

# Big Data Analysis

Lecture 10

2019/12/19

# Review Session

---

- Databases
- Hadoop
- Linear regression with regularization
- Logistic regression
- Gradient descent methods

# About Final Projects

---

- A good project should include
  - Motivation – justify which big data related issues to tackle
  - Data processing
  - Data analytics
  - Conclusion
- No more than 5 people in a group, total 7 or 8 groups
- Project presentations to be held on 1/9 and 1/13
- Each group has 25 minutes (20 min presentation + 5 min Q&A)
- Attendance will be checked

# Data Resources – News Information

---

- Datago: (only accessible from campus)
  - Username: phbs
  - Password: phbs\_datago
  - Address: <https://datago.com.hk/auth/login>
- ChinaScope:
  - To be announced