

## **SDSC2102 Statistical Methods and Data Analysis**

### **Project**

#### **Part I: Project description**

1. Data collection: please collect a set of publicly available data you are interested in via internet or other resources.
2. Problem formulation: based on the data collected by you and the related background, define your objective and formulate your problem of data analysis.
3. Data pre-processing: analyze the data properties, their distributions and statistical properties. Determine dimensions of parameters to be used in the analysis.
4. Data modeling: apply taught data analytics methods/algorithms to develop data-driven models for describing the relationships among your data. Justify the reason for selecting the method or algorithm and validate its effectiveness.
5. Results interpretation: relate your results of data analysis back to the practical problem and generate insights about reality based on the results.

#### **Part II: Requirement and evaluation**

1. Project team: you are required to form teams for the class project. Each team should have 3 or 4 members.
2. Project presentation: each team is required to give a presentation about their work and submit the slides in the last week. Details on this will be provided later.