Department of City and Regional Planning CPLN 505 Planning By Numbers/ Spring 2019

Professor: Megan Ryerson TA: Joshua Davidson

Assignment 2: Comparing Groups and Testing for Association and Correlation

(due February 22, 2019, 11:59 pm)

Philadelphia is served by a Combined Sewer System (CSS) that was originally constructed during the Industrial Revolution, over 100 years ago. In an effort to upgrade its infrastructure to meet 21st century environmental standards, the city requires significant investment in stormwater and wastewater infrastructure. The previous administration, headed by Mayor Michael Nutter, was successful in adopting one of the most aggressive green infrastructure plans in the country as a means to reduce loading to the system, while also providing economic and environmental benefit to the city. Mayor Kenney has brought you on as a consultant of the City of Philadelphia to evaluate how Philadelphia compares to other cities in water infrastructure needs. To do this, you need to characterize and compare the water infrastructure needs across the entire United States.

The Clean Watersheds Needs Survey (CWNS) is a publicly available dataset containing wastewater utilities' self-reported facility budgetary needs. Survey results have been released for 2004, 2008, and 2012. You are provided with cleaned data from the 2012 survey, which has been spatially joined with county-level data from the US Census.

For students who would like to map this data in GIS software, latitude and longitude data are provided. Export the provided .csv (file to .dbf format using the R package foreign. In ArcGIS, Add the .dbf file, then in the Layer tree, right click on the tabular dataset, and select 'Display XY Data'. When selecting the spatial reference, choose: 'Geographic Coordinate System NAD 1983 (2011)'. From here, you can do spatial joins to other datasets (you don't have to for this assignment, but this is often very useful).

Data Key:

Column Name	Description	Source
STCOU	State and county identifier	US Census County-level
		data
CWNS_NUMBER	Facility identifier	EPA 2012 CWNS
TOTAL_OFFICIAL	Total reported need for water infrastructure	EPA 2012 CWNS
_NEED		
PRES_RES_REC_	Present (2012) residential population	EPA 2012 CWNS
COLLCTN	receiving collection	
PRES_RES_REC_	Present (2012) residential population	EPA 2012 CWNS
TRMT	receiving treatment	
PRES_N_RES_RE	Present (2012) non-residential population	EPA 2012 CWNS
C_COLLCTN	receiving collection	

PRES. REC. CTRTM PROJ. RES. REC COLLCTN PROJ. RES. YR Projected residential population receiving PROJ. RES. YR Projected residential population receiving PROJ. RES. YR Projected residential population receiving PROJ. N. RES. REC TRMT PROJ. N. RES. RE C. TRRMT PROJ. N. RES. RE C. TRRMT PROJ. N. RES. RE C. TRRMT Projected non-residential population PROJ. N. RES. RE C. TRRMT EPA REGION PROJ. RES. RE C. TRRMT State where facility is located Projected non-residential PA 2012 CWNS PROJ. N. RES. RE C. TRRMT State where facility is located PROJ. N. RES. RE C. TRRMT State where facility is located PROJ. N. RES. RE PROJ. N. RES. RE C. TRRMT State where facility is located PRA 2012 CWNS PROJ. N. RES. RE C. TRRMT State where facility is located PRA 2012 CWNS STATE PA 2012 CWNS PROJ. N. RES. RE C. TRRMT State where facility is located PPA 2012 CWNS STATE PACILITY CITY PROJ. N. RES. RE PRA 2012 CWNS PROJ. N. RES. RE C. TRRMT State where facility is located PPA 2012 CWNS PROJ. N. RES. RE PRA 2012 CWNS PROJ. N. RES. RE State where facility is located PPA 2012 CWNS PROJ. N. RES. RE PRA 2012 CWNS PPA 2012	DDDG M DDG DD		TD 4 CVIDIO
PROJ. RES_REC_ COLLCTN PROJ. RES_YR Projected residential population receiving COLLCTN PROJ. RES_YR Projected residential population receiving TRMT Treatment PROJ. RES_REC_ COLLCTN PROJ.N_RES_RE C_COLLCTN PROJ.N_RES_RE Projected non-residential population receiving collection PROJ.N_RES_YR Projected non-residential population receiving collection PROJ.N_RES_YR Projected non-residential population receiving treatment PROJ.N_RES_RE C_COLLCTN PROJ.N_RES_RE Projected non-residential population receiving treatment PROJ.N_RES_RE Projected non-residential population receiving treatment PROJ.N_RES_RE Projected non-residential population receiving treatment PROJ.N_RES_RE PROJ.N_RES_RE Projected non-residential population receiving treatment PROJ.N_RES_RE PROJ.N_RES_RE PROJ.N_RES_RE Projected non-residential population PROJ.N_RES_YR Projected non-residential population PPROJ.N_RES_YR Projected non-residential population PPR 2012 CWNS PPR 2012 CWNS PPA 2012 CWNS PP	PRES_N_RES_RE	Present (2012) non-residential population	EPA 2012 CWNS
COLLCIN PROJ_RES_YR PROJ_RES_YR PROJ_RES_REC_TRAMT Treatment PROJ_N_RES_RED_Trojected residential population receiving treatment PROJ_N_RES_RED_Trojected non-residential population receiving collection PROJ_N_RES_RED_Trojected non-residential population receiving collection PROJ_N_RES_YR Projected non-residential population receiving collection PROJ_N_RES_YR Projected non-residential population receiving collection PROJ_N_RES_YR Projected non-residential population receiving treatment PROJ_N_RES_RED_Trojected non-residential population receiving treatment EPA_REGION PROJ_N_RES_RED_Projected non-residential population receiving treatment EPA_REGION		receiving treatment	
PROJ RES YR Projected residential population receiving treatment EPA 2012 CWNS PROJ N. RES RED rejected residential population receiving treatment EPA 2012 CWNS PROJ N. RES Projected non-residential population population EPA 2012 CWNS PROJ N. RES PR Projected non-residential population population EPA 2012 CWNS PROJ N. RES RE Projected non-residential population EPA 2012 CWNS PROJ N. RES RE Projected non-residential population EPA 2012 CWNS PROJ N. RES RE Projected non-residential population EPA 2012 CWNS PROJ N. RES RE Projected non-residential population EPA 2012 CWNS PROJ N. RES REPA 2012 CWNS EPA 2012 CWNS PACILITY NAME Name of facility EPA 2012 CWNS PACILITY CITY City where facility is located EPA 2012 CWNS LANTITUDE Latitude of facility EPA 2012 CWNS PATOTITUDE Longitude of facility EPA 2012 CWNS PATOTITUDE<		Projected residential population receiving	EPA 2012 CWNS
PROJ_RES_REC TRMT TRMT PROJ_N_RES_RE Projected non-residential population receiving receiving collection PROJ_N_RES_RE C_COLLCTN PROJ_N_RES_YR Projection year for non-residential population REPA 2012 CWNS PROJ_N_RES_YR Projected non-residential population REPA 2012 CWNS PROJ_N_RES_RE Projected non-residential population C_TRMT PROJ_N_RES_RE Projected non-residential population C_TRMT REPA_REGION EPA Region (1-10), see map below EPA 2012 CWNS STATE State where facility is located EPA 2012 CWNS FACILITY NAME Name of facility WNER_TYPE Public, Private, Federal EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS FACILITY_CITY Latitude of facility EPA 2012 CWNS EPA		I .	
TRMT PROJ_N_RES_RE Projected non-residential population COLLCTN PROJ_N_RES_YR Projection year for non-residential population PROJ_N_RES_YR Projected non-residential population PROJ_N_RES_RE Projected non-residential population PROJ_N_RES_RE Projected non-residential population receiving treatment EPA 2012 CWNS PROJ_N_RES_RE Projected non-residential population receiving treatment EPA 2012 CWNS PROJ_N_RES_RE PROJ_N_RES_RE Projected non-residential population receiving treatment EPA 2012 CWNS EPA 2012 CWNS STATE State where facility is located FPA 2012 CWNS STATE PA 2012 CWNS FACILITY_NAME PROJ_N_RES_RE FACILITY_CITY City where facility EPA 2012 CWNS FACILITY_CITY City where facility EPA 2012 CWNS FACILITY_CITY City where facility EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS TMDL_INDICATO R COUNTYFP County code of location of facility EPA 2012 CWNS TATEFP State code US Census County-level data (via spatial join) US Census County-level data (via spatial join) US Census County-level data (via spatial join) WE COUNTYNAME County name of location of facility US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county WS Census County-level data (via spatial join) MEDINC9 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1969 in county US Census County-level data (vi			
PROJ_N_RES_RE C_COLICTN PROJ_N_RES_YR PROJ_N_RES_YR PROJ_N_RES_YR PROJ_N_RES_YR PROJ_N_RES_RE C_TRMT Projected non-residential population C_COUNTYNAME PROJ_N_RES_RE FACILITY County name of location of facility C_COUNTYNAME C_COUNTY-Level data (via spatial join) MEDINCGG			EPA 2012 CWNS
C COLLCTN PROJ_N_RES_YR Projection year for non-residential population PROJ_N_RES_RE Projected non-residential population receiving treatment PROJ_N_RES_RE PA Projected non-residential population receiving treatment PROJ_N_RES_RE PA Region (1-10), see map below EPA 2012 CWNS STATE State where facility is located FPA 2012 CWNS PACILITY_NAME PHOLIC, Private, Federal PA 2012 CWNS PACILITY_CITY City where facility is located EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS FACILITY_CITY City where facility EPA 2012 CWNS FACILITY_CITY City EPA 2012 CWNS FACILITY_C			
PROJ_N_RES_YR projection year for non-residential population population population population receiving treatment receiving re			EPA 2012 CWNS
PROJ_N_RES_RE C TRMT			
PROJ_N_RES_RE C_TRMT Projected non-residential population receiving treatment EPA 2012 CWNS EPA REGION EPA Region (1-10), see map below EPA 2012 CWNS STATE State where facility is located EPA 2012 CWNS FACILITY_NAME Name of facility EPA 2012 CWNS GWNER_TYPE Public, Private, Federal EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS MEDINCATO **Ye-contributing to a TMDL receiving water body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1989 in county US Census County-level data (PROJ_N_RES_YR		EPA 2012 CWNS
C_TRMT receiving treatment EPA_REGION EPA Region (1-10), see map below EPA 2012 CWNS STATE State where facility is located EPA 2012 CWNS FACILITY_NAME Name of facility EPA 2012 CWNS OWNER_TYPE Public, Private, Federal EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water EPA 2012 CWNS R body STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1989 in county US Census County-level data (via spatial join)			
EPA_REGION EPA Region (1-10), see map below EPA 2012 CWNS STATE State where facility is located EPA 2012 CWNS FACILITY_NAME Name of facility EPA 2012 CWNS OWNER_TYPE Public, Private, Federal EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) COUNTYPP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1999 in county US Census County-level data (via s		Projected non-residential population	EPA 2012 CWNS
STATE State where facility is located EPA 2012 CWNS FACILITY_NAME Name of facility EPA 2012 CWNS OWNER_TYPE Public, Private, Federal EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water body STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC9 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC9 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 2000 US Census County-level data (via spatial join)	C_TRMT	receiving treatment	
FACILITY_NAME Name of facility EPA 2012 CWNS OWNER_TYPE Public, Private, Federal EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS LATITUDE Longitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water R State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) WEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Total population in county in 1980 US Census County-level data (via spatial join) POP00 Total population in county in 1990 US Census County-level data (via spatial join) US Census County-level data (via spatial join) US Census County-level data (via spatial join)	EPA_REGION	EPA Region (1-10), see map below	EPA 2012 CWNS
OWNER_TYPE Public, Private, Federal EPA 2012 CWNS FACILITY_CITY City where facility is located EPA 2012 CWNS LATITUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1969 in county US Census County-level data (via spatial join)	STATE	State where facility is located	EPA 2012 CWNS
FACILITY_CITY City where facility is located EPA 2012 CWNS LATTUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water R body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Total population in county in 1980 US Census County-level data (via spatial join) DOPO0 Total population in county in 2000 US Census County-level data (via spatial join) DOPO10 Total population in county in 2010 US Census County-level data (via spatial join) US Census County-level	FACILITY NAME	Name of facility	EPA 2012 CWNS
FACILITY_CITY City where facility is located EPA 2012 CWNS LATTUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water R body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Total population in county in 1980 US Census County-level data (via spatial join) DOPO0 Total population in county in 2000 US Census County-level data (via spatial join) DOPO10 Total population in county in 2010 US Census County-level data (via spatial join) US Census County-level	OWNER_TYPE		EPA 2012 CWNS
LATITUDE Latitude of facility EPA 2012 CWNS LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level da			
LONGITUDE Longitude of facility EPA 2012 CWNS TMDL_INDICATO Y=contributing to a TMDL receiving water body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC09 Median income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 198		· · ·	
TMDL_INDICATO R Y=contributing to a TMDL receiving water body EPA 2012 CWNS STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 2000 US Census C		·	
R body STATEFP State code US Census County-level data (via spatial join) COUNTYFP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC90 Median income in 2009 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Total population in county in 1980 US Census County-level data (via spatial join) POP80 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level			
STATEFP State code			2111 2012 0 11110
COUNTYFP County code of location of facility US Census County-level data (via spatial join)			US Census County-level
COUNTYPP County code of location of facility US Census County-level data (via spatial join) COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC09 Median income in 2009 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level data (via spatial join)		State code	
COUNTYNAME County name of location of facility US Census County-level data (via spatial join) ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC09 Median income in 2009 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	COUNTYFP	County code of location of facility	
COUNTYNAME County name of location of facility ALAND Land area (m²) of county Water area (m²) of county US Census County-level data (via spatial join) WEDINC69 Median income in 1969 in county WE Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county WE Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county WE Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC99 Median income in 2009 in county US Census County-level data (via spatial join) WEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2000 US Census County-level data (via spatial join)	COUNTIFI	County code of location of facility	
ALAND Land area (m²) of county US Census County-level data (via spatial join) AWATER Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC99 Median income in 2009 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level data (via spatial join)	COUNTYNAME	County name of location of facility	
ALAND Land area (m²) of county Water area (m²) of county Water area (m²) of county US Census County-level data (via spatial join) MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join) MEDINC79 Median income in 1979 in county US Census County-level data (via spatial join) MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC09 Median income in 2009 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP00 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level data (via spatial join)	COUNTINAME	County name of location of facility	•
AWATER Water area (m²) of county Water area (m²) of county-level data (via spatial join) Water area (m²) of county Water area (m²) of county-level data (via spatial join) Water area (m²) of county Water area (m²) of county-level data (via spatial join) Water area (m²) of county-level data (via spat	ALAND	Land area (m2) of county	
AWATER Water area (m²) of county MEDINC69 Median income in 1969 in county MEDINC79 Median income in 1979 in county MEDINC89 Median income in 1989 in county MEDINC99 Median income in 1989 in county MEDINC99 Median income in 1999 in county MEDINC99 Median income in 1999 in county MEDINC99 Median income in 1999 in county MEDINC99 Median income in 2009 in county MEDINC09 Median income in 2009 in county MEDINC09 Median income in 2009 in county MEDINC09 Median income in 1969 in county MEDINC09 Median income in 1969 in county MEDINC09 Mean income in 1969 in county MEDINC09 MEDINC09 Median income in 2009 in county US Census County-level data (via spatial join) MEDINC09 Total population in county in 1980 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level data (via spatial join)	ALAND	Land area (iii) or county	
MEDINC69 Median income in 1969 in county US Census County-level data (via spatial join)	ΔΜΛΤΕΡ	Water area (m2) of county	US Consus County-level
MEDINC69Median income in 1969 in countyUS Census County-level data (via spatial join)MEDINC79Median income in 1979 in countyUS Census County-level data (via spatial join)MEDINC89Median income in 1989 in countyUS Census County-level data (via spatial join)MEDINC99Median income in 1999 in countyUS Census County-level data (via spatial join)MEDINC09Median income in 2009 in countyUS Census County-level data (via spatial join)MEANINC09Mean income in 1969 in countyUS Census County-level data (via spatial join)POP80Total population in county in 1980US Census County-level data (via spatial join)POP90Total population in county in 1990US Census County-level data (via spatial join)POP00Total population in county in 2000US Census County-level data (via spatial join)POP10Total population in county in 2010US Census County-level data (via spatial join)	AWAIEK	water area (iii-) or county	
MEDINC99 Median income in 1979 in county MEDINC89 Median income in 1989 in county MEDINC99 Median income in 1989 in county MEDINC99 Median income in 1999 in county MEDINC99 Median income in 1999 in county MEDINC99 Median income in 1999 in county MEDINC99 Median income in 2009 in county MEDINC99 Median income in 2009 in county MEDINC99 Median income in 2009 in county MEDINC99 Mean income in 1969 in county MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP90 Total population in county in 2000 MEDINC09 Median income in 1989 in county MEDINC09 US Census County-level data (via spatial join) MEDINC09 Median income in 1989 in county MEDINC09 Median income in 1989 in county MEDINC09 Median income in 1989 in county MEDINC09 Median income in 1999 in	MEDINC60	Modion income in 1060 in county	
MEDINC79Median income in 1979 in countyUS Census County-level data (via spatial join)MEDINC89Median income in 1989 in countyUS Census County-level data (via spatial join)MEDINC99Median income in 1999 in countyUS Census County-level data (via spatial join)MEDINC09Median income in 2009 in countyUS Census County-level data (via spatial join)MEANINC09Mean income in 1969 in countyUS Census County-level data (via spatial join)POP80Total population in county in 1980US Census County-level data (via spatial join)POP90Total population in county in 1990US Census County-level data (via spatial join)POP00Total population in county in 2000US Census County-level data (via spatial join)POP10Total population in county in 2010US Census County-level data (via spatial join)	MEDINCOG	Median income in 1909 in county	
MEDINC89 Median income in 1989 in county US Census County-level data (via spatial join) MEDINC99 Median income in 1999 in county US Census County-level data (via spatial join) MEDINC09 Median income in 2009 in county US Census County-level data (via spatial join) MEDINC09 Mean income in 1969 in county US Census County-level data (via spatial join) MEANINC09 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level data (via spatial join)	MEDINCEO	Modion in come in 1070 in country	LIC Congress County level
MEDINC89Median income in 1989 in countyUS Census County-level data (via spatial join)MEDINC99Median income in 1999 in countyUS Census County-level data (via spatial join)MEDINC09Median income in 2009 in countyUS Census County-level data (via spatial join)MEANINC09Mean income in 1969 in countyUS Census County-level data (via spatial join)POP80Total population in county in 1980US Census County-level data (via spatial join)POP90Total population in county in 1990US Census County-level data (via spatial join)POP00Total population in county in 2000US Census County-level data (via spatial join)POP10Total population in county in 2010US Census County-level data (via spatial join)	MEDINC/9	Median income in 1979 in county	data (via anatial inin)
MEDINC99 Median income in 1999 in county MEDINC09 Median income in 2009 in county MEDINC09 Median income in 2009 in county MEDINC09 Mean income in 1969 in county MEANINC09 Mean income in 1969 in county MEANINC09 Mean income in 1969 in county MEANINC09 Total population in county in 1980 POP90 Total population in county in 1990 POP00 Total population in county in 2000 Total population in county in 2000 MEDINC09 US Census County-level data (via spatial join) US Census County-level data (via spatial join) US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level data (via spatial join) US Census County-level data (via spatial join)	MEDINGO	Modion in course in 1000 in country	
MEDINC99Median income in 1999 in countyUS Census County-level data (via spatial join)MEDINC09Median income in 2009 in countyUS Census County-level data (via spatial join)MEANINC09Mean income in 1969 in countyUS Census County-level data (via spatial join)POP80Total population in county in 1980US Census County-level data (via spatial join)POP90Total population in county in 1990US Census County-level data (via spatial join)POP00Total population in county in 2000US Census County-level data (via spatial join)POP10Total population in county in 2010US Census County-level	MEDINC89	Median income in 1989 in county	
MEDINCO9 Median income in 2009 in county MEANINCO9 Mean income in 1969 in county MEANINCO9 Mean income in 1969 in county DS Census County-level data (via spatial join) WEANINCO9 Total population in county in 1980 POP90 Total population in county in 1990 POP00 Total population in county in 1990 Total population in county in 2000 DS Census County-level data (via spatial join) WEANINCO9 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level data (via spatial join) US Census County-level data (via spatial join)	MEDINGO	Malia i a sa sa i a sa sa i a sa s	
MEDINC09Median income in 2009 in countyUS Census County-level data (via spatial join)MEANINC09Mean income in 1969 in countyUS Census County-level data (via spatial join)POP80Total population in county in 1980US Census County-level data (via spatial join)POP90Total population in county in 1990US Census County-level data (via spatial join)POP00Total population in county in 2000US Census County-level data (via spatial join)POP10Total population in county in 2010US Census County-level	MEDINC99	Median income in 1999 in county	· ·
MEANINCO9 Mean income in 1969 in county US Census County-level data (via spatial join) POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 1990 US Census County-level data (via spatial join) US Census County-level data (via spatial join) US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	MEDING	Malianian and a constitution	
MEANINC09Mean income in 1969 in countyUS Census County-level data (via spatial join)POP80Total population in county in 1980US Census County-level data (via spatial join)POP90Total population in county in 1990US Census County-level data (via spatial join)POP00Total population in county in 2000US Census County-level data (via spatial join)POP10Total population in county in 2010US Census County-level	MEDINC09	Median income in 2009 in county	
POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	AFRANTS		
POP80 Total population in county in 1980 US Census County-level data (via spatial join) POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	MEANINC09	Mean income in 1969 in county	
POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	DODO		
POP90 Total population in county in 1990 US Census County-level data (via spatial join) POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	POP80	Total population in county in 1980	
POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	707		
POP00 Total population in county in 2000 US Census County-level data (via spatial join) POP10 Total population in county in 2010 US Census County-level	POP90	Total population in county in 1990	· ·
POP10 Total population in county in 2010 US Census County-level			
POP10 Total population in county in 2010 US Census County-level	POPoo	Total population in county in 2000	•
data	POP10	Total population in county in 2010	
			data

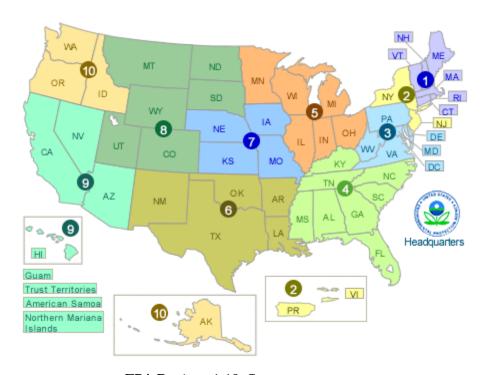
PCTWHITEoo	Percent of county population identifying as	US Census County-level
	white in 2000	data (via spatial join)
PCTWHITE10	Percent of county population identifying as	US Census County-level
	white in 2010	data (via spatial join)
CSS	1=facility associated with a combined sewer	EPA 2012 CWNS (via T.
	system (CSS)	Lim recategorization)
MS4	1=facility associated with a municipal	
	separate sewer system (MS4)	Lim recategorization)

Data References

US EPA. "Clean Watersheds Needs Survey 2012: Report to Congress," January 2016. https://www.epa.gov/cwns/clean-watersheds-needs-survey-cwns-2012-report-and-data.

USA Counties Data File Downloads.

http://www.census.gov/support/USACdataDownloads.html#IPE



EPA Regions 1-10. Source: usepa.gov

I. CREATE METRICS TO CONTEXTUALIZE REPORTED NEED

Calculate the following metrics:

- Projected change in residential population receiving collection
- Projected percent change in residential population receiving collection
- Projected change in residential population receiving treatment
- Projected percent change in residential population receiving treatment

- Projected change in non-residential population receiving treatment
- Projected percent change in non-residential population receiving treatment
- Projected change in non-residential population receiving collection
- Projected percent change in non-residential population receiving collection
- Residential population density (using county area and population)
- Change and percent change in population density
- Percent change in median income

II. PROVIDE SUMMARY STATISTICS OF THE DATA

Provide summary tables and figures that describe the trends that you see in the data. You may include summary tables of the 10 best and worst performing MSAs in each resilience measure, density plots, histograms, box plots, or any other means to communicate the data to your reader.

- Top and Bottom 10 facilities with greatest/least need (perhaps separating them out by region)
- Tables with min, max, quartiles, median, and mean reported for relevant variables
- Boxplots, density plots and histograms of need by EPA region, climatic region (NOAA), state, or other grouping (**hint:** log transformations of data can help in identifying relationships in data when they appear to be highly skewed, non-normally distributed or otherwise difficult to differentiate.

Hint for R: for categorical data ("factor" data type in R), the summary() command will display frequencies, rather than the typical min, max, median, mean, and quantile summary.

Hint for R: To find specific facilities in this dataset (which has more than 8,000 entries), try using the grep command (https://stat.ethz.ch/R-manual/R-devel/library/base/html/grep.html) for string pattern matching. Setting the ignore.case option in the grep command to TRUE is particularly helpful since this dataset contains 'mixed case' entries (R is typically case-sensitive).

Example for finding 'Philadelphia' facilities:

III. TESTS OF ASSOCIATION

As a group, is there a significant difference in the magnitude of overall budgeted need for facilities with CSS versus facilities with MS4s (Municipal Separate Sewer Systems)? Use a t-test of difference in means to compare these two groups.

As a group, is there a significant difference between facilities' needs for those that are associated with a TMDL? Does this vary by region?

Is there evidence of significant regional differences in need? Use the chi-square test (CrossTable() function).

Choose five variables in the dataset, or from the variables that you created in Question I that you suspect may be associated with reported infrastructure need. Using the CrossTable()function to determine if there is evidence of association between each of these variables and total overall need. Where do the Philadelphia facilities fall amongst the categories that you created for each variable?

IV. CORRELATION

Choose five variables in the dataset, or from the variables that you created in Question I. Create scatterplots between these variables and reported need, and between each other. Use the cor.test() function to determine evidence of correlation between variables.

V. CONCLUSION

Draft a conclusion to the report to Mayor Kenney summarizing your findings. What additional variables would be helpful to include that might better explain the variation you observe if you had additional scope?