Algorithms, Big Data, and Online Marketplaces

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

How many of you have Taobao/JD/PDD installed on your phone?

What's an algorithm?

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

Kevin Zhang

kevinzyx@hku.hk

Kevin is also an alumnus of our MSMK programme.

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

FT中文网

首页　中国　全球　经济　金融市场　商业　创新经济　教育　**观点**　管理　生活时尚　视频　音频　每日英语　换脑ReWired

Lex专栏　**专栏**　分析　评论　社评　书评　读者有话说

专栏　**明德商论**　＋ 关注

自愿数据共享，真的可以保护消费者？

李曦：笔者的研究发现，基于自愿的数据分享更可能会推高产品的价格，进一步损害消费者的权益。

李曦　＋ 关注

HKU
BUSINESS
SCHOOL
港大經管學院

46

电商下半场 如何从 "流量至上" 走向 "品质为王" ？

2025年11月19日19:46 | 来源：人民网　　　　　　　　　　　　　　　 T_T 小字号

对话现场。人民网记者 任峰涛摄

Teaching Materials

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

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

Password is: 26ATC

Grading

Grading

The overall grading structure is as follows.

Group work (50%), including:

- One group data analytics project, 20%.
- One group AI project, 20% (new!) -- it replaces a data project.
- One mini-presentation in class, 10%.

Individual work (50%), including:

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

Group Data Project

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

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

Group AI Project (New!)

Students work in groups on one AI project. You will use a large language model to generate some interesting and innovative content.

- No presentations for the group project;
- No reports for the group project;
- You only need to submit the prompts you used to communicate with AI and the output of your AI.

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 your own explanation. The presentation only takes around 5 minutes.

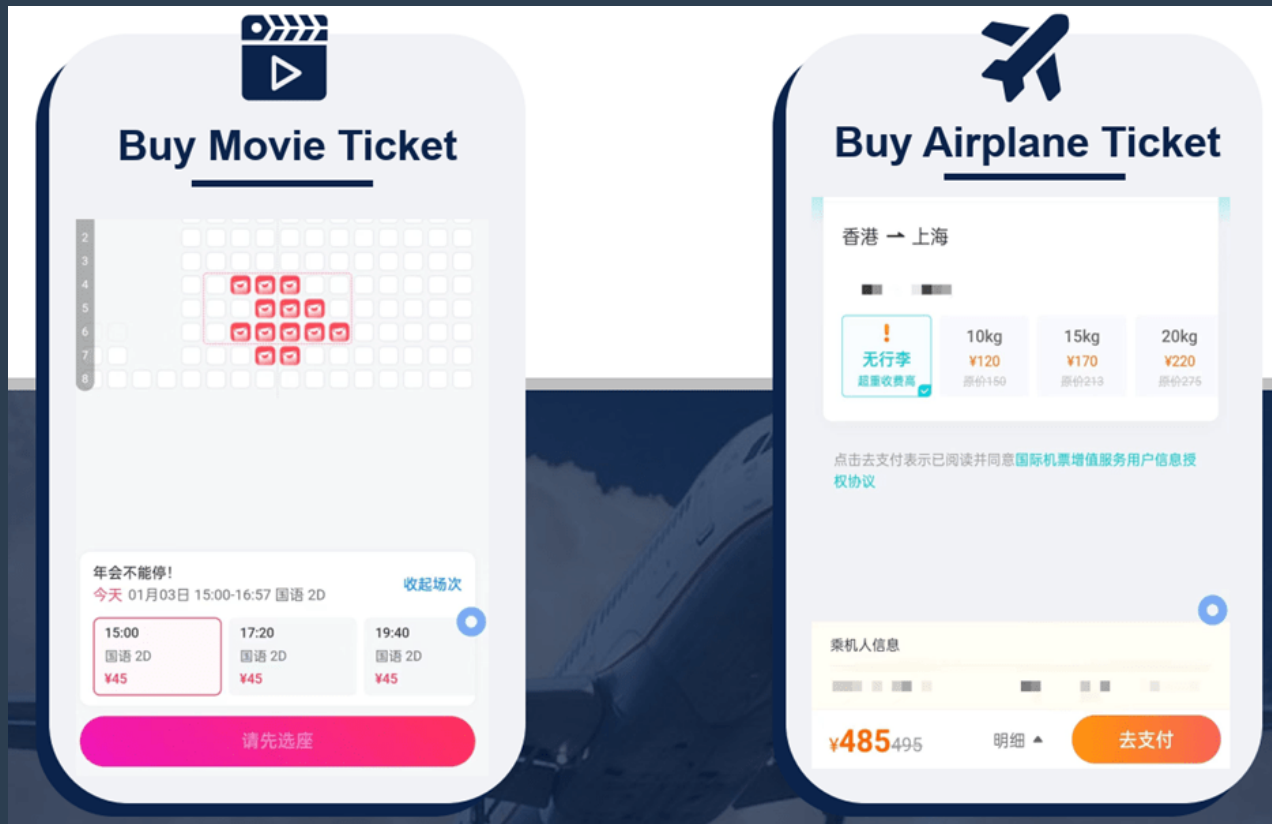I will make a few mini-presentations myself first so that you know how to present.

Mini Presentation: Example

Some online marketplaces display the sales, some display reviews, and some display viewership, why?
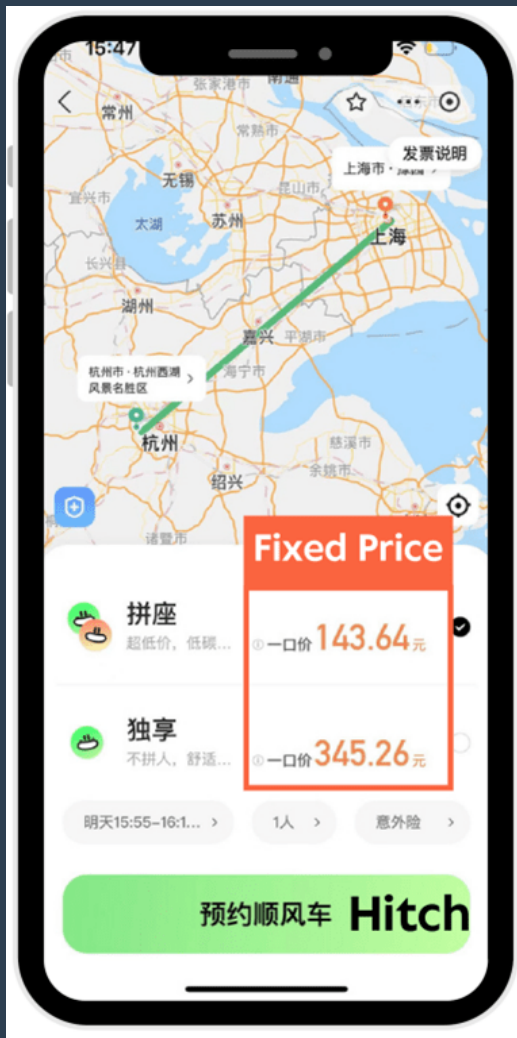
Why fixed price (一口价) for Hitch (顺风车) but dynamic price (实时计价) for Express (专车)?

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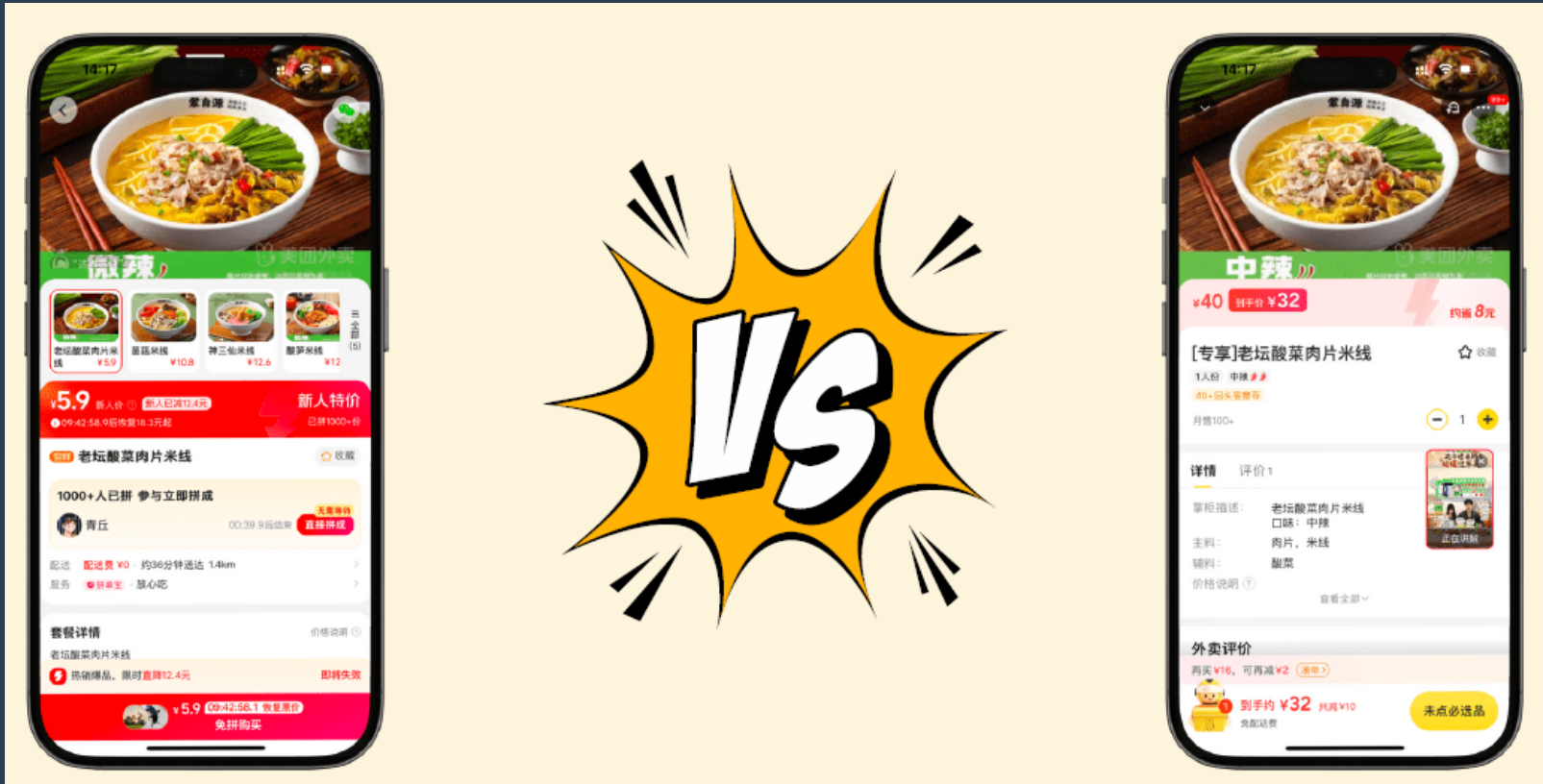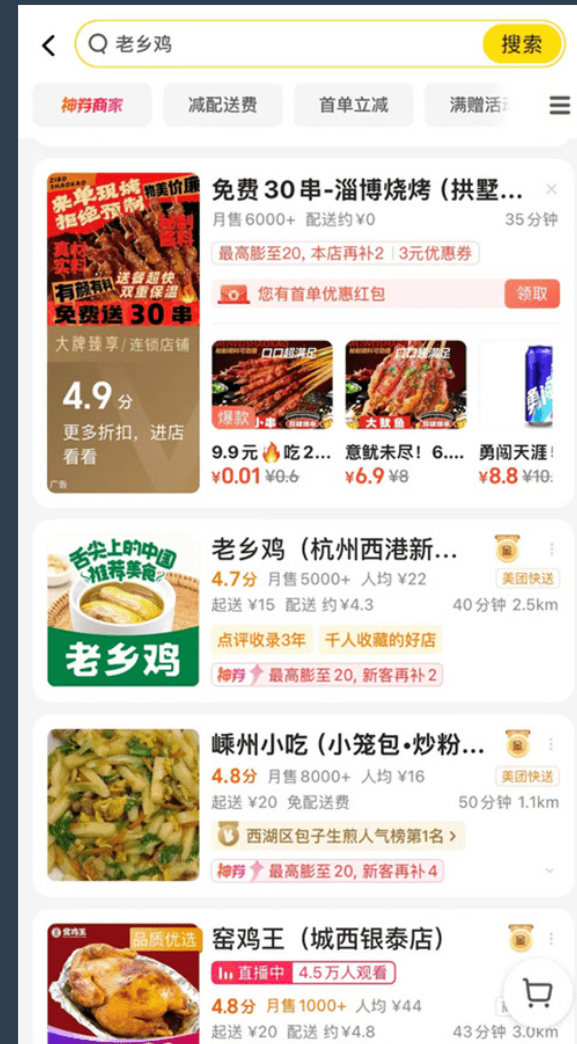
cookies

whiskey

Why does Fresh Hippo (盒马鲜生) charge lower prices offline?
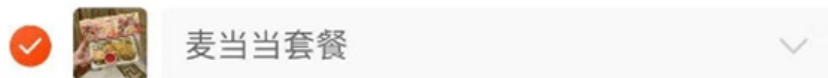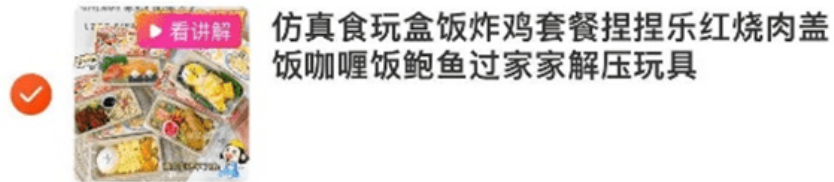
Why is the same product cheaper on Pinhaofan (拼好饭)?

Meituan ranks restaurants while Pinhaofan ranks products. Why?

Why does the same movie receive so different ratings on Douban and Maoyan?

义乌市七界贸易有限公司 >

仿真食玩盒饭炸鸡套餐捏捏乐红烧肉盖饭咖喱饭鲍鱼过家家解压玩具

▶看讲解

麦当当套餐

¥8.00   − 1 +

芒果糯米饭
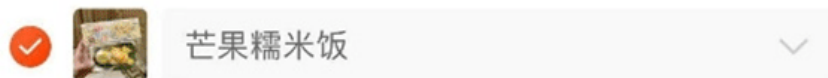
¥8.00   − 1 +

有2种商品因不符合商家混批下单要求，未被勾选

AI帮你挑

全选   优惠:¥5.00 共计:¥14.80   结算
2种2件 | 含运费 金额明细 ∧

新   年   齐⁵   奔²   富
首页   会员店   消息   采购车   我的

On 1688, shipping cost is listed on product page.

On Taobao, shipping cost is listed on the checkout page only.

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. This is out of my control (I don't care about your attendance myself).

Individual Participation

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

- The number of questions answered,
- The difficulty of that question, and
- The quality of your answer.

The TA takes notes whenever you answer questions in class. If you believe your participation was overlooked, please check with the TA right away—ideally during the next class break.

Individual Project

Choose one of the following three 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).

(C) Use AI to generate some content (tentative, TBC).

No presentations. Submit slides for (A), report for (B) or your output and prompts for (C).

Anything unclear? Ask me now.

Overview

Overview

We have 10 lectures, including

- Today's Introduction
- One half-class data workshop and one guest lecture (to be confirmed)
- Eight regular lectures

Overview

The topics to be covered include

- Data visualization with Tableau (fine for beginners)
- Programming with R (with the help of AI; 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

- Large language models: prompt engineering and RAG
- AI Agents
- Causality: Does X really cause Y?
- 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 new AI contents this year.
- We add AI components to classical content.
- The slides are all updated.

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

TODO List

TODO List #1

We have group assignments, so please form groups with your classmates. Deadline: Dec 13, 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 #1

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 #2

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

- Download R here
- Download RStudio here
- Download Tableu Public 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！

TODO List #3

Join Github Copilot (Free Education Version). You need to sign up for a Github copilot account.

Follow the instrcutions here to add Copilot to your RStudio. Thanks to our TA for writting the detailed instructions.

TODO List #4

You probably need a VPN to visit the official ChatGPT website [the HKU version is problematic]. You can choose any VPN (many provide a free trial period).

If you cannot visit the official ChatGPT website, you may use other tools such as Deepseek and Gemini instead, but the quality of the answers varies a lot.