

Data Analytics

Assignment 9 – Data Visualization

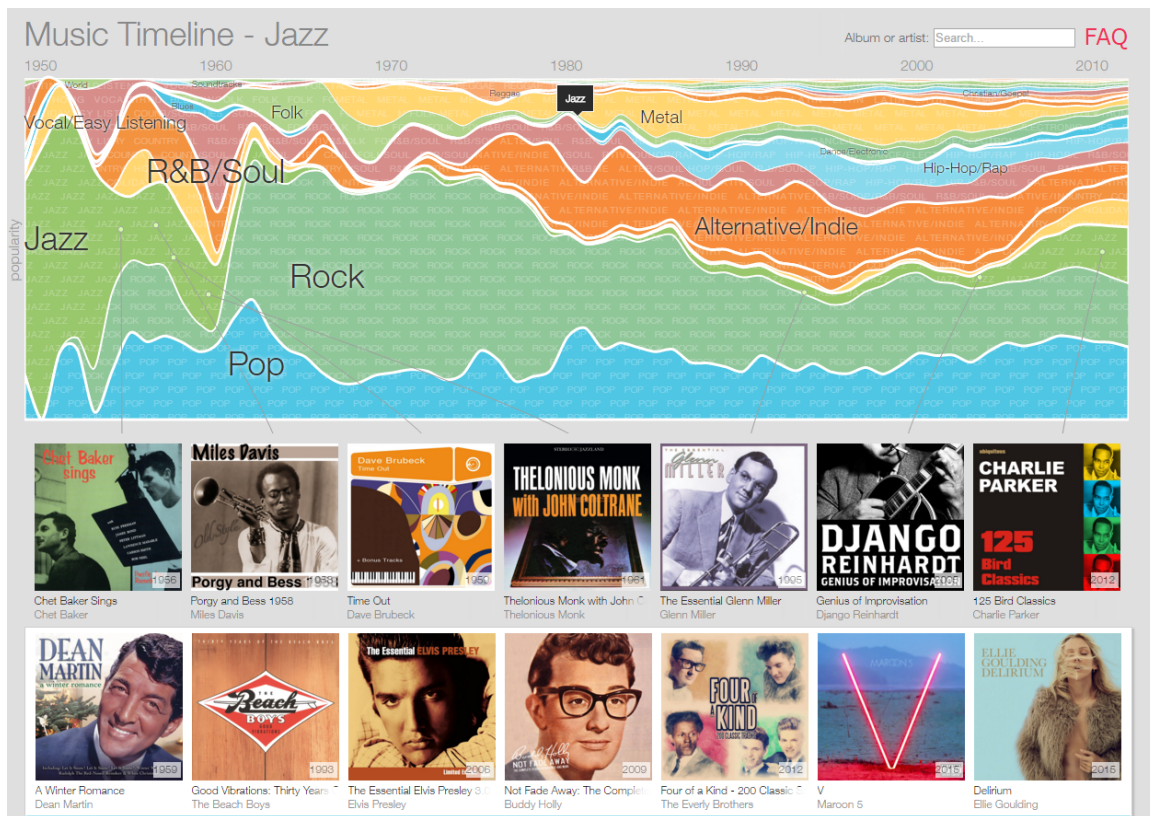
Fengxi Liu (FEL34), Jinrong Liu (JIL181), Tong Wei (TOW6), Zhenyu Peng(ZHP6)

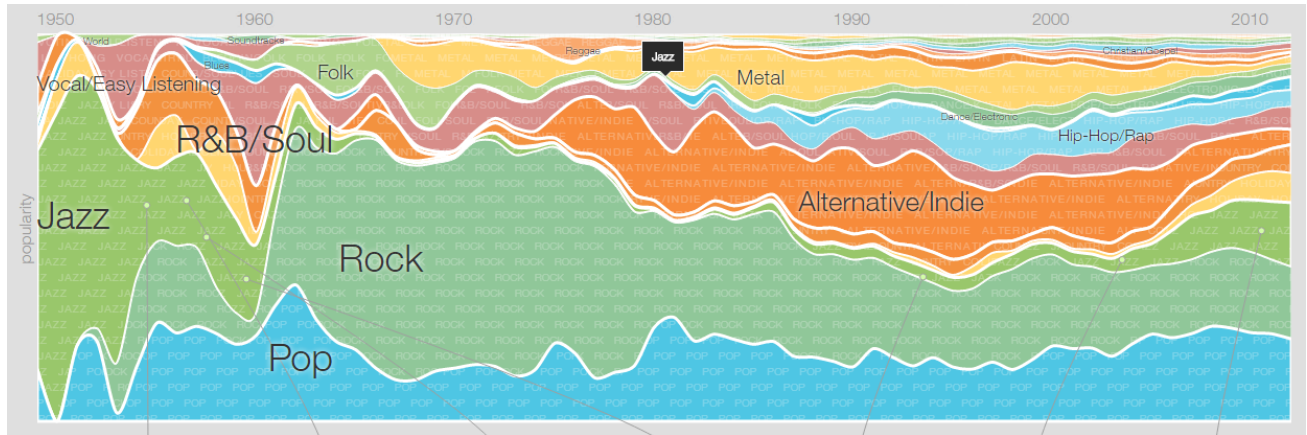
Example 1: The music timeline

Source:

Graph: <http://research.google.com/bigpicture/music/>

Explanation: <https://music-timeline.appspot.com/about.html>





Introduction:

Above are screenshots of Google's music timeline. This graph shows the development of genres of music through time. The horizontal axis represents the timeline from 1950 to 2010. And the vertical axis represents the level of popularity of each genre of music. Different color block represents different genre of music. This level of popularity is calculated by the genre of music banning and waxing based on how many Google Play Music users have an artist or album in their music library, and other data^[1]. Thicker the color stripe is, means more popular (more people have albums of this genre of music of this time).

Take Jazz music as an example, we can find that Jazz was the main genre of music in 1950s but almost disappear since the middle of 1960s.

When users' mouse hanging on each stripe, the detail information about this genre of music (albums, artists) will be show below the graph as the first screen shot.

Comments:

Follow the principles of excellent graphical data and graphical integrity, we can comment Google's music timeline graph from the following several parts:

1. It's a good graph presenting an interesting data of the popularity of different music genres, and those data is based on the large amount of data collected from google music users. This guarantees the truth of data.
2. This graph gives users a clear view of a complex idea. From this single graph, we can not only know what the major genre is in a specific time, but can also know the development of each genre. There is no "chartjunk". And it maximize the "data-ink" ratio. This follows the principle of "giving viewers greatest number of ideas in the shortest time and least ink".
3. This graph is multivariate.
4. This graph follows the principle that reveal the trend and structure from the board and show fine details (for example, detail information about the albums and artists while clicking).

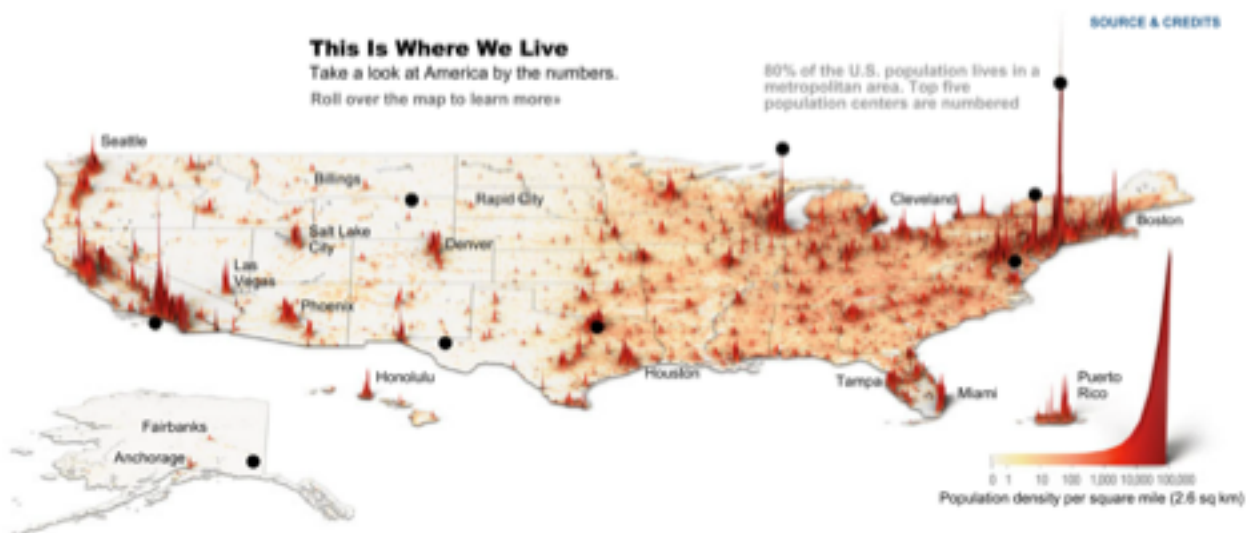
5. The number of variable dimensions depicted doesn't exceed the number of dimensions in the data. And this data has been normalized, which helps promote the graphical integrity.
6. Different color strips make it easy for viewers to distinguish different genres of music. And they are attractive. It achieves graphical elegance.

Although this graph is a good example of data visualization, there are still several aspects to be promoted. Although for non-data-ink, less is more. But we don't think it's a good idea to totally separate the description text with the graph. Viewers need to click FAQ for guidance to understand this graph.

Example 2: Where America Lives

Source:

<http://content.time.com/time/interactive/0,31813,1549966,00.html>



Introduction:

This final visualization from Time is a simple, but incredibly effective presentation of how population differs across the United States. The picture shows that 80% of the U.S. population lives in a metropolitan area. Top five population centers are New York City, Los Angeles metro area, Chicago metro area, Philadelphia metro area, and Dallas-Fort Worth metro area.

Comments:

The greatest part of the picture is that the population density is quite visualized. You can directly distinguish the dense area and sparse area. It is very clear that the east coast and west coast have more population than the Midwest. One of the advantages of this presentation is that even an 8 years old kid can understand it.

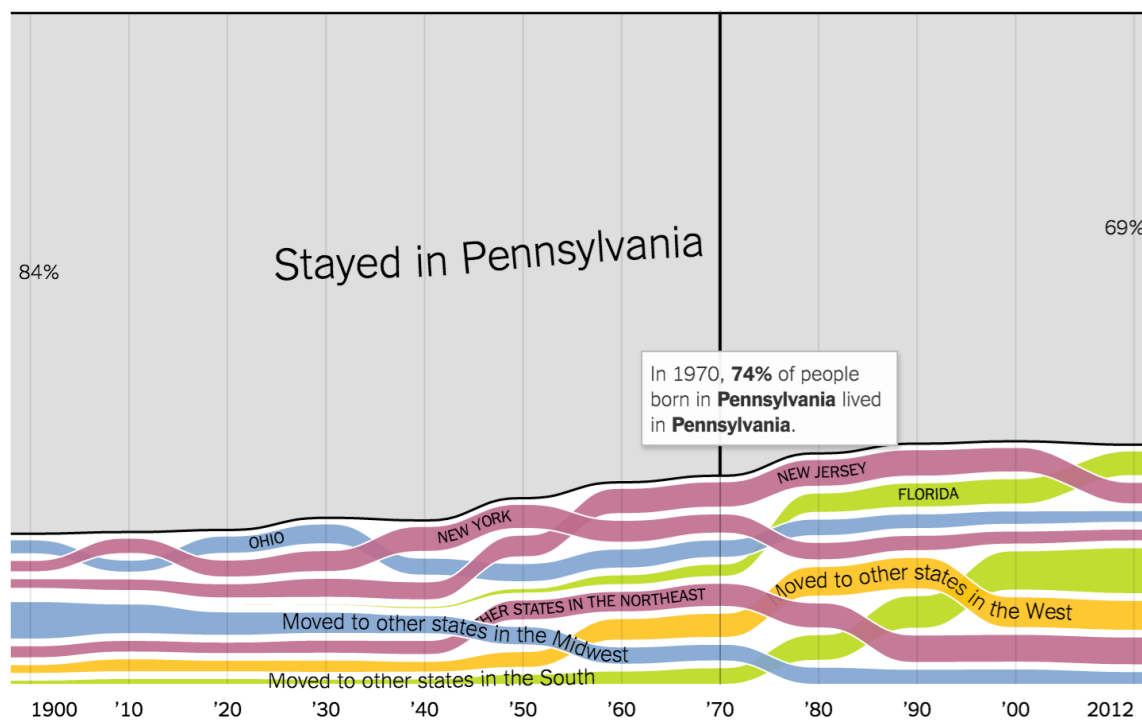
Example 3: U.S. Migration Patterns

Source:

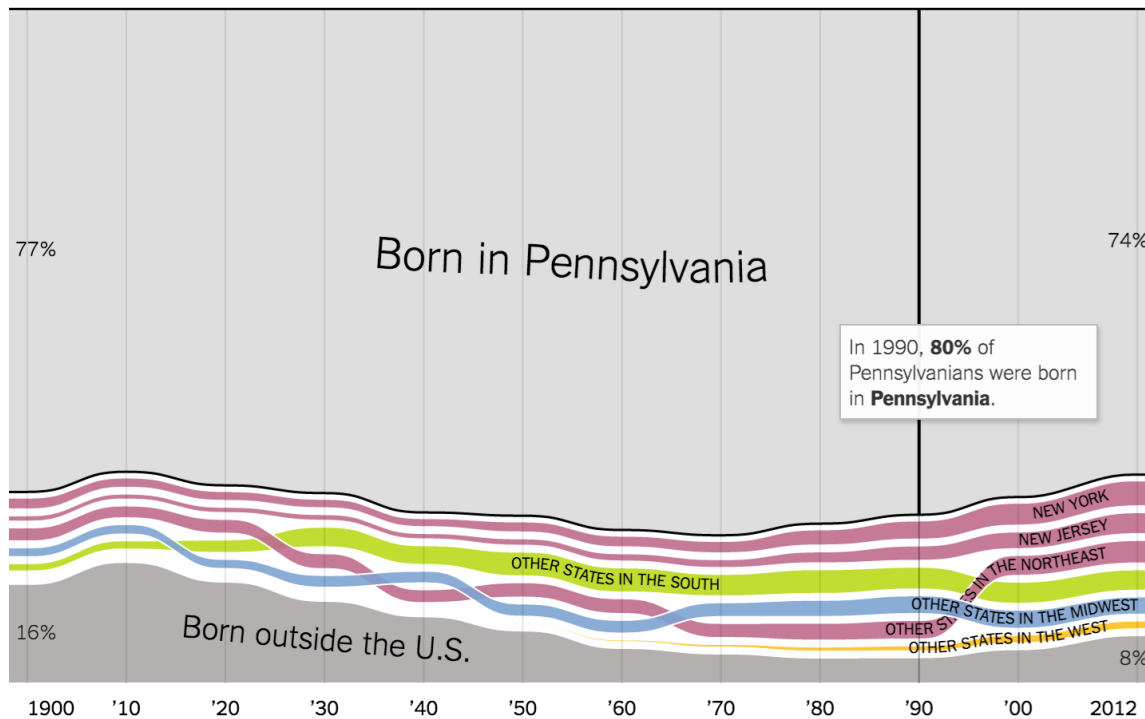
Where We Came From and Where We Went, State by State

(<http://www.nytimes.com/interactive/2014/08/13/upshot/where-people-in-each-state-were-born.html?abt=0002&abg=0>)

Diaspora Out of Pennsylvania:



Migration Into Pennsylvania:



Introduction :

In this visualization, the New York Times illustrates how Americans have moved around the country since the 1900s. It breaks down migration state by state.

Comments:

Some math about the graphs:

- Diaspora out of Pennsylvania

Since 1980, the number of state natives living in their home state, New York or New Jersey is down by half a million. The number living in the South is up by about the same amount.

- Migration into Pennsylvania

Since 1980, Pennsylvania has lost more than a quarter-million residents born in the state, but has made up for it with an immigration of more than half a million residents from New York, New Jersey and Maryland. Still, it has one of the nation's highest percentages of residents born in the state. ^[2]

Tufte's principles tell us that the visualization being good should tell the truth and do it effectively clarity and precision. [3] This visualization is great because it not only shows the percentage, but also shows the variation tendency combined with timeline. It has clear, detailed labels to defeat distortion and ambiguity.

References:

- [1]. <https://www.maptive.com/8-great-examples-complex-data-visualized/>
- [2]. <http://www.nytimes.com/interactive/2014/08/13/upshot/where-people-in-each-state-were-born.html?abt=0002&abg=0#Pennsylvania>
- [3]. <http://classes.engr.oregonstate.edu/eecs/spring2015/cs419-001/Slides/tufteDesign.pdf>