

Algorithms, Big Data, and Online Marketplaces

Scan the QR code using WeChat to post live
comments on our screen:



Applications of Big Data

Demand Forecasting

Cinelytic is a Los Angeles based start-up. It collects data on title name, synopsis, logline, genres, rating, per-territory distributors, production budget, keywords, associated production companies, per-territory theatrical revenues, social media metrics, and other relevant data to predict the box office revenue of a new movie.



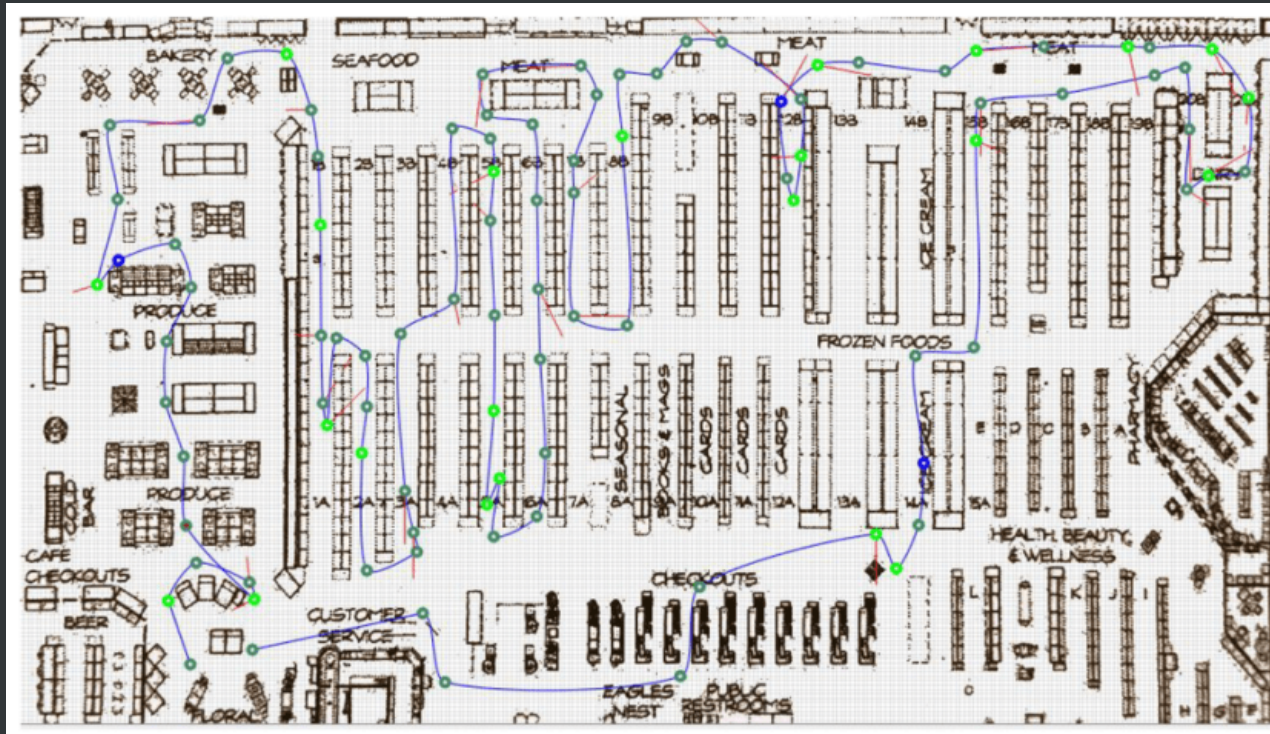
Product Management

Prof. [Raymond Burke](#) of Indiana University was hired by a chain supermarket in the US, where he manages a number of shelves. He installed video cameras in store and invited users to wear eye-tracking system to analyze store flow and eye movement. Based on the data, he optimized the layout of the store and product, improving store revenue by 30%. You can find his video [here](#).

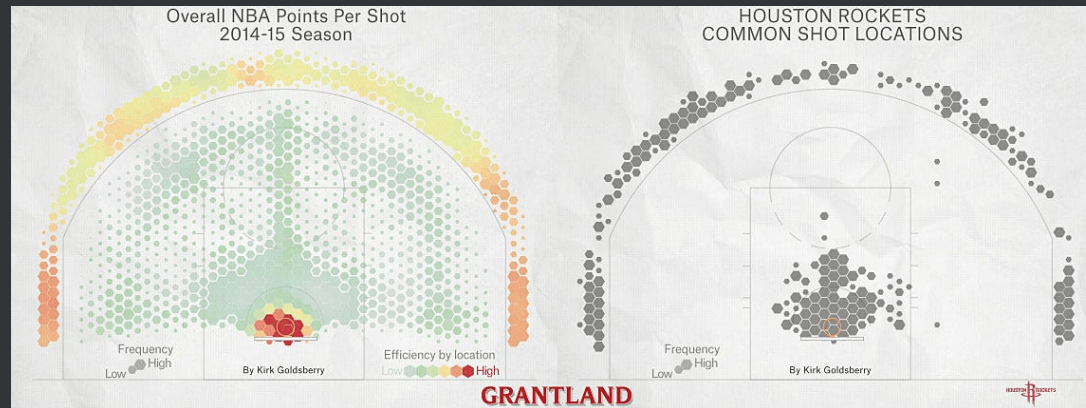
Product Management



Product Management



Opponent Analysis



Every NBA team employs data scientists who analyze their own strengths and weaknesses, as well as those of their opponents. These professionals collect various types of data, including information on ball and player location, speed and distance, touches, passing, shots, and behavioral habits, among others.◦

Data Analytics

Data champions: How the Golden State Warriors are turning on-court data into a competitive advantage

March 5, 2021

Consumer Analytics



Nemesysco is an Israeli start-up. It helps call centers monitor phone calls to detect emergencies and irregularities. It also helps insurance companies detect fraud through analyzing phone calls.

Personalized Recommendation

有人只是去参加了一次北京婚博会，晚上回到家打开微博和微信，发现信息流广告全部变成了婚纱照、婚庆公司、婚礼礼服等。令他感到恐怖的是在此之前从未在手机进行过结婚相关的任何搜索。这一切发生改变的原因仅仅是因为他去了一次婚博会这个地方而已。

有人在知乎看到除甲醛的相关问题，只是百度了一下，结果连一个美食app都开始推荐除甲醛公司。在百度上打开某理财网站，不到半小时推销电话就打过来了。

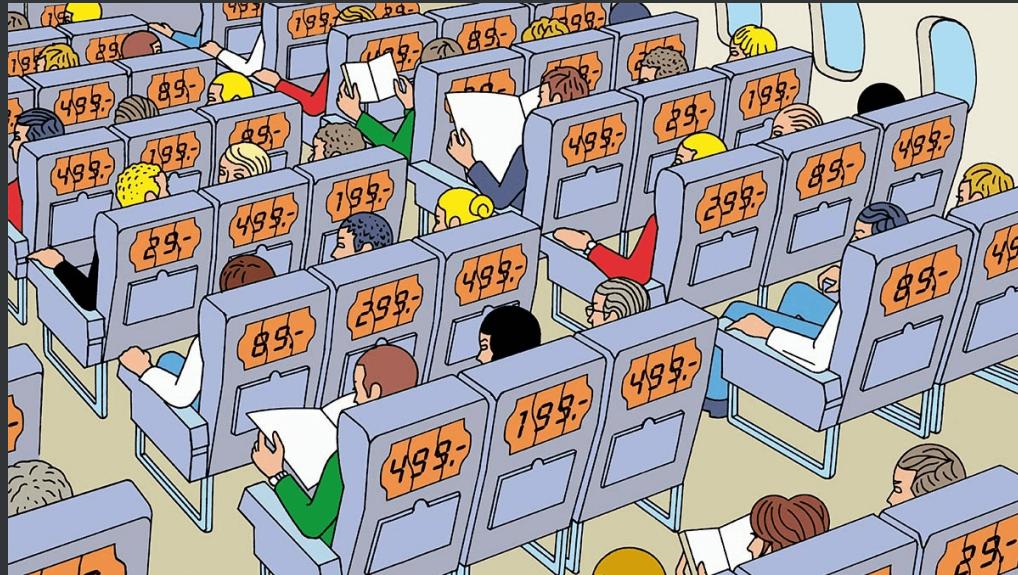
有时候你在网上搜了一本小說，然后突然就会有很多假网站在百度上显示他们网站有这本小说可以下载，然后让百度把他排在首页，你打开链接一看其实里面没有，但是有其他东西的广告。

Personalized Pricing



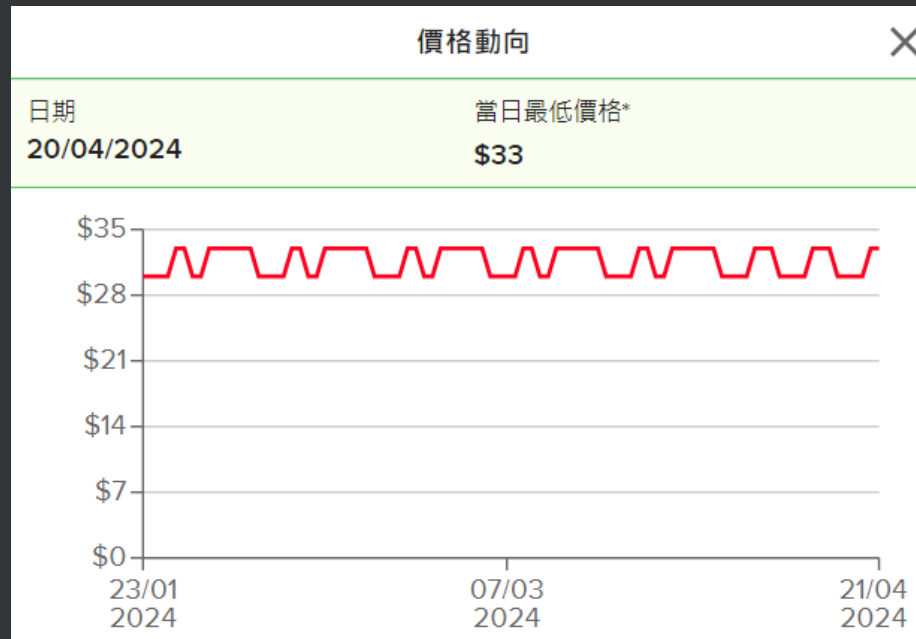
The online traveling agency Orbitz charges users accordingly to their device. iOS users pay 30% more than Windows users do.

Personalized Pricing



Airlines and OTAs have long been known as varying prices according to your geographic locations (IP addresses).

Traditional Pricing



Platforms such as Amazon automates their prices even every minute.

Wait, let me ask you another question.

Instructor

Xi Li (李曦)

Professor of Marketing

Professor of Information and Innovation Management

Director, Asia Case Research Centre

Associate Director, Institute of Digital Economy and Innovation

Ph.D. (Management), University of Toronto

M.Phil. (Operations Research), HKUST

B.E. (Computer Science), Tsinghua University

Email: xili@hku.hk

Office hour: by appointment

Teaching Assistant

Ray Song
soongray@hku.hk

BTW, Prof. Vincent Zhang and I will be co-teaching an exciting new course on Python programming (offered in Module 5). You can try the course if you want to explore more about programming.

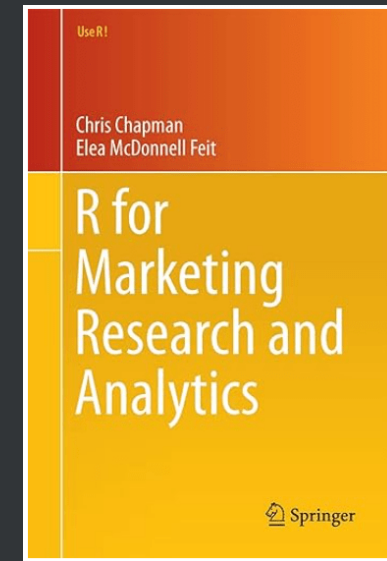
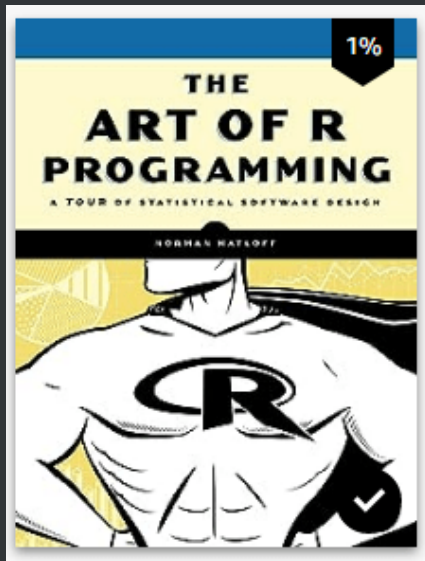
Teaching Materials

All the course materials will be posted on the
following website:

https://ximarketing.github.io/_pages/teaching/

Password is: 2425

No textbooks. However, we suggest some additional reading if you want to explore more.



Grading

Grading

The overall grading structure is as follows.

Group work (50%), including:

- Two group data analytics project, 20% each.
- One mini-presentation in class, 10%.

Individual work (50%), including:

- Individual class attendance, 15%.
- Individual class participation, 10%.
- Individual course project, 25%.

Group Project

Students work in groups on two data projects on the dataset provided by the instructor. You will analyze data using the techniques you learned in (and also outside) the class.

- No presentations for the group project;
- No reports for the group project;
- You only need to submit a few pages slides showing how you analyzed the dataset and the conclusions that you have reached from it.

Mini Presentation

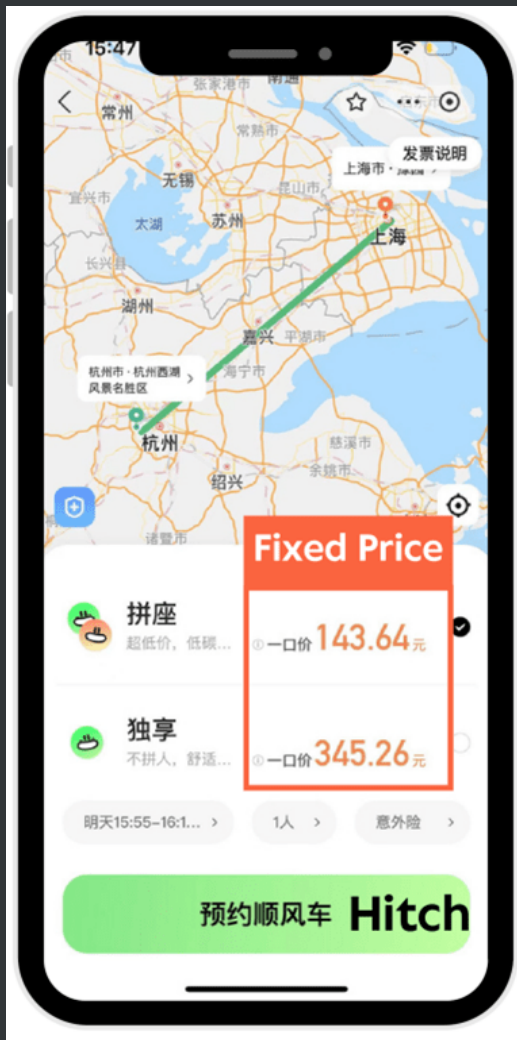
Each group needs to propose an interesting observation from online (or even traditional) marketplaces. You will invite other groups to discuss about the observation and explain the possible causes. The group also offers their own explanation. The presentation only takes around 5 minutes.

I will make a few mini-presentations myself first so that you know how to present.

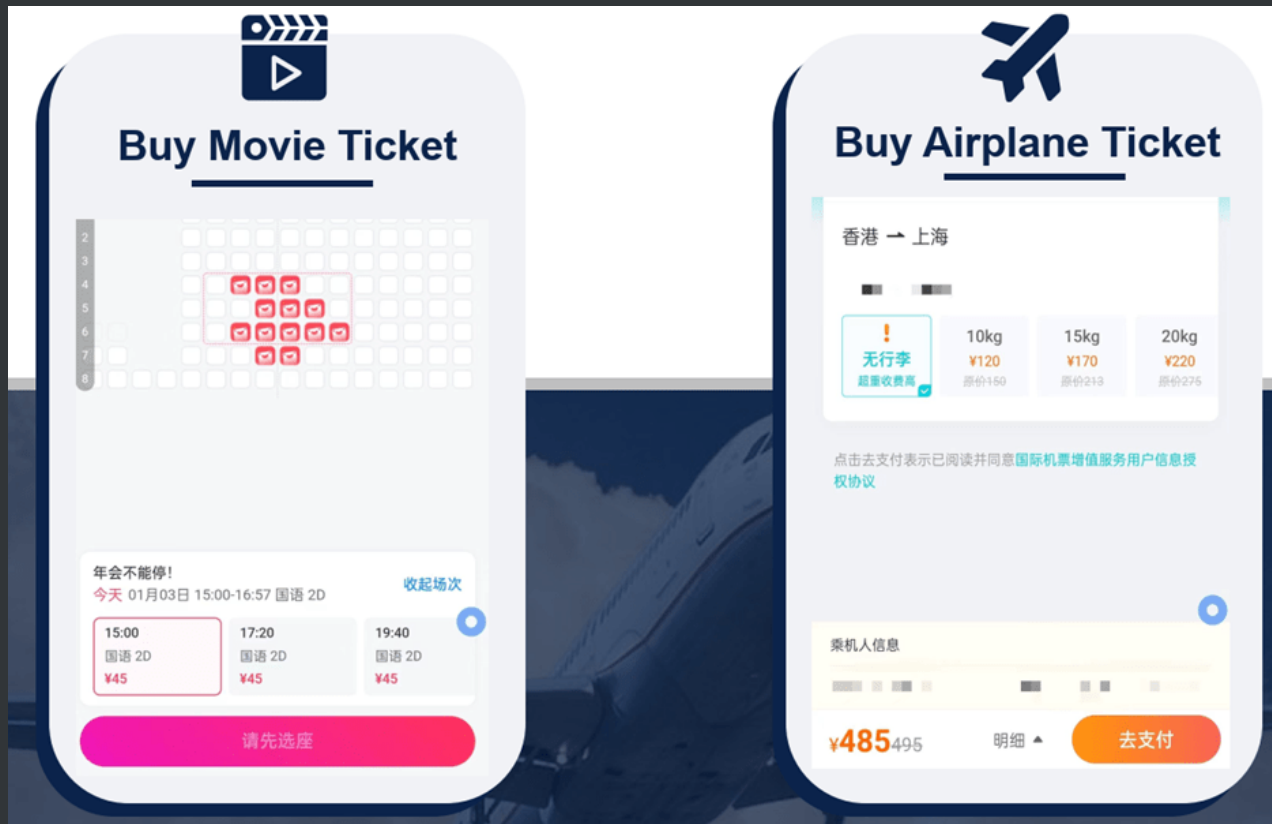
Mini Presentation: Example



Some online marketplaces display the sales, some display reviews, and some display viewership, why?



Why fixed price (一口价) for Hich (顺风车) but dynamic price (实时计价) for Express (专车)?



When booking movie tickets, you select seat first and pay later; when booking flights, you pay first and select seats later. But why?



cookies



whiskey

Why does Fresh Hippo (盒马鲜生) charge lower prices offline?

Individual Attendance

10 points, 1 point for each class.

Scan your HKU ID to check-in each class. If you forget to bring your HKU ID, confirm your attendance manually with the TA during the class.

Otherwise, you miss the point even if you attended the class.

Individual Participation

Answer questions in class to get your participation points.
Your points depend on:

- The number of questions answered,
- The hardness of the question, and
- The quality of your answer.

The TA takes notes whenever you answer questions in class.
If you suspect that TA ignored your participation, check with TA immediately, ideally during the next class break.

Individual Project

Choose **one** of the following two tasks:

(A) Find a dataset yourself and conduct your own analysis, just like what you do in your group project;

(B) Find a company that uses data cleverly, and discuss the company's data strategy (no data analysis needed).

Again, no reports, no presentations, only slides.

Anything unclear? Ask me now.

Overview

Overview

We have 10 lectures, including

- Today's Introduction
- Two half-class data workshops (totalling 1 lecture)
- Eight regular lectures

Overview

The topics to be covered include

- Programming with R (fine for beginners)
- Data visualization with Tableau (fine for beginners)
- Web scraping: How to download text and images from the web
- Modeling discrete choice: How do consumers make choice in the supermarket?

Overview

The topics to be covered include

- Text mining: Getting insights from text
- Causality: Does X really cause Y ?
- Tips in data analytics
- Personalized prices and recommendation systems

Overview

Note that you may have heard about this class from our alumni. However, this year we have an overhaul of the class and the contents are very different from previous years:

- We have brand-new slides this year;
- We include a new data analytics technique, “conditional logit model” in the class;
- We include data analytics techniques for AB tests;
- We updated teaching materials on R and Tableau.

Overall, the class is not harder but becomes more interesting and useful!

TODO List

TODO List

We have group assignments, so please form groups with your classmates. **Deadline: Dec 3, 23:59.**

- Each group can have up to 8 students.
- Create a name for your group.
- Submit your group information to your TA.
- Let me (or the TA) know if you do not have a group.

If you do not have a group and do not let us know, we assume that you prefer to work as a one-person group.

TODO List

Here are some past group names for your reference (listed in alphabetical order):

- 25 hours
- CEOs
- Data Geeks
- Data Killers
- Doctor Malou
- GGBond
- Gold Miners
- Sense 8

TODO List

Our class uses R and Tableau for teaching. Please install R and RStudio on your laptop and bring it with you for the next class. You can

- Download R [here](#)
- Download RStudio [here](#)

Note that your installation path should not contain any non-English letters, otherwise you will be unable to use some functions.

安装路径必须为纯英文，否则某些功能无法正常使用！