

# Business Intelligence Techniques and Applications

---

## Session 1a. Introduction

---

Renyu (Philip) Zhang

# About Me



张任宇

2007年北京大学

2008年

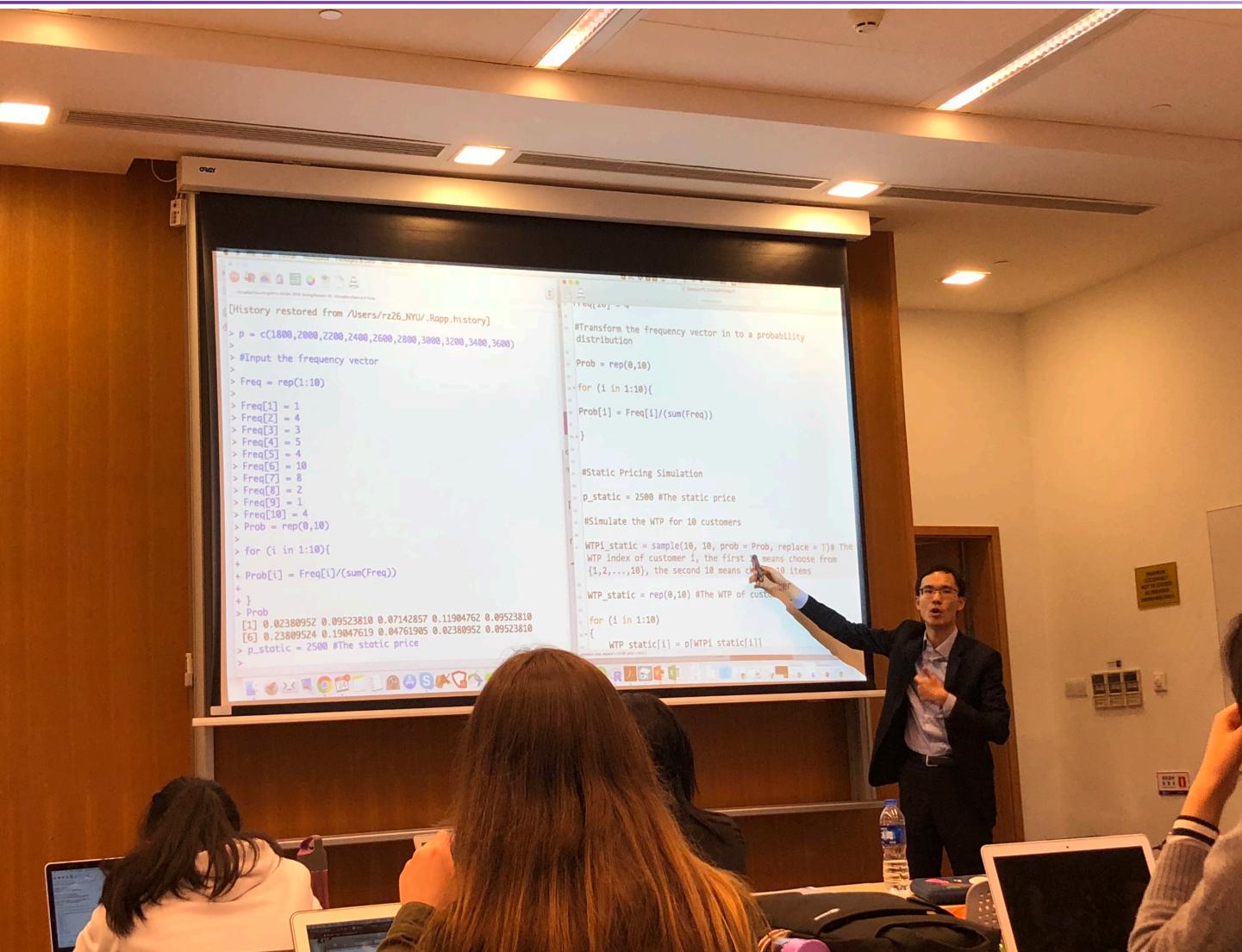
# About Me



# About Me



# About Me



# About Me



# Who Am I?

- I am a scholar, a teacher, and a practitioner in data science/AI and operations research.
- Research:
  - How to use data analytics to improve business decision making, especially for digitalized online platforms.
- Teaching:
  - Data science/AI for business to undergraduate, master, EMBA and PhD students.
  - CUHK Business School, Associate Professor (with tenure), since 2022
  - NYU Shanghai, Assistant Professor, 2016-2022; Visiting Scholar, since 2022
  - Washington University in St. Louis, PhD, 2011-2016
  - Peking University, BS, 2007-2011
- Data Science Practitioner:
  - Economist and Tech Lead, Kuaishou (快手; <https://www.kwai.com/>).
  - Evaluating and optimizing the ecosystem of Kuaishou.



# Who Are You?

---



# What is Business Analytics?



# What is Business Analytics?

---

"Big data is like teenage sex; everyone talks about it, no one really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it."

-Dan Ariely, Duke University



# What is Business Analytics (in this Course)?

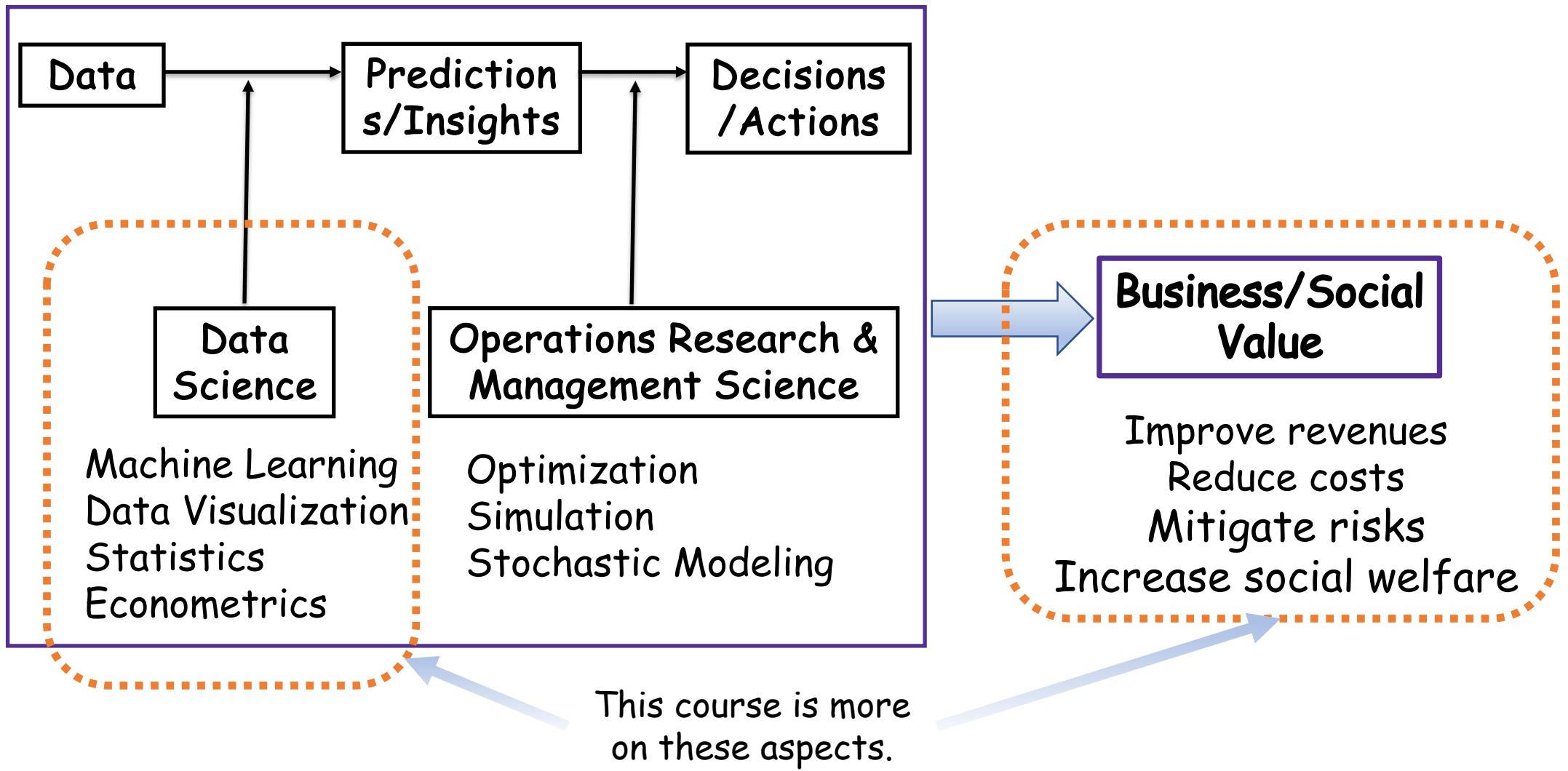
---

- Exciting applications of analytics:
  - Autonomous driving
  - AlphaGo
  - Image and sound recognition
  - Autonomous translation
  - Recommender system
  - Online advertising
  - And many more

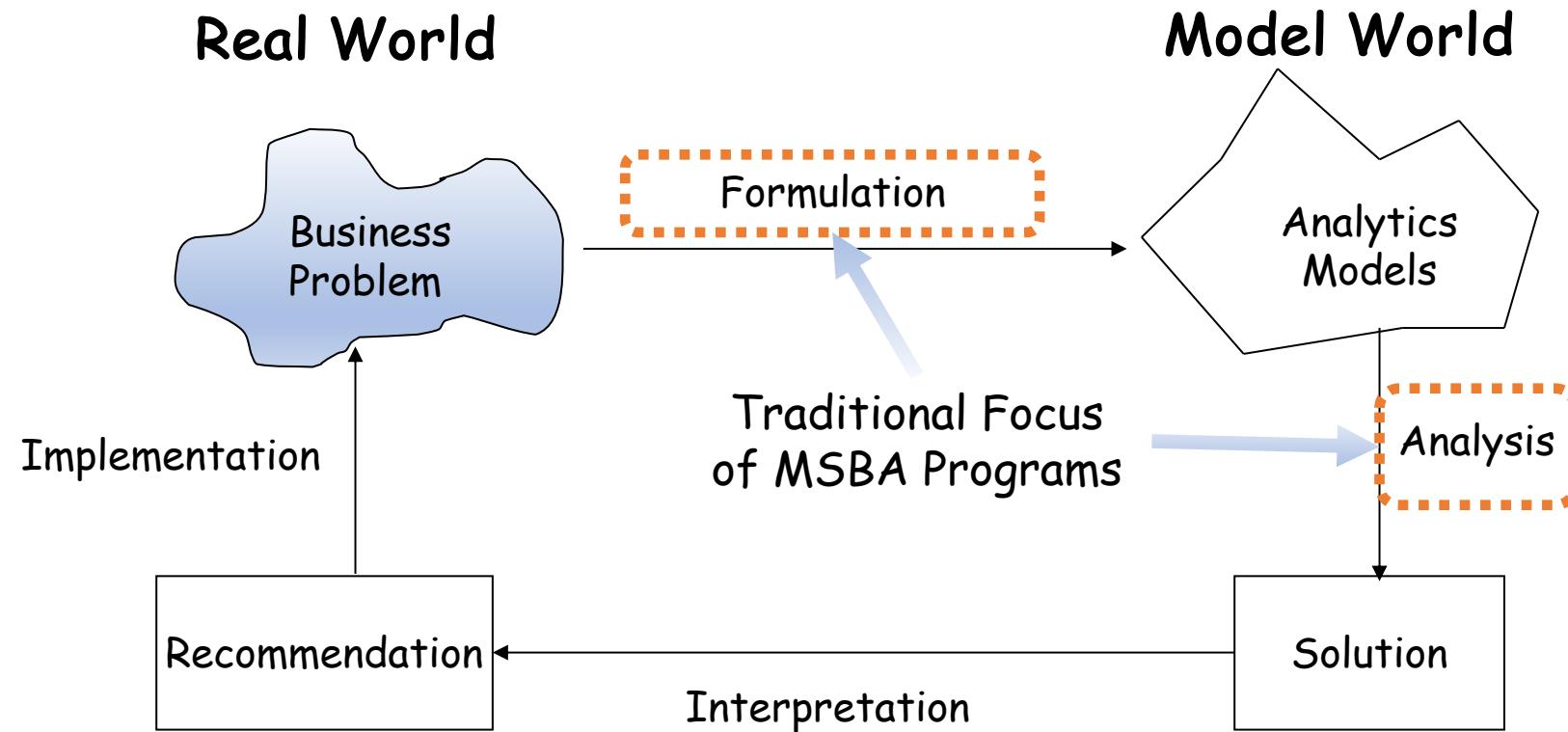
The **automated** scientific process of transforming **data** into **insights** for making better **decisions** and adding **values** to individuals, companies, and the society.

[https://www.youtube.com/watch?v=5adE\\_cxtSbY](https://www.youtube.com/watch?v=5adE_cxtSbY)

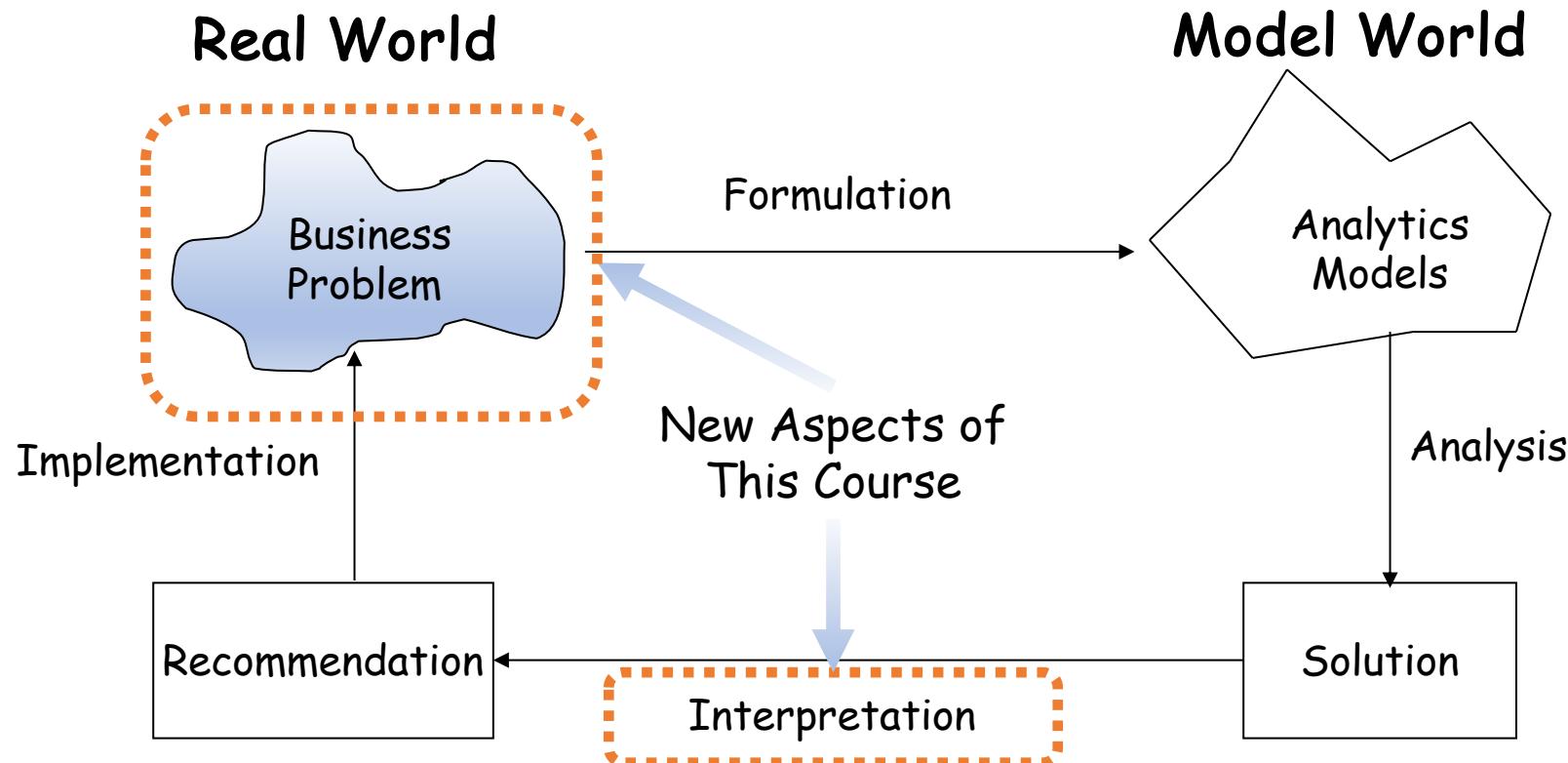
# Data-Driven Decision Making



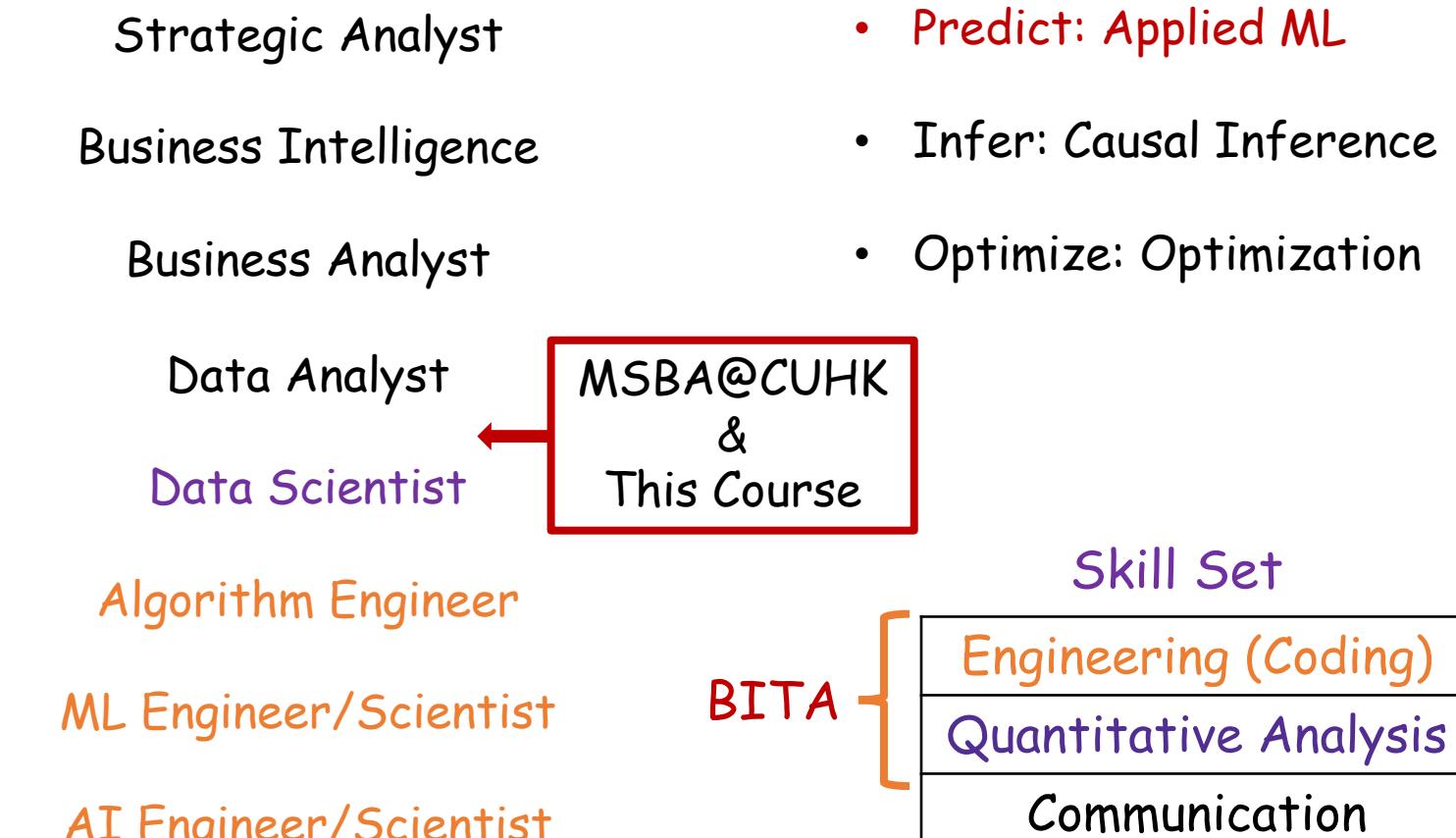
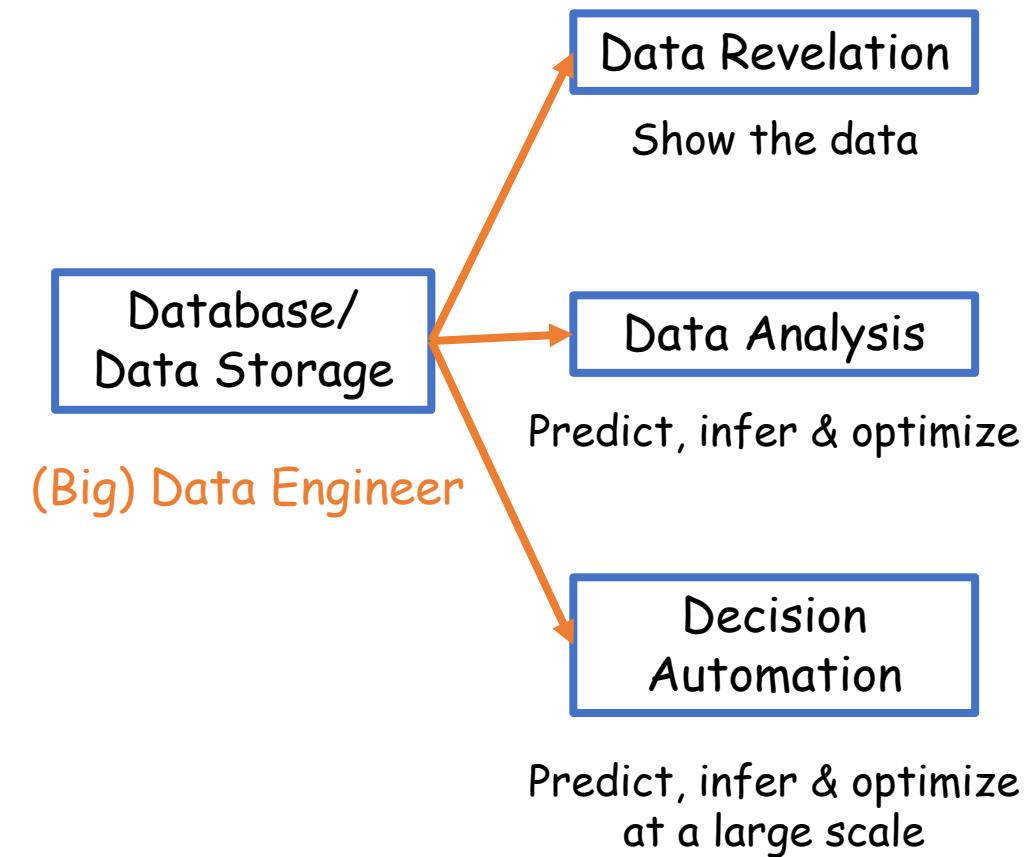
# Business Analytics Procedure



# Business Analytics Procedure



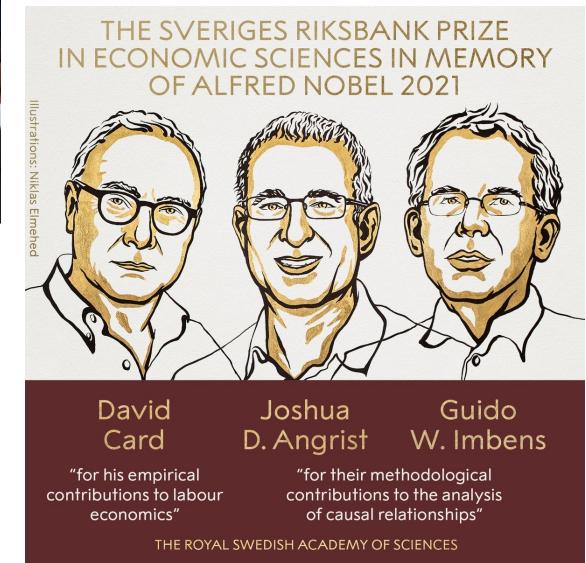
# Full Spectrum of Data Analytics Jobs



This course: Mostly **harmless** and hopefully **useful** data analytics.

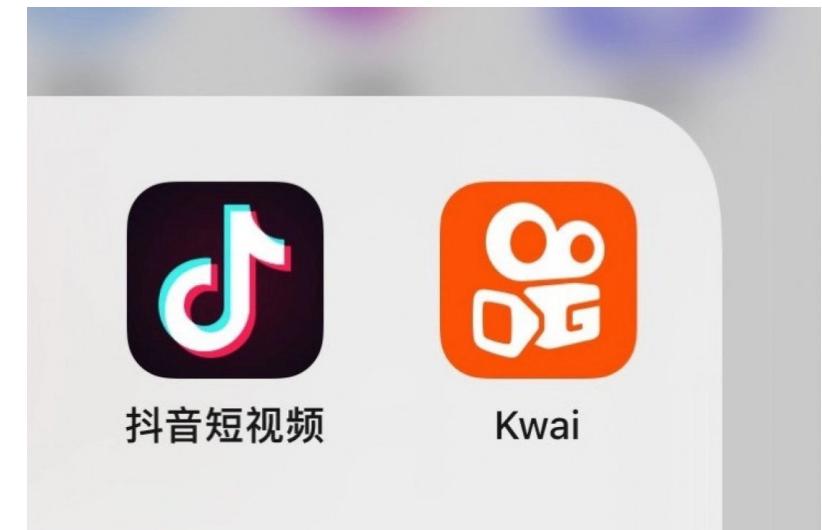
# (Causal) Inference

- Inference (mostly causal inference): To establish the **cause-effect** relationship between an intervention and an **outcome**.
- What is the social welfare impact of increasing the minimum wage?



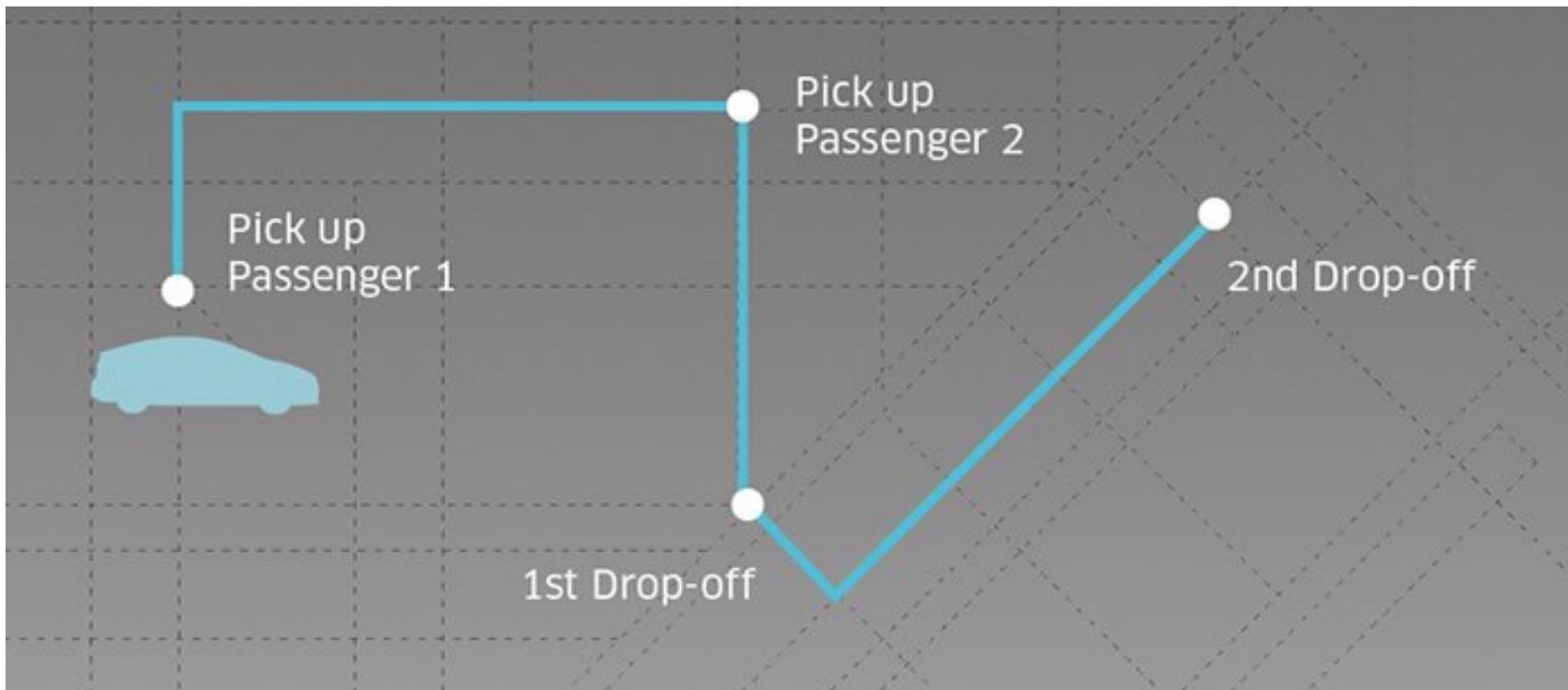
# Prediction

- **Prediction (mostly machine learning):** To predict an important unobserved outcome using features.
- What is the probability for you to click/like/follow/share a video if it is shown to you?



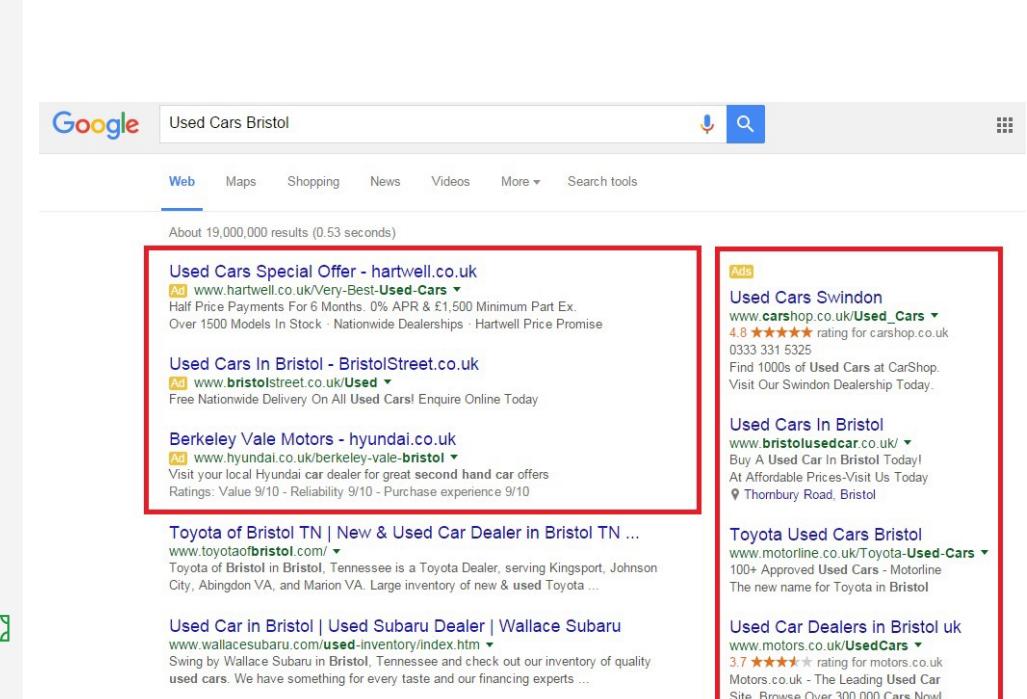
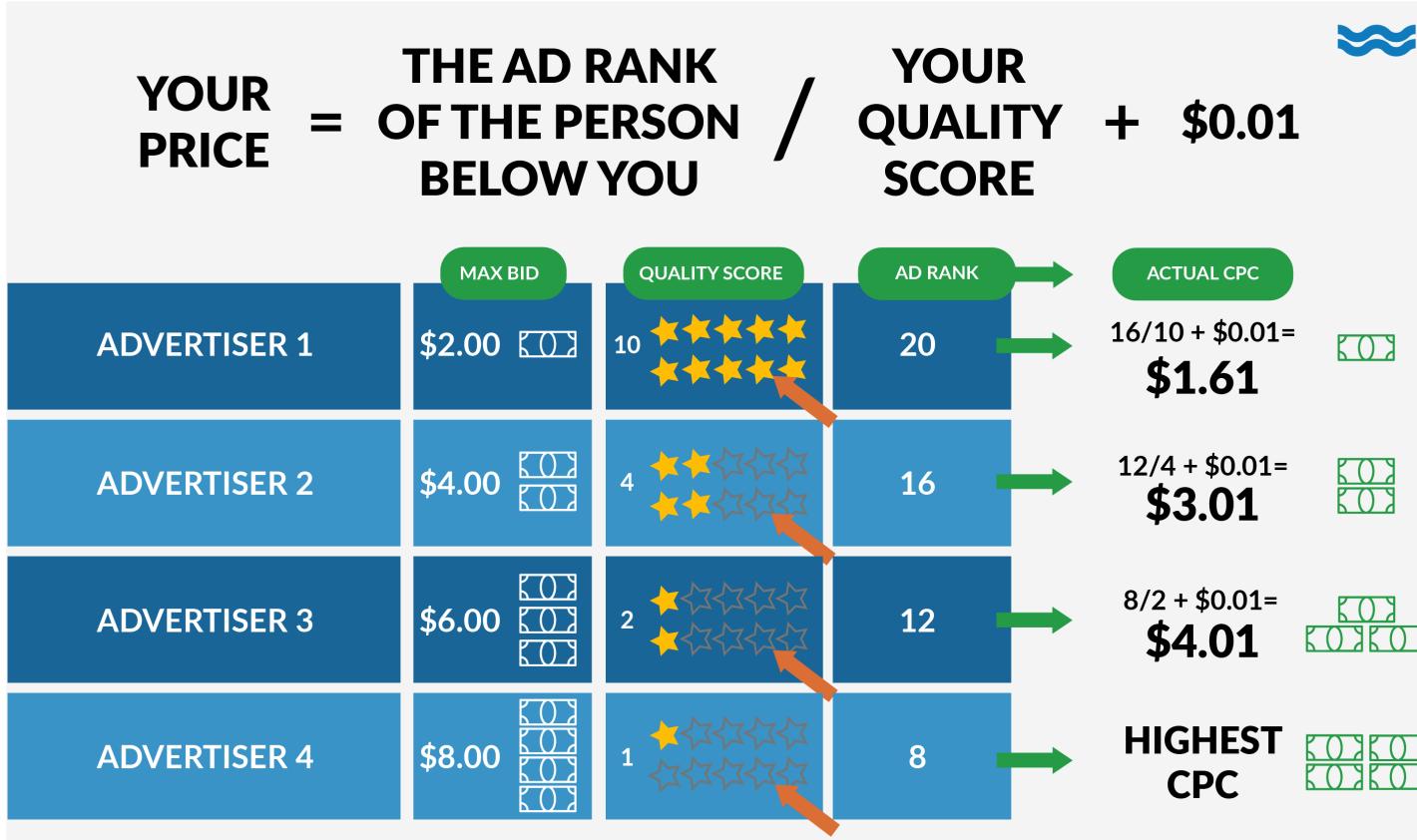
# Optimization

- **Data-Driven Optimization:** To directly control the system and improve the objective.
- How should Uber match the drivers with the riders and route the trips to minimize service time?



# Mechanism Design

- **Data-Driven Mechanism Design:** To design the mechanisms of a **marketplace** so that the participants would **automatically** play the **desired strategies**.
- How should Google design the auction mechanisms to sell its ad spaces?



# Data Analytics Classified by Expertise

---

- **Inference (mostly causal inference):**
  - What is the value of reducing the quarantine time from 4+3 to 0+3?
- **Prediction (mostly machine learning):**
  - Given a consumer query at Amazon, what is the most suitable search result?
- **Data-Driven Optimization:**
  - Given a car and 10 restaurants to deliver food, what is the optimal route for Food Panda?
- **Data-Driven Mechanism Design:**
  - How should YouTube pay its content creators on different ads?

# Analytics Covered in this Course

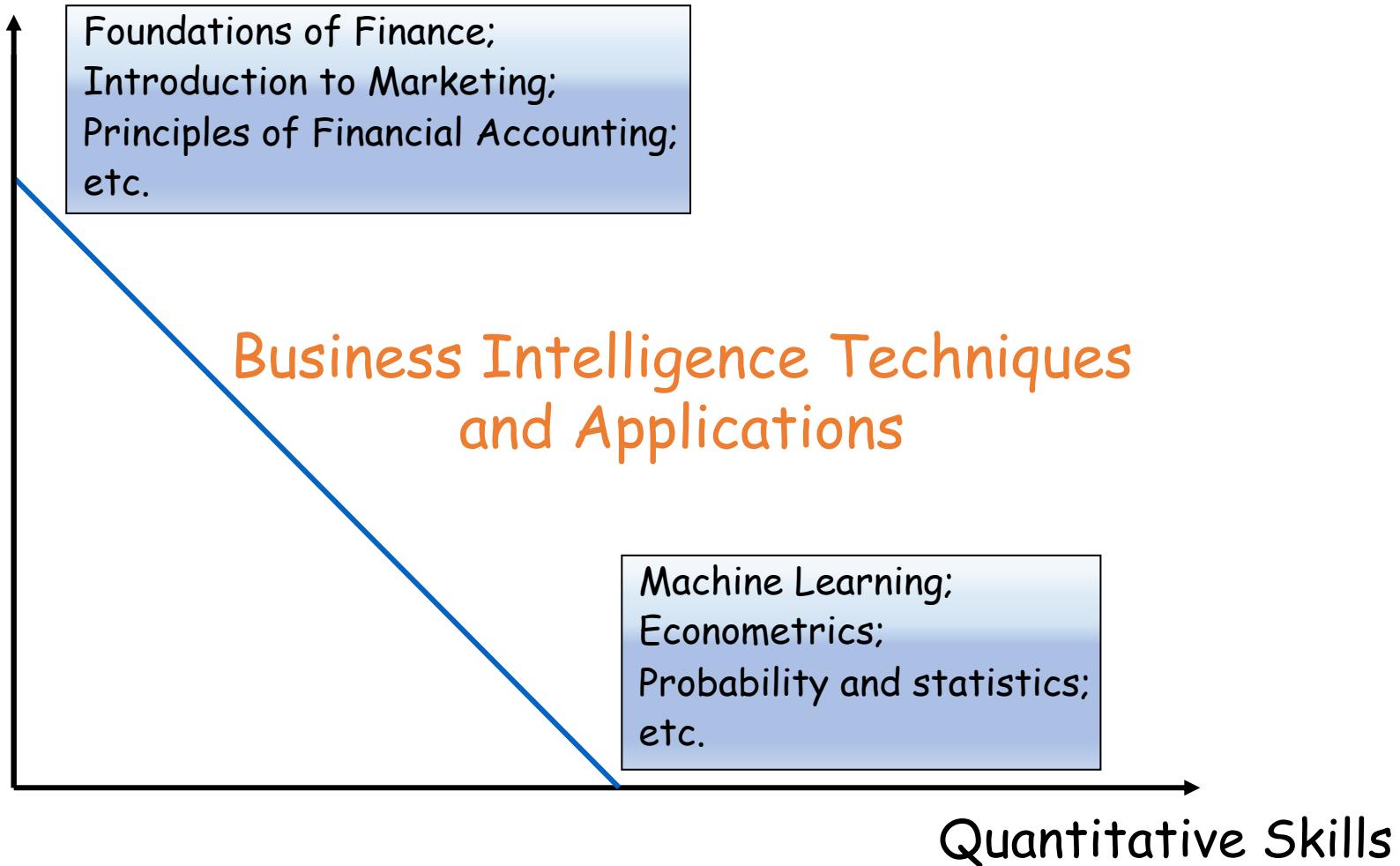
---

- **Prediction**
  - Predicts *future* probabilities and trends
  - Finds relationships in data that may not be readily apparent with descriptive analysis
  - Tools: Machine learning, statistics, etc.
- **Other interesting (and important) stuffs**
  - Coding (copiloted by AI) basics
  - Data (pre-)processing
  - Generative Artificial Intelligence

# This Course

---

Business Sense



# Course Objective

---

- Our goal: Convince you of the tremendous business (and social) value of analytics and further inspire you to use it in your career and life.
  - At the end of this course, hopefully, you will
    - Have an open mind about data;
    - Be ready to be convinced by data and quantitative analysis;
    - Be ready to solve a real problem using data and analytics tools (including the AI-powered copilots for coding);
    - Be well-prepared to study more advanced analytics courses.
  - This is not a math course, not a data science course, not a computer science course, and not even a business course, but an inter-disciplinary course that bridges business applications and analytics methodologies/toolboxes.
  - Connection to real business and job referral opportunities.
-

# Course Content

---

- **Module 0:** Python and data analysis basics
  - Python coding, descriptive analysis, data visualization
- **Module 1:** Prediction with Machine Learning
  - Supervised learning, unsupervised learning, reinforcement learning, generative AI

# Course Prerequisites

---

- No prerequisites, but some knowledge of statistics and coding will be useful.
- Highschool math is also assumed.
- Not adverse to programming (everything implemented in Python).
- Not adverse to analytical thinking and quantitative analysis in general.

# Learning by Doing

---

- Lectures and demonstrations (recordings available on GitHub)
- Extensive cases and data: End-to-end analysis and **problem-solving**
- In-class discussions
- Weekly problem sets and 2 projects
- Laptop
  - Bring a laptop to every class. **Close your laptop until you are asked to use it.**
  - Install the required applications (Python and Anaconda).
  - Download Jupyter Notebooks (from GitHub) to your laptop before each class.
- Attendance is required.

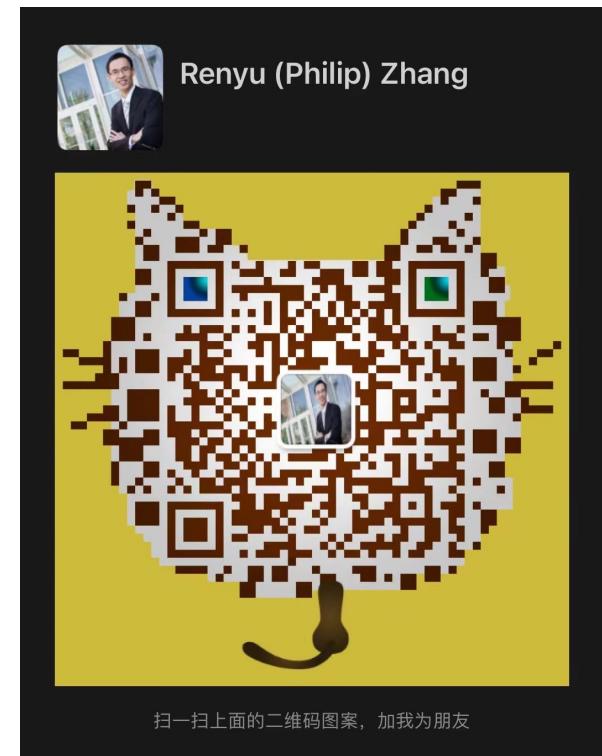
# Course Materials

---

- Blackboard
  - Homework Submission
- GitHub:
  - <https://github.com/DSME6756-2023/BA-W2023>
- Anonymous Survey:
  - [Link to the survey.](#)
  - You are more than welcome to submit any feedback throughout this course.
- No required text books
  - Reference books given in the syllabus and provided at GitHub.

# Course Communications

- **Class Meeting:** Monday, 9:30AM-12:15PM (@WMY\_508) or 2:30PM-5:15PM (@YIA\_LT6)
  - You can choose to go to either Section.
- **Office hour:** Monday, 1:00PM-2:00PM, @CYT\_911, or by appointment
- **WeChat group:** Online discussion forum.
- **Instructor contact**
  - Office: CYT\_911
  - Email: [philipzhang@cuhk.edu.hk](mailto:philipzhang@cuhk.edu.hk)
  - Tel: 852-3943-7763
  - WeChat: rphilip\_zhang
- **Teaching Assistant:** Qinlu Hu
  - Office hour: By appointment.
  - Email: [qinlu.hu@link.cuhk.edu.hk](mailto:qinlu.hu@link.cuhk.edu.hk)



# Grading

---

- Class participation, 10%
  - Baseline 6%; lose 2% for each class missed; failing grades for missing 4 or more classes
- Problem sets,  $30\% = 6\% * 5$ 
  - Due every Monday at 9:30AM before class
  - 5 problem sets with the highest scores will count
- Projects,  $20\% = 10\% * 2$ 
  - Project 1: A Kaggle competition, due on Monday, February 26
  - Project 2: TBD, due on Sunday, March 3
- Final Exam, 40%
  - 2:00PM-5:00PM, Saturday, March 2
  - Close-book, close-notes, electronic devices NOT allowed

# Grading

---

- Problem sets and projects submitted via Blackboard:
  - Everyone should **individually** finish and submit his/her **own solutions**, though discussions with others are allowed.
- Regrading:
  - Submit your requests within **7 calendar days** after receiving your grade.
- “Zero-tolerance” policy
  - Any violation of academic integrity is strictly prohibited and will be treated seriously.

# Who Will Find This Course a Nightmare?

---

- You want an easy A.
- You hate the quantitative/analytical way of thinking and solving problems.
- You hate coding/programming.
- You hate me.

Otherwise, you are very much welcome joining me to enjoy the excitements and challenges of **Business Analytics**!

# Data Scientist vs. Alchemist

