0. By the feeling

凭感觉

1. The Emergence of Statistics

凭人分析

essentially numbers in a spreadsheet that were manually examined

2. The Early Days of Modern Data Storage

开始使用计算机如excel

**1928**

[Fritz Pfleumer](http://en.wikipedia.org/wiki/Fritz_Pfleumer), a German-Austrian engineer, invents a method of storing information magnetically on tape. The principles he develops are still in use today, with the vast majority of digital data being stored magnetically on computer hard disks.

3. The Start of Large Data Centers

使用数据库，1台电脑完成分析

**1970**

IBM mathematician Edgar F Codd presents his framework for a “relational database”. The model provides the framework that many modern data services use today, to store information in a hierarchical format, which can be accessed by anyone who knows what they are looking for. Prior to this accessing data from a computer’s memory banks usually required an expert.

4. Big Data analytics

Nowadays, average US company with over 1,000 employees is storing more than 200 terabytes of data, one computer cannot provide enough computing resource to handle that amount of data, so finally, Big Data Analysis emerges. Let me explain the term Big Data analytics more precise.

5. What is big data

Actually, Big Data Analytics is an aggregation of a set of modern technologies. They are Data management, Data mining, Hadoop, In-memory analytics, Predictive analytics, Text mining and many other modern technologies.

With these amazing technologies, Big data analytics examines large amounts of data to uncover hidden patterns, correlations and other insights. It’s possible to analyze your data and get answers from it much more quickly and efficiently than traditional technologies. For example, Netflix can find it. Google can find it. Nowadays, Big data analytics have become the dominant design in the behavior analysis industry.

Netflix

1. Netflix introduction

Netflix is an American entertainment company, specializes in providing streaming media and video-on-demand online.

In April 2017, Netflix reported having over 98 million subscribers worldwide, including more than 50 million in the United States.

2. How Netflix keep old subscribers

留住老用户的办法就是不断加入新content

Netflix doesn’t blindly pick which movies to stream. Licensing movies from studios is expensive, so Netflix must find some technologies to help them decide. And of course, the answer is big data analytics.

As mentioned, Netflix has over 98 million subscribers worldwide. Having this large user base allows Netflix to gather a tremendous amount of data.

Netflix may track

The details of the contents you watch(type, director, actor, etc)

* When you pause, rewind, or fast forward
* What day you watch content (Netflix has found people watch TV shows during the week and movies during the weekend.)
* The date you watch
* What time you watch content
* Where you watch (zip code)
* What device you use to watch (Do you like to use your tablet for TV shows and your Roku for movies? Do people access the Just for Kids feature more on their iPads, etc.?)
* When you pause and leave content (and if you ever come back)
* The ratings given (about 4 million per day)
* Searches (about 3 million per day)
* Browsing and scrolling behavior
* Netflix also looks at data within movies. They take various “screen shots” to look at “in the moment” characteristics. Netflix has confirmed they know when the credits start rolling; but there’s far more to it than just that. [Some have figured](http://gigaom.com/2012/06/14/netflix-analyzes-a-lot-of-data-about-your-viewing-habits/) these characteristics may be the volume, colors, and scenery that help Netflix find out what users like.

With analyzing this big data, Netflix can make better decisions and ultimately make users happier with their limited budget.

3. How Netflix attract new subscribers

吸引新用户的最佳方法就是创造竞争对手没有的content

In 2011 Netflix made one of the biggest decisions they’ll ever make. It wasn’t anything material, but rather it was about content. They [outbid top television channels like HBO and AMC to earn the rights for a U.S. version of *House of Cards*](http://www.deadline.com/2011/03/netflix-to-enter-original-programming-with-mega-deal-for-david-fincher-kevin-spacey-drama-series-house-of-cards/), giving them 2 seasons with 13 episodes in each season.

At a cost of $4 million to $6 million an episode, this 2-season price tag is over $100 million. So why did they make such a big bet?

With big data analytics, Netflix knew:

* A lot of users watched the David Fincher directed movie *The Social Network* from beginning to end.
* The British version of “*House of Cards*” has been well watched.
* Those who watched the British version “*House of Cards*” also watched Kevin Spacey films and/or films directed by David Fincher.

This combination of [factors had a lot of weight in Netflix’s decision to make the $100 million investment in creating a U.S. version of *House of Cards*](http://www.salon.com/2013/02/01/how_netflix_is_turning_viewers_into_puppets/). Jonathan Friedland, Chief Communications Officer, [says](http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html?pagewanted=all&_r=0) *“Because we have a direct relationship with consumers, we know what people like to watch and that helps us understand how big the interest is going to be for a given show. It gave us some confidence that we could find an audience for a show like* House of Cards*.”*