**Project Proposal – DATA 602**

I propose to analyze a dataset of spirits purchase information of Iowa Class “E” liquor licensees by product and date of purchase from January 1, 2012 to current. The dataset will allow me to analyze total spirits sales in Iowa of individual products at the store level. Since the dataset has store level detail information I will attempt to map the resulting analysis, as well as graph other descriptive statistics regarding the dataset. I hope to discovery some unique trends over time or determine a general sense of the liquor market place in Iowa. Also, the data set is large at 791 MB so I also hope to learn techniques to efficiently process large amounts of data using Python. https://data.iowa.gov/Economy/Iowa-Liquor-Sales/m3tr-qhgy sample data in json format { "DATE": [ "02/26/2015" ], "CONVENIENCE.STORE": [ "" ], "STORE": [ 2515 ], "NAME": [ "Hy-Vee Food Store #1 / Mason City" ], "ADDRESS": [ "2400 4TH ST SW" ], "CITY": [ "MASON CITY" ], "ZIPCODE": [ "50401" ], "STORE.LOCATION": [ "2400 4TH ST SW\nMASON CITY 50401\n(43.148463097000047, -93.236272961999987)" ], "COUNTY.NUMBER": [ 17 ], "COUNTY": [ "Cerro Gordo" ], "CATEGORY": [ 1022100 ], "CATEGORY.NAME": [ "TEQUILA" ], "VENDOR.NO": [ 434 ], "VENDOR": [ "Luxco-St Louis" ], "ITEM": [ 87937 ], "DESCRIPTION": [ "Juarez Tequila Silver" ], "PACK": [ 12 ], "LITER.SIZE": [ 1000 ], "STATE.BTL.COST": [ "$6.92" ], "BTL.PRICE": [ "$10.38" ], "BOTTLE.QTY": [ 48 ], "TOTAL": [ "$498.24" ] }