

**MIS 0855 Data Science (Section 005) – Fall 2016**  
**Answer Sheet for Assignment #1 – Data Analysis Plan**

Name: CONG XUAN NGO

Part 1: Develop Hypotheses

<b>Hypothesis (1 sentence)</b>	<b>Rationale (2-3 sentences)</b>
Hypothesis 1: UK Women living in a neighborhood that has higher density of population experience lower birth weight when giving births than other women.	Women living in a highly populated neighborhood are more likely to be exposed with toxic air pollutants from traffic roads, smoking, and fuels burning, etc. These pollutants are so harmful to individuals that they can debilitate women's health, which leads to births with lower birth weight.
Hypothesis 2: In EU Referendum, UK residents living in an area that has higher density of population cast more votes REMAIN than others in less populated areas.	UK residents living in more populated areas tend to have more access to international economic integration. For example, they rely on multinational firms to get jobs, enjoy goods imported from other EU countries, and earn money from services for EU foreigners. This idea encourages them to value the importance of EU, which results in more votes REMAIN.
Hypothesis 3: A Philadelphia district that is more sanitary experiences lower school dropout rates than others.	The more sanitary a district is, the more educated and informed families live in it. Because educated parents are more likely to ensure that their children get high education, school dropout rates are lower.
Hypothesis 4: In the UK, an area with higher local smoking rates experiences higher suicide mortality rates.	When people are stuck in dilemmas or difficulties, they tend to find smoking as solutions to their problems. Moreover, cigarettes contain a great amount of nicotine, which is harmful to mental health of smokers, and suicide cases are likely to happen when people are in bad moods such as depression or melancholy.
Hypothesis 5: A London neighborhood with more public toilets experiences higher happiness score.	Toilets are essential for people and the demand of using toilets takes place everywhere in every time. Therefore, a neighborhood with more public toilets makes people feel more comfortable and

	prevents the inconvenience and trouble of find toilets when they go outside, which leads to higher happiness score.
--	---

## Part 2: Identify Data Sets

<b>Dat ase t</b>	<b>Name</b>	<b>URL</b>	<b>Instructions (if URL is not available)</b>
1	UK Population Density	<a href="https://data.gov.uk/dataset/population_density">https://data.gov.uk/dataset/population_density</a>	
2	UK Emission of Air Pollutants	<a href="https://data.gov.uk/dataset/emissions_of_air_pollutants">https://data.gov.uk/dataset/emissions_of_air_pollutants</a>	
3	UK Births with Low Birth Weight	<a href="https://datahub.io/dataset/births-low-birthweight">https://datahub.io/dataset/births-low-birthweight</a>	
4	EU Referendum Result	<a href="https://data.gov.uk/dataset/eu-referendum-23-june-2016-wdc">https://data.gov.uk/dataset/eu-referendum-23-june-2016-wdc</a>	
5	Philadelphia Sanitation Districts	<a href="https://www.opendataphilly.org/dataset/sanitation-districts">https://www.opendataphilly.org/dataset/sanitation-districts</a>	
6	Philadelphia School Dropout Rates	<a href="https://www.opendataphilly.org/dataset/school-dropout-rates-2004-2012">https://www.opendataphilly.org/dataset/school-dropout-rates-2004-2012</a>	
7	UK Smoking Status	<a href="https://files.datapress.com/london/dataset/smoking-indicators-borough/2016-09-02T10:49:48/smoking-status.xls">https://files.datapress.com/london/dataset/smoking-indicators-borough/2016-09-02T10:49:48/smoking-status.xls</a>	
8	UK Suicide Mortality Rates	<a href="http://data.london.gov.uk/dataset/suicide-mortality-rates-borough">http://data.london.gov.uk/dataset/suicide-mortality-rates-borough</a>	
9	London Happiness	<a href="http://data.london.gov.uk/dataset/subjective-personal-well-being-borough">http://data.london.gov.uk/dataset/subjective-personal-well-being-borough</a>	In the dataset, data of London happiness are above and emboldened.
10	London Public Toilets	<a href="https://datahub.io/dataset/public-toilets12149">https://datahub.io/dataset/public-toilets12149</a>	

## Part 3: Mapping Data to the Hypotheses

<b>Hypothesis</b>	<b>Source</b>
Hypothesis 1: UK Women living in a neighborhood that has higher density of population experience lower birth weight when giving births than other women.	1
	2
	3
Hypothesis 2: In EU Referendum, UK residents living in an area that has higher density of population cast more votes IN than others in less populated areas.	1
	4
Hypothesis 3: A Philadelphia district that is more sanitary experiences lower school dropout rates than others.	5
	6
Hypothesis 4: In the UK, an area with higher local smoking rates experience higher suicide mortality rates.	7
	8
Hypothesis 5: A London neighborhood with more public toilets experiences higher happiness score.	9
	10