

An Introduction to R
R Workshop 3: World Bank Data
QS311 Section E: Baker University
October 29nd, 2019

Contents

1	Introduction	2
2	Lab Procedures	2
2.1	Obtaining the Variables of Interest	2
3	Analyzing the Data	3
3.1	Questions	3

1 Introduction

Exploratory Data Analysis with One and Two Variables

R Workshop in NAVCO/World Bank Conflict Data

QS311 Section E: Baker University

Goal: To explore development indicators in countries with conflict.

In this workshop, we will use a merged dataset between the NAVCO database and World Bank development indicators.

A few highlights to review from the previous R sessions:

Loading World Bank data: This panel is already constructed for you, so you do not need to load in World Bank data with the `wbstats` package. Rather, you should import the *full panel.dta* file on Moodle. This is a Stata dataset, so when you go to import on R, make sure you choose the Stata option rather than, say, excel.

2 Lab Procedures

1. In this workshop we want to focus on what development looks like in countries with conflict. Following literature on economic development across countries, we use World Bank indicators on the services, agriculture, trade, and industry in addition to indicators on life expectancy, infant mortality, the number of telephone subscriptions per 100 people, population density, and the percent that live in rural areas.
2. The "treatment" variables in the dataset include *treated*, *nv_treated* and *v_treated*.
 - (a) *treated*: this takes the value 1 if a country ever had a violent or non-violent revolution and is zero otherwise.
 - (b) *nv_treated*: this takes the value 1 if a country ever had a non-violent revolution and is zero otherwise.
 - (c) *v_treated*: this takes the value 1 if a country ever had a violent revolution and is zero otherwise.
3. Outcome Variables of Interest
 - (a) *unemployment*: unemployment rate in a country.
 - (b) *male_unemployment*: unemployment rate of men in a country.
 - (c) *female_unemployment*: unemployment rate of women in a country.

Summary of Variables

These variables are summarized in the following table below.

Variable	Description
<i>CountryCode</i>	The ISO3 Code for Country
<i>year</i>	Year, 1960 - 2018
<i>ag</i>	Share of GDP which is agriculture
<i>services</i>	Share of GDP which is services
<i>trade</i>	Share of GDP which is trade
<i>industry</i>	Share of GDP which is manufacturing
<i>telephone_subscriptions</i>	Number per 100 people
<i>infant_mortality</i>	Number per 1,000 births
<i>life_expectancy</i>	Life expectancy at birth
<i>population_density</i>	People per sq. km of land
<i>rural_pop</i>	Percent living in rural areas
<i>revolution</i>	binary outcome; 1=revolution of any kind
<i>nv_treated</i>	binary outcome; 1=non-violent revolution
<i>v_treated</i>	binary outcome; 1 = violent revolution
<i>unemployment</i>	General unemployment rate
<i>male_unemployment</i>	Male unemployment rate
<i>female_unemployment</i>	Female unemployment rate

3 Analyzing the Data

3.1 Questions

Please answer the following questions and submit them on Moodle for Homework Assignment 5 (Due: October 29). In order to answer the question, you should choose one of the countries which has an active and ongoing conflict (major war or war) which was presented in class.

1. Obtain scatterplots for each of the primary variables (ag, services, trade, industry, telephone subscriptions, infant mortality, life expectancy, population density, and rural population) with year on the x-axis.
2. From these macro variables, what impact does the war seem to have had on economic development in the country?
3. Merge in the wb classifications data and include in your homework submission the code used.
4. Use a column plot to get the number of country-year combinations that had a conflict by region. Approximately what percentage of the country-year combinations had conflicts? Note: the total number of country-year combinations is the length of the data frame.
5. Which regions of the world have seen the most conflict since the 1960s? Are there moral implications that some regions have been more afflicted than others? What would the utilitarians and cosmopolitans say about the unequal distribution of warfare?