ELSA Populations' Census Data Summary

This dataset contains population census results of a total of 4175 cities in 163 countries over the years of 1993 to 2014. The dataset is spread very unevenly such that different countries and cities report varying information for different time frames. For example, Switzerland shows 5 cities' statistics for 2000-01 and 11-13 cities's statistics between 2006-12 as can be seen on page 1 of the report.

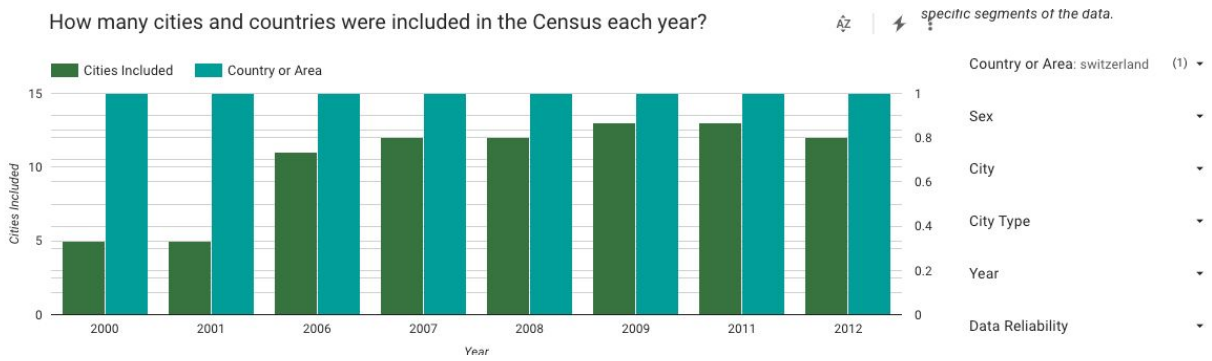

Figure 1. Switzerland data quality

From the report page 3 we can see that there are 120 cases (City-Year combinations) where only Female populations statistics are included. There are further 157 cases where Female population proportion is between 60-99% which is also unrealistic. There are 143 cases where only male population statistics were collected.

The most populous city at any one time is Ciudad de Mexico in 2010 with a total population 20,1M people. Tlalnepantla is showing the same population figure, but it's most likely a fluke in the data because Mexico City is the biggest city in Mexico. The least populous city in the dataset is Vaduz in 2010 which is also a data fluke as is shown by the World Population Review . In 2019, there were approximately 5450 people living in the capital of Liechtenstein.

Pages 4-5 of the report show how (un-)evenly are populations spread across different countries. We can measure this by calculating the population Gini coefficients for each country. Gini coefficient is a measure of inequality and it ranges from 0 to 1. If a given country's Population Gini coefficient is 1, it means that all of that country's population lives in only one city while all other cities have zero inhabitants. If a country's Gini is 0, it means that the population is perfectly evenly spread across all the cities.

The highest Gini coefficient shows up for Peru (0,79) in 2007. At the time, Peru had a total population of 9,8M and we can verify on page 3 that 86% of the people lived in its capital Lima which is indeed a very high proportion. Even though many cities in Peru are missing either female or male population statistics, quick googling shows that we can more or less trust this conclusion.

City Populations by Year

PS! Population Proportion figure is only meaningful if you choose ONLY one Year from the filters on the right!

	City	Year	Population	Population Proportion ▼	% Female
1.	lima	2007	8,472,935	86.38%	51.15%
2.	chiclayo	2007	250,081	2.55%	0%
3.	ica	2007	219,856	2.24%	51.52%
4.	juliaca	2007	216,716	2.21%	50.77%
5.	cuzco	2007	168,077	1.71%	0%
6.	tacna	2007	119,055	1.21%	0%
7.	pucallpa	2007	102,972	1.05%	100%
8.	huaraz	2007	100,931	1.03%	51.3%
9.	pisco	2007	99,550	1.01%	49.57%
10.	tarapoto	2007	58,215	0.59%	0%

Figure 2. Peru cities' populations

Page 5 of the report is asking the question whether people over the world are moving from smaller cities to larger cities. If we average countries' Ginis over the years and the Global Gini would be significantly increasing, then we could conclude that populations gravitate towards bigger cities.

However, given the inconsistent data, we can notice that the changes in the global Gini are more affected by adding and removing some countries than anything else. We can notice a more stable period of data between 2005-2011 and zooming in to this period we can see that the Gini stays rather constant between 0,29-0,31 so we cannot find any evidence for changing structures of population dynamics.

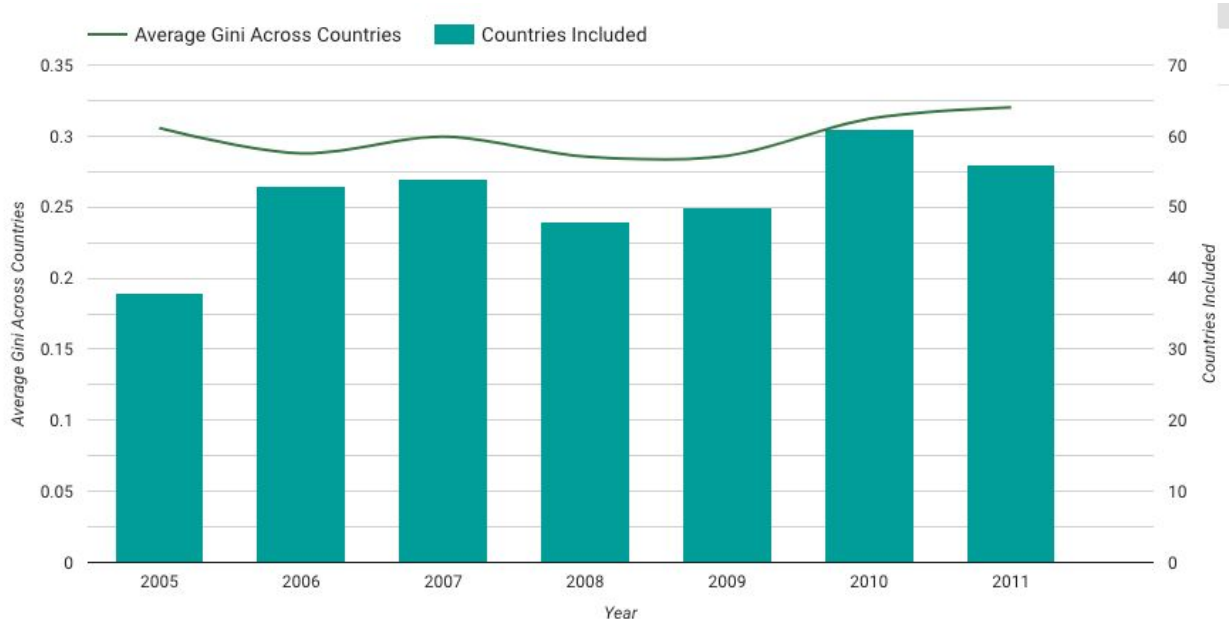

Figure 3. Global Average Cities' Population Gini Index between 2005-2011

Author: Edgar Rootalu