Assignment from Session 9: District RCT Exercise

Megha Joshi

1. If all goes well with randomization, the treatment and the comparison groups should not differ in terms of any observed or unobserved variables. Such balance in the groups on these variables is key to be able to infer the causal effect of treatment. If any important confounding variables are not balanced between the treatment and comparison groups, we may not be able to say that the treatment caused the change in outcome. Therefore, it is important to look at distribution of baseline characteristics to ensure treatment and comparison groups are balanced on all important confounding variables.
2. The table only has race/ethnicity and gender variables. Other confounding variables may be important. If those variables are imbalanced, then we may not be able to validly infer that the math tutoring program caused the change in the student math achievement scores. They include prior math achievement, socio-economic status (SES), special education status. For example, if students in treatment group had higher SES (perhaps due to unlucky randomization), it may be because they have better financial resources that caused them to score higher in the math achievement test rather than due to the treatment - the tutoring program.
3. The introductory paragraph mentions that the study that this data is based on was conducted in a district in a different state. It is important to figure out whether the results would be similar for students in the UPS in Washington state where I want to decide whether to implement the program. The district that the original study is from may have different population (e.g., in terms of racial composition or prior ability) than the students in my district.
4. Before I make the decision, I would like other information including details about the randomization, the research design, level of randomization (students, teachers, schools), any issues after randomization like attrition or joiners and compliance, analytic strategy used, missing data etc. Further, I would probably recommend analyzing how similar the population of the original study is to the population of UPS. If ALL my concerns are met, then I would recommend UPS to implement the program.