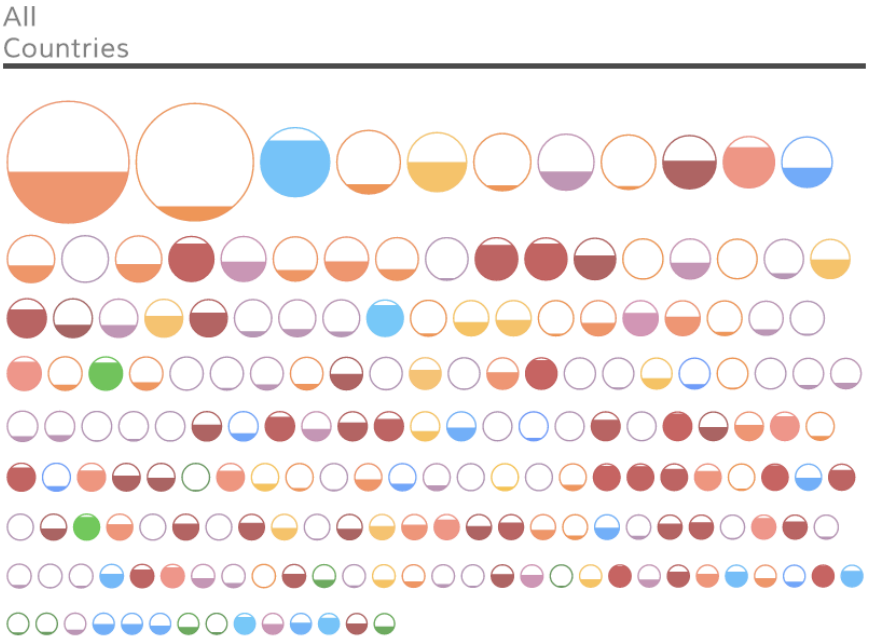
World Population & Internet Usage Growth (1990-2012)



Bowls Visualization of Internet Usage Over Time

Josh Feldman
Lincoln Cohen
David Hau

CS/INFO 3300 - Data Driven Web Applications

Project 2

Data Description


```
[{"Date":"12/31/12","Percent":42.3001,"Country":"China","Continent":"Asia","Population":1350695000,"Year":2012},
{"Date":"12/31/11","Percent":38.3,"Country":"China","Continent":"Asia","Population":1344130000,"Year":2011},
{"Date":"12/31/10","Percent":34.3,"Country":"China","Continent":"Asia","Population":1337705000,"Year":2010},
{"Date":"12/31/09","Percent":28.9,"Country":"China","Continent":"Asia","Population":1331260000,"Year":2009},
{"Date":"12/31/08","Percent":22.6,"Country":"China","Continent":"Asia","Population":1324655000,"Year":2008},
{"Date":"12/31/07","Percent":16,"Country":"China","Continent":"Asia","Population":1317885000,"Year":2007},
{"Date":"12/31/06","Percent":10.5232,"Country":"China","Continent":"Asia","Population":1311020000,"Year":2006},
{"Date":"12/31/05","Percent":8.5233,"Country":"China","Continent":"Asia","Population":1303720000,"Year":2005},
{"Date":"12/31/04","Percent":7.3,"Country":"China","Continent":"Asia","Population":1296075000,"Year":2004},
{"Date":"12/31/03","Percent":6.2,"Country":"China","Continent":"Asia","Population":1288400000,"Year":2003},
{"Date":"12/31/02","Percent":4.5957,"Country":"China","Continent":"Asia","Population":1280400000,"Year":2002},
{"Date":"12/31/01","Percent":2.6397,"Country":"China","Continent":"Asia","Population":1271850000,"Year":2001},
{"Date":"12/31/00","Percent":1.7759,"Country":"China","Continent":"Asia","Population":1262645000,"Year":2000},
```


Our data consisted of the continent, population, and percentage of Internet users for 188 countries over a span of 22 years, from 1990 to 2012. We got Internet use statistics from Quandl (http://www.quandl.com/society/internet-users-by-country) and population data from the World Bank website, filling in any data gaps with the last known population of the country (http://data.worldbank.org/indicator/SP.POP.TOTL). We manually combined the data from the two sources into an Excel file, extracted the year from each line of data, and sorted the data first by population and then by country in order to make it easier to display in the final graphic. To get the data for whole continents, we simply combined the data of each of its countries. For population, these are just summed, and for Internet use, we summed the Internet use percentage multiplied by the population of each country and then divided by the population we computed for the whole continent. Finally, we used a data converter tool (http://shancarter.github.io/mr-data-converter/) to obtain the data in a JSON file.

The range of our data was constrained by the available data on Quandl - the widest set of Internet use data spanned the years 1990-2012, so this became our range. For any countries that lacked data going back to 1990, we simply filled in "0%" as the usage statistic for each missing year. This seemed like a reasonable assumption because the site's data layout was such that it generally only displayed non-zero values. Similarly, for any country that had a "gap" in the Internet use data, we simply filled in the last known statistic from previous years. The country background images are clippings taken from a world map: (http://www.mynchoice.com/images/world.png).

Mapping & Transformations

In the final display each country is sorted by its respective continent (note that Antarctica was not included in our data because it has no permanent population, which also means it has no measurable Internet use statistics.) The data is pre-sorted by population so a user can easily compare similarly sized countries.

We used a few different scales in out project. One such scale was color. The color of each circle depends on which continent it is in. This becomes more relevant in the sandbox at the bottom of the page, where countries from various continents are all in the same svg. The color key is available directly above "All Countries" SVG rendering.

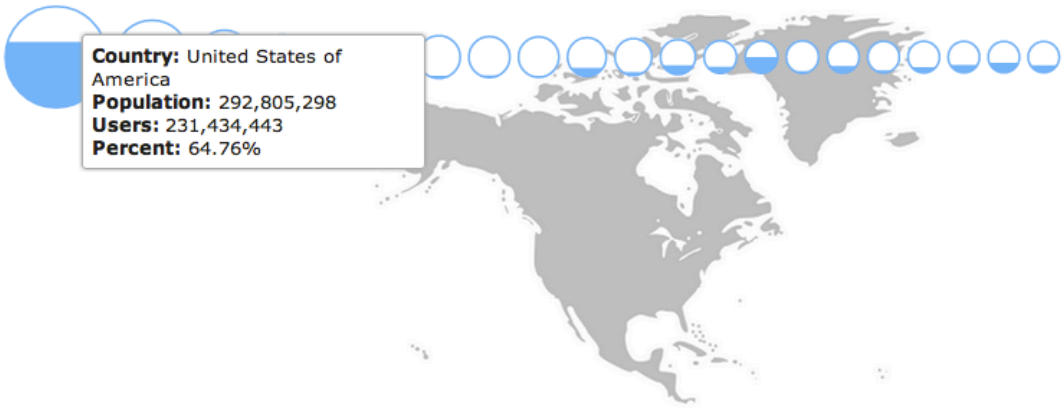
Another scale is the level of circle fill. This corresponds to the percentage of people in each country/continent who use the Internet. The fraction of the circle that is filled is the percentage, so a full circle would represent 100% Internet use. This "liquid bowl" styling was inspired by an example provided on the D3 Gallery Examples page (http://invision-web.net/ticket-status/). In our project we modified and expanded on the functionality of the project on this site.

Finally, the size of each circle was scaled according to the population of the country/continent. This was accomplished via an rScale, which transformed a country's population statistic into the radius of its corresponding circle. This is a square root scale, so that the area of each circle corresponds to the population rather than the radius.

Visualization Stories

Our visualization yielded many interesting results. One basic and expected result was that the proportion of Internet users in a country/continent tended to grow over time, as did the population (with the exception of a few countries afflicted by war / famine / disease in the past 22 years). Beyond that, each continent seems to have its own traits regarding Internet use of time. It is worth mentioning that the data is sorted by decreasing population.

North America

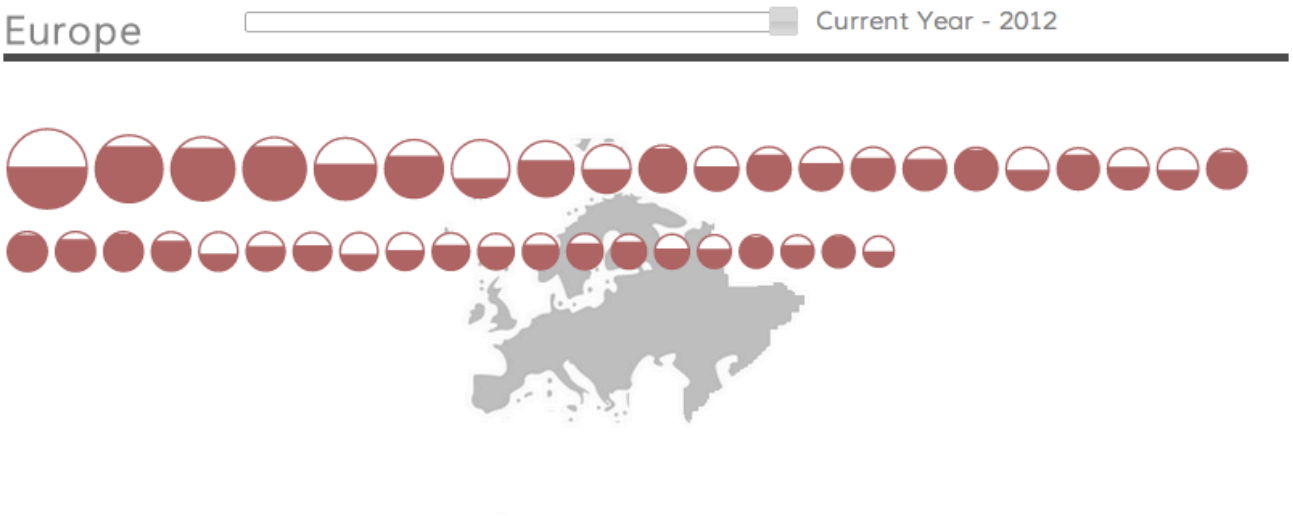


North America has two clear standouts, the US and Canada, both of which have high populations and Internet use. Other than that, the other North American countries are fairly varied in terms of population and Internet use. One interesting fact, however, is that as of 2012, no North American country has an Internet use proportion lower than 10%. There are a few islands with very high Internet use, the highest of which are Antigua and Saint Kitts and Nevis, both of which have around 80% usage. Looking at the increase of Internet use in North America over time, we see that only the US and Canada are high in Internet use by 2000; Internet use in other North American countries begins to rise only after 2000.

Asia is the population king, with China and India both topping a billion people. In terms of Internet use, the continent is hit or miss. There are a decent amount of countries with over 70% Internet use, such as Japan and Qatar. There are a few countries around 50% Internet use, but in general all countries in Asia either have very high

Internet use or very low Internet use. China is an outlier, with a huge population and a usage statistic near 50%. India has surprisingly low Internet use considering their standing in the world economy. Finally, an interesting fact about Asia is that the countries with higher Internet use tend to be lower population countries.

Africa is very clearly the most depressing continent in terms of Internet use. A majority of the countries have little to no Internet use, even in 2012. Nigeria, South Africa, and Malaysia are among the few reassuring outliers, although it is clear that much still needs to be done to improve Internet accessibility in this continent. Progress has already been made, however: in 2000, practically no countries in Africa had any Internet use at all.



Europe is by far the continent with the most consistently high Internet use. By 2012 there are very few countries with less than 50% Internet use. Over time, Internet use in select countries seems to rise somewhat sporadically. For example, Ukraine and Albania are very low in 2006 (possibly because of their former membership in the Soviet Union), but by 2012 both countries possess a solid base of Internet users.

Australia, like North America, has two main standouts: Australia and New Zealand. These countries have over 80% Internet use while the next highest country is Tuvalu, with 35%. This could result from the fact that these countries are island nations, but recall that islands did fairly well in North America in terms of Internet use. Like the trend over time in North America, Australia and New Zealand were the only two coun-

tries on the Australian continent to have Internet early on, with the other countries picking up the slack around 2008.

The "All Countries" section at the bottom of the page is what we unofficially call "the Sandbox." The controls in the pop-up div at the bottom allow uses to play around with the data and decide which countries, population brackets, etc., they want to see. This population-ordered list of all the countries in the world makes it abundantly clear that there is no strong correlation between population and Internet use.

One interesting thing to do is look at all the countries over time. The US stands out, reaching a high proportion of Internet use far before any other country. This is reasonable - after all, the Internet has its roots in various institutions across America. After that, we see that a selection of countries follows: Canada, Australia, New Zealand, and a decent number of European countries, including Sweden and Finland, shoot up after the US.

Another interesting view occurs when you set the minimum Internet use percent to a very high value. At 85% we have a good amount of European countries, 3 Asian countries, then Canada and New Zealand. At 90%, it's only European countries left. Just a few years earlier there were no countries this high. When we do the opposite and set the maximum to be very low, we see a good amount of countries, a majority from Africa and Asia, as well as a couple from Australia.

One further observation comes from deselecting all continents except for Europe and North America. The result is that North America sandwiches Europe. Especially early on, the US and Canada seem to be the best of the group, and the rest of North America comes up as the worst. This trend comes to an end in recent years, with the US and Canada seemingly plateauing before 90%, while European countries rise past them.