

THE BRITISH INVASION



Final Project Report

Course i247

Information Visualization

UC Berkeley, School of Information

Alex Smolen

Michael Porath

Table of Contents

About the Visualization	3
Inspiration	4
Brainstorm	5
Related work	7
Process	10
Data Collection	11
Design Iterations	12
Initial Concept	12
First design iteration	13
Second design iteration	15
Focus Group (Mid-Term presentation)	16
First Interactive Prototype and User Testing	16
Third design iteration	17
Fourth design iteration	18
Final Interactive Version	19
Future work	22
Conclusion	23

About the Visualization

Our project is called “The British Invasion”. We set out to visualize record chart data for the Beatles in the UK and the US over time, shedding light on their amazing cross-over success that led to a tidal wave of successful UK rock bands in the US - *The British Invasion*. We hope that the visualization is compelling and enjoyable to Beatles fans. The final interactive prototype is available at <http://bit.ly/britishinvasion>.

The US record chart data is based on *Billboard Hot 100* singles chart positions¹. These positions take into account several factors including radio airplay and sales. Chart positions are for individual songs, so A & B sides may chart differently.

The UK record chart data is based on *The Official Charts Company* chart positions². Chart positions are determined by sales data exclusively, and single A & B sides are not represented individually.

The singles chart data is displayed in three ways. First, the single is represented by a circular mark on the timeline, representing the release date, and color-coded by country. The peak position for the single is shown in the central bar chart, where the peak position refers to the highest chart position attained by the single. Finally, a spark graph of single positions is viewable when hovering over the timeline slider.

In addition to singles data, album (LP) releases are displayed both in the timeline as square marks and next to the singles bar charts as album covers. The goal is to provide context for the singles chart positions.

¹ see http://en.wikipedia.org/wiki/Billboard_Hot_100

² see http://en.wikipedia.org/wiki/The_Official_Charts_Company

Inspiration

Our visualization was inspired by a project named “Charting the Beatles”, led by a graphic artist named Michael Deal. He has created several information visualizations about the Beatles, including depictions of lyrical structure, chords, and song authorship contributions. The project had grown via a public Flickr pool of contributions. Some of the original visualizations were very compelling, and because both of the group members of this project were lifelong Beatles fans, we decided to create a visualization to contribute to the project.

We had noted in the discussion of the Charting the Beatles project that most of the visualizations were static, and that people wanted to see interactive visualizations. Both of the group members had programming experience, as well as some background in design, and so we decided that an interactive visualization would be a good showcase for our skills. We had discussed several programming frameworks for interactive visualization during the course, including *Protovis* and *Flare*. We knew that we could use one of these frameworks to quickly create a compelling interactive visualization.

Brainstorm

The first step was determining the purpose of the visualization and more specific, what data we wanted to visualize. We knew that there was some existing data about the Beatles in the Billboard chart dataset that one of the group members had used in a previous project, but we wanted to perform a more exhaustive search to see what sort of data we might try to visualize. This has involved both online and library research as well as several discussions about potential visualization ideas.

First, we looked online for data sets about the Beatles. There was a significant amount of data about the lyrics and chords of Beatles songs, including both audio chord and beat tempo analysis. We thought briefly about analyzing the lyrics, but thought the data seemed too subjective and qualitative to be of deeper interest for a data analysis. Many of the information visualizations which contributed to the Charting the Beatles project used lyrics as the data source, and the results weren't visually or conceptually engaging. The audio and tempo data was in a format that was difficult to understand for someone unfamiliar with music theory or digital music analysis, and we intended to create a visualization for all Beatles fans, not music experts. We also found additional record chart data for Beatles music in a variety of countries. This data was available in a wide variety of formats and difficult to piece together, but potentially interesting because of its quantitative nature and because it had a geographic component which could lead to a map-based visual analysis.

Next we went to the campus music library to see if there was additional data about the Beatles that could be collected from analog sources. Due to the organizational system of the library, there were several sections that contained books about the Beatles. We examined each section to see what data, if any, was published. Most of the books that we found were biographical, although some were organized in a chronological format labeled by date. Some contained data about instruments or recording equipment that the Beatles used. There was one compilation of research projects about the Beatles music that contained some bizarre

visualizations of data about chord sequences and lyrics pronunciation. We concluded our research by checking out the relevant works to include in our discussions.

Based on our research, we discussed a number of ideas during our brainstorming process. We considered doing a timeline of geographic details about where the Beatles were, or concerts they performed. We also thought about compiling chord data and showing changes in the complexity and simplicity of the Beatles song structure.

We finally decided to take the record chart positions from international locations and map them over time. While talking about this idea, we struck on the concept of the British Invasion, a well-known phenomenon where British rock artists surged in popularity during the mid-60s. We thought this was an intriguing idea because it allowed us to form hypotheses like:

The Beatles were popular in the UK only, until they became known in America and quickly became popular there.

Certain songs, like "I Want to Hold Your Hand", had the most immediate popularity in the US

The British Invasion started in 1964 and ended in 1967.

The US market got flooded with Beatles Singles after they made the step across the Atlantic.

We also felt the visualization might tell a compelling story. Watching the songs rise in popularity and "invade" the US would be visually appealing. We did some preliminary research around the visualization of "invasions" to come up with relevant examples.

Related work

In the classic board game *Risk*, players try to hold on to and acquire territory (see Figure 1). There is an invasion when a player's troops enter a region, and this invasion is represented as numbers in colored circle overlaid on a world map. This gave us our initial idea of using circles to represent songs and to show these songs invading other countries, from the UK.



Figure 1 Screenshot of online game Risk. The coloring of the countries and the circles showing army strengths are the two elements that relate to our British Invasion project

Armsflow (Figure 2) is an interactive visualization that shows arms transactions between countries. We liked the animation of the transaction that showed the source and destination clearly. We thought we could apply animation techniques to clearly demonstrate that songs were coming from the UK to the US and other countries.

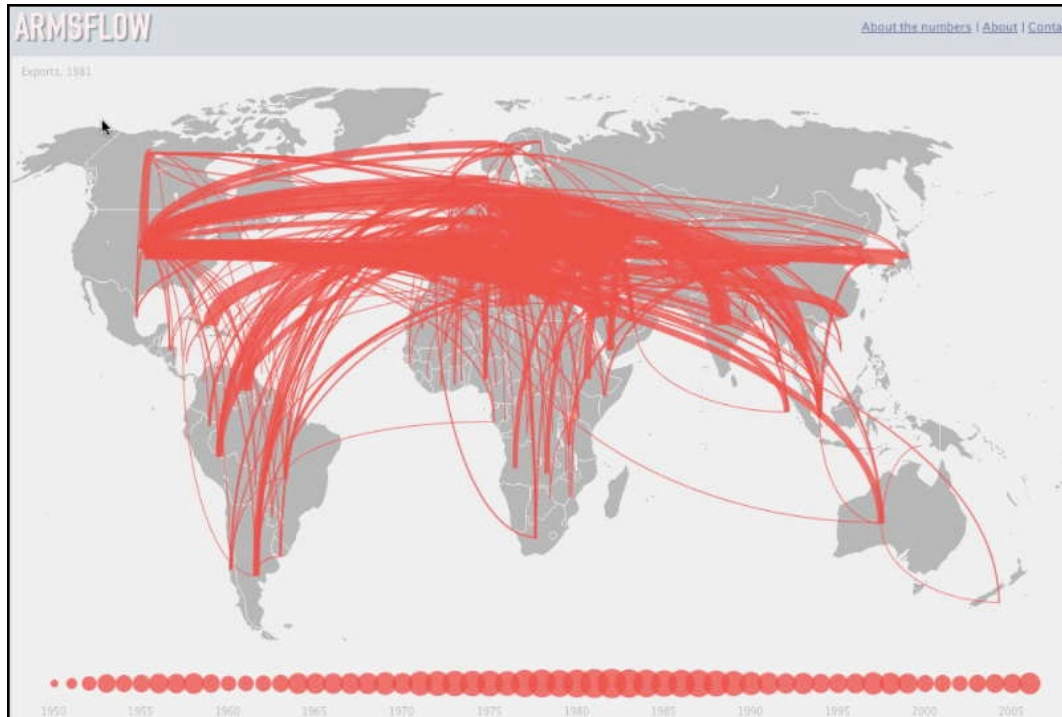


Figure 2 Visualization showing arms transactions between countries. Differently sized circles at the bottom are a feasible way to visualize quantitative data over time

The classic info graphic by Charles Joseph Minard (Figure 3) showing Napoleon's march into Russia was another visualization of an invasion that we considered. While the visualization isn't interactive, we liked the composition of elements that included relationships between them. The line graph at the bottom shows the temperature of the brutal winter march back to France, and the relationship between that graph and the upper graph demonstrating the army size gives important additional context to the visualization. We wanted to consider applying this technique of combining multiple graphs into one visualization to provide opportunities for contextual analysis.

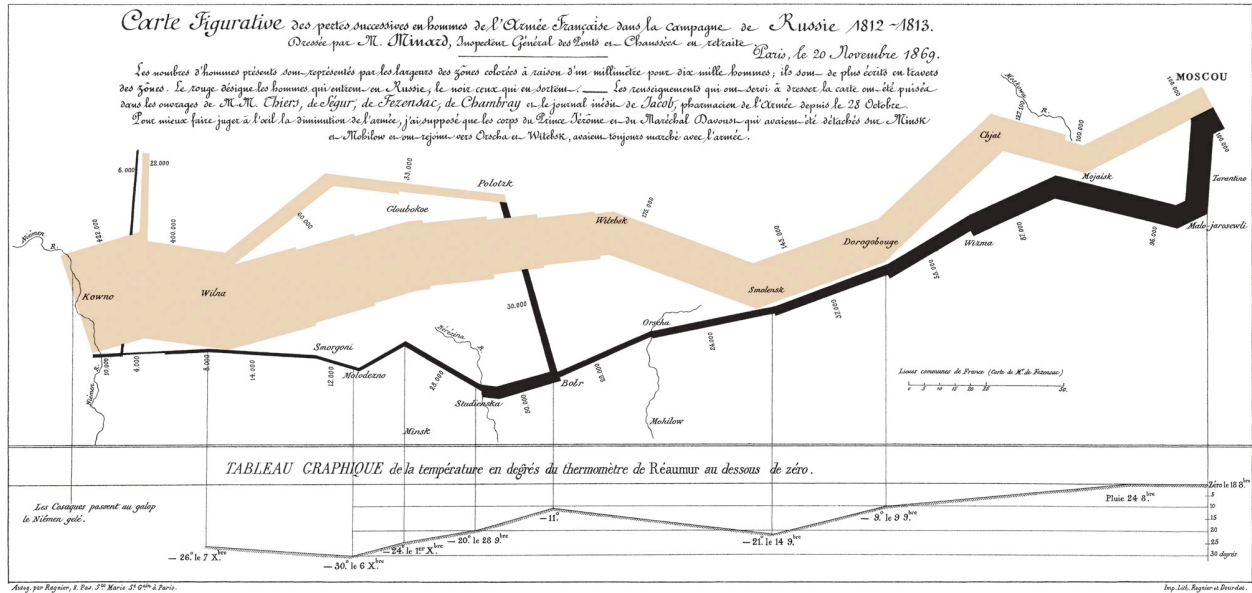


Figure 3 Visualization by Charles Joseph Minard portraying the losses suffered by Napoleon's army in the Russian campaign of 1812. Edward Tufte often references this as a good example for a visualization.

Process

For our design and development work and due to the small group size, we used an iterative process where we quickly built, tested and refined ideas. At the initial stages, we built low-fidelity prototypes to communicate design elements that would drive early activities like data collection and technology choice. We continued to re-imagine parts of our low-fidelity prototypes, and it was fairly time consuming up front to settle on a design that we felt conveyed all aspects of the data in an intuitive and aesthetically appealing way. It took four iterations until we felt comfortable with our design choices and implemented them in code.

We developed the visualization using *Adobe Flash Builder*, which is an *Eclipse*-based IDE. We integrated the *Subclipse* plugin and kept all of our code in an SVN server. Since we worked on the project in the same physical space, we informally divided the work along visualization areas and merged our changes several times each day, which worked very well. We kept a small list of bugs, to-dos, and user testing results in a simple text file.

Data Collection

We set out to collect the data for international chart positions and found that the data was not as easily comparable as could be expected. Many countries song chart positions were available as <Entry Date, Weeks on Chart, Peak Position on Chart> triples as opposed to statistics for each week on the chart. These different types of data couldn't be easily integrated or compared in the visualization as we had envisioned. Consequently, we decided to focus on the US and UK initially, since collecting and cleaning all of that data would have been too labor-intensive for the scope of this project.

We obtained the US chart positions from a spreadsheet created by *The Whitburn Project*, which is a newsgroup that curates the same data that Joel Whitburn sells about Billboard chart positions. This information would also be available on microfilm of Billboard magazine in the UC Berkeley Music Library.

The UK chart data was much more easily accessible, at <http://www.chartstats.com> and <http://www.everyhit.com>.

International record chart data was found in a variety of sources, the best of which were <http://www.lanet.lv/misc/charts> and <http://top40-charts.com>. However, these sites had questionable data quality and the format was often inconsistent with our design goals.

Design Iterations

Initial Concept

Our initial concept (Figure 4) was a world map overlaid with circles that represented Beatles songs. We would use the record chart data that we would collect online to show the position of each song in each country, week after week. At the bottom, there would be an animated timeline, which you could manually adjust or set to “play”, like an online video. The color of each of the dots represented the chart position of the song as a heat map style gradient. The user would be able to select which songs and countries to include, and hovering over each song would bring up a tool tip that contained additional information.



Figure 4 Initial concept sketch focusing on an international comparison of chart data. The position in the charts is indicated by a heat map. A navigation allows for brushing and filtering.

While we liked several elements of this concept, when we presented this design as part of our project proposal we gathered external as well as our own feedback and identified several significant challenges.

First, the limitation of a map became apparent when we realized that very small countries would be unable to fit several songs in a clear manner. This left us with two options – make each country the same size, which would have led to distortion, or only include the US and

UK. We decided to follow the second strategy, since we were fairly sure that the chart data for both countries was more readily available and comparable than other countries chart data. Also, we this gave us the opportunity to limit the scope of our visualization initially so we could iterate our design more quickly.

We also realized that a gradient heat map might be hard to make visually appealing and intuitive, and it could be hard to tell what the exact chart positions were based on the color. We decided to using a simple number label, which was more exact.

First design iteration

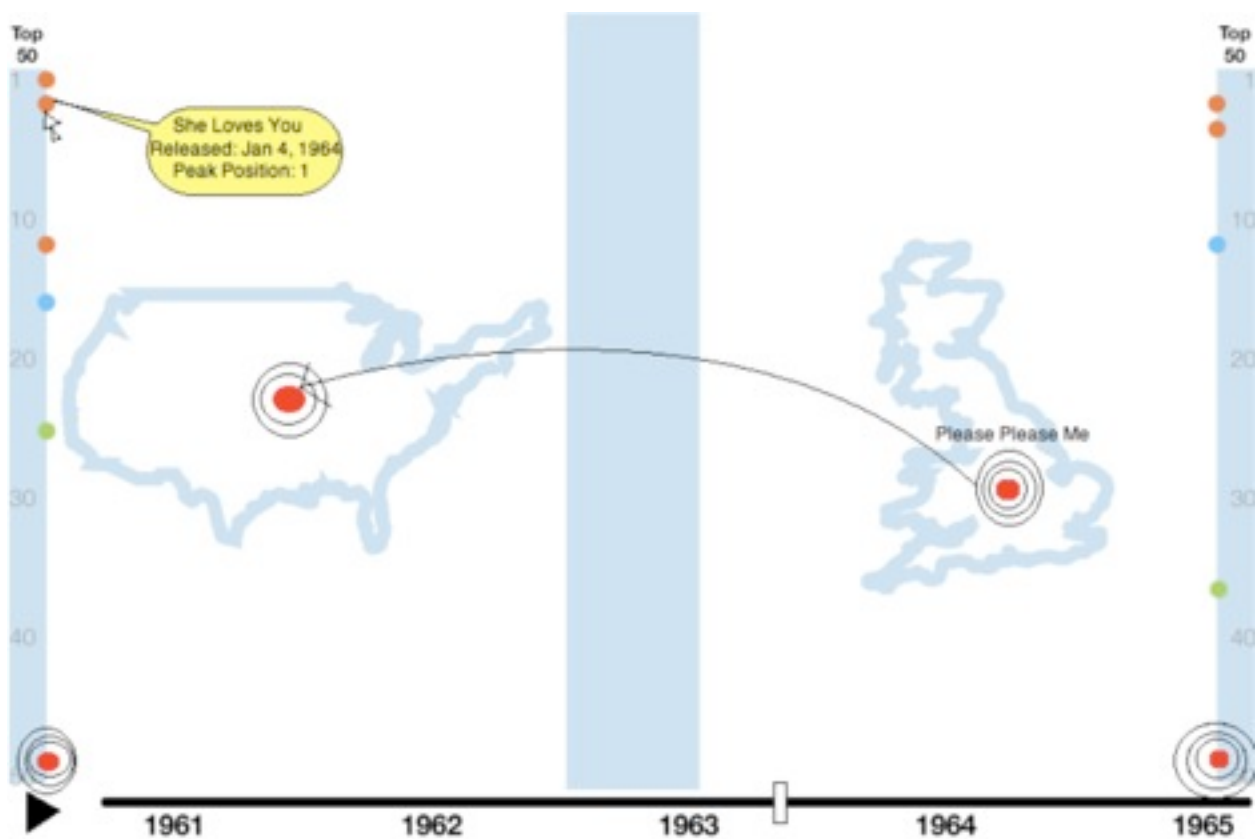


Figure 5 Graph restricted on UK and US data showing an animated chart graph on both sides.

For our first design iteration, we got rid of the map and simply showed the US and the UK, as represented by the country shapes. We had animated charts on the left and right to show

songs that were on the charts and their chart positions. Part of our motivation for this design was the need to display dense amounts of data, such as the beginning of 1964 when The Beatles held all 5 songs at the top of the US charts. We used circles in the countries to represent songs, which would transition from the UK to the US when they entered the charts in the US. We continued to use the playable timeline at the bottom along with tooltips for each song.

The biggest changes in this visualization were the shift from an international to a focus on the US and UK as well as the mirrored chart position indicators on the left and right.

Second design iteration



Figure 6 Introduced “Bubble Graph” for both countries showing bigger circles for higher peak positions. This visualization also introduced contextual information in the form of album releases, as well as tooltips for each data element.

Our second design (Figure 6) iteration elaborated on the first design iteration, and included several graphical enhancements. We added circles representing each song on the countries’ charts to represent the peak position. These circles changed in size based on the peak position of the song, so bigger hits were bigger circles. We continued to use the animation to show these songs crossing from the US to the UK. Our timeline now included contextual events - in this case, album releases. Each album was color coded to show which singles came from which albums. We expanded on the amount and type of data which might be available in the

tooltip and added brushing/linking where hovering over a song highlighted the song in the other country.

Focus Group (Mid-Term presentation)

We presented our concept, as well as demonstrated our design and initial functional prototype, to the class for our mid-term presentation. Up until this point, the feedback we had received was mostly informal. After our demonstration, we received several helpful pieces of feedback from the class.

First, people noted that in the mirrored chart position strips, where the songs moved up and down over time, it was difficult to see the "shape" of a song's popularity. Did a song start out at number one? Did it stay there long? These kinds of questions were difficult to answer without intently watching the chart over time, which drew the focus from the rest of the visualization panel. In our design, one piece of feedback we received was that the inclusion of the country outlines was unnecessary and added clutter. This was the last vestige of our map idea, and seemed ready to be removed. Also, people asked if we could provide more context for the chart data, such as by including historical events on the timeline. We considered each of these critiques in our subsequent designs.

First Interactive Prototype and User Testing

We incorporated the results of the mid-term presentation into an initial interactive prototype (Figure 7) and conducted basic user testing. People seemed to like the idea of the project, and we felt confident that we had the major elements. We experimented with adding a spark graph, that shows the single chart position over time. Users seemed to agree that the spark graph was interesting for context but not necessary. They also talked about how they would like to the singles on the timeline itself. We came up with the following experimental design to test some of these elements.

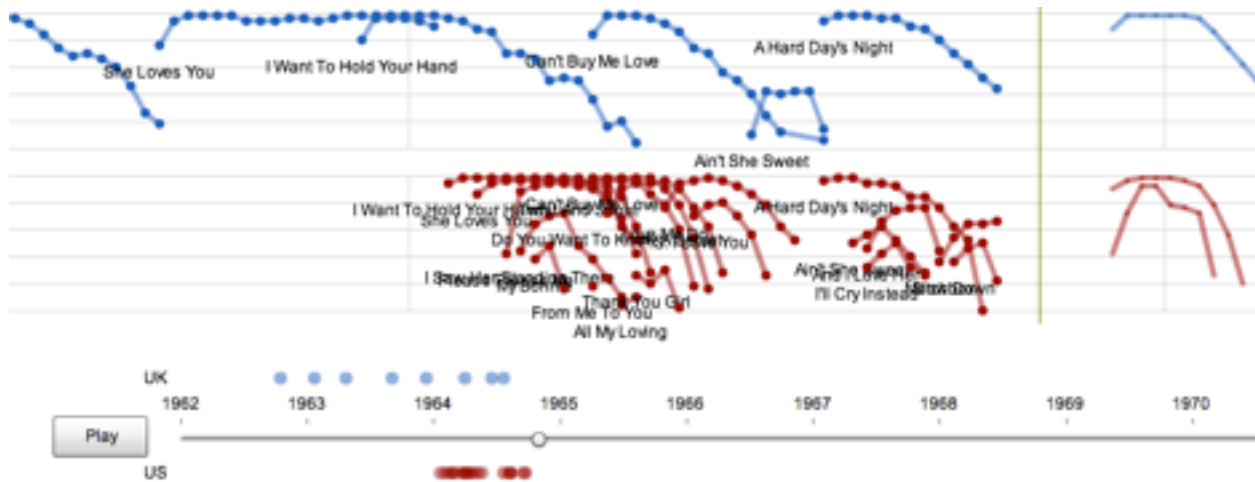


Figure 7 First interactive prototype associating single release dates with the timeline and providing more details about the progress of a song in the charts.

Third design iteration

The third design iteration (Figure 8) was quickly assembled to test an idea of showing song chart positions as bar graphs, rather than as bubbles. We liked the idea that the title of the song only had to appear once, and the bar graphs could go to either side for the US and the UK. This bar graph would show the singles in the visualization permanently and across the timeline, rather than temporarily as in the previous chart position strip. We demarcated each single release date with a dot on the timeline. We also added a "zoom-in" function for viewing spark graphs of singles that showed the rise and fall of its popularity. This allowed users to view the data without always presenting it and obscuring other data. We also dropped the country shapes due to feedback from our midterm presentation.

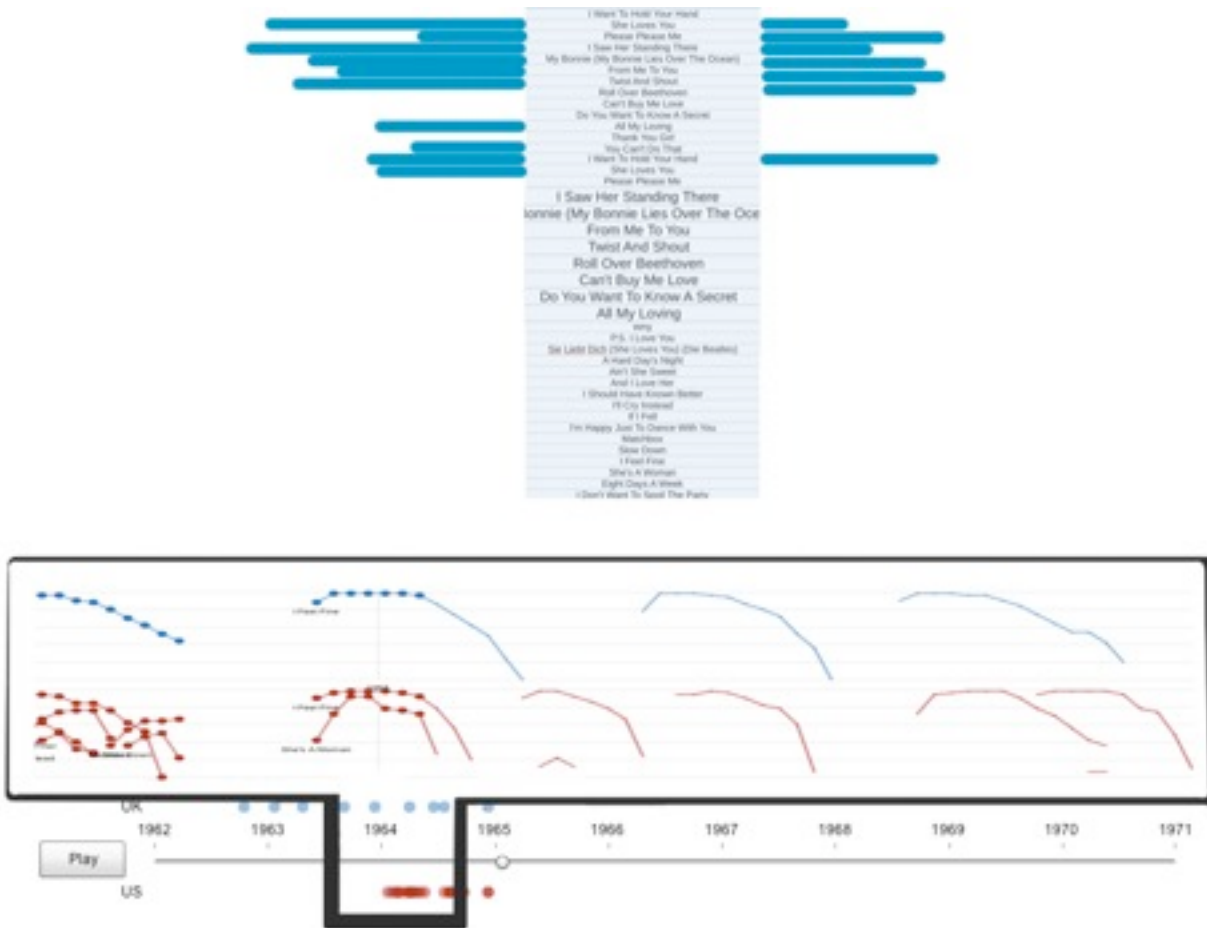


Figure 8 Third design iteration introducing a bar graph to compare a song’s success between the US and the UK. The list like arrangement also provides an overview over the song titles. The spark graph provides a zoom-in function that shows the progress of a song’s success within a limited time span in detail.

Fourth design iteration

Our fourth design iteration (Figure 9) was an enhancement of the previous iteration, with higher fidelity and some additional content. First, we added albums to the visualization to provide context to the era that the timeline was in. It also added some visual interest to the design. We tied them together with the single chart by including both as square or circle icons to the timeline. We introduced the ability to step through, rather than play, the timeline using the plus and minus buttons on the lower right side. We refined the spark

graphs by including them as a moveable window over the timeline. Furthermore, we tried to separate the singles from the albums by introducing a gray border that included the album timeline annotations and the album covers, while keeping a white fill that included the timeline annotations and the bar chart with their focus on singles.

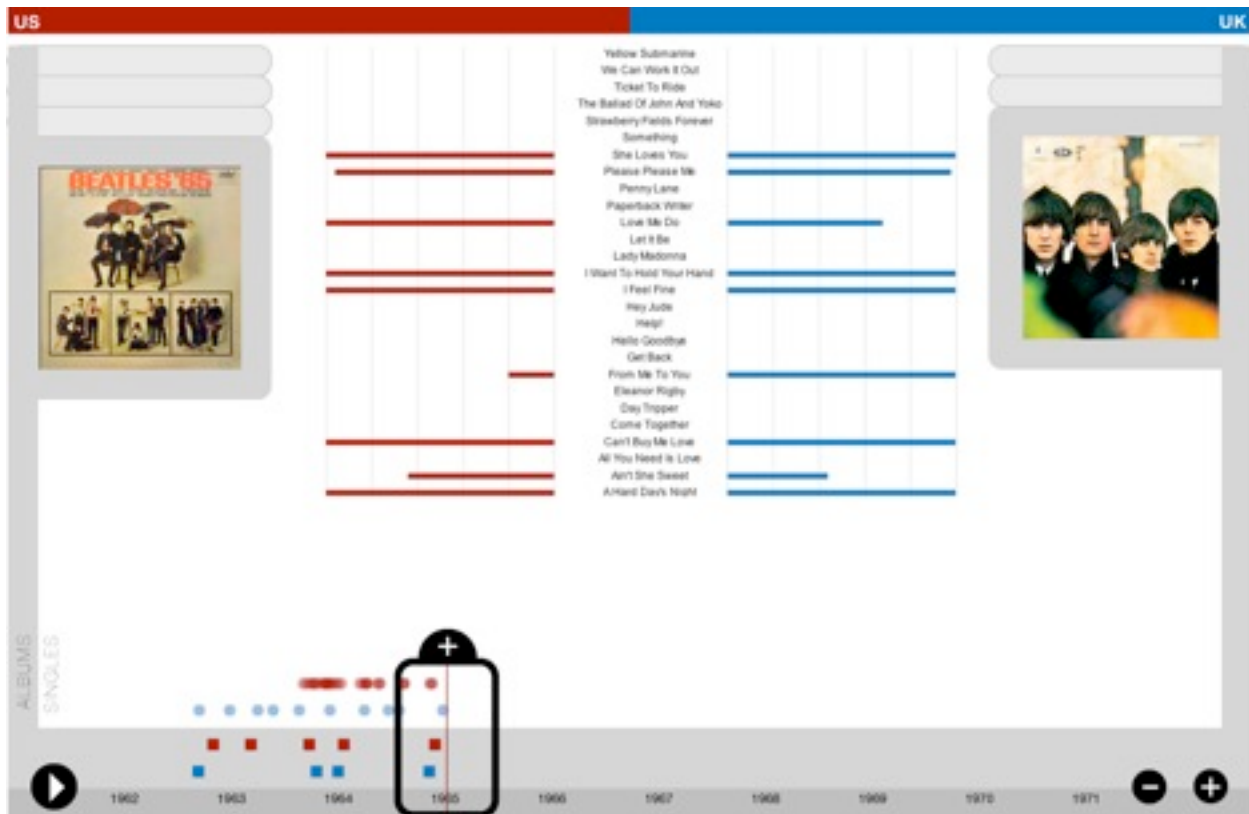


Figure 9 Iteration focusing on graphical sophistication and consolidation of all previous ideas and introducing more context information about.

Final Interactive Version

We eventually implemented the fourth design iteration in our final interactive prototype (Figure 10). It contains a singles bar graph, which scrolls to accommodate the large number of songs and the potential limited resolution of browsers. We included a small white label with the peak position to give the user exact information. The timeline is more accurately annotated with US and UK singles and album release dates, and there is a simple legend on

the lower right hand side that explains the use in color and builds on the cognitive perception of the white and grey areas. The spark graph (Figure 11b) is available by hovering over the "plus" area on top of the scrolling timeline window.

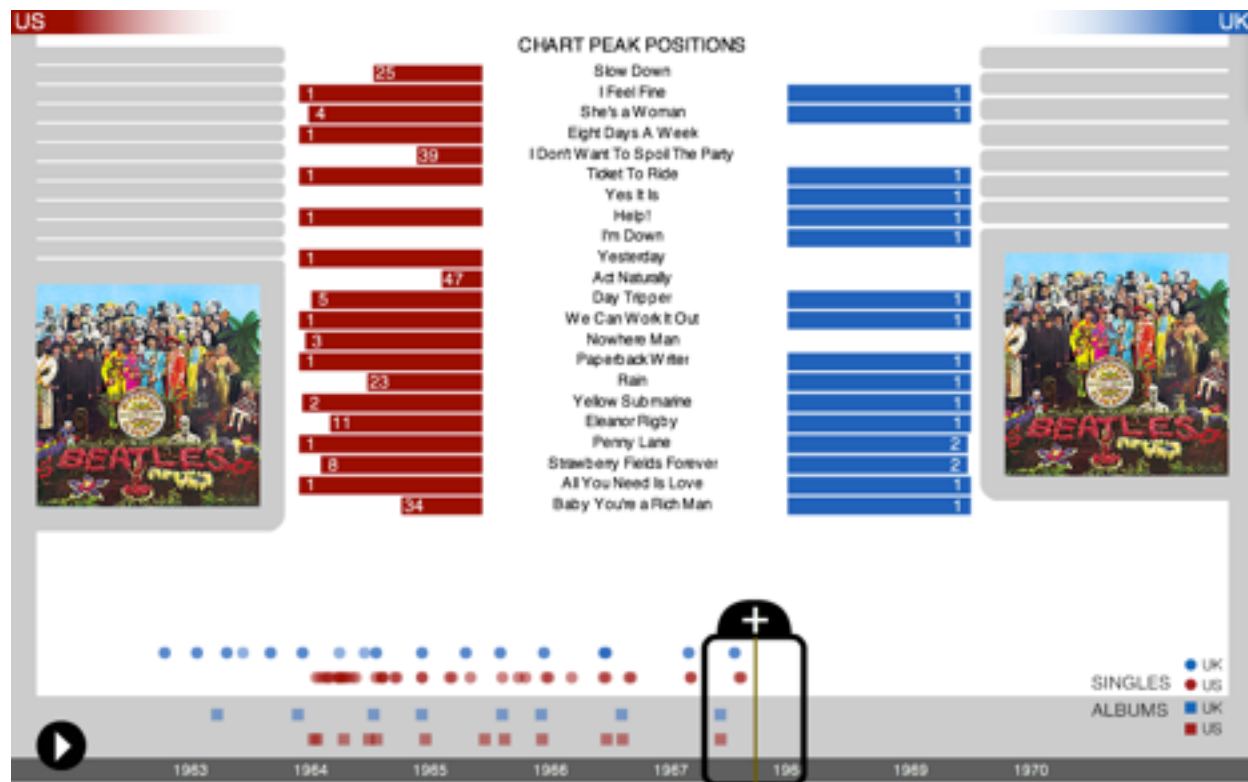


Figure 10 Iteration focusing on graphical sophistication and consolidation of all previous ideas and introducing more context information about.

Many of these design elements were the result of extensive but informal user testing where we asked friends and family for feedback. Their misunderstandings helped us to fix several small usability issues, such as adding a label to the bar chart for the peak position, or identifying critical ones, like dealing with reduced resolutions and making the play button very clear (Figure 11a) to the user initially by increasing its size when the visualization is loaded.

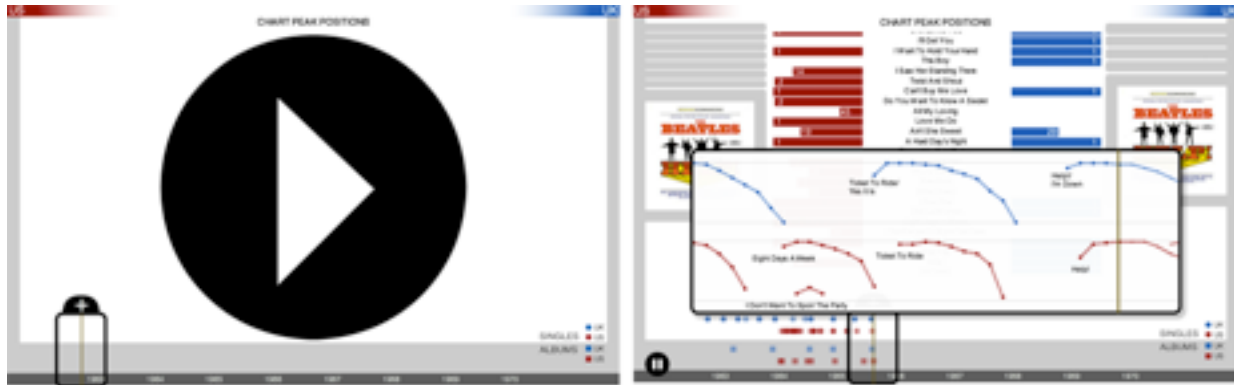


Figure 11 (a) left side; enlarged play button to draw the user's attention on the action. (b) right side: Spark graph providing more information about the progress of a song after it has entered the chart.

We dropped the idea of the previous/next buttons because the feedback we received indicated they were unnecessary. We also made the timeline more distinctive by making it darker.

Future work

When we gave our final interactive prototype demonstration, people seemed to be enthusiastic about our design. We heard some suggestions that echoed things we had heard in our user testing. One of these ideas is to include audio, since the visualization is about a music group. There are a variety of approaches for doing this, but we think it would be a natural fit for this visualization to include the songs in some playable format along with more detailed information about the songs. We also heard feedback about adding contextual events, like news items about the Beatles, which we think could be a good addition to the visualization. We would like to extend our visualization to support tooltips, as well as relating songs in different graphs to each other via brushing and linking as well as relating songs to albums.

Additionally, we would like to make this visualization suitable for public use by improving the quality of the code to reduce load time and resource usage. We would like to make it friendly for smaller screen sizes in allowing to resize the visualization. Furthermore, we aim toward providing additional contextual information, such as an "About the Visualization" section that describes the work, which builds on the introduction provided in this document. Finally, releasing the visualization to the public will require us to fix remaining bugs and further refine some of the graphical elements.

Conclusion

Overall, we are extremely happy with the outcome of this project. We felt that our process encouraged a lot of idea generation and trial, and that we received good feedback that helped us make strong design and implementation decisions. We feel our visualization addresses an area that hasn't been covered by other visualizations and gives more insight to the raw data of chart position. The visualization is also engaging and enjoyable to watch.

Finally and not least important, this project enhanced our understanding of the information visualization practice, and hopefully the project similarly enhances the user's understanding of the British Invasion of The Beatles.