

## Worksheet for Assignment #1

Use this worksheet to complete the assignment. The completed worksheet can be longer than one page. There are some sample responses (the shaded cells) provided to you as a guide. Remove those before you begin!

### Part 1: Develop Hypotheses

Hypothesis (1 sentence)	Rationale (2-3 sentences)
People living in the Philadelphia metropolitan area have lower average incomes than people living in the San Francisco, CA metropolitan area	People in San Francisco are earning higher incomes because of they have a better job opportunities in the tech industry. More and more educated people are moving to the San Francisco area.
Hypothesis 2: San Francisco receives more sunshine average per day than Philadelphia	San Francisco is at 37.7833 degrees North while Philadelphia is at 39.95 degrees North. Areas at lower latitude values typically have higher sunshine values while higher latitude values are much cloudier.
Hypothesis 3: The cost of housing in San Francisco is higher than the cost of housing in Philadelphia	San Francisco has more expensive housing and properties because of San Francisco's gentrification. The growing population of well-paid workers has led to development of wealthier housing.
Hypothesis 4: San Francisco has a younger population compared to Philadelphia	With the tech industry offering higher paying jobs, more and more college graduates are moving to San Francisco. In addition, San Francisco is more liberal and diverse which attracts more young people.
Hypothesis 5: Philadelphia has a lower chance of experiencing an earthquake compared to San Francisco	San Francisco is on the edge of the Pacific plate while Philadelphia is a long distance away from a plate boundary. Every time the tectonic plates move and the rock underground breaks along the fault, the areas along the edge are at risk of an earthquake.

### Part 2: Identify Data Sets

Data Set	Name	URL	Instructions
1	Median Household Income in Metropolitan Areas	<a href="http://www.census.gov/content/dam/Census/library/publications/2014/acs/acsbr13-02.pdf">http://www.census.gov/content/dam/Census/library/publications/2014/acs/acsbr13-02.pdf</a>	Scroll down to Page 4
2	Median Household Income	<a href="http://www.city-data.com/">http://www.city-data.com/</a>	Scroll down to Interactive Data Map. Zoom in to

MIS 0855 Data Science – Spring 2015 – Assignment #1

			get median household income in county
3	Annual Possible Sunshine	<a href="http://www1.ncdc.noaa.gov/pub/data/ccd-data/pctposrank.txt">http://www1.ncdc.noaa.gov/pub/data/ccd-data/pctposrank.txt</a>	Press Control + F to search for cities
4	Average Annual Sunshine	<a href="http://www.currentresults.com/Weather/US/average-annual-sunshine-by-city.php">http://www.currentresults.com/Weather/US/average-annual-sunshine-by-city.php</a>	
5	Housing and Transportation as a Percent of Income	<a href="http://www.citylab.com/housing/2014/08/what-does-it-really-cost-to-live-in-san-francisco/379166/">http://www.citylab.com/housing/2014/08/what-does-it-really-cost-to-live-in-san-francisco/379166/</a>	Scroll down to Figure 6
6	Median House Listing Price	<a href="http://www.trulia.com/home_prices/">http://www.trulia.com/home_prices/</a>	
7	Median Age by County	<a href="http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf">http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf</a>	Scroll down to Figure 7 on Pg.11
8	Median Age for U.S. Counties	<a href="http://www.governing.com/gov-data/census/median-age-county-population-map.html">http://www.governing.com/gov-data/census/median-age-county-population-map.html</a>	
9	Damaging Earthquakes in the U.S. (1750-1996)	<a href="http://earthquake.usgs.gov/earthquakes/states/us_damage_eq.php">http://earthquake.usgs.gov/earthquakes/states/us_damage_eq.php</a>	
10	Earthquake Intensity	<a href="http://www.ngdc.noaa.gov/hazard/icons/usaint.jpg">http://www.ngdc.noaa.gov/hazard/icons/usaint.jpg</a>	

Part 3: Mapping Data to the Hypotheses (add rows as needed)

Hypothesis	Data	Source
Hypothesis 1: People living in the Philadelphia metropolitan area have lower average incomes than people living in the San Francisco, CA metropolitan area	Median Household Income in Metropolitan Areas	Data Set #1
	Median Household Income	Data Set #2
Hypothesis 2: San Francisco receives more sunshine average per day than Philadelphia	Annual Possible Sunshine	Data Set #3
	Average Annual Sunshine	Data Set #4
Hypothesis 3: The cost of housing in San Francisco is higher than the cost of housing in Philadelphia	Housing and Transportation as a Percent of Income	Data Set #5

MIS 0855 Data Science – Spring 2015 – Assignment #1

	Median House Listing Price	Data Set #6
Hypothesis 4: San Francisco has a younger population compared to Philadelphia	Median Age by County	Data Set #7
	Median Age for U.S. Counties	Data Set #8
Hypothesis 5: Philadelphia has a lesser chance of experiencing an earthquake compared to San Francisco	Damaging Earthquakes in the U.S. (1750-1996)	Data Set #9
	Earthquake Intensity	Data Set #10