

資料分析與學習基石

(Fundamental of Data Analytics and Learning)

Homework 1 – First visit in Kaggle data

Hung-Yu Kao (高宏宇)

Intelligent Knowledge Management Lab



Master Program of Artificial Intelligence

Institute of Medical Informatics,

Dept. of Computer Science and Information Engineering

National Cheng Kung University, Tainan, Taiwan



Goal

- Survey a dataset in Kaggle
 - Study the data
 - What information is released?
 - What information you can find and mine from it?
 - Any kind of integration between datasets? (if you survey multiple datasets)
 - Study the code
 - Notebook (Kernel) survey
 - Comparison
 - Why? How?



Report & Schedule

- Dataset description: 資料集與資料欄位描述
- Data analysis
 - 現有分析方法與產出
 - 資料分析價值與可能產出
- Scoring
 - 基本資料集描述, 資料特性描述 (35%)
 - Notebook方法介紹與比較 (35%)
 - **Your insight** (20%)
 - 資料特殊性 (10%)
- Schedule
 - Report Due: **3/1** 23:59pm

