Course Overview

MSBA7001 Business Intelligence and Analytics

HKU Business School
The University of Hong Kong

Instructor: Dr. DING Chao

About Me

- Dr. DING Chao (丁超)
- My experiences
  - Chongqing: where I was born and raised
  - Guangzhou: where I studied physics
  - Florida, US: where I obtained my doctoral degree
  - Hong Kong: where I started my academic career
- My hobbies
  - Hiking
  - Photography
  - Traveling
  - Tennis

Agenda

- Value of Data
- Business Intelligence & Business Analytics
- Course Structure
- Survey

Value of Data

Data

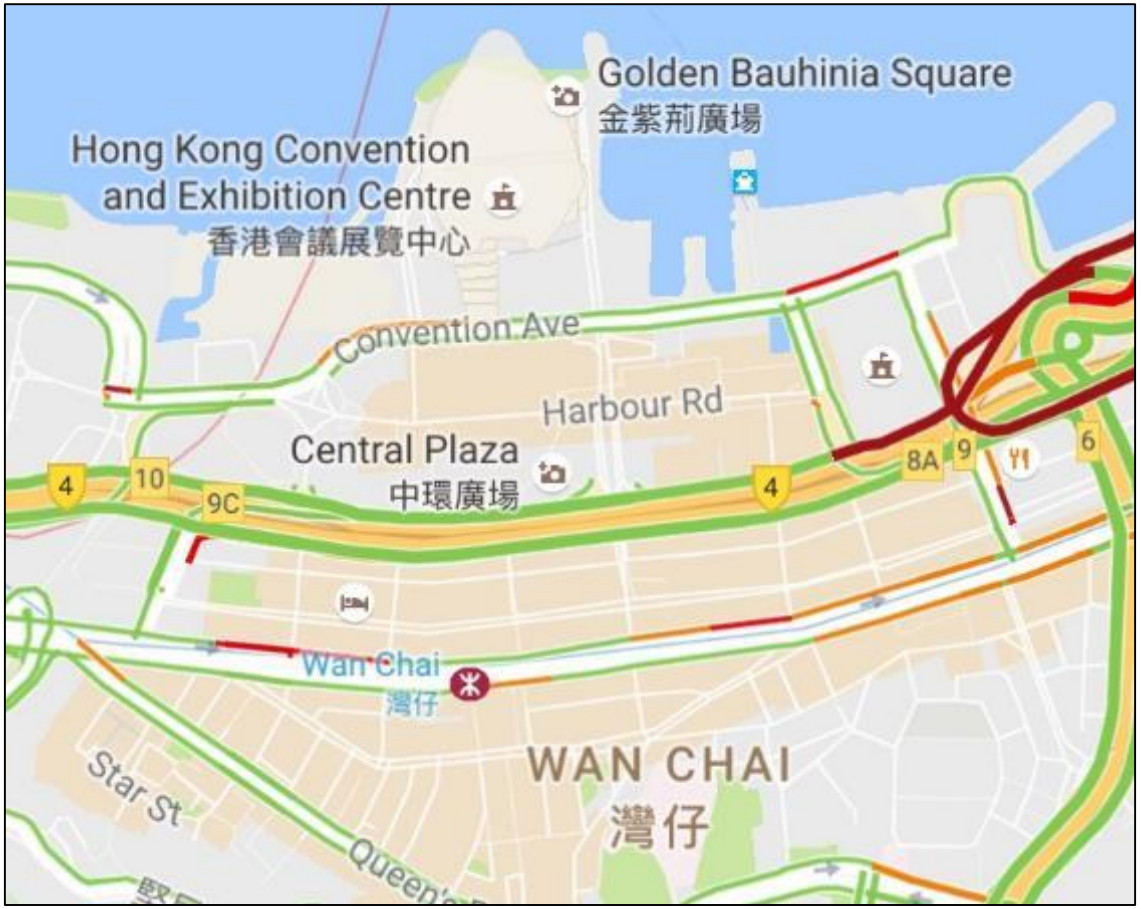
A vast amount of data generated every second

- Sales records
- Online searches
- Web visits
- Social posts
- Bank transactions
- Online conversations
- Facial ID
- Product inventory
- …

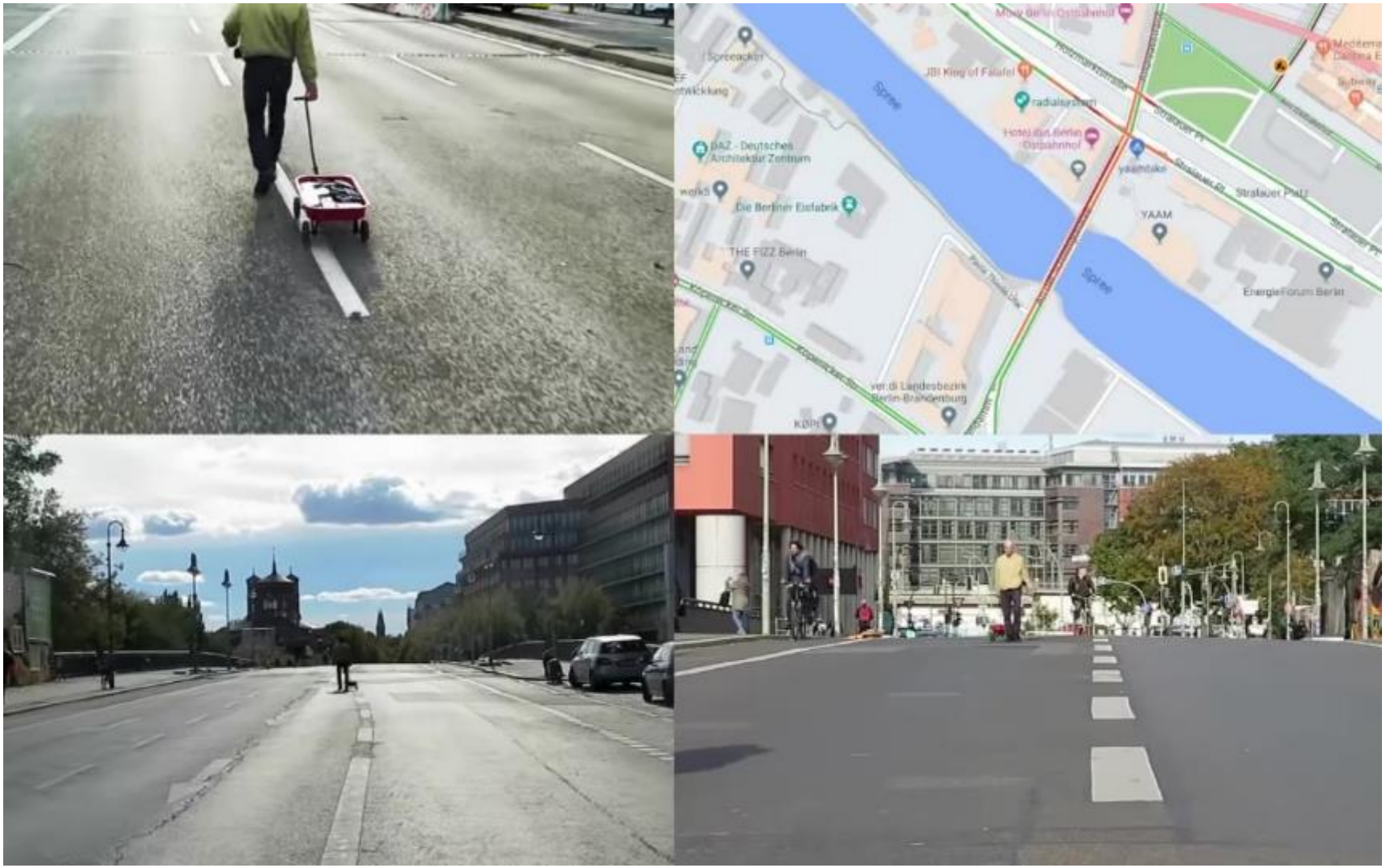
Then what??

Google Maps

How do you think Google Maps gather real-time traffic data?



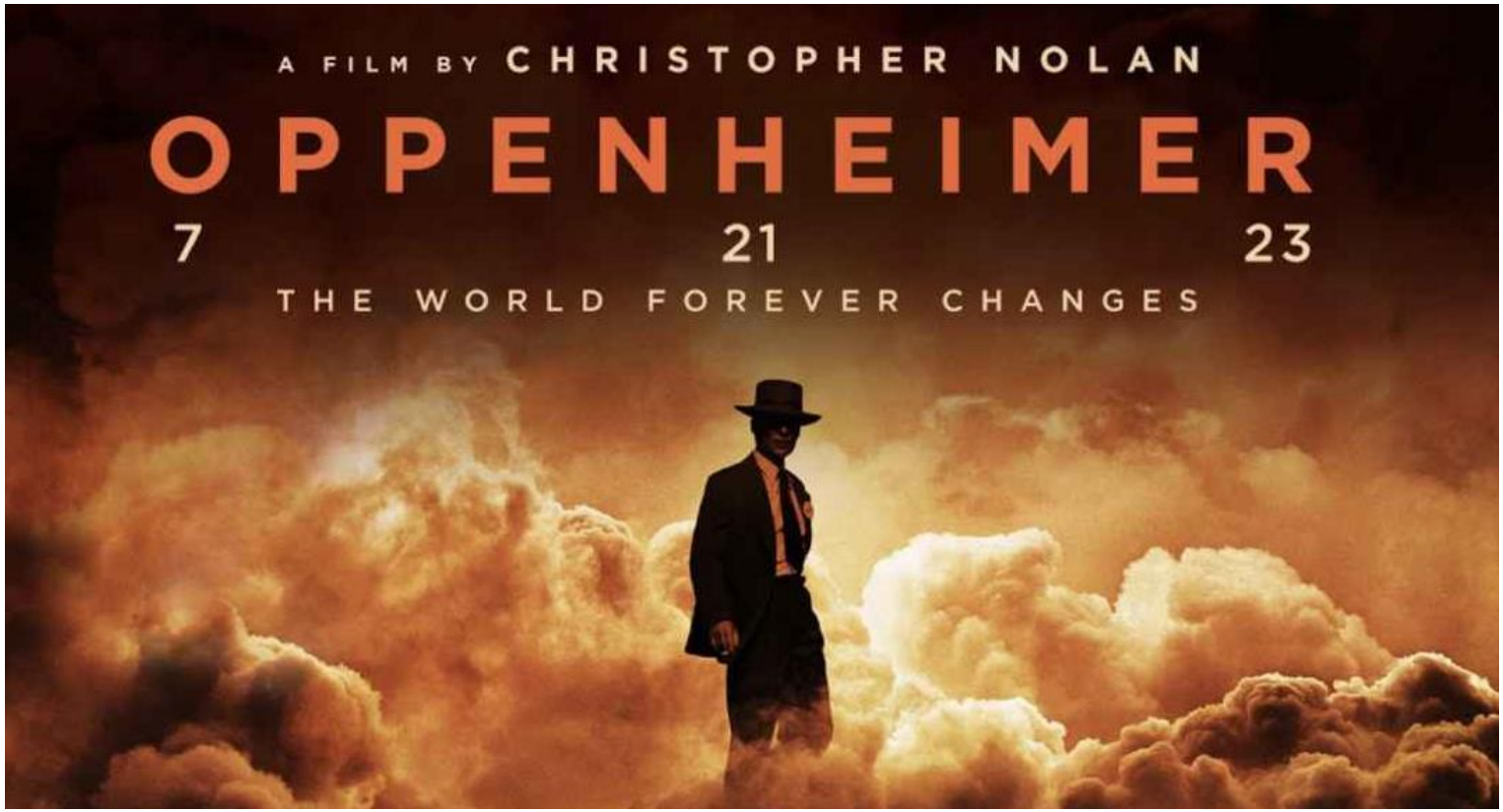
Google Maps



Source: Simon Weckert, http://www.simonweckert.com/googlemapshacks.html

Box Office

How to predict movie ticket sales?

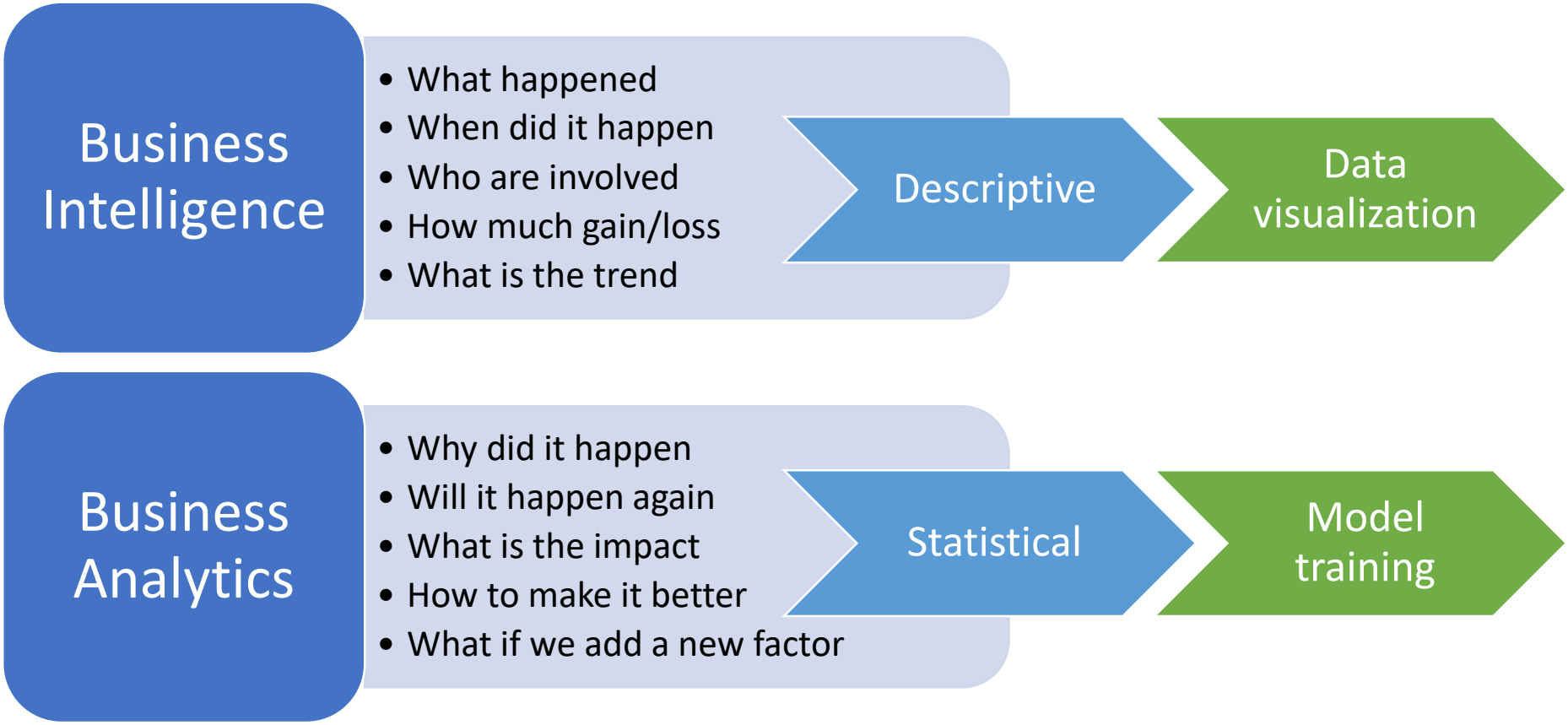


Business Intelligence& Business Analytics

Business Intelligence & Business Analytics

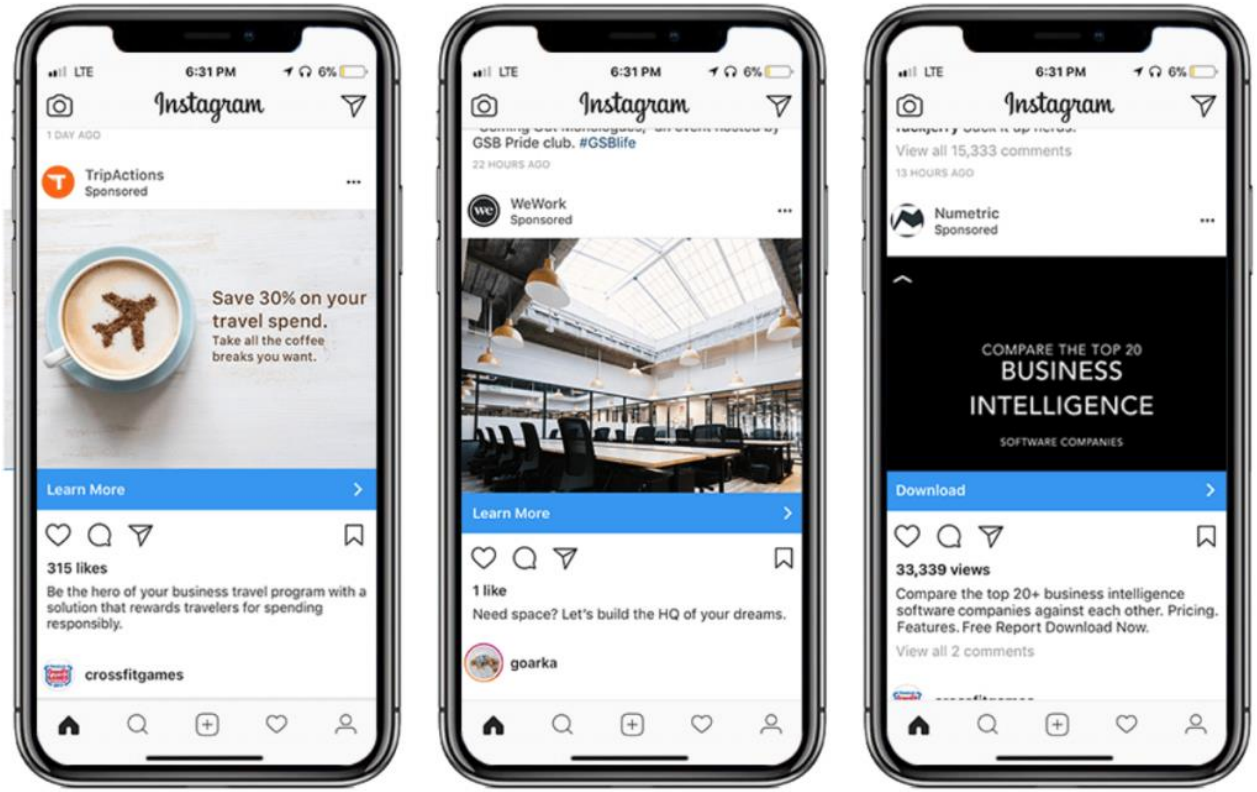
- **Business Intelligence (BI)** is used to gather, analyze and report data, so as to manage day-to-day operational management within a business.
- **Business Analytics (BA)** is used to gain insights from past business activities, to inform future business decisions, and to automate and optimize business processes.

Business Intelligence & Business Analytics



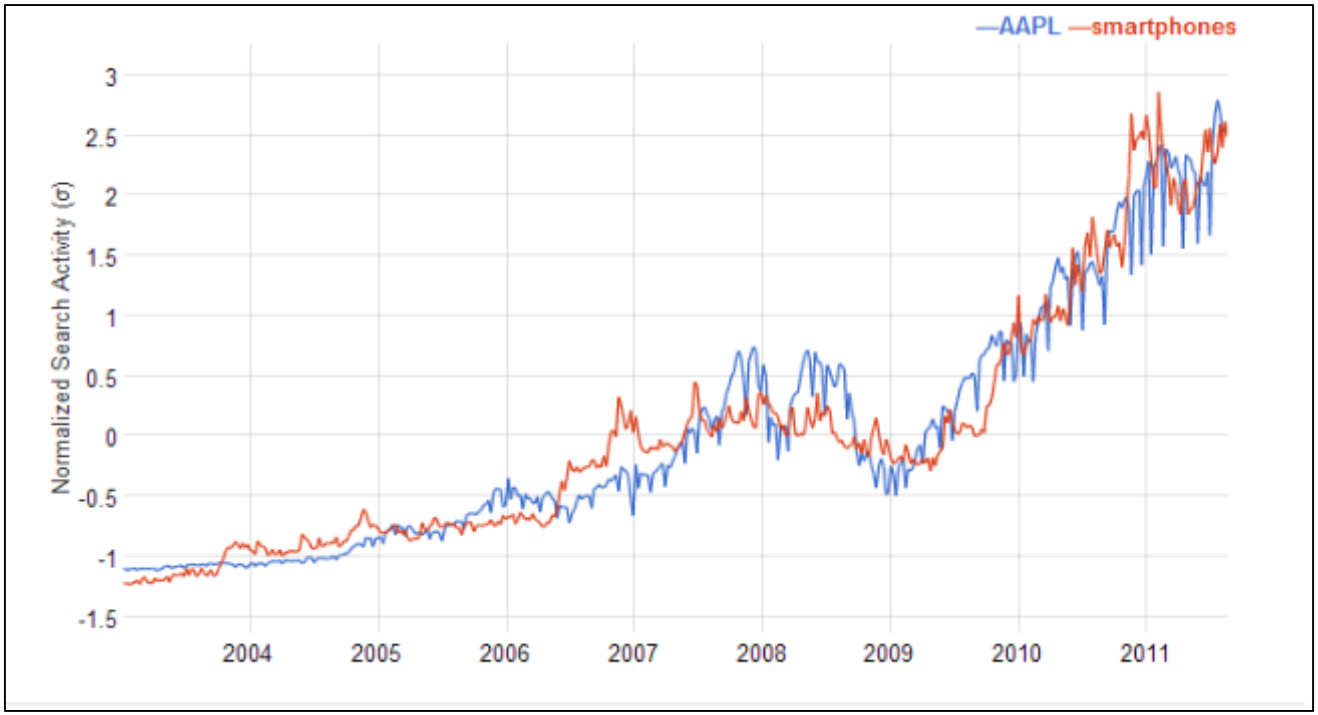
Targeted Advertisement

Use user's browsing record to present relevant advertisement.



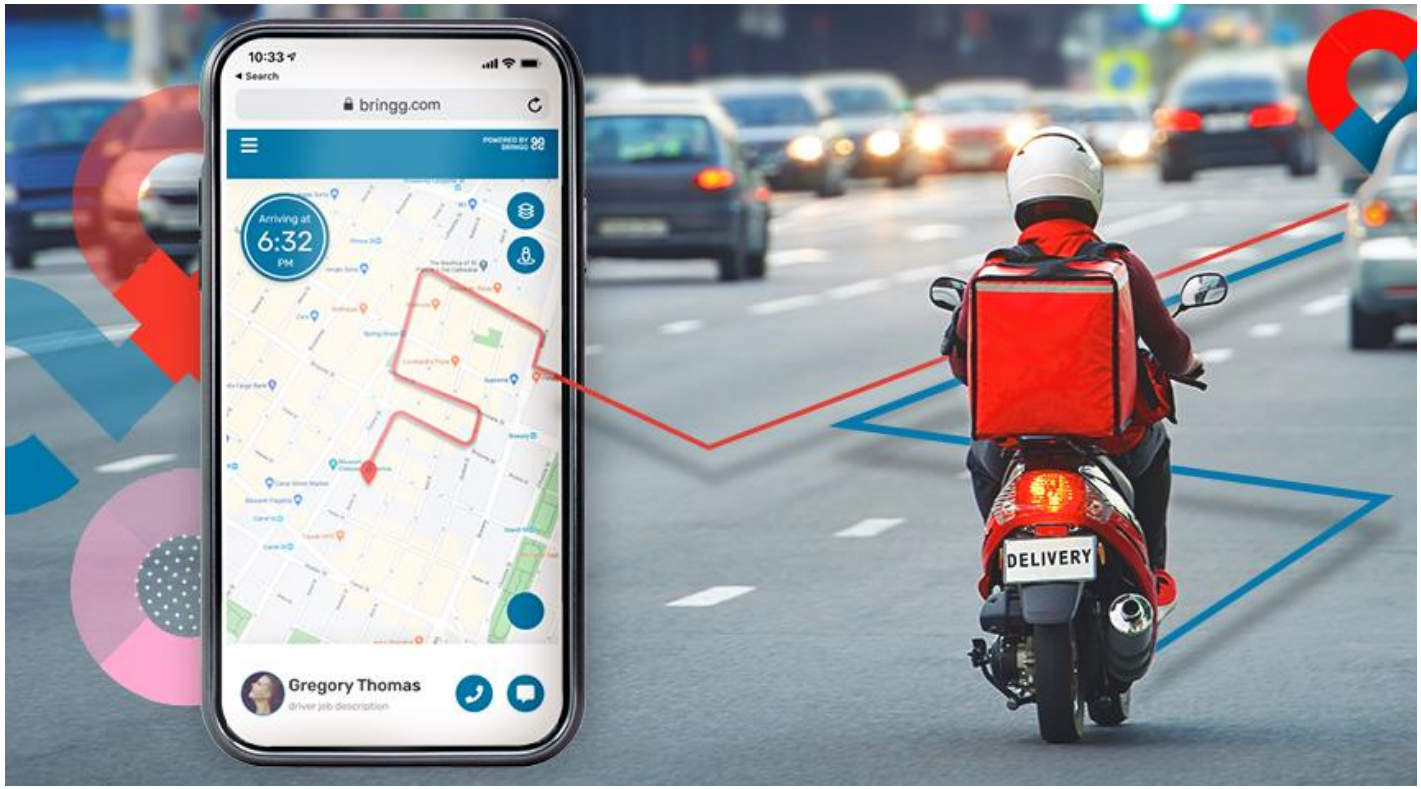
Market Trend

Use Google trend to “guess” stock price.



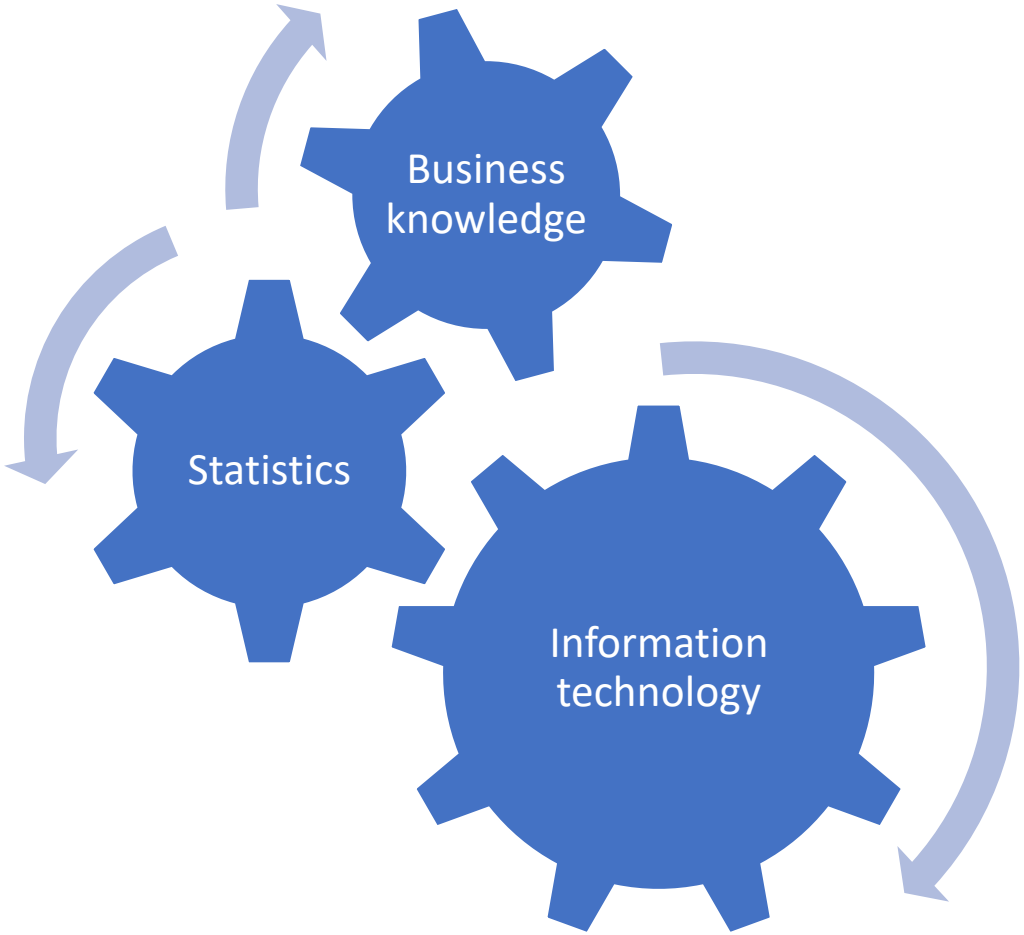
Delivery Route Optimization

Utilize traffic and order information to optimize routing.

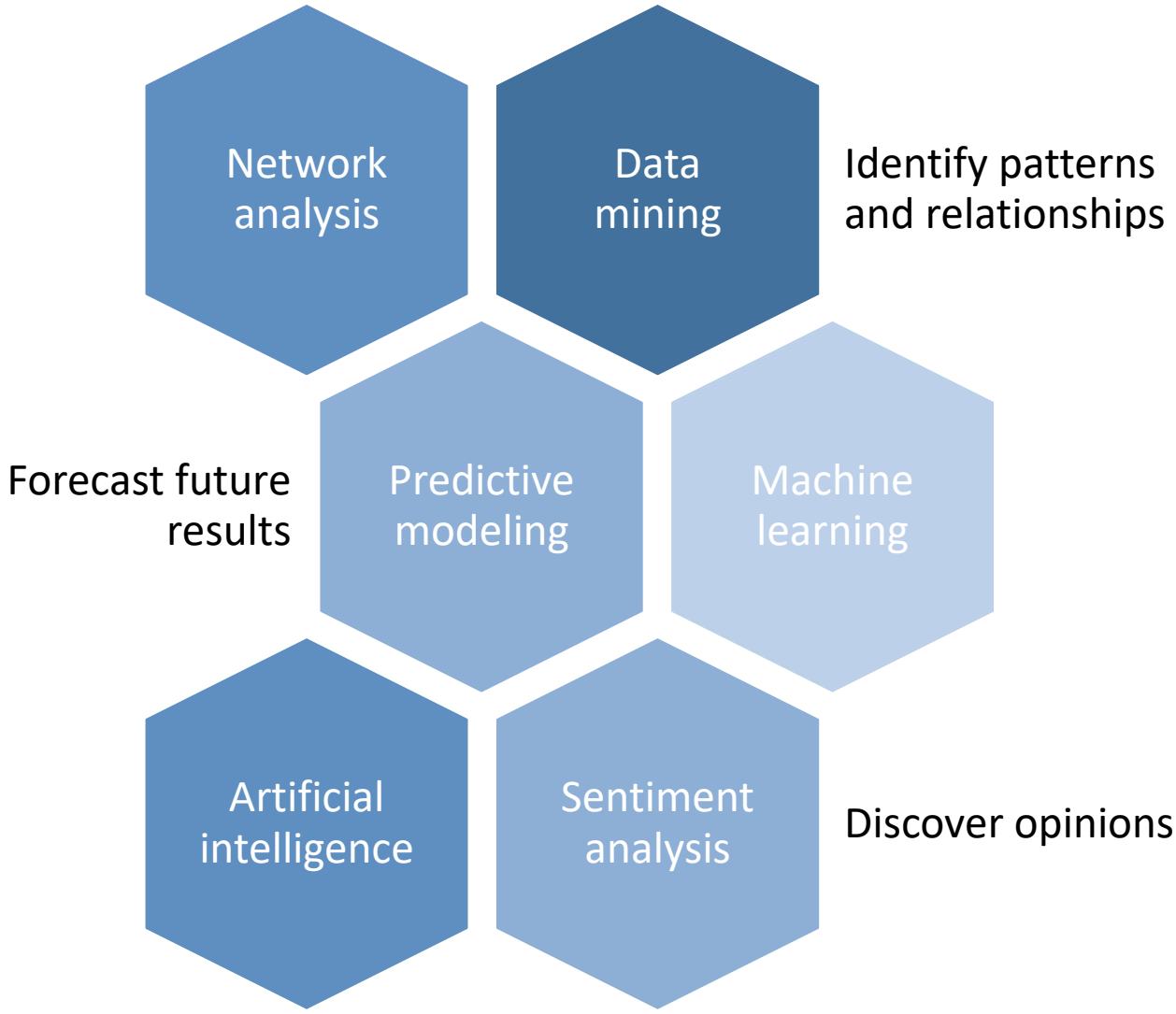


Source: BRINGG

Required Domain Knowledge



Related Fields



MSc(BA) Courses (partial)

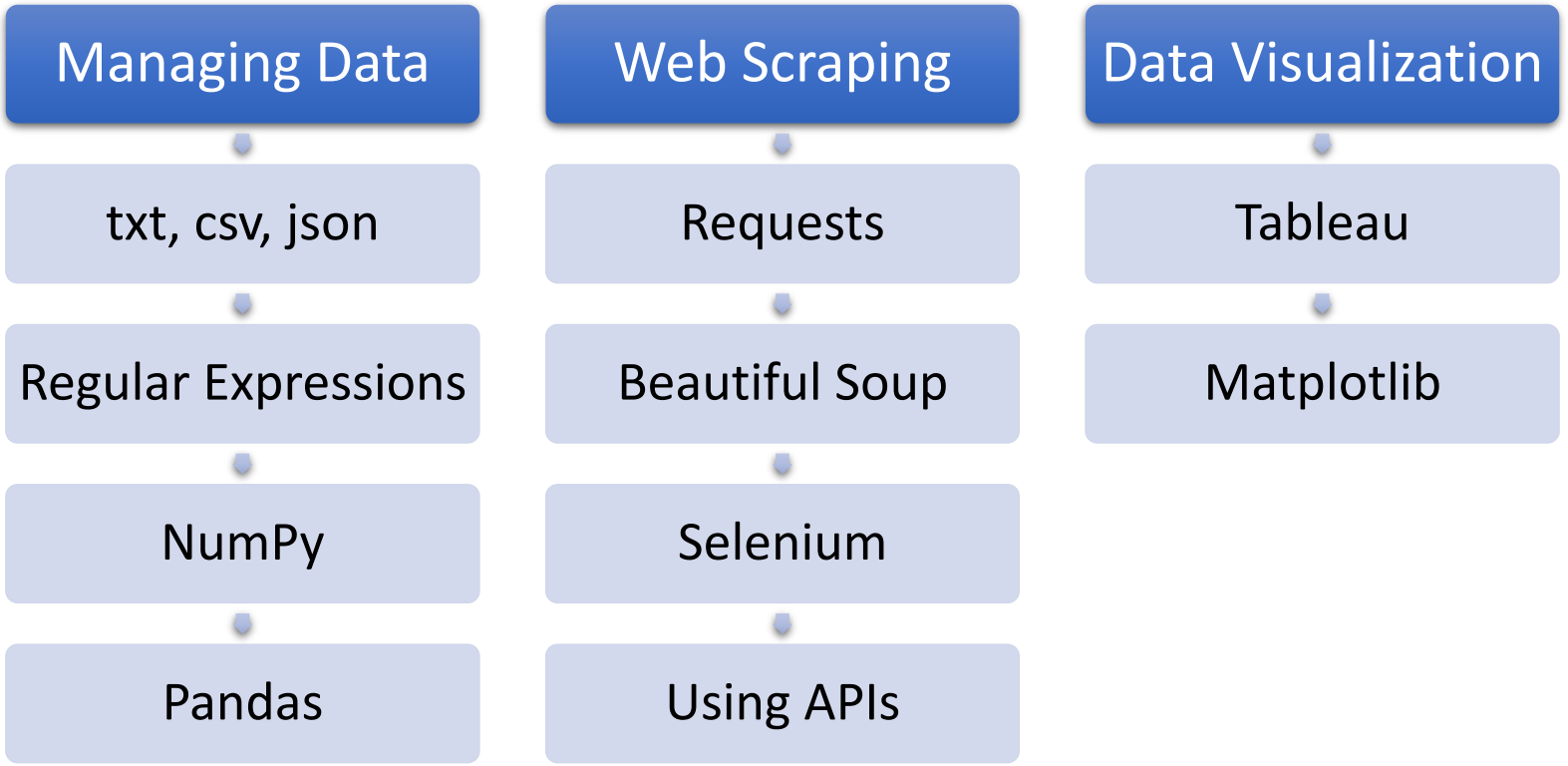
Course Title	Language
MSBA7002 Business Statistics	R
MSBA7003 Quantitative Analysis Methods	R, Python
MSBA7004 Operations Analytics	R, Python
MSBA7005 Capstone	R, SQL, Python
MSBA7012 Social Media and Digital Marketing Analytics	SQL, Python
MSBA7013 Forecasting and Predictive Analytics	R
MSBA7021 Prescriptive Analytics	Python
MSBA7023 Geospatial & Business Analytics	Python
MSBA7024 Database Design and Management	SQL, Python
MSBA7025 Digital Experimentation Methods	Python
MSBA7026 Big Data Analytics on the Cloud	SQL, Python
MSBA7027 Machine Learning	R
MSBA7028 Deep Learning	Python
MSBA7031 Blockchain Cybersecurity Risk Analytics	R, Python
MSBA7032 Quantitative Trading	Python

Course Structure

About This Course

- This course mainly covers three areas.
- Aims to train students on two essential tools.
- **Python** for model training.
- **Tableau** for data visualization.

Scope

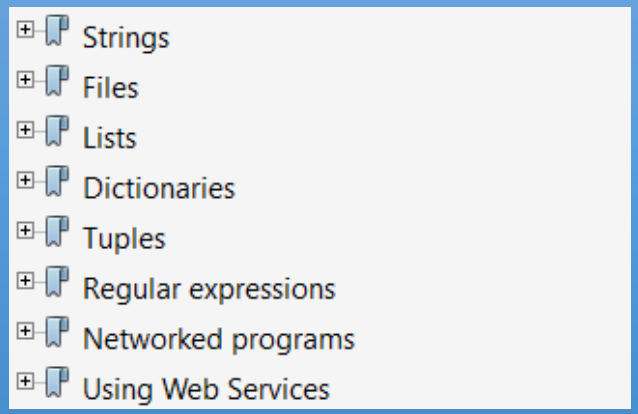

Managing Data
txt, csv, json
Regular Expressions
NumPy
Pandas
Web Scraping
Requests
Beautiful Soup
Selenium
Using APIs
Data Visualization
Tableau
Matplotlib

Required Textbooks

Python for Everybody

-- Charles R. Severance

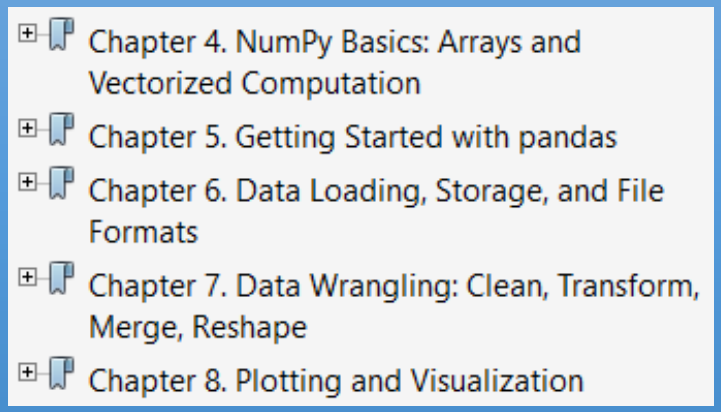
- Basic syntax and functions



Python for Data Analysis

-- Wes McKinney

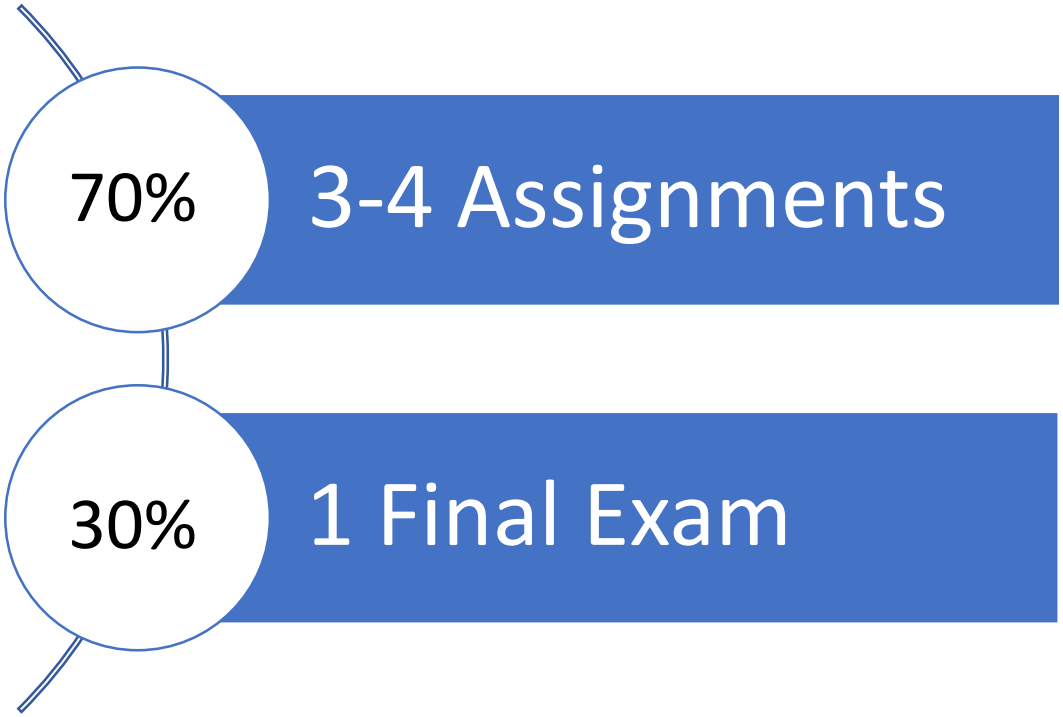
- Data analytical skills



What Should You Expect?

- Very technical
- Self-learning
- Practise in and outside classroom
- Heavy workload
- Expect to learn fast
- Will take a lot of your time after class
- Get help from your classmates

Assessments



Penalties for Late Assignment Submission

- 30% deduction for any late submission within 24 hours
- 60% deduction for one day overdue
- No grades for any submission after 48 hours.

Tips

- Obtaining the correct output is more important.
- We are not in pursuit of code efficiency or elegance.
- You will not be tested on algorithms. It's not a CS course.
- The primary task you will be repeating:
  - obtain data → clean data → save data → visualize data
- Do keep a good coding habit:
  - ✓ Divide and conquer
  - ✓ Add comments
  - ✓ Clean up testing code
  - ✓ Use autocomplete

Is there a way to fail this course?

```
if (attendance < 70%) or (grade // 10 < 4):
      return 'YES'
```

Policies on the Use of GenAI

- University policy
  - ❑Student access to ChatGPT 3.5 and Dall•E 2:
    - ➢ https://chatgpt.hku.hk/
    - ➢ https://dalle.hku.hk
  - ❑Students will have 500,000 tokens per month to access any topic they like, with up to 20 prompts per topic.
  - ❑Once all their tokens have been used, there will be no further access until the next month.
- Course policy
  - ❑Use GenAI in assignments: optional and assistive
  - ❑Use GenAI in final exam: **strictly forbidden**
  - ❑Use GenAI in other learning activities: recommended

Teaching Assistants

- There are two TAs assigned to this course:
  - Mr. Neil Cheung (neilc@hku.hk)
  - Mr. Anthony Wong (aswwong@hku.hk)

- Their primary responsibilities are:
  - Taking your attendance
  - Grading your assignments
  - Answering your inquiries
  - Handling logistics

Other Things to Note

- After each session, I will stay behind for a short while to answer your questions.
- Or, send me an email to make an appointment

chao.ding@hku.hk

- My office is KK807 on the main campus.
- All electronic materials are on Moodle.
- Your feedback is always welcomed.
- Please bring your laptop/tablet to class.
- Carnets, a free mobile app, provides Jupyter notebooks.

Optional Reference Books

- Allen Downey (2012). Think Python, O'Reily/Green Tea Press, free book at http://greenteapress.com/wp/think-python/
- Mark Lutz (2013). Learning Python, 5th Edition, O'Reily
- Joel Grus (2015). Data Science from Scratch: First Principles with Python, O'ReilyRyan
- Mitchell (2015). Web Scraping with Python: Collecting Data from the Modern Web 1st Edition, O'Reily
- Jake Vanderplas 2016. Python Data Science Handbook, O'Reily
- Matthew A. Russell (2013). Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More, 2nd Edition, O'Reily

Survey