The requirement for the project is to design, execute and analyze data from a sample
drawn from some real-world population. This is a group assignment: teams of three people
will form by March 30 using procedures described in the next paragraph, and each team
will be responsible for independently conducting its own study. Cooperation between
teams is allowed, but each team must collect its own sample dataset independently of every
other team and prepare its own report on its work.

Teams will form through a combination of voluntary actions and random assignment.
Any three people who wish to form a team should let me know the membership
of the team by noon on March 29.\footnote{Email me the team members’ names, being sure to CC all proposed team members on the message you send me.} Anyone who is not on a team by noon
March 29 I will assign to a team using a pseudorandom number generator.\footnote{When randomly forming the teams I will try to take into account any timezone information I have by that time.}

The plan before the pandemic was to devote the class meeting of March 30 to in-person
discussions so you could develop your initial ideas about what data you want to use for
your project and what kinds of analysis you want to do. Depending on conditions on
March 30, we might still have such a class meeting. But if in-person conversations in the
classroom seem unwise at the time, instead the alternative plan is for you to set up team
e-meetings on your own using Zoom, Hangouts or whatever method you prefer. I will
circulate the email addresses of the members of randomly constructed teams to facilitate
communication within each team. If you would like me to set up an e-meeting your team
can use, please let me know as soon as you can so I can do that. It’s possible that several
teams will want to meet together to formulate ideas about their separate projects.

By the end of classtime on March 30 (or by 6pm the same day), each team will
email me a paragraph describing in a preliminary way the project the team expects to
complete. By noon on April 1, each team will send me a formal proposal describing the
design and data gathering plan in more detail (earlier submission of these proposals is
encouraged). I will promptly let each team know whether the plan is approved. The lab
sessions of March 31 will be dedicated to work on the projects: the labs as such may or
may not meet, depending on conditions, but if not in person then you’re encouraged during
those times to create team e-meetings; GSI Josh will be available during the respective
labs’ times to join your e-meeting to offer advice and assistance.

A written report about the project is due by noon on April 20. Each team will submit a
single joint report (best is to email me a document in pdf format), and all team members
will receive the same grade for the report. The report should describe and justify the
sampling design, and it should report inferences (with bounds) about at least three
quantities of interest, including at least one ratio or regression estimator. All estimates
must be based on the same sample (or perhaps on subsamples of one sample). All aspects
of the design and analysis should be explicitly described and justified.

On April 13 or 18, each team will make a presentation to the class during classtime
about the team’s project. The quality of the presentation will factor into the overall grade
for the project. The specific date on which each team is to present will be randomly assigned on April 11.

On the final exam, at least one question will refer to and be based on the work done for the project. For example, you may be asked to justify your sampling design and explain why one or more other designs were not used. Or you may be asked to discuss, very specifically, the relative efficiency of estimates produced in your sample.

Walter Mebane