Regression Modeling in Practice

**Assignment – Week 1**

**Writing About Your Data**

**By** [mapolarbear@gmail.com](mailto:mapolarbear@gmail.com)

**About My research**

This is my third course and my major interest is demographics combined in **GAPMINDER** dataset [www.gapminder.org](http://www.gapminder.org) project.

My main research/study question is which factors affect suicide rates in different countries and possible affect and correlation of variables like: employment, breast cancer, urban rates, geographical region, and if there is a difference between female and male statistics.

**Sample**

The sample is from the GAPMINDER dataset combined various collections of observational data provided by [www.gapminder.org](http://www.gapminder.org) for 2013 countries (N=213). The GAPMINDER project collects various country-level demographic data on health, wealth and development like: geographical region income per person, urban rate, employment rate, breast cancer rate, female and mail suicide ratings, short and long unemployment rates, politic score and many other. The combined final GAPMINDER data to study female and male suicide rating, contains sample data for 2002 and 2004, with the most reported countries across the globe.

**Procedure**

Data were collected during 2002 and 2004 by various sources, including the US Census Bureau’s International Database, Institute for Health Metrics and Evaluation, United Nations Statistics Division, International Labor Organization and the World Bank. Each country presented with corresponding rating, including breakdown by female and male gender.

**Measures**

Country represented by unique name (N=213) and represent data for 2002 and 2004.

Region ID (explanatory variable) created form country geographical locations based on World Bank grouping and representing following regions:

ASIA (EX. NEAR EAST) - group 1

WESTERN EUROPE - group 2

WESTERN EUROPE - group 3

ASTERN EUROPE - group 4

LATIN AMER. & CARIB - group 5

NEAR EAST - group 6

NORTHERN AFRICA - group 7

NORTHERN AMERICA - group 8

OCEANIA - group 9

SUB-SAHARAN AFRICA - group 10

C.W. OF IND. STATES - group 11

Total Female or male suicides per country for 2002 and 2004 combination of time series collected by WHO Global Bureau of diseases and Valance and Injury Prevention.

Female or male suicides 2002 and 2004 ratings per 100,000 standard population, age adjusted, collected by WHO Global Bureau of diseases and Valance and Injury Prevention.

Female or Male 2002 and 2004 employment rates (% of total population) of employed female or male for age 15+, that has been employed during the given year collected by International Labor Organization.

Breast cancer per 100th represents 2002 breast cancer new cases per 100,000 female residents.

Urban rate (% of total country population) for 2002 and 2004, population living in urban area based and defined national statics offices collected by World Bank

Urban rate Group categorical (explanatory variable) created from urban ratings, representing % of urban population in country and grouped as:

Group 1 – Less or equal % 36.84 of Total percent of population

Group 2 – Less or equal % 59.88 of Total percent of population

Group 3 – Less or equal % 77.20 of Total percent of population

Group 4 – Greater then % 77.20 of Total percent of population

Employment rate categorical (explanatory variable) created from 2004 female employment rate and grouped as:

Group 1 – Less or equal % 37.7 of Total percent of female country population

Group 2 – Less or equal % 46.2 of Total percent of female country population

Group 3 – Less or equal % 53.5 of Total percent of female country population

Group 4 – Less or equal % 53.5 of Total percent of female country population

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* End of week 3 assignment \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*