

Title: GDP (Current Prices, US Dollars) and Average Annual Growth Rates for Countries in the United Nations (UN)

Part One: Introduction

I. Research Interest

The project aims to make gross domestic product (GDP) and its breakdown for country members of the United Nations internationally comparable at current prices in nominal US dollars. Based on the information, we were interested in identifying the poorest and richest countries in 2014 by their GDP and countries with fastest and lowest growth. The final objective was to predict GDP in 2020 for the United States based on its average GDP growth in recent years.

II. Description of the Data Sources

Two data sets were selected for this project. Both are publically accessible at the UN website.

Dataset #1 GDP is accessible through the following link:

<http://unstats.un.org/unsd/snaama/downloads/Download-GDPcurrent-NCU-countries.xls>

The data set contains information on GDP and its breakdown at current prices in nominal National currency for all countries compiled by the United Nations. The raw data are in EXCEL format.

Dataset #2 ExRates is accessible through the following link:

<https://treasury.un.org/operationalrates/OperationalRates.php>

The data set contains information of operational exchange rates for one United States Dollar (USD) for all UN countries as the day of August 1, 2016. Even though the website claims the raw data should be downloaded in EXCEL format, the downloaded file was in tab-delimited Text format.

Part Two: Methods for Managing the Data

I. Reading Data into SAS

Data preparation involved in this project included reading data into SAS by using IMPORT and INFILE procedures, renaming variables, defining variables' length and format, and adding necessary labels.

II. Data Validation

Data were checked for errors through the PRINT, FREQ, and MEANS procedures. Data errors were cleaned by using IF-THEN/ELSE statements in the DATA step.

III. Merging Data

In the project, the two data sets could be merged based on the variable Country or Currency variable. Before merging the two data sets, the SQL procedure was used to compare the two data sets to see which variable could be a more efficient choice. The final decision was to merge the two files based on Country by using the SORT procedure and MERGE statement.

IV. Creating New Variables

With the merged data, new variables were created to convert GDP and its breakdown from its national currency to nominal US dollars.

V. Sub-setting Data

GDP only observations from the year of 2009-2014 were selected by using the WHERE statement to create a subset for future analysis.

VI. Sorting Data

The subset were then sorted by data in 2014 in ascending or descending order and a list of first ten observations were printed as the poorest or richest countries.

VII. Predicting

Annual growth rates from 2009-2014 were calculated and the average annual growth rates (AAGR) were obtained. The data were sorted and countries with the fastest and lowest AAGR were printed. GDP over the next five years were also estimated based on AAGR by using the DO loop.

Part Three: Results and Summary

I. Data Validation

The first data set GDP contains information from 220 countries. It contains missing values but it is understandable since some countries joined the UN later than 1970. The data set has 3,709 records with 48 variables with three character variables: Country; Currency; Indicator Names. The other 45 variables are numeric, listing data from Year 1970 to Year 2014.

The second data set ExRates has 220 records from 220 countries with five variables. The first three variables (Country; Currency; Currency Code) are character variables. The fourth variable Effective Date is August 1, 2016 in our project. The last variable (Operational Rate) is numeric.

The two data sets were successfully merged into a data file named as 'Merge_New' and GDP and its breakdown were thus converted from its national currency to nominal US Dollars based on the exchange rate as of the date of August 1, 2016 (data file named as 'GDPDollar'. A subset was created to contain only GDP data in nominal US Dollars from 2009 to 2014 (data file name 'GDP_Only').

II. Data Analysis

The richest and poorest countries in 2014 were identified as the following tables showed.

Richest Countries in 2014

Obs	Country	Y2014
1	Belarus	\$391,183,676,884,422.00
2	United States	\$17,348,071,500,000.00
3	China	\$9,624,457,795,133.68
4	Japan	\$4,654,914,558,472.55
5	Germany	\$3,236,015,538,290.79
6	United Kingdom	\$2,390,051,315,789.47
7	France	\$2,366,758,046,614.87
8	India	\$1,870,705,250,596.66
9	Italy	\$1,791,186,574,028.86
10	Brazil	\$1,688,974,020,816.46

Poorest Countries in 2014

Obs	Country	Y2014
1	Liberia	\$25,112,426.04
2	Tuvalu	\$31,347,418.36
3	Montserrat	\$63,085,185.19
4	Kiribati	\$150,481,036.12
5	Nauru	\$151,680,304.74
6	Marshall Islands	\$208,814,181.00
7	Palau	\$233,500,000.00
8	Cook Islands	\$265,820,786.52
9	Sao Tome and Principe	\$280,122,161.61
10	Micronesia (FS of)	\$308,000,000.00

Annual GDP growth rates from 2009 to 2014 and their averages were calculated. Countries with the fastest and lowest average annual GDP growth rates were identified as following.

Countries with Lowest GDP Growth Rates

Obs	Country	Average Annual Growth Rate	Growth Rate from 2009 to 2010	Growth Rate from 2010 to 2011	Growth Rate from 2011 to 2012	Growth Rate from 2012 to 2013	Growth Rate from 2013 to 2014
1	Greece	(5.62%)	(4.84%)	(8.41%)	(7.64%)	(5.66%)	(1.57%)
2	San Marino	(3.90%)	(5.03%)	(8.53%)	(5.15%)	0.23%	(1.00%)
3	Somalia	(3.08%)	(13.15%)	(0.36%)	0.49%	(0.65%)	(1.73%)
4	French Polynesia	(2.13%)	(2.75%)	(2.92%)	(0.95%)	(1.81%)	(2.23%)
5	Andorra	(1.21%)	(3.83%)	(2.42%)	(0.68%)	(0.06%)	0.96%
6	Bermuda	(1.15%)	(1.41%)	(4.01%)	(0.64%)	0.99%	(0.69%)
7	Cyprus	(1.13%)	3.44%	2.25%	(0.40%)	(7.21%)	(3.71%)
8	Barbados	(1.09%)	(3.36%)	(1.98%)	(0.61%)	0.90%	(0.42%)
9	Spain	(0.70%)	0.17%	(0.97%)	(2.57%)	(1.11%)	0.96%
10	Portugal	(0.19%)	2.55%	(2.09%)	(4.41%)	1.11%	1.87%

Countries with Fastest GDP Growth Rates

Obs	Country	Average Annual Growth Rate	Growth Rate from 2009 to 2010	Growth Rate from 2010 to 2011	Growth Rate from 2011 to 2012	Growth Rate from 2012 to 2013	Growth Rate from 2013 to 2014
1	Belarus	44.23%	19.67%	80.67%	78.48%	22.39%	19.93%
2	Venezuela (Bolivarian Republic of)	35.55%	43.77%	33.50%	20.48%	37.32%	42.69%
3	Sudan	34.31%	11.93%	40.31%	24.43%	57.84%	37.05%
4	Ethiopia	26.39%	14.18%	35.86%	45.09%	15.70%	21.13%
5	Argentina	25.52%	28.29%	27.68%	19.62%	23.17%	28.85%
6	Ghana	25.24%	25.81%	29.92%	25.91%	24.03%	20.55%

Obs	Country	Average Annual Growth Rate	Growth Rate from 2009 to 2010	Growth Rate from 2010 to 2011	Growth Rate from 2011 to 2012	Growth Rate from 2012 to 2013	Growth Rate from 2013 to 2014
7	Nauru	24.66%	(-2.04%)	22.20%	39.90%	35.96%	27.29%
8	Uzbekistan	24.04%	26.35%	26.25%	24.33%	21.50%	21.75%
9	Mongolia	23.74%	27.67%	35.02%	26.68%	14.90%	14.41%
10	Iran (Islamic Republic of)	23.11%	21.88%	31.03%	13.75%	31.77%	17.12%

Based on the average annual GDP growth rates, GDP in 2020 for the United States was predicted.

Predicting US GDP in 2020 based on AAGR

Obs	Country	Y2014	Average Annual Growth Rate	PRED
1	United States	\$17,348,071,500,000.00	3.77%	\$18,001,906,996,295.10

III. Summary

We were able to achieve all of our objectives by practicing the data management methods and techniques listed above. For future data exploration, the data can be subset to compare the financial statistics among different regions of the world. Also the data can be merged with other data files such as population to calculate GDP per capita.

The variable “Label” is read in as character variable. Other variables are read in as numeric variables. To validate “Label”, we used the “proc freq” statement; while to validate other numeric variables, we used “proc means” statement with modifier “min max nmiss n”. “min” and “max” are to make sure pixels are within the range of gray-scale values (from 0 to 255); “nmiss” and “n” are to check for missing values.