

Problem Set #1

MACS 30200, Dr. Evans

Due Wednesday, Apr. 19 at 11:30am

Table 1: Data Sets and student assignments

?	?	?	?
Cheng Yee Lim	Julian McClellan	Alice Chung	Weijia Li
ChihYu Chiang	Benjamin Rothschild	Luxi Han	Rodrigo Valdes Ortiz
Erin Ochoa	Yuqing Zhang	Huanye Liu	Dongping Zhang
Bobae Kang	Chelsea Ernhofer	Ariel Xu	Sushmita V. Gopalan
Reid McIlroy-Young	Wenxi Xiao	Wanlin Ji	Haylee Ham
Yiqing Zhu	Jingyuan Zho	Esha Banerjee	Xinshu Sun
Hyung Jin Cho	Soowan Kim	Xingyun Wu	Yang Hou
Joanna Tung		Zhuo Leng	Tong Ju

1. **Write a data section for your assigned data set.** Your assignment is to write a data description section of a paper for your assigned dataset from Table 1 above. Your write up should have the following properties.
 - (a) Describe how to access data, where it is stored, who curates it.
 - (b) Cite other key papers that have used this data.
 - (c) Describe how the data were collected.
 - (d) Include a table that gives descriptive statistics for at least 8 key variables (you can do more).
 - (e) Include at least one key visualization of the data that exhibits an interesting characteristic.
 - (f) Show at least one conditional (slice) description of the data (e.g., all variable descriptive statistics by nationality of survey respondent). This can be a table or visualization.
 - (g) Items (a) through (f) are minimum requirements. Feel free to go beyond these if you feel you can describe some interesting characteristics of the data.
 - (h) A PDF of your data description section should be uploaded into your “ProblemSets/PS1” folder of your GitHub project repository by 11:30am on Wednesday, April 19.