

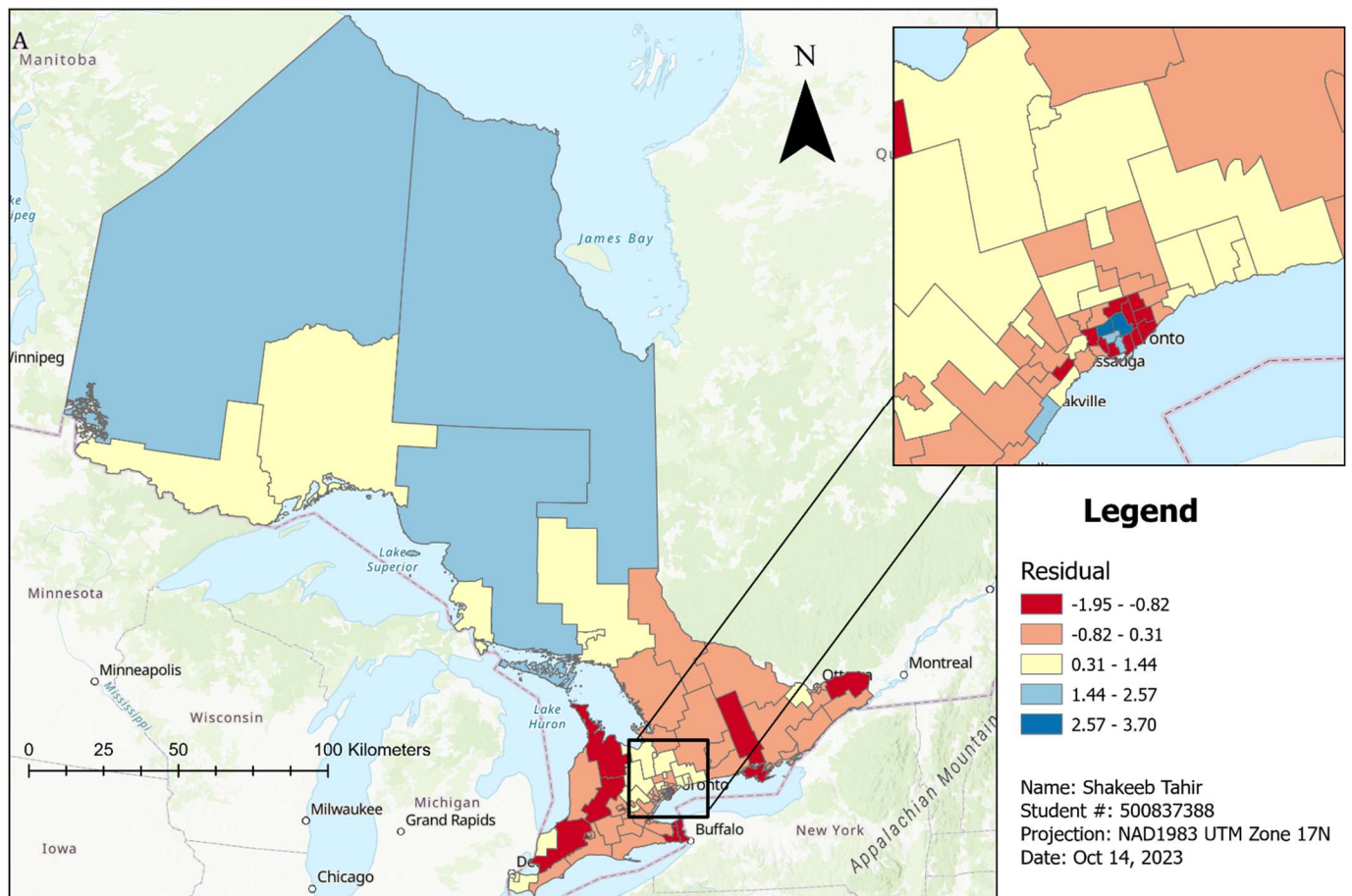
Shakeeb Tahir

500837388

SA8903 – Applied Spatial Statistics

Dr. Brian Ceh

2006 Ontario Census Divisions, Rated (per 1000) Log Average Family Income Regression Residuals



The pattern we can see from the map is that a lot of the positive residuals tend to be in the northern area of Ontario while the negative residuals tend to be in the southern area of Ontario. However, there is one exception which we can see is the GTA, specifically the Toronto area which has a few positive outliers in the center of Toronto. It is possible that the reason the northern areas of Ontario are overestimated is because the average family income could be higher in those areas through other means than our predictors. Since we are using variables

such as university degrees, finance & insurance, arts & recreation, these tend to be lower in remote areas like northern Ontario along with unemployment which is higher in those areas. This might be causing the model to underestimate how much income these areas have since it would assume that less university degrees, finance & insurance, arts & recreation, and higher unemployment would lead to a lower family income. However, the areas in northern Ontario likely get their income from other sources such as hunting or trades work which we only have one variable (Agriculture/Forest/Fishing/Hunting) to measure with. Thus, the average income is more likely to be higher than predicted since the income is achieved through different means which are not quantified. The vice-versa is true for southern area of Ontario as these variables are good at predicting the economy of the area as they relate to how most people earn their income. The one exception we can see on the map is the Toronto area, specifically in the center where it is dark blue with two light blues. The dark blue division is Don Valley West and the two light blues are St. Paul's and Toronto center. This is interesting as there is possibly something unique going on in these division with regards to how people earn this income which is why the residuals are positive. It could be that since these areas represent downtown Toronto and the core Toronto area, this is where higher average income families are likely to move so the data could be skewed by those high earners while the other variables are still at their normal levels. There are numerous reasons why this could be and a deeper dive into the demographics of those census divisions would be required to find out why.