README: Emissions Levels and Life Satisfaction:

Collaborative Data Analysis

Spring 2016

By Katie Levesque, Meerim Ruslanova, and Sarah Unbehaun

1. Background

This project focuses on the relationship between objective and subjective aspects of climate on the reported well-being. The research centers on the German Bundeslaender and their residents from 1990 to 2012. The aim of the project is to analyse the relationship between the objective measures of GHG emissions (CO_2 per squared kilometer) of each State and subjective individual concerns about the environment on one hand, and individual life satisfaction levels on the other. The demographic characteristics (age, gender, employment status, income, family status), as well as the time-component, also present an interest for this research. In addition to the simple Fixed Effects Model, a multilevel analysis is applied in order to analyse the relationships between the climate-factors and well-being on both State and individual levels.

2. Data

The final data set is composed of 1) State-level data of German Bundeslaender (Emissions in total, CO_2 tons per capita & squared kilometer) and 2) individual-level data (life satisfaction, gender, age, employment status, family status, environmental concerns). The former is composed via web-scraping from the following sources:

- Statista.com

- Environmental-Economic Accounting of the Bundeslaender

- Agency for Renewable Energy of North Rhine-Westphalia

- Federal Statistical Office and Statistical Offices of the Bundeslaender

- Laenderarbeitskreis

The individual-level data is provided by the GSOEP conducted by the German Institute for Economic Research (DIW). Due to confidentiality restrictions, DIW could only supply a shortened sample with specified variables in a *.dta* format in the *Data* folder of the repository. The short dataset contains the information on reported levels of life satisfaction, subjective concerns about the environment, age, gender, income, employment, family status, and state residence of a respondent.

All raw files are distributed among the Data sub folders (Energy, Emissions, GSOEP), while the new intermediary and final data frames are saved in the parent *Data* folder. The code for data gathering and transformation is located in the *Data[SourceFile.R]*. SourceFile.R combines both State- and individual-level data into a final data frame (All_Merged_Data.csv). This data frame is used for descriptive statistics and analysis, whose R code is located in the *Data[Analysis.rmd]*, which was the initial file submitted after the proposal.

3. Final output

Final_Paper folder contains all the analysis files, graphics, and bibliography that construct the Final Research Paper, "Emissions Levels and Life Satisfaction: A Study on the German States". The main file linking all the parts is the *Final_Paper*[*Final_Paper.rmd*] file.

Final_Presentation_Files contain the files used to create a Presentation , but the html file is located in the parent folder.

Website for the project can be found here .

4. File structure

Climate-Happiness

|- Data (raw & clean files, SouceFile for transforming data, final dataset)

    ||- Energy (data on State renewables)

    ||- Emissions (data on State CO2 per capita)

    ||- GSOEP (.dta file with life satisfaction & demographic factors; codebook for variables)

    |||- Analysis files (rmd, pdf, html) > Assignment 3

|- Final_Paper

|- Final_Presentation_Files

|- Index.rmd & Index.html (*gh-pages* branch) > Website

- note: Folder "Data" also contains all raw files for the time being, until we resolve our issue of local paths on a Mac.