

---

ISyE 2028 – Basic Statistical Methods - Fall 2015  
Bonus Project: “Big” Data Analytics  
Proposal (or Final Report)

---

Mary Latimer  
GTID 902868975

Since I started working for Coca-Cola, the drink preferences of people has been a real interest to me. I grew up around Atlanta, and all of my friends and family growing up preferred Coke. Coming to college, I met a lot of people who preferred Pepsi, who were also from out of state. This made me wonder if Coke vs Pepsi is a regionally induced preference, since the birthplace of Coke is in Atlanta. I think that a greater proportion of people who grew up in Southeastern states prefer Coke as compared to people who grew up in non Southeastern states.

I will make a survey collecting data on: drink preferences growing up, drink preferences today, and the state in which they grew up. I will collect data from my peers ( I will send it to multiple email groups at Georgia Tech). To make sure that I have collected unbiased and balanced data(because we are in Atlanta and there are more people from the Southeast than anywhere else, I will cap the data collection at 50 samples for each set of data: Southeast and non-Southeast.

I will do Hypothesis testing to analyze the results. My original Hypothesis is that the proportion of Southeasterners that prefer Coke ( $p_1$ ) is greater than the proportion of non-Southeasterners that prefer Coke( $p_2$ ). I think that this hypothesis will be true.

$$H_0 = p_1 > p_2$$

$$H_a = p_1 \geq p_2$$