisions. At the heart of their value proposition is business analytics.

Leaders using new business models to disrupt traditional markets

Products are offered better faster, cheaper

Amazon: cloud-based hosting storage and processing (client include AirBnb and Uber), early dominance of the cloud computing sector

* Internet companies offer information and services for free
* Use analytics to make their services attractive
* Identify users toward whom to target specific advertising campaigns



Sharing economy business models： Airbnb (provide owners with predictive analytics tools to optimize their revenue based on location, day of the week, time of the year)and Uber

# Wk3

## Web Marketing

1. Web Marketing: Metrics

Revenue metrics & dynamic: marketing is the deliberate, measurable process of creating and increasing positive awareness and interest for a specific brand, product or service offering within a target demographic

Demographic: a group of people with common characteristics, age, region,

Pizzeria in Trotters: Switching cost so high, easy to develop a loyal following of repeat customers

In a bigger city: 70000 competitors, match market differentiation to target demographic

* Special freezing process, shipping box, franchise network, ship anywhere in 2 days=>
* People who love pizza, preference or need for vegetarian, no vegetarian pizza locally, can afford luxury products, planning means in advance=>
* Target people who like vegetarian community on Facebook, buy names and address, people who buy luxury meals for others (office managers, enterprise sales rep, party planners)

1. Web Marketing: AdWords

* Offered by Google, priced by auction
* Identifies ad words=> terms advertisers placed bids on=> match topic(emergency pet care=>veterinary hospital and animal hospital)
* Maximum cost per click-through(max CPC) vs actual cost per click-through(actual CPC)
* Most expensive AdWords: insurance$54, Attorney$47, Mortgage$47

Note that Google only get paid when someone clicked through a sponsor link

* CPC bid X Quality Score = Ad Rank, Quality Score = weighted combination of: expected click-through rate, ad relevance and landing page experience, prevent exploitative misdirection that would hurt Googled brand

Landing page: product information=> good; sell ads, direct to other links=> bad, keep your landing page text directly relevant to your adwords

* Conversion rate, actual CPC / Conversion rate = acquisition cost
* LTV Lifetime value, represents the present value of all future revenues from that customer

In theory: $500 LTV, 2% Conversion rate, when CPC / C rate < LTV=>profitable, so we can pay $10CPC (10/2%=500), still profitable

Business goal: acquisition cost < fist year’s average revenues per customer,

$100 acquisition > $85, potential revenue is in risk, CPC < $1.7(1.7/2%=85) is safer

1. Web Marketing: Segmentation

* High conversion to revenue rates, high recurring revenue and lifetime value
* Figure out what kind of visitors become our best customers, focus on attracting more like them and engaging in a high-level form of segmentation
* Where do they come from:

1. Sponsored search

2. Organic search: unpaid link

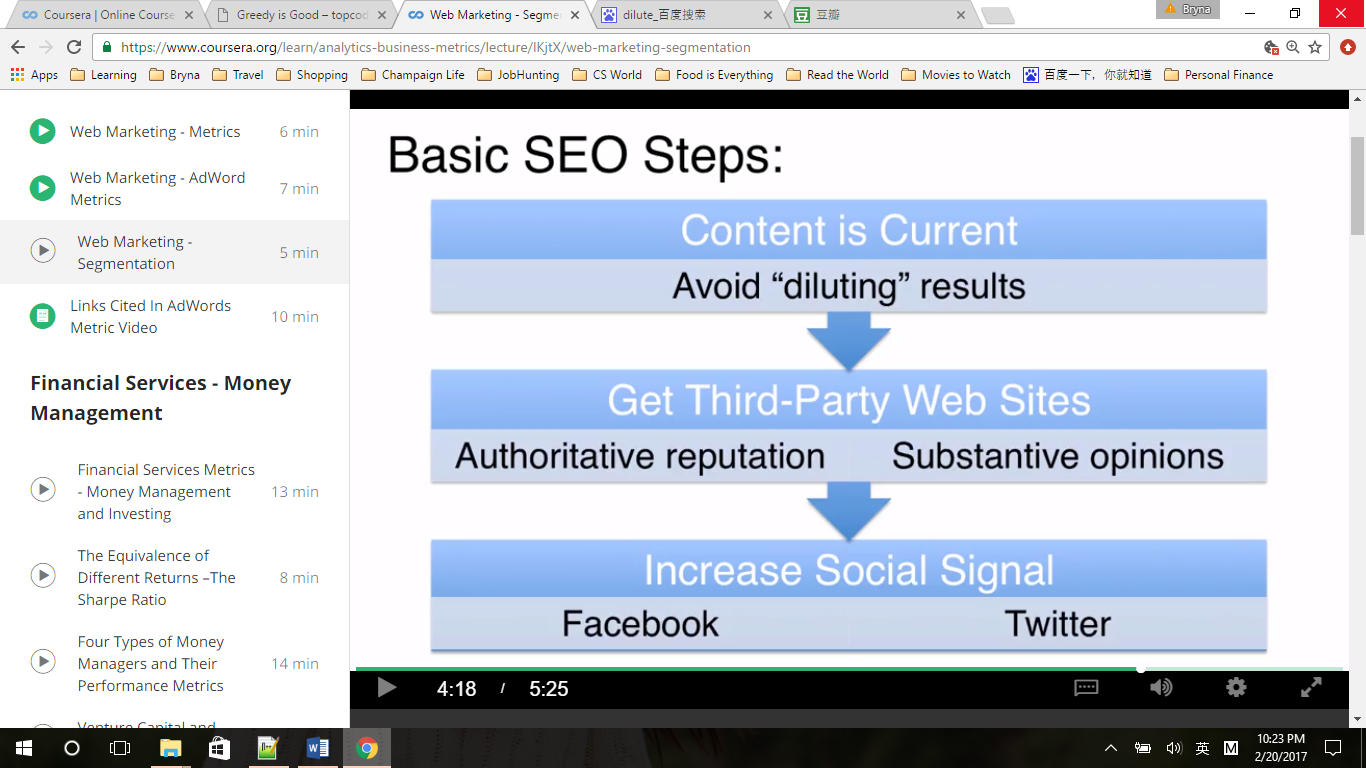
3. Clicked link: group email or tweet

4. Third-party website: blog post or article

5. Direct type in

* Devices used: laptop or mobile, browser
* Where in the world: country and state-IP address
* New or returning customer
* What happened at the site: bounce( leave immediately
* Duration of visit
* How many pages visited
* Click stream

If organic link have lower bounce rate=> prove SEO(search engine optimization)



## Finance Services – Money Management

1. Financial Services Metrics – Money Management and Investing

Absolute rate of return, annual rate of return

- Continuously compounded return: ln(final price/first price), ln(130/100)=0.2624(2y), 13.12% per y

- Discreet rate of return: (final price/first price)-1, (130/100-1)=30%(2y),(130/100)1/2-1=14.02% per y

- IRR: internal rate of return cashed invested at different times=> same fixed discrete annual rate of return for each payment

Example: 1 m at year 0, 1 m at year 1, IRR => (1+x)4+(1+x)3=5 =>x=29.62%

- series of annual returns: 25%,-18%,10%,-4% for 4 years

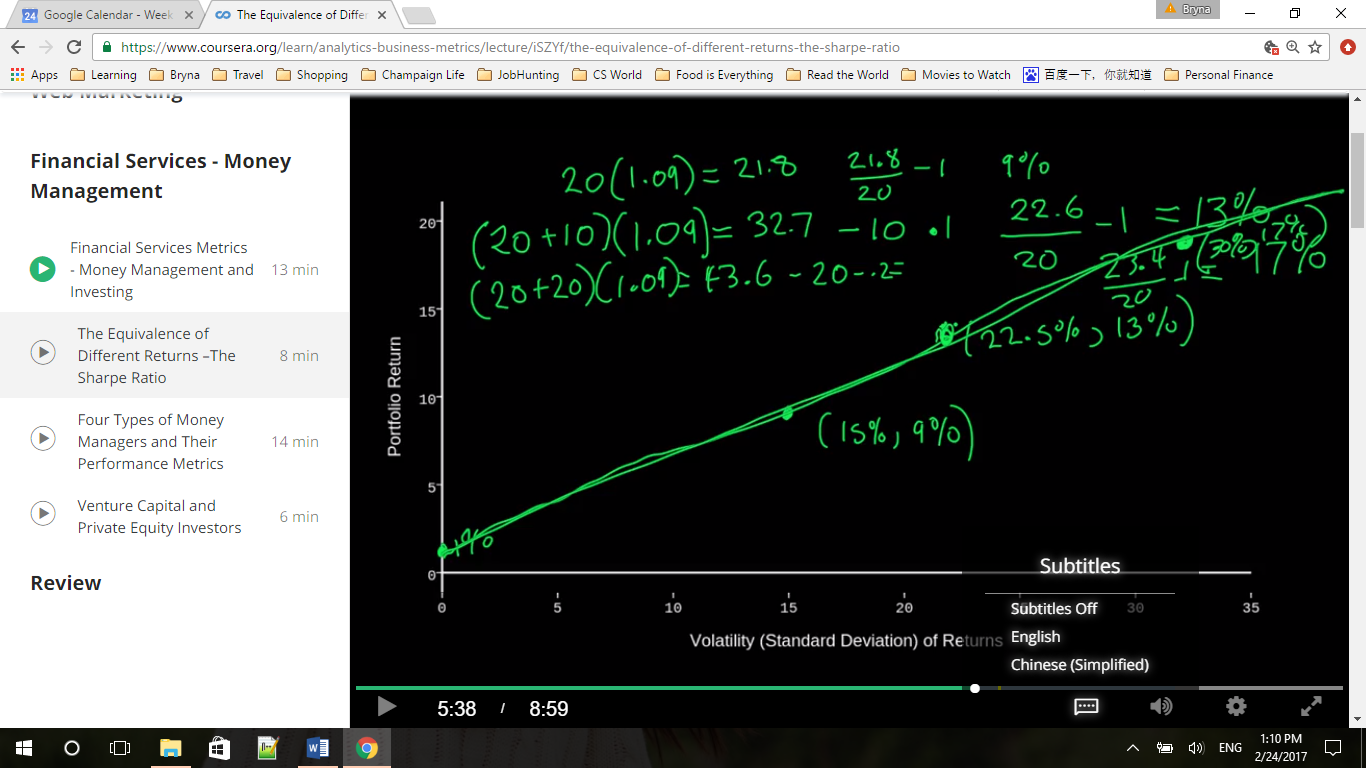
geometric means=> (1.25x0.82x1.1x0.96)1/4-1=2% ! long-term return without huge range or scattered of returns is more preferred

Standard Deviation: mean=3.25%, sd=(((21.75%)2+(-21.25%)2+(6.75%)2+(-7.25%)2)/4)1/2=16%

Volatility of returns

Zero chance of loss is a risk-free investment, US government three-month treasury bill rate 2015 =0.08%

1. The Equivalence of Different Returns – the Sharpe Ratio



Annual return=9%, volatility of return =15%, borrow money rate=1%

Borrow 10, annual return=>13%; but volatility increase 50% = 22.5% (because 10/20=50%)

Borrow 20, annual return=>17% (40x1.09-20.2)/20-1=17%, but volatility =30%

(15%,9%)=>(22.5%,13%)=>(30%,17%)

Volatility of return will increase just as quickly as the as the return

Sharpe Ratio = (return-risk free rate)/volatility of return, a revenue metric/a risk metric representing how many unites of revenue can be achieved for how many units of risk

(9%-1%)/15%=(13%-1%)/22.5%=(17%-1%)/30%, have to know the volatility to compare the manager skill

1. Four Types of Money Managers and Their Performance Metrics
   1. Index Fund Managers

* Stock indexes track performances of entire market, for example S&P 500 (500 largest companies in the UR ranked by their market capitalization)
* Market capitalization = (stock price/per share)x(number of shares tradable), $52x5,000,000=$260 million
* Market-capitalization weighted index( not an equal weighted index)
* Apple $666 billion / $15 trillion = 4.44%, its stock price changes are weighted by how much they contributes to the total
* Passive investment: buy and hold an investment attracts the performance of an index as closely as possible for the long term, SPDR S&P 500 ETF investment trust “SPY”
* Index fund managers do not need to pick stocks, manage a passive investment, rebalance and maintain the correct ratios of each stock to match a paper index at the same time minimizing transaction costs
* Judged by 1) How closely their fund performance matches that of their index, 1 tenth of 1% in a year is unusual

2) how low their expense ratio is( primary factor)

Expense Ratio = money spent on operating fund/ total market value of fund assets, operating costs include marketing, administrative expenses, mostly what the fund managers pay themselves, but does not include brokerage fees the fund pays to buy and sell its assets. SPY expense ratio = 0.0945%, over $ 127 million paid out every year; iShares Trust- iShares Core S&P 500 ETF(IVV) = 0.07%, Vanguard S&P 500 ETF(VOO)=0.05%

* 80% of professional active money managers performed worse than SPY, worse than random (if pick by random, 50% managers would outperform the index)
  1. Mutual Fund Managers

Specialize: buy stocks of a particular company size, by market cap, or in a particular market sector or in a particular country or group

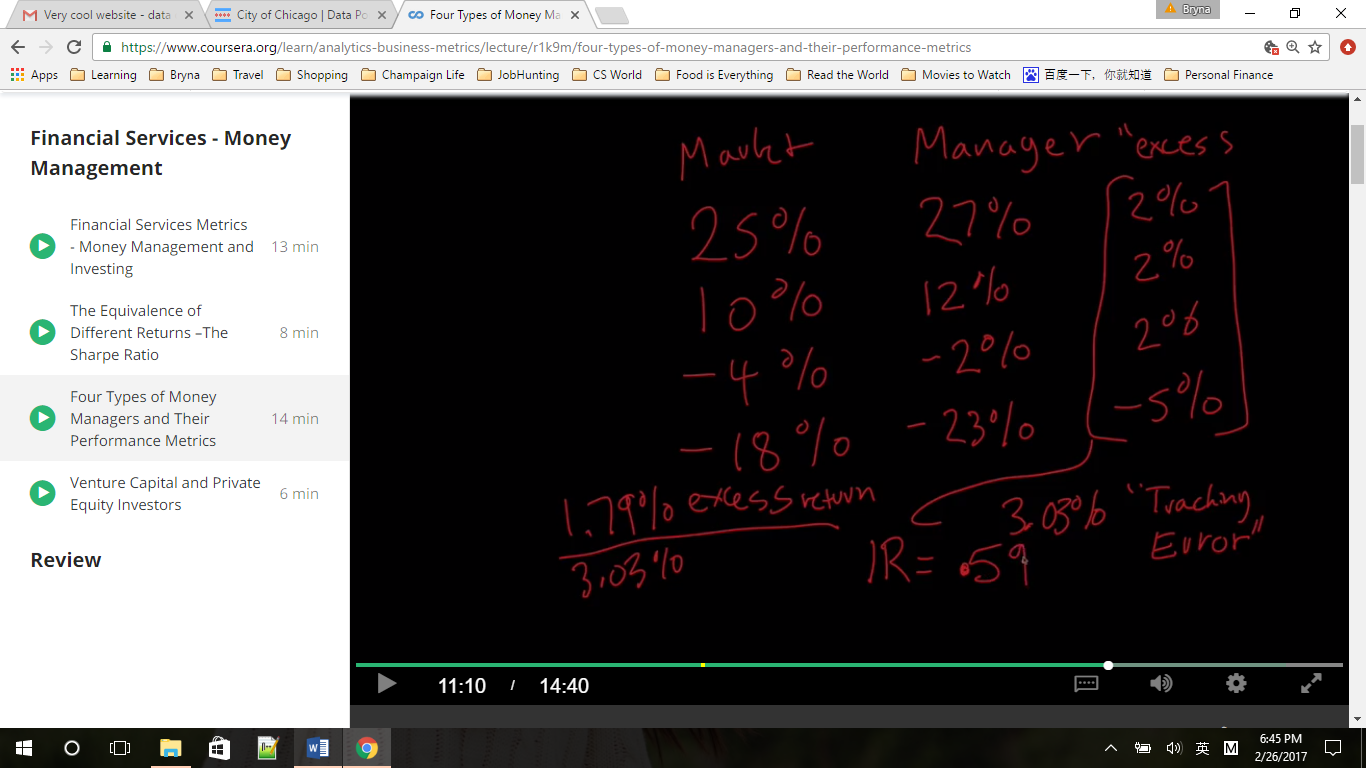
* Universe: select 30 or so stocks from the large pool of 500 or 1000 investments
* The mean return of all the stocks in the universe is the relevant index => benchmark for evaluating the performance
* Judged on 3 measures 1) excess return = portfolio return – benchmark return

12%-(S&P 500) 8% = 4%, what if S&P = -20%, excess return 4% means portfolio return =-16%

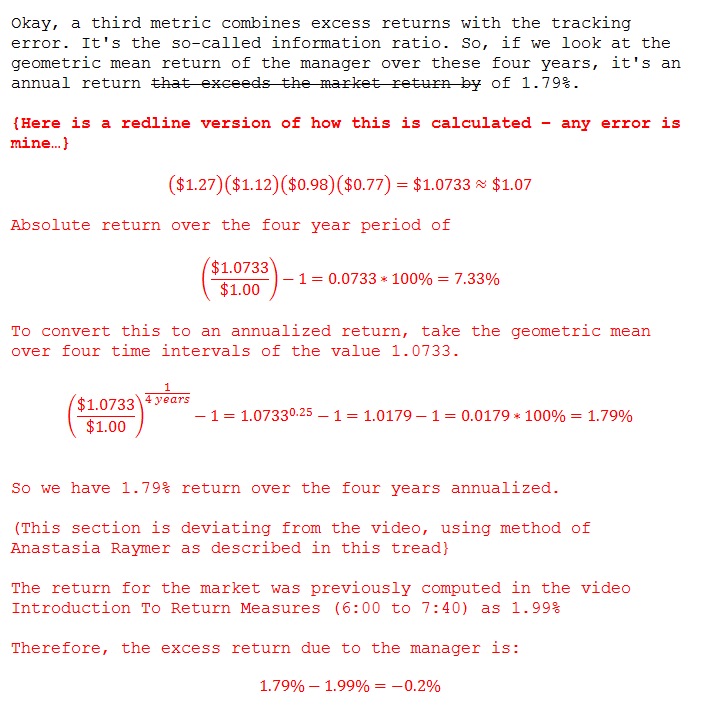
2) tracking error = standard deviation of excess return, risk metric

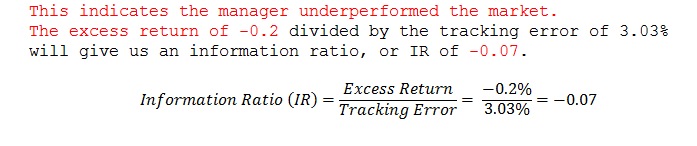
3) information ratio = excess return / tracking error

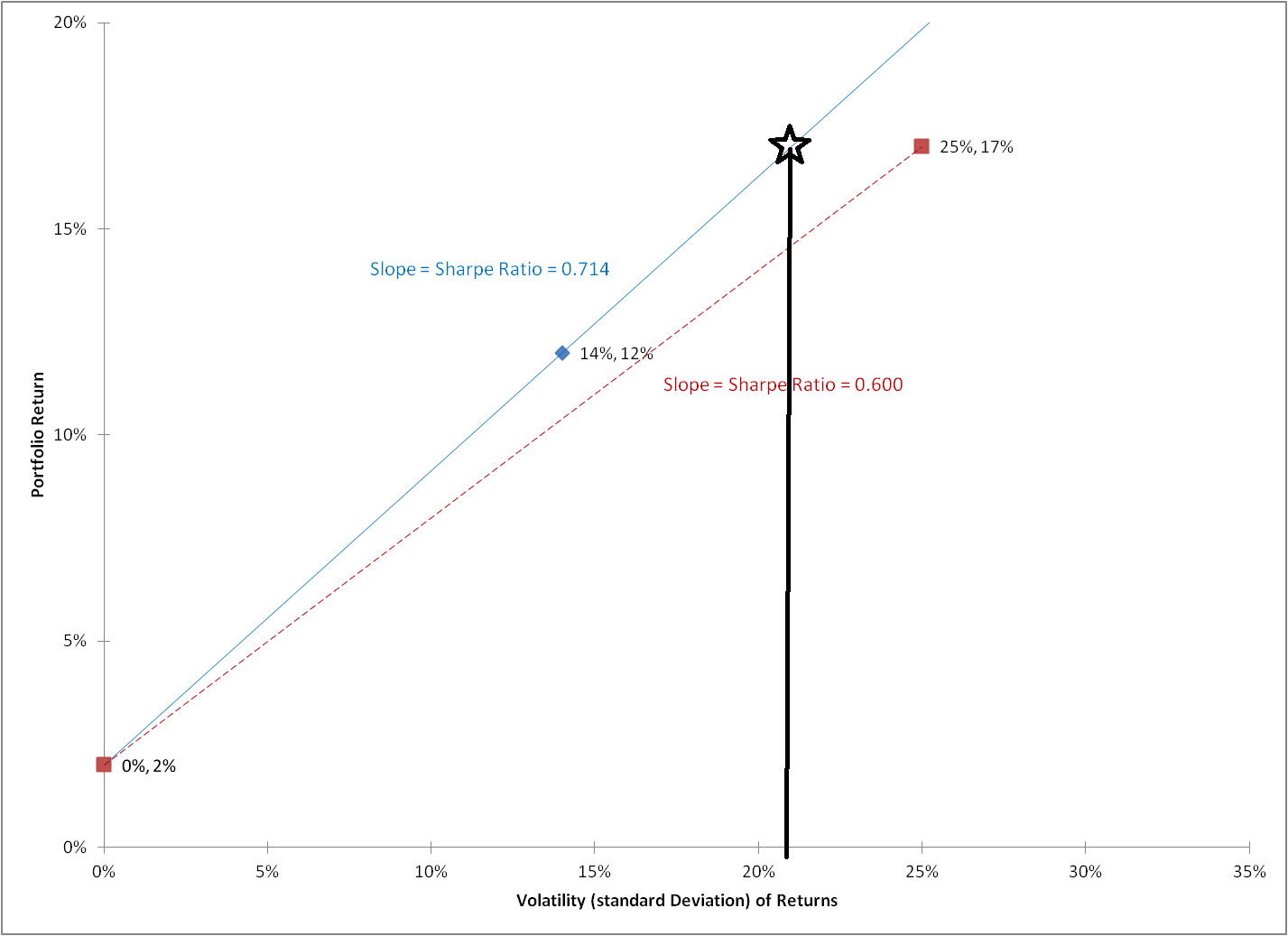
4) Sharpe ratio compared to the Sharpe ratio of their benchmark

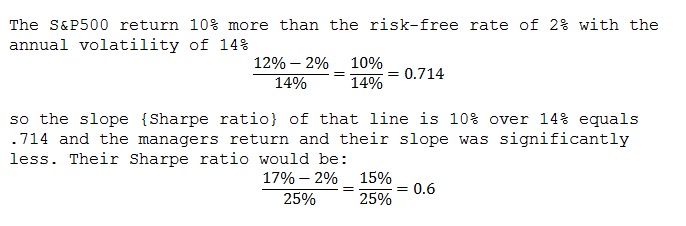


Tracking error = 3.03%









* 1. Venture Capital & Private Equity Investors
* Do not collect cash upfront
* Investors commit money over the life of fund
* $20 million commitment = 8 “calls” of 2.5 million spread out over 4-5 years
* Time to source, develop and structure; multiple investment, same company; investor don’t want cash sitting idle
* Internal rate of return is the best metric: discrete, compounded rate of return on just the money actually drawn by the funds
  1. Hedge Fund Managers
* Fewer restrictions
* Permitted to short: which means structuring a deal so that they can make money when a stock price goes down and to invest in many different types of assets, including options and derivatives.
* Invest in different asset classed, not tied to single sector universe
* Maximum drawdown from high water mark



Unlucky one would have lost 57% of their money

* Years to break even: the worst high to low loss / annualized return = 57% /5% = 11.4 which is the number of years at the average return that an investor would recover from the worst lost(assuming they enter the fund at the worst time)
* A fund is generating a steady compounded return every day, its volatility of return =0 and its internal lineal correlation of time and log wealth =1.

1. Venture Capital and Private Equity Investors