UNIVERSITY OF CALIFORNIA

SANTA CRUZ

Population in 2100. What the millennials need to know

A report submitted in partial satisfaction of the

requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER ENGINEERING

by

Bharath Nagesh

March 2017

The Project

is approved:

_____

Professor Suresh K. Lodha, Chair

_____

Professor Jishen Zhao

Table of Contents

Web links

Visualization: https://bharathn18.github.io/project/

GitHub Repository: https://github.com/bharathn18/project

List of tables and figures

Acknowledgement

I have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals. I would like to extend my sincere thanks to all of them.

I would like to express my special thanks of gratitude to Professor Suresh K Lodha, who gave me the golden opportunity to do research on the topic “Data Driven Visualization”, and helped me in gaining a lot of insights towards Data Journalism and how it affects the current era. I am thankful for all the guidance provided during working on this proposal.

I would also like to thank Professor Jishen Zhao for her willingness to read my project report and provide valuable feedback on this project.

Abstract

In the current era of data explosion, almost anything can be expressed in numbers. Data Driven Visualization uses sophisticated libraries and frameworks to process a seemingly abstract set of data and transform it into a compelling story that could not be told otherwise. This has become the basis for Data Journalism in presenting abstract data into compelling news that attract readers and helps them gain sufficient understanding of the underlying data.

Over the last century the world has witnessed a massive surge in the human population. The changing population and demographic outline of the planet has been the prime reason for many of the major decisions today. Environmental policies, in many cases, are formulated based on the impact of population growth. Companies make investment decisions based on estimated and current population figures. Local Governments need population figures in order to implement sustainable policies and provide necessary awareness through education. Though predictions aren't always accurate, they provide a good indication of where the situation is headed.

In order to make predictions, the primary ingredient is good data. We need good data which takes into account as many scenarios as possible and then gives acceptable numbers. One such dataset is the one provided by the United Nations – Population Division. The dataset provides data beginning from 1950 and extrapolates it all the way to 2100. Since we are interested in the trends that one can expect to see and not a year by year variation, the dataset was reduced to contain data for every 2 decades. Using the dataset other relevant information such as the percentage share for every continent and percentage increase in population was obtained.

The objective of the project is to provide a convincing visualization on as to where the world is headed by the end of the 21st century using the JavaScript library D3.js. D3.js Version 3 offers flexibility in building visualizations in any way the user desires and is not constrained to a set of visualization forms. It has the capabilty to handle large amount of DOM functions which helps in making the visualization much more interactive.

A multi-line chart, a set of donut graphs and a bar graph was designed to drive home the fact that massive growth can be witnessed in Africa over the coming decades. Colorful sections and innovative shapes have been used in order to make the chart more appealing to the viewer.

1. Introduction

In the current era of data explosion, Data Science has become one of the most pursued fields. As quoted by Stephen Few, “Numbers have an important story to tell. They rely on you to give them a clear and convincing voice.”. This can mainly be attributed to the fact that excess data available to us needs to be processed and filtered as per requirements. Data Science helps us discover useful data and visualization transforms it into an exciting story.

As more and more data is available, we need to modify the data into a more simplified way. Processing of large data sets in short amount of time has become possible due to the increase in processing power of computers. This caused the people to turn to visualization techniques to process data efficiently. Visualization helps in generating simple charts to help viewers understand the story of the underlying data. As data is visualized, the underlying aspects become visible to the viewers and data can be processed more easily.

Hence, Data Driven Visualization has become a major solution for data analysis and information extraction. This in turn helps in creating beautiful storylines that attract the attention of viewers and helps them relate to the data easily.

The major reasons that state the importance of visualization include:

- It helps in making the data into an integral part in decision making at various levels of any organizations.
- It transforms the way people see the world, by using the data available and converting it into interesting visualizations that can be understood by viewers even without technological expertise.
- It improves the value of data by helping data scientists uncover hidden patterns and gain insights on the available data.

2. Background and Related Work

2.1. Computational Social Science

With the rapid advances in technology, unprecedented volumes of data have been created revealing both individual and group’s behavior in our society. For instance, people send emails to communicate with their colleagues in daily work, make online transactions to get the merchandise that they need, connect with each other and maintain friendship through social networks. All of such activities would generate digital traces that can be translated to comprehensive pictures to assist the understanding of our lives and societies. With thrilling new opportunities brought by increasing volumes of data, a new discipline known as Computational Social Science has been turning into more prominent to move us towards a quantitative comprehension of our intricate social systems.

Computational Social Science is defined as "the interdisciplinary investigation of the social universe on many scales, ranging from individual actors to the largest groupings, through the medium of computation". In general, it refers to an interdisciplinary that makes computational approaches to social sciences. It empowers collecting and analyzing data with an epoch-making breadth and depth and scale. Computational Social Science involves with a board range of fields such as computational economics, computational sociology, anthropology, human geography, psychology and law, and focuses on investigation of social and behavioral relationships through the combinations of algorithmic information extraction, network models, social complexity analysis and various computer simulations.

Numerous studies have been done in the field of computational social science. Onnela et al. have figured out the local and the global structure of a society-wide communication network by exploring the communication patterns of millions of phone users. Balcan et al. demonstrated the relationships between human mobility with the spatiotemporal pattern of a global epidemic by analyzing mobility data from 29 countries and integrating a timescale-separation technique in a worldwide-structured meta-population epidemic model. Batty provides new insights into how cities evolve by linking urban economics and transportation behavior to developments in network science, allometric growth and fractal geometry. Chainey et al. summarized the relations between Geographical Information Systems (GIS) and practical criminal justice issues and demonstrated crime mapping can play an important role in crime reduction process.

Lillo et al. identified groups of players on the market by subtracting individual strategies from detailed financial data and also demonstrated the players' role in stabilization and destabilization. Dzogang et al. discovered periodic changes in the behavior of a large population by analyzing historical news of United States and United Kingdom from 1836 to 1922. In general, computational social science has demonstrated powerful capability to assist unveiling the inner patterns of various problems such as society communication network, international epidemics, urban development, crime issues, trends of stock market and the periodic patterns of human behaviors. It is essential to apply the insights of computational social science to get a better understanding of the increasingly complexity of our interconnected global society.

2.2. Data Journalism

Now we are living in a digital world that almost everything can be described with numbers, which is transforming traditional journalism into a new media approach – Data Journalism. Data Journalism can be simply explained as journalism done with data, which can help a journalist tell a compelling story through engaging info-graphics. Raw data is always confusing, boring and puzzling to the public's mind. Therefore, It has a growing value to gathering, filtering, visualizing the raw data and re-organizing it into a more comprehensive way. Currently, pioneers like the New York Times, the Guardian and the Texas Tribune have demonstrated how data- driven stories can provide deeper insights into the inner patterns of our rapidly changing global society. Many visualization tools have been created and accelerated the development of Data Journalism.

2.2.1 Data Driven Documents – D3

Data-Driven Documents (D3) is a JavaScript library for making representation-transparent approach to dynamic and interactive data visualizations for the web. According to the official definition, "D3.js is a JavaScript library for manipulating documents based on data. D3 helps you bring data to life using HTML, SVG, and CSS. D3's emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM manipulation". Specifically, D3 enables users to directly inspect and manipulate DOM by selectively binding input data to arbitrary document elements and dynamically transforming the generation and modifications of content. The feature of representational transparency makes D3 more expressive and better integrate with developer tools. D3 also retains features of high efficiency, powerful declarative components and simplified debugging. D3 is the successor of Protovis. In 2009, Bostock et al. presented Protovis - a graphic toolkit for visualization by composing simple graphical primitives. In 2011, Bostock et al. released the first version of D3 with comparable notational efficiency. Currently, many data visualization sites use D3 as the core toolkit to build compelling visualizations. The New York Times is one of the best examples of D3 usage in digital media. Fig. 2 demonstrated a D3 example of an interactive map of the paths to the white house.

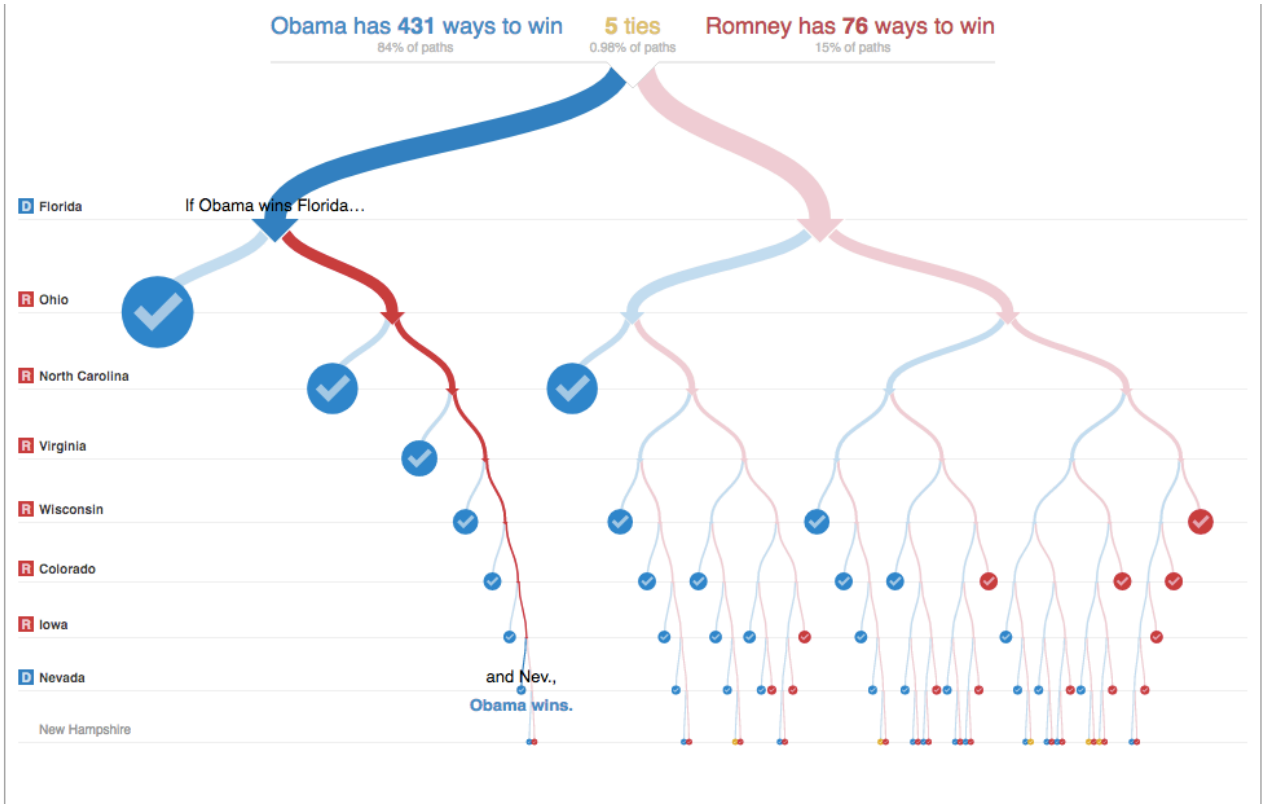

Fig.1. 512 Paths to the White House Visualization by New York Times

2.2.2 Visualization Tools

As more and more data is being available over the web, we need visualization tools to create interactive visualizations to make it more readable. Also, we can observe interesting patterns that would help us to build machine learning algorithms that can use the data effectively. This means that the need for data visualization tools and designers to transform the data is constantly increasing.

As author, data journalist and information designer David McCandless said in his TED talk: "By visualizing information, we turn it into a landscape that you can explore with your eyes, a sort of information map. And when you're lost in information, an information map is kind of useful."

The major advantages of visualizations include:

- Ease in explaining the data to people
- Comprehend the data effectively
- Find useful patterns and relationships
- Discover trends that could help in customer satisfaction

We have a wide variety of visualization tools available, some of which could be used by even non-programmers to make interesting visualizations. Below is a comparison of few popular visualization tools.

Advantages	• Has more than 1000 example visualizations • It is available for free • Has	• It is available for free • Easy to use • Has clear documentation	• Easy to use • Can be used by non-programmers • It is available for free • Has personalized tech support and community forum	• Can also work on old browsers • Has personalized tech support and community forum • Has more
Disadvantages	• No pre built charts • No personalized tech support • Time	• Only has 6 chart types	• Hard to make visualizations attractive	• It is not a free software

Table 1. Comparison of Popular Visualization tools

2.2.3 Best Data Visualization Tools

There are many tools available for visualization. But, not all tools can provide to the requirements of a particular visualization. Following tools are taken from the web site titled "The 38 Best Tools for Data Visualization".

The tools that cost money and therefore are not suitable for us include:

- InstantAtlas - only useful for mapping visualizations.
- FusionCharts - offers different kinds of charts with total control over visualizations. Also has more than 800 live examples.
- ZingChart - JavaScript library with over 100 pre-built charts
- Modest Maps - interactive mapping tool with wide range of visualizations
- Excel - mainly used for offline static charts.

The tools that are not open source and not interesting include:

- Tableau - real time updating with drag and drop of data, with option for team collaboration.
- HighCharts - rendered using SVG, broad range of chart options and free for non-commercial use
- Google Charts - highly flexible and includes built-in animation and user interaction controls.

The tools that are cheap and can be used to make quick visualizations, but do not allow flexibility to the visualizations include:

- Chart.js - equipped with only 6 chart types and allows for quick interactive visualizations.
- RAW - built on top of D3 and enables quick visualizations with limited control.
- Digraphs - interactive charts with good flexibility, but only for dense datasets.
- Timeline - only used for making interactive timelines
- Exhibit, Leaflet - interactive mapping tools
- JqPlot - JQuery plugin for line and point charts

The tools that we are mostly interested in are those that are free, open-source, allow interactive visualization on the web and also allows users to have greater control on the details of the visualization. This category of tools includes:

- D3 - JavaScript library that uses HTML, CSS, and SVG to create interactive visualizations with full control over them. It can work on even complex datasets to create attractive visualizations.
- R - statistical package to parse large data sets. It is complex tool but provides great control over the visualizations.
- Processing - Interactive visualizations that are compiled using Java. It also has a JavaScript library to use on websites without need for Java applets.
- Gephi - visualization and exploration tool. Allows users to crunch large data sets, clean the data and sort data

2.3. Background

As per the demographics definition, World Population is the total number of humans living on the planet. After the great famine of 1315 the world has seen a continuous growth in numbers reaching to the modern-day figure of 7.5 billion. By the end of the century, this number is expected to go even further and reach a figure of 12.6 billion with Asia slowing down it's growth and Africa speeding up its population expansion.

Population figures raise awareness on critical issues such as environmental sustainability, over-population and scarcity of resources. Following are some of the situations where knowledge of population is necessary,

- A planner has to look into the correlations between the size of population and the economic development of the country. How are the changes in population size influencing the economic development and vice a versa?

- From the political perspective, one has to look into the size of population while determining the boundaries of constituent areas as well as the number of members of the National Assembly, the Senate, and Provincial Assemblies. The information about the composition as well as the distribution of population helps in the determining the number of voters, the type of voters, and the location of voters

- A businessman will be interested in the population because he uses the manpower in the production of goods and services and ultimately look for consumer of goods and services. At various stage of business operation interests will be in size, composition and distribution of population. Production of goods and services are in accordance with the characteristics of the ultimate consumer.

- An educationist has to look into the educational needs of the population and arrange the facilities accordingly. The size of population, school going age population, gender, occupational background of their parents,

All the above reasons lead us to extrapolate population figures and understand where the millennials might see themselves by the end of the century. Some of the conclusions that have been formulated are,

- Population growth in Asia will slow down and the growth is expected to take a downward turn very soon
- By the end of the century Africa and Asia will be home to about 80 percent of the world's population
- Countries like Nigeria, Congo and Tanzania will see an exponential growth.

2.4. Related Visualizations

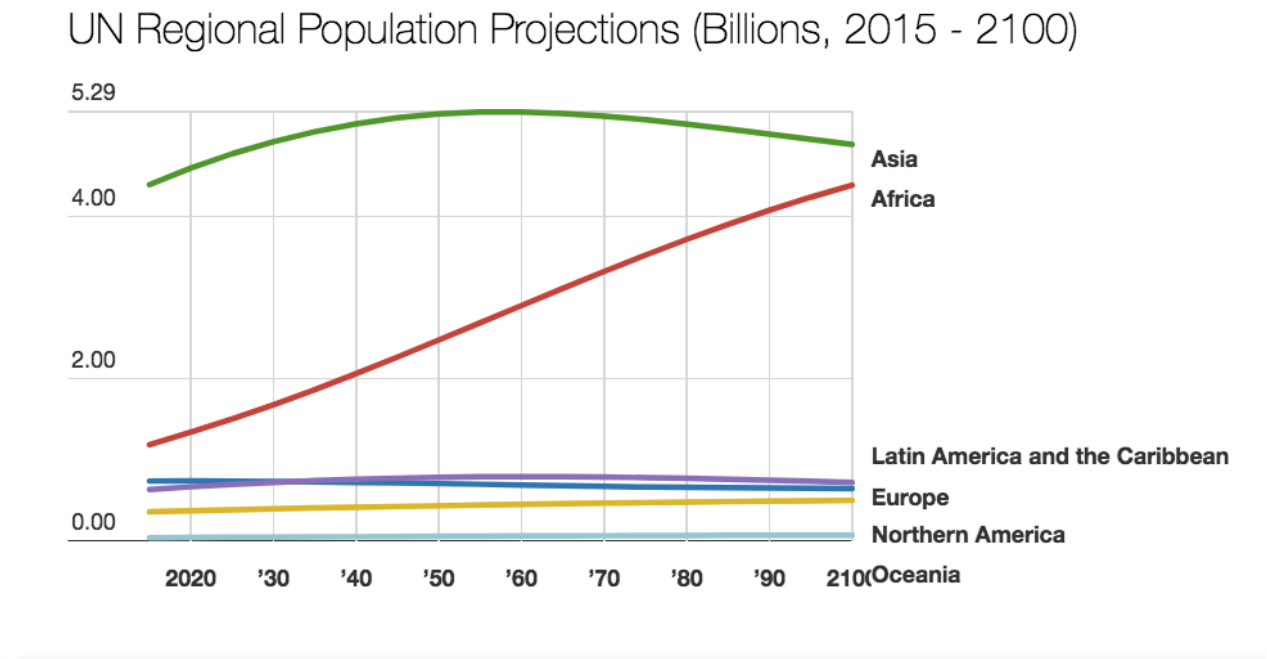

Fig.2. http://blogs.worldbank.org/opendata/future-world-s-population-4-charts

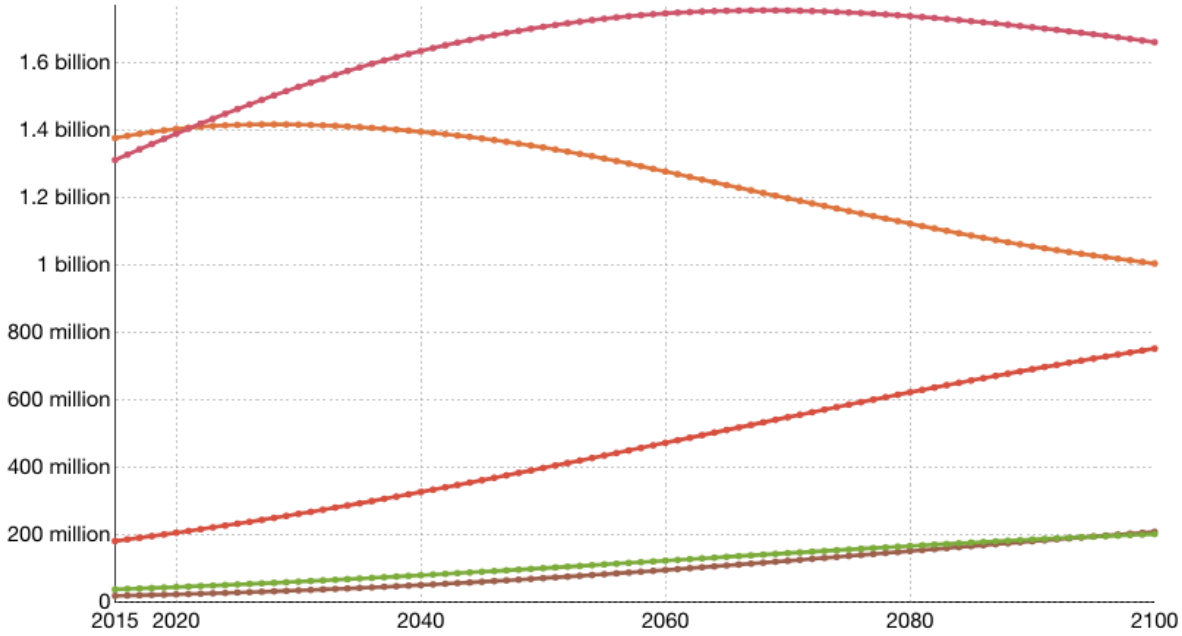

Fig.3. https://ourworldindata.org/future-world-population-growth/

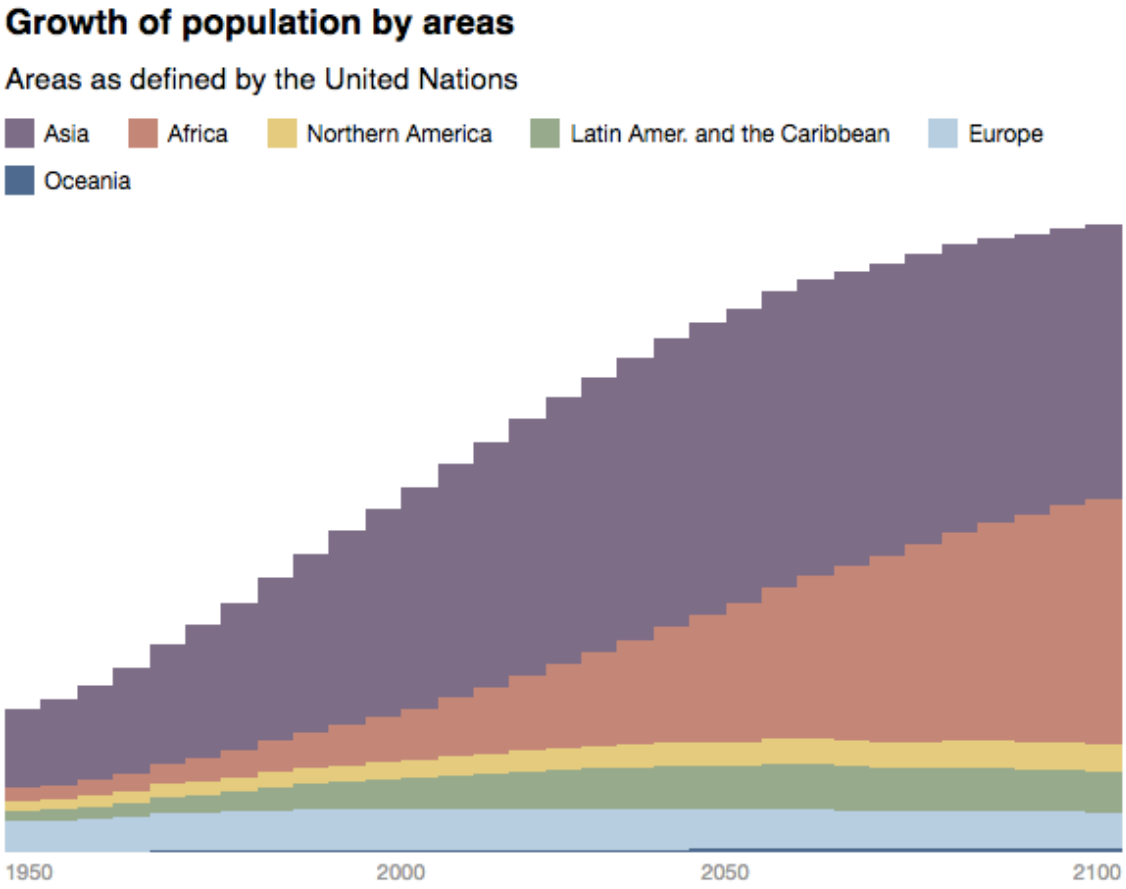

Fig.4. http://news.nationalgeographic.com/2015/07/world-population-expected-to-reach-9-7-billion-by-2050/

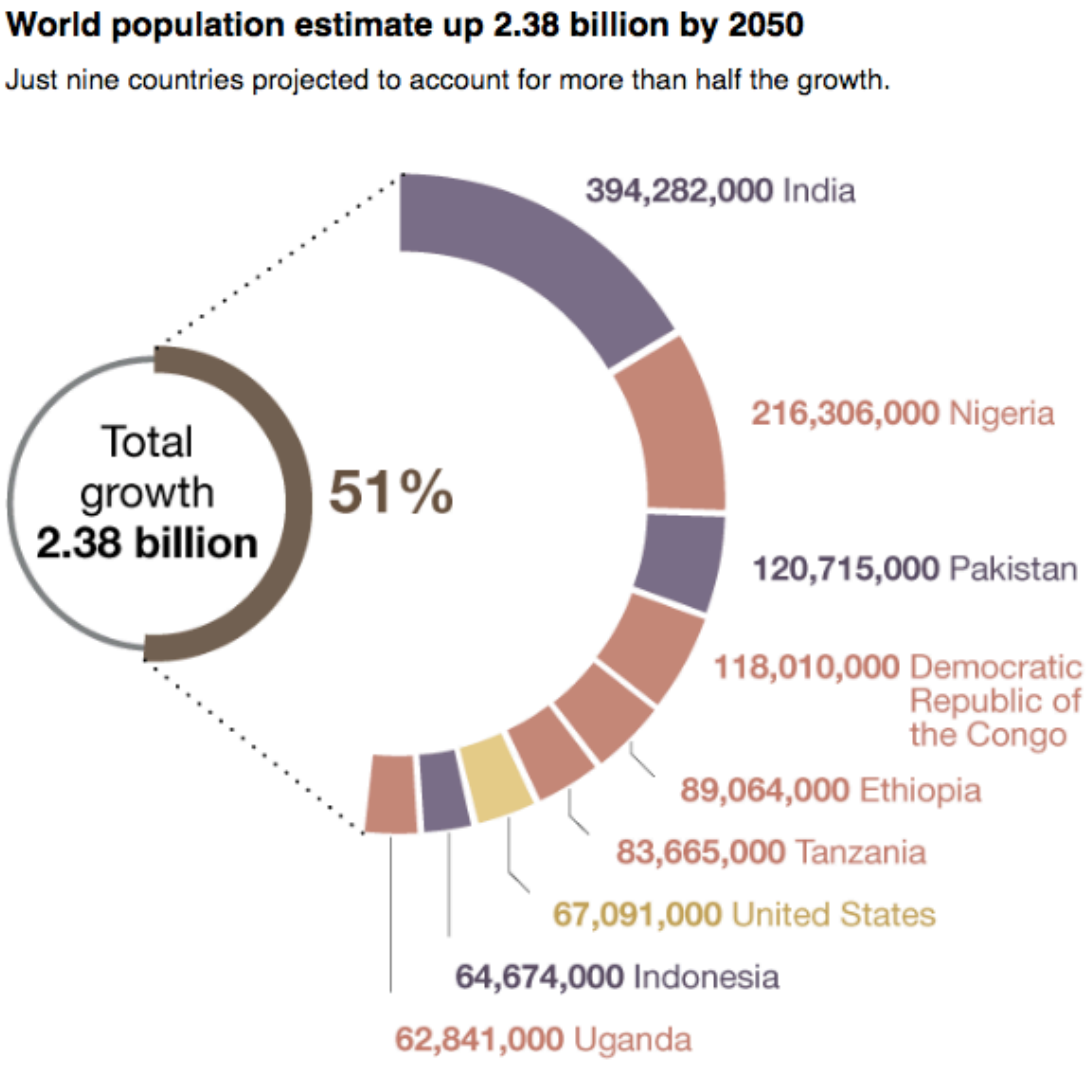

Fig.5. http://news.nationalgeographic.com/2015/07/world-population-expected-to-reach-9-7-billion-by-2050/

3. Research

3.1. Purpose of Visualization

The world is growing constantly. Currently there are about 7.5 billion on the planet and the number is set to go up to approximately 12 billion. Asia and Africa will be the domicile of 80 percent of the planets population.

The information that has been visualized in this project will drastically impact on the lives of millennials who will be living through the century. It will better assist them in understanding how they need to shape decisions and how policies need to be put in place so that it can avoid problems such as exhaustion of resources.

3.2. User Profile

Our target audience is the millennial generation which refers to everyone born in the mid-1980s and early 1990s who are interested in understanding the direction in which the world is headed and as to where they can see themselves in coming years of the 21st century.

3.3. Data Design

3.3.1. Data Curation

Data curation is a broad term used to indicate processes and activities related to the organization and integration of data collected from various sources, annotation of the data, and publication and presentation of the data such that the value of the data is maintained over time, and the data remains available for reuse and preservation. The data was obtained from the website of UN's Department of Economic and Social Affairs. The datasets used for the project are 'Population Expectancy.xls' and 'Population-Growth.xls'.

Index	Variant	Major area, region, country or area *	Notes	Country code	Total population, both sexes combined, as of 1 July (thousands)							
					1950	1951	1952	1953	1954	1955	1956	1957
1	Estimates	**WORLD**		900	2 525 149	2 571 868	2 617 940	2 664 029	2 710 678	2 758 315	2 807 246	2 857 663
2	Estimates	More developed regions	a	901	812 989	822 320	832 149	842 294	852 613	863 004	873 402	883 775
3	Estimates	Less developed regions	b	902	1 712 161	1 749 547	1 785 792	1 821 735	1 858 064	1 895 310	1 933 845	1 973 884
4	Estimates	Least developed countries	c	941	195 725	199 494	203 322	207 278	211 417	215 777	220 378	225 225
5	Estimates	Less developed regions, excluding least developed countries	d	934	1 516 436	1 550 053	1 582 470	1 614 458	1 646 648	1 679 534	1 713 467	1 748 660
6	Estimates	Less developed regions, excluding China		948	1 158 315	1 180 688	1 204 646	1 230 102	1 256 979	1 285 212	1 314 748	1 345 543
7	Estimates	High-income countries	e	1503	800 383	810 059	820 445	831 323	842 519	853 902	865 388	876 935
8	Estimates	Middle-income countries	e	1517	1 593 830	1 628 936	1 662 526	1 695 456	1 728 429	1 761 999	1 796 568	1 832 384
9	Estimates	Upper-middle-income countries	e	1502	824 937	846 501	865 620	883 202	899 993	916 581	933 386	950 670
10	Estimates	Lower-middle-income countries	e	1501	768 893	782 434	796 906	812 254	828 436	845 419	863 182	881 714
11	Estimates	Low-income countries	e	1500	130 103	132 018	134 093	136 353	138 810	141 467	144 318	147 344
12	Estimates	**Sub-Saharan Africa**	f	947	179 680	183 055	186 598	190 308	194 187	198 234	202 455	206 854
13	Africa	**AFRICA**		903	228 902	233 449	238 231	243 244	248 488	253 963	259 671	265 614
14	Africa	**Eastern Africa**		910	66 923	68 374	69 875	71 432	73 050	74 734	76 490	78 321
15	Africa	Burundi		108	2 309	2 359	2 404	2 446	2 488	2 532	2 578	2 628
16	Africa	Comoros		174	156	160	164	167	170	173	176	179
17	Africa	Djibouti		262	62	63	65	66	68	70	71	74
18	Africa	Eritrea		232	1 142	1 163	1 185	1 208	1 233	1 259	1 286	1 315
19	Africa	Ethiopia		231	18 128	18 467	18 820	19 184	19 560	19 947	20 348	20 764
20	Africa	Kenya		404	6 077	6 240	6 412	6 593	6 782	6 980	7 186	7 401
21	Africa	Madagascar		450	4 084	4 168	4 256	4 348	4 444	4 544	4 647	4 754
22	Africa	Malawi		454	2 954	3 008	3 065	3 125	3 187	3 252	3 320	3 391
23	Africa	Mauritius	1	480	493	506	521	537	554	571	588	605
24	Africa	Mayotte		175	15	16	16	17	18	19	20	21
25	Africa	Mozambique		508	6 313	6 405	6 503	6 607	6 717	6 832	6 953	7 079
26	Africa	Réunion		638	248	259	268	277	284	292	300	308
27	Africa	Rwanda		646	2 186	2 251	2 313	2 378	2 449	2 526	2 608	2 694
28	Africa	Seychelles		690	36	37	37	38	38	39	39	40
29	Africa	Somalia		706	2 264	2 308	2 352	2 397	2 444	2 492	2 541	2 592
30	Africa	South Sudan		728	2 583	2 601	2 625	2 652	2 685	2 721	2 761	2 805
31	Africa	Uganda		800	5 158	5 309	5 456	5 601	5 748	5 899	6 056	6 221
32	Africa	United Republic of Tanzania	2	834	7 650	7 847	8 056	8 275	8 503	8 741	8 988	9 244
33	Africa	Zambia		894	2 317	2 376	2 439	2 504	2 573	2 644	2 718	2 795
34	Africa	Zimbabwe		716	2 747	2 830	2 918	3 010	3 105	3 204	3 305	3 411

Fig.6. Raw Data obtained from UN

Following are the fields of significance:

- Region : The area under consideration
- Total Population : This contains data for every country beginning from 1950 all the way up to 2100.

	A	B	C	D
1	Country			Percentage Change in Population (2015-2100)
2	Niger			951.9784493
3	Zambia			546.8689872
4	Burundi			460.5366028
5	United Republic of Tanzania			459.436206
6	Angola			454.4628653
7	Somalia			440.5616373
8	Mali			428.3076683
9	Uganda			419.7419153
10	Malawi			405.6889503
11	Democratic Republic of the Congo			403.1045502
12	Senegal			396.005512
13	Chad			391.018105
14	Congo			376.4833681
15	Mozambique			356.2473017
16	Iraq			349.9982498
17	Burkina Faso			347.3209902
18	Gambia			346.8073618
19	Côte d'Ivoire			345.5823997
20	Madagascar			335.3113773
21	Nigeria			312.8645766
22	Guinea			289.0145686
23	Togo			281.5770055
24	Liberia			254.773442
25	Equatorial Guinea			253.165219
26	Cameroon			252.8998128
27	Kenya			240.6191734
28	South Sudan			238.3482909
29	State of Palestine			232.3614866
30	French Guiana			231.7602734
31	Benin			226.6973497
32	Mauritania			221.0426683
33	Sudan			216.4605379
34	Mayotte			213.1733433
35	Eritrea			198.7055144
36	Guinea-Bissau			197.6331178
37	Comoros			192.6052096
38	Sao Tome and Principe			182.3981843
39	Timor-Leste			172.9461961

Fig.7. Percentage Change Data

3.3.2. Data Scraping and Wrangling

After careful analysis of the 'Population Expectancy.xls' dataset and to figure out which years we will be focusing on. We decided to keep the data from the following years:

1950, 1970, 1990, 2000, 2015, 2025, 2050, 2075 and 2100.

The Population data for the years 1950–2015 was obtained from the sheet 'Estimates' and the Remaining data was obtained from the 'Medium Variant' sheet. To make the parsing process easier, data for all the selected years was manually copied to a new sheet named 'Modified Set' in which the population data was entered for the selected years. In the 'Modified Set' sheet we have some rows which contain Population Data for regions like Northern Africa, Northern Asia, Micronesia etc. These rows were redundant in the population count as they will be accounted for when we consider the whole continent as one.

After this manual work, we wrote a script 'Excel.py' which reads the 'Modified Set' sheet and removes the redundant rows in the data set and also creates a new CSV File called 'Project-Data.csv' file which contains the population of all the countries and the continents for the selected years. In the next step we separate the data of all the continents and place it in another sheet called 'Continent'. The first sheet is named 'Country'. This is followed by some manual processing of data to our get the data of percentage population changes between 1950–2015 and 2015–2100. This prepares our dataset which is to be used for our Visualization.

Country	1950	1970	1990	2000	2015	2025	2050	2075	2100
Burundi	2 309	3 457	5 613	6 767	11 179	15 177	28 668	45 868	62 662
Comoros	156	228	415	548	788	981	1 502	1 978	2 307
Djibouti	62	160	588	723	888	1 003	1 186	1 201	1 126
Eritrea	1 142	1 805	3 139	3 535	5 228	6 585	10 421	13 685	15 616
Ethiopia	18 128	28 415	48 057	66 444	99 391	125 044	188 455	231 879	242 644
Kenya	6 077	11 252	23 446	31 066	46 050	58 610	95 505	131 802	156 856
Madagascar	4 084	6 576	11 546	15 745	24 235	31 728	55 294	82 037	105 499
Malawi	2 954	4 604	9 409	11 193	17 215	23 134	43 155	66 667	87 056
Mauritius	493	826	1 056	1 185	1 273	1 304	1 249	1 104	952
Mayotte	15	37	95	150	240	308	497	657	752
Mozambique	6 313	9 262	13 372	18 265	27 978	36 462	65 544	99 401	127 648
Réunion	248	462	611	737	861	921	989	946	870
Rwanda	2 186	3 755	7 260	8 022	11 610	14 377	21 187	25 223	25 692
Seychelles	36	52	71	81	96	100	100	90	81
Somalia	2 264	3 445	6 322	7 385	10 787	14 344	27 030	43 139	58 311
South Sudan	2 583	3 647	5 762	6 693	12 340	15 951	25 855	35 446	41 752
Uganda	5 158	9 446	17 384	23 758	39 032	53 497	101 873	157 089	202 868
United Republic of Tanzania	7 650	13 606	25 458	33 992	53 470	72 033	137 136	219 332	299 133
Zambia	2 317	4 185	8 143	10 585	16 212	21 892	42 975	72 182	104 869
Zimbabwe	2 747	5 206	10 485	12 500	15 603	19 370	29 615	37 128	40 263
Angola	4 355	6 301	11 128	15 059	25 022	34 016	65 473	103 977	138 738
Cameroon	4 466	6 771	12 070	15 928	23 344	29 530	48 362	67 876	82 382
Central African Republic	1 327	1 829	2 938	3 726	4 900	5 942	8 782	11 216	12 515
Chad	2 502	3 645	5 958	8 343	14 037	19 075	35 131	53 205	68 927
Congo	808	1 335	2 386	3 109	4 620	5 983	10 732	16 527	22 015
Democratic Republic of the Congo	12 184	20 010	34 963	48 049	77 267	104 536	195 277	300 399	388 733
Equatorial Guinea	226	291	377	531	845	1 102	1 816	2 517	2 984
Gabon	473	590	952	1 232	1 726	2 116	3 164	4 024	4 466

Fig. 8. Final Data used for the visualization

3.4. Visualization Design

The visualization design is a combination of 3 sets of charts. Each of them have been described in the following sections.

3.4.1. Multi-line Chart

A multiline chart was drawn on the logarithmic scale to indicate the changing trend in the top countries contributing towards world population. This shows how the population growth in India and China will slow down in the next few years while African countries will see a massive increase in their population figures.

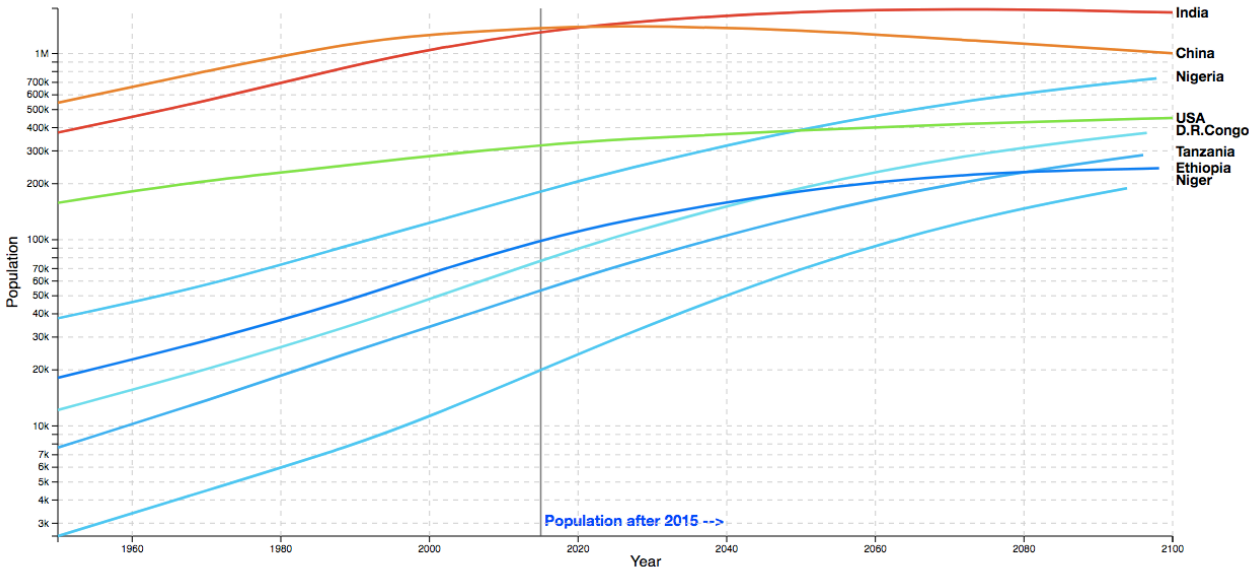

Fig. 9. MultiLine Chart

3.4.2. Donut Charts

A series of donut charts have been created in order to demonstrate as to how the percentage share of Africa is increasing since 1990 while that of Asia has been decreasing.

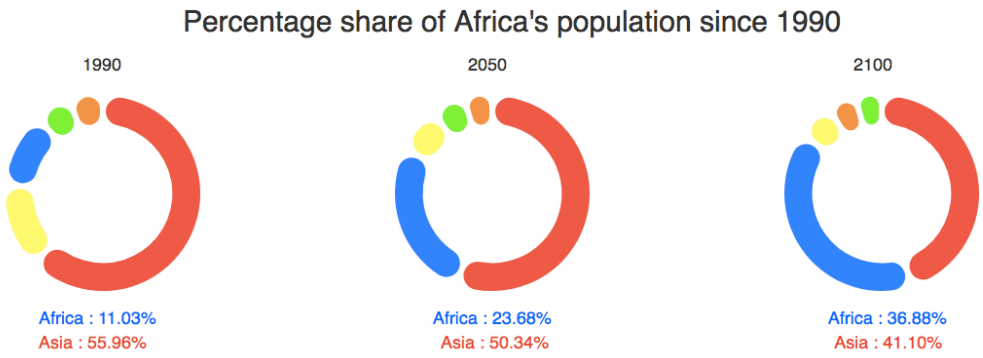

Fig.10. Donut Charts

3.4.3. Bar Chart

A series of bar charts were made in order to emphasize the growth of Africa. Countries like Niger which are seeing over 950 percent growth were made visible with the help of the bar charts.

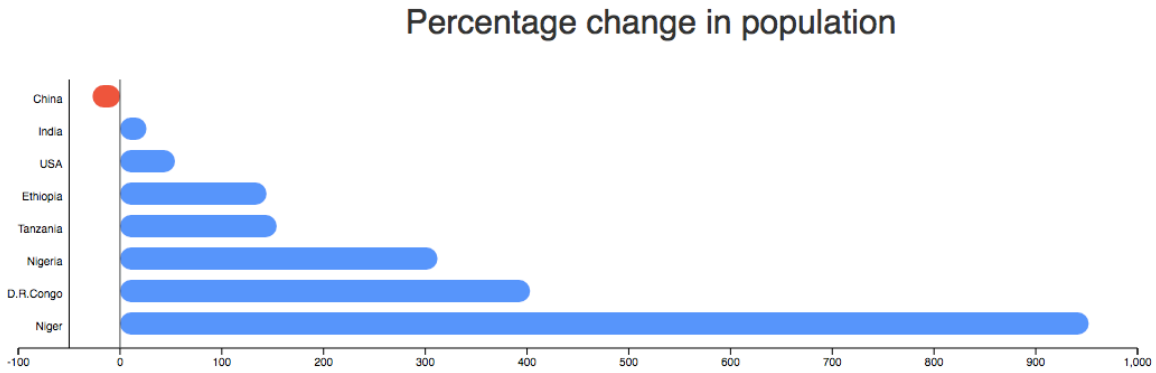

Fig.11. Bar Charts

3.4.4. Color schemes

A consistent color scheme was used throughout the visualization.

- Africa and related countries were colored using shades of blue
- Asia and related countries were colored using shades of red
- North America and related countries were colored using shades of green
- South America and related countries were colored using shades of orange
- Europe and related countries were colored using shades of yellow

4. Conclusions and Future Work

The knowledge of population is vital to the millennial generation as it will help them in making important decisions in many sectors. The visualization gives an outline as to what one can expect in the coming years. Over time the visualization can be extended with the addition of population pyramids of various countries to understand the reasoning behind changing population trends in the countries.

5. Bibliography

[1] Cairo, Alberto. The functional art: An introduction to information graphics and visualization. New Riders, 2012.

[2] Tufte, Edward R., and P. R. Graves-Morris. The visual display of quantitative information. Vol. 2. No. 9. Cheshire, CT: Graphics press, 1983.

[3] Huff, Darrell. How to lie with statistics. New York: Norton 1954.

[4] Yau, Nathan. Data points: Visualization that means something. John Wiley & Sons, 2013.

[5] Nathan, Yau. Visualize this: The flowing data guide to design, visualization, and statistics. Leipzig: John Wiley & Sons Ltd, 2011.

[6] Kirk, Andy. Data visualization: A successful design process. Packt Publishing Ltd, 2012.

[7] Few, Stephen. Show me the numbers: Designing tables and graphs to enlighten. Analytics Press, 2012.

[8] Few, Stephen. Now you see it: Simple visualization techniques for quantitative analysis. Analytics Press, 2009.

[9] Ware, Colin. Visual thinking: For design. Morgan Kaufmann, 2010.

[10] Ware Colin. Information visualization: Perception for design. Elsevier, 2012.

[11] Bertin, J. Semiology of graphics, 1967. Reprinted by University of Wisconsin Press, Madison, WI, 1983.

[12] Carpendale, M. S. T. Considering visual variables as a basis for information visualization. The University of Calgary, Department of Computer Science, 2001-693-16, 2003.

[13] Cleveland, William S., and Robert McGrill. Graphical perception and graphical methods for analyzing scientific data. Science 229.4716 (1985): 828-833.

[14] Mackinlay, Jock. Automating the design of graphical presentations of relational information. ACM Transactions On graphics (Tog) 5.2 (1986): 110-141.

[15] Rogowitz, Bernice E., and Lloyd A. Treinish. Why should engineers and scientists be worried about color. IBM Thomas J. Watson Research Center, Yorktown Heights.

[16] Borland, David, and Russell M. Taylor . Rainbow color map (still) considered harmful. IEEE computer graphics and applications 27.2 (2007): 14-17.

[17] Heer, Jeffrey and Michael Bostock. Crowdsourcing graphical perception: Using mechanical turk to assess visualization design. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2010.

[18] Jeffrey, Heer, Bostock Michael and Ogievetsky VADIM. A tour through the visualization zoo. Communications of the ACM 53.6 (2010): 56-67.

[19] Judelman, Greg. Aesthetics and inspiration for visualization design: Bridging the gap between art and science. Information Visualization, 2004. V 2004. Proceedings. Eighth International Conference on, (pp.245-250). IEEE, 2004.

[20] Brian Suda, Sam Hampton-Smith. The 38 best tools for data visualization http://www.creativebloq.com/design-tools/data-visualization-712402

[21] Jonathan Gray, Liliana Bounegru and Lucy Chambers. The Data Journalism Handbook http://datajournalismhandbook.org/1.0/en/

[22] Troy Thibodeaux. 5 tips for getting started in data journalism http://www.poynter.org/2011/5-tips-for-getting-started-in-data-journalism/147734/

[23] Scott Murray. Interactive Data Visualization for the Web
http://chimera.labs.oreilly.com/books/ 1230000000345/ch02.html

[24] Wikipedia. Data journalism https://en.wikipedia.org/wiki/Data_journalism

[25] Mike Bostock. Data Driven Documents d3js.org

[26] Wikipedia. Computational Social Sciences

[27] WorldBank Statistics
http://blogs.worldbank.org/opendata/future-world-s-population-4-charts

[28] Our World in Data Visualizations
https://ourworldindata.org/future-world-population-growth/

[29] National Geographic
http://news.nationalgeographic.com/2015/07/world-population-expected-to-reach-9-7-billion-by-2050/

Appendix 1:Datasets

Index	Variant	Major area, region, country or area *	Notes	Country code	Total population, both sexes combined, as of 1 July (thousands)							
					1950	1951	1952	1953	1954	1955	1956	1957
1	Estimates	**WORLD**		900	2 525 149	2 571 868	2 617 940	2 664 029	2 710 678	2 758 315	2 807 246	2 857 663
2	Estimates	More developed regions	a	901	812 989	822 320	832 149	842 294	852 613	863 004	873 402	883 779
3	Estimates	Less developed regions	b	902	1 712 161	1 749 547	1 785 792	1 821 735	1 858 064	1 895 310	1 933 845	1 973 884
4	Estimates	Least developed countries	c	941	195 725	199 494	203 322	207 278	211 417	215 777	220 378	225 225
5	Estimates	Less developed regions, excluding least developed countries	d	934	1 516 436	1 550 053	1 582 470	1 614 458	1 646 648	1 679 534	1 713 467	1 748 660
6	Estimates	Less developed regions, excluding China		948	1 158 315	1 180 688	1 204 646	1 230 102	1 256 979	1 285 212	1 314 748	1 345 543
7	Estimates	High-income countries	e	1503	800 383	810 059	820 445	831 323	842 519	853 902	865 388	876 936
8	Estimates	Middle-income countries	e	1517	1 593 830	1 628 936	1 662 526	1 695 456	1 728 429	1 761 999	1 796 568	1 832 384
9	Estimates	Upper-middle-income countries	e	1502	824 937	846 501	865 620	883 202	899 993	916 581	933 386	950 670
10	Estimates	Lower-middle-income countries	e	1501	768 893	782 434	796 906	812 254	828 436	845 419	863 182	881 714
11	Estimates	Low-income countries	e	1500	130 103	132 018	134 093	136 353	138 810	141 467	144 318	147 344
12	Estimates	**Sub-Saharan Africa**	f	947	179 680	183 055	186 598	190 308	194 187	198 234	202 455	206 854
13	Africa	**AFRICA**		903	228 902	233 449	238 231	243 244	248 488	253 963	259 671	265 614
14	Africa	**Eastern Africa**		910	66 923	68 374	69 875	71 432	73 050	74 734	76 490	78 321
15	Africa	Burundi		108	2 309	2 359	2 404	2 446	2 488	2 532	2 578	2 628
16	Africa	Comoros		174	156	160	164	167	170	173	176	179
17	Africa	Djibouti		262	62	63	65	66	68	70	71	74
18	Africa	Eritrea		232	1 142	1 163	1 185	1 208	1 233	1 259	1 286	1 315
19	Africa	Ethiopia		231	18 128	18 467	18 820	19 184	19 560	19 947	20 348	20 764
20	Africa	Kenya		404	6 077	6 240	6 412	6 593	6 782	6 980	7 186	7 401
21	Africa	Madagascar		450	4 084	4 168	4 256	4 348	4 444	4 544	4 647	4 754
22	Africa	Malawi		454	2 954	3 008	3 065	3 125	3 187	3 252	3 320	3 391
23	Africa	Mauritius	1	480	493	506	521	537	554	571	588	605
24	Africa	Mayotte		175	15	16	16	17	18	19	20	21
25	Africa	Mozambique		508	6 313	6 405	6 503	6 607	6 717	6 832	6 953	7 079
26	Africa	Réunion		638	248	259	268	277	284	292	300	308
27	Africa	Rwanda		646	2 186	2 251	2 313	2 378	2 449	2 526	2 608	2 694
28	Africa	Seychelles		690	36	37	37	38	38	39	39	40
29	Africa	Somalia		706	2 264	2 308	2 352	2 397	2 444	2 492	2 541	2 592
30	Africa	South Sudan		728	2 583	2 601	2 625	2 652	2 685	2 721	2 761	2 805
31	Africa	Uganda		800	5 158	5 309	5 456	5 601	5 748	5 899	6 056	6 221
32	Africa	United Republic of Tanzania	2	834	7 650	7 847	8 056	8 275	8 503	8 741	8 988	9 244
33	Africa	Zambia		894	2 317	2 376	2 439	2 504	2 573	2 644	2 718	2 795
34	Africa	Zimbabwe		716	2 747	2 830	2 918	3 010	3 105	3 204	3 305	3 411

[1] https://www.dropbox.com/s/tk0xdj2u8jmb9wn/Population-Expectancy.xls?dl=0

[2] https://www.dropbox.com/s/tk0xdj2u8jmb9wn/Population-Expectancy.xls?dl=0

	A	B	C	D
1	Country			Percentage Change in Population (2015-2100)
2	Niger			951.9784493
3	Zambia			546.8689872
4	Burundi			460.5366028
5	United Republic of Tanzania			459.436206
6	Angola			454.4628653
7	Somalia			440.5616373
8	Mali			428.3076683
9	Uganda			419.7419153
10	Malawi			405.6889503
11	Democratic Republic of the Congo			403.1045502
12	Senegal			396.005512
13	Chad			391.018105
14	Congo			376.4833681
15	Mozambique			356.2473017
16	Iraq			349.9982498
17	Burkina Faso			347.3209902
18	Gambia			346.8073618
19	Côte d'Ivoire			345.5823997
20	Madagascar			335.3113773
21	Nigeria			312.8645766
22	Guinea			289.0145686
23	Togo			281.5770055
24	Liberia			254.773442
25	Equatorial Guinea			253.165219
26	Cameroon			252.8998128
27	Kenya			240.6191734
28	South Sudan			238.3482909
29	State of Palestine			232.3614866
30	French Guiana			231.7602734
31	Benin			226.6973497
32	Mauritania			221.0426683
33	Sudan			216.4605379
34	Mayotte			213.1733433
35	Eritrea			198.7055144
36	Guinea-Bissau			197.6331178
37	Comoros			192.6052096
38	Sao Tome and Principe			182.3981843
39	Timor-Leste			172.9461961

Appendix 2:Code

- Full code available at https://github.com/bharathn18/project

```
<!DOCTYPE html>
<html>
   <head>
        <meta charset="utf-8">
        <title>Final Viz</title>
        <link rel="stylesheet" type="text/css" href="stylesheet.css">
        <script type= "text/javascript" src="https://d3js.org/d3.v3.js"></script>
        <script src="https://d3js.org/topojson.v1.min.js"></script>
        <link rel="stylesheet" href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.6/css/
            bootstrap.min.css">
        <link rel="stylesheet" href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.6/css/bootstrap-
            theme.min.css">
        <script src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.3/jquery.min.js"></script>
        <script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.6/js/bootstrap.min.js"></script>
        <script src="https://ajax.googleapis.com/ajax/libs/jquery/1.10.2/jquery.min.js"></script>
        <script src="moment.min.js" charset="utf-8"></script>
        <script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.6/js/bootstrap.min.js" integrity=
            "sha384-0mSbJDEHialfmuBBQP6A4Qrprq5OVfW37PRR3j5ELqxss1yVqOtnepnHVP9aJ7xS" crossorigin=
            "anonymous"></script>


        <style type="text/css">
            dummydeclaration { padding-left: 4em; }
            tab1 { padding-left: 13em; }
            tab2 { padding-left: 15.5em;}
            tab3 { padding-left: 16em;}
            tab4 { padding-left: 16.5em;}
            tab5 { padding-left: 19.5em;}
        </style>
   </head>
   <body>
       <h1 style = "text-align:center; font-family: sans-serif; font-size:
           35px ;color:#004D4D"><bold>Population in 2100! What the millenials need to know</bold></h1>
       <h1 style = "text-align:center; font-family: sans-serif; font-size: 25px">By : Bharath Nagesh</
           h1>
       <p style="text-align:center; font-family:sans-serif; font-size:20px"> Africa will be the center
           of population growth in the coming century.</p>
       <p style="text-align:center; font-family:sans-serif; font-size:20px"> Countries like Nigeria,
           Tanzania and Ethiopia will witness an exponential increase in population.</p>
       <p style="text-align:center; font-family:sans-serif; font-size:20px"> India's and China's
           population will stabilize around 2020 and 2060 respectively and then decline by the end of
           the century.</p>
       <p>
           <script type="text/javascript" src="Multiline.js"></script>
       </p>
       <p style="text-align:center; font-family:sans-serif; font-size:30px"> <br><br>Percentage share
           of Africa's population since 1990</p>
       <p style="font-family:sans-serif; font-size:18px">
            <tab2>1990</tab2>
            <tab5>2050</tab5>
            <tab5>2100</tab5>
       </p>
       <p><script type="text/javascript" src="Pie.js"></script></p>
       <p style="font-family:sans-serif; font-size:18px">
           <span style="color:#0358FF;"><tab1>Africa : 11.03%</tab1>
           <tab2>Africa : 23.68%</tab2>
           <tab2>Africa : 36.88% </tab2></span><br>
           <span style="color:#FF0D00;"><tab1>Asia : 55.96% </tab1>
```