



Data Mining

資料探勘

Project 1

Hung-Yu Kao, Fall 2017

Association Analysis

2

- Dataset1: Select from kaggle.com
- Dataset2: Use IBM Quest Synthetic Data Generator
 - ▣ <https://sourceforge.net/projects/ibmquestdatagen/>
 - ▣ Generate different datasets
- Implement Apriori Algorithm and apply on these datasets
 - ▣ Hash? Tree?
 - ▣ FP-growth
- Use association analysis tools (e.g., WEKA) to generate association rules from the datasets you generate
- **Compare** your results

- Due: 11/7 9am