**Group 32: Research on U.S. Citizen Innovativeness**

**Project Proposal**

**Principal Investigators: Zheng Su (zsu78@wisc.edu), Yanni Zhang**

**Questions to be addressed**

[The first section describes the purpose of your study. Make it very clear what relationships you are studying and/or what questions you would like to answer. You may also want to talk about why these questions are interesting.]

Technological innovation has always been a huge push for economic growth and its importance has also elevated as economic growth has slowed down since the onset of the pandemic. We want to study the main factors that correlate with making an individual an innovator. After a brief literature review, we believe that factors, including birthplace, gender, race, and socioeconomic status, may influence citizens' ability to innovate. We want to apply the data to verify these assumptions. Also, since the dataset contains children who have total patent citations in the top 5% of their birth cohort, we can use parallel studies to see if the factors influencing the emergence of leading innovators and general innovators are the same.

**Data**

The data is from the dataset of *Who Becomes an Inventor in America? The Importance of Exposure to Innovation* in Opportunity Insights(<https://opportunityinsights.org/data/>). Most columns are percentages in specific populations, so there are no units. The dataset covers patenting outcomes of children who were born from 1980 to 1984, categorized by geography, gender, and parent income. One shortcoming of the data is that innovators are a tiny percentage of the total, resulting in each percentage being too small for visual analysis. One way to solve this is to transform the data into the number of innovators in 10,000 people.

**Preliminary results**

[Discuss the answers you have found to your questions. These answers may change, and there may still be some questions you have not answered yet.]

**Work plan**

1. Continue to examine differences in innovation rates for the effects of multiple variables, such as across socioeconomic classes, race, and gender, college, discipline fields.

We have analyzed XXX so far, and we will continue to complete the analysis of XXX subsequently. (取决于第三问做了啥).

1. Apply findings mentioned above for generalized innovation results to individuals whose total patent citations ranked in the top 5% of their birth cohort.

This step is generally not too difficult because the data set also shows the percentage of top innovators (individuals whose total patent citations ranked in the top 5% of their birth cohort) for different classes of variables, so this step of the study can be implemented with only minor modifications to the code from the previous step.

1. Use the factors that have the most significant impact on innovation to conduct case studies.

As shown in the pre-proposal, there are significant differences in innovation capacity across regions. We intend to use the factors we find that influence innovation to analyze the regions that currently have the highest innovation rates and find the reasons for their high innovation rates. Also, we prepare to examine where Wisconsin stands on each of these factors to suggest where Wisconsin should further improve