Team 2 (Austin White, Rylan Hunter, Mayank Jaggi)
10/8/2019

Project 1 White Page

We will be addressing how the opioid crisis has affected the different socio-economic classes in the United States. We will accomplish this by looking at the amount of each drug per capita that was prescribed in each zip code in 2010, and then comparing this data with the zip code's population density and median household income. We got this information by pulling from the ARCOS database, as well as a couple of other data files of 2010 census data we found online. We will process our ARCOS data into a CSV file that is organized by zip code, with columns for the total number of pills of each drug type in each zip code and the average dosage of each type of pill in each zip code. This data will then be merged with the census data, which will be used to create another column for the pills per capita and the amount of each drug in milligrams per capita.

The median income and population density should give us enough information to find any correlation between socio-economic class and opioid use, as the population density will give us information on the rurality of the zip code, and the median income will give information about the economic status of the zip code. We can also combine median income and population density through a simple equation to determine whether or not there is a stronger correlation when these two data points are looked at together (for example, we will be able to tell if an upper-class urban area has a large or smaller opioid problem than a lower-class rural area).

Figure 1 below shows the first 9 zip codes and the corresponding census data that we are using. We will be adding the data from ARCOS (Figure 2 shows a sample of the ARCOS data we have pulled) to this database, at which point we will be able to create charts and plots to investigate any correlation between our data sets.

ZIP	MEDIAN_HOUSEHOLD_INCOME	POPULATION	POPULATION_DENSITY_PER_SQ_MILE
1001	56663	16769	1465.565461
1002	49853	29049	527.751031
1003	28462	10372	14587.90436
1005	75423	5079	114.8004159
1007	79076	14649	278.2706153
1008	63980	1263	23.47758198
1009	51452	741	929.7365119
1010	75625	3609	103.87704
1011	63476	1370	43.33248988

Figure 1: Census Data Organized by Zip Code

	BUYER_CITY	BUYER_STATE	BUYER_ZIP	BUYER_COUNTY	DRUG_NAME	QUANTITY	TRANSACTION_DATE	Revised_Company_Name	dos_str	YEAR
0	GILBERT	AZ	85233	MARICOPA	HYDROCODONE	20	2010-03-23	Apotheca Inc.	5.0	2010
1	PHOENIX	AZ	85085	MARICOPA	HYDROCODONE	3	2010-09-15	Apotheca Inc.	5.0	2010
2	PHOENIX	AZ	85085	MARICOPA	HYDROCODONE	3	2010-12-09	Apotheca Inc.	7.5	2010
3	PULASKI	NY	13142	OSWEGO	HYDROCODONE	1	2010-12-02	Allergan, Inc.	10.0	2010
4	PULASKI	NY	13142	OSWEGO	HYDROCODONE	5	2010-11-05	Allergan, Inc.	10.0	2010

Figure 2: ARCOS Data