

INFORMATION VISUALISATION REPORT

WWI BOMBING OPERATIONS

Group 2

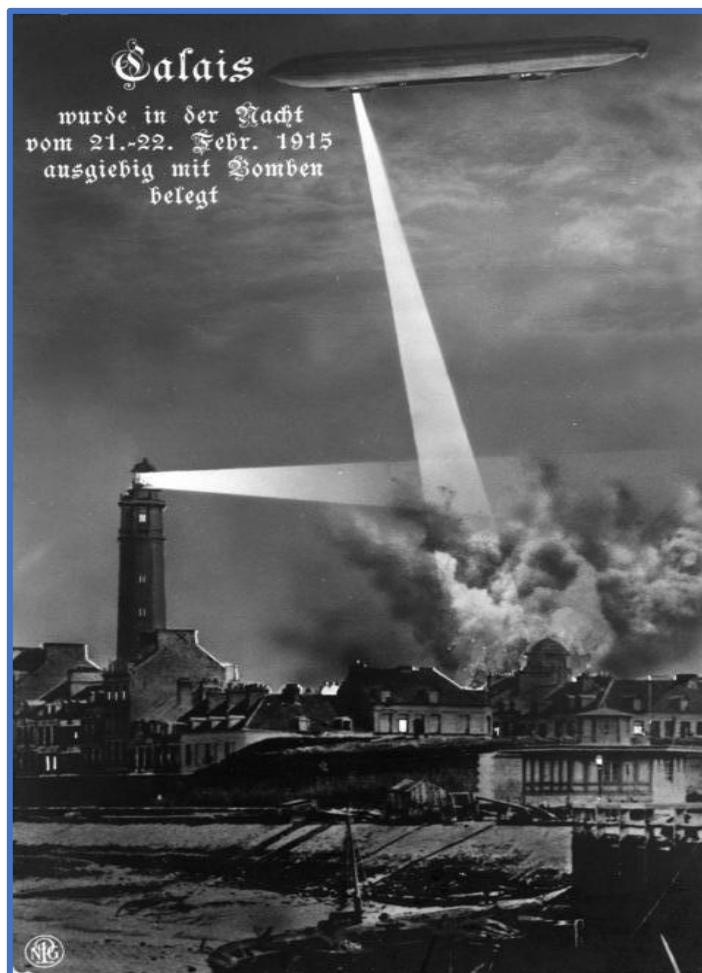
Lau, Wing Chau

Li, Yuman

Suvarna, Harsh

Taylor, Raphael

Watson, Cameron Stewart



Demo video:

Link - <https://www.youtube.com/watch?v=SLvpmp7mZ8>

Table of Contents

Demo video:	1
The data	2
Visualisation Tasks	3
The core systems	4
Generalised Selection	4
Design comparison	5
User evaluation comparison	8
Future work	10
References	11
Image Reference	11
Work Log:	12
Appendix	13
SYSTEM A Screenshots	13
SYSTEM B Screenshots	14
System C Screenshots	15
Test Results	17

The data

This data set encapsulates the strategic details of 1,441 Allied aerial bombing operations during World War I, digitised and compiled in the THOR database, a resource valuable for historical analysis, ordinance discovery, and tactical education. The link is as follows: <https://www.kaggle.com/datasets/usaf/wwi-bombing-operations>

The THOR (Theatre History of Operations Reports) database, specifically the WWI Bombing Operations dataset, is an example of a temporal and geospatial dataset. It is temporal because it includes time-specific data, such as the date of each mission (MSNDATE) and the takeoff time (TAKEOFFTIME). It is geospatial as it contains geographic coordinates for each operation, including latitude and longitude of the target (LATITUDE, LONGITUDE) and the takeoff base (TAKEOFFLATITUDE, TAKEOFFLONGITUDE).

The dataset is categorised as a historical military operational dataset. It is structured as a tabular data set, with each row representing a single bombing operation and each column representing a different attribute of that operation. The attributes include categorical data such as the type of operation (OPERATION), the country responsible (COUNTRY), and the type of target (TGTTYPE). It also contains quantitative data such as the number of planes attacking (NUMBEROFPLANESATTACKING), the weight of weapons expended (WEAPONWEIGHT), and the altitude at which the operation took place (ALTITUDE).

Moreover, the dataset has attributes for qualitative analysis, such as textual descriptions of enemy action (ENEMYACTION) and route details (ROUTEDetails), as well as for quantitative analysis, such as the number of friendly casualties (FRIENDLYCASUALTIES) and bomb load (BOMBLOAD). It also provides data for relational analysis, connecting bombing operations to external data sets such as weather reports (WEATHER), and can be utilised for longitudinal studies over the course of the war.

In general, this dataset is a temporal, multivariate collection comprising structured quantitative data—such as bomb load, aircraft numbers, and payload weight—and qualitative data, like mission type and target location. Data points are categorically and sequentially organised, enabling intricate chronological and spatial analysis. Utilising examples from the dataset: one could track the frequency of bombings over the course of the war, identify the most targeted locations via geospatial mapping, or correlate the number of planes in a mission with the bomb load to deduce tactical changes over time. The dataset's quantitative aspect allows for statistical analysis and predictive modelling, potentially revealing trends in the data like increases in bomb loads as the war progressed.

Visualisation Tasks

The visualisation systems A, B, and C will be designed to facilitate a diverse range of visualisation tasks by providing comprehensive answers to a multitude of questions. The following are the key visualisation tasks which the system aims to support:

Analyse

- What trend is seen in terms of bombing operations over the period of WW1 as per the visualisation (Discover)
- Calculate the percentage of mission carried out by each attacking country on each attacked country (Derive)

Correlation

- Is there a correlation between Bombload and weapon weight. Are there any outliers?
- Is there a correlation between Bombload and altitude. What does it say about the missions?

Query

- Identify which country has the greatest number of targeted locations (Identify)
- Find which are the top 3 Allied countries with the greatest number of operations (Summarise)
- Find the top 3 most popular choice of airplane model for missions in Italy (Summarise)

Trend

- What has been the trend of number of missions through the war

Search

- How many operations did USA have on September 15, 1918 (Lookup)
- Which country had the most operation on Germany (Locate)
- Which day had the greatest number of missions (Explore)
- Which country had a declining number of missions as the war progressed (Explore)
- Which country was targeted the least and the most (Browse)

The core systems

We created 3 systems which we submitted as zip files.

Generalised Selection

Semantic structure

The semantic structure of this visualisation system is determined based on several levels of abstraction. The data is mainly categorised by attacking country (COUNTRY) and targeted country (TGTCOUNTRY). Charts uses each bombing incident or the count of the bombing incidents as a data point with attributes including targeted country, attacking country, date, weapon weight, and bomb load. Temporal details of each bombing incidents are also specified using the date attribute (MSNDATE), year and month which are converted from MSNDATE.

Traversal policy

The traversal policy contains temporal traversal and categorical traversal. For temporal traversal, the visualisations provide interactive features for users to traverse the temporal dimensions. Users can perform dragging, scaling on the graph to traverse the date to check bombing incidents through query relaxation which is utilised to adjust the scope of the original query [1], in this case, on a specific date, a specific month, or a specific year. The occurrence date of the bombing incidents is displayed to the users in a specific adjustable time frame. For categorical traversal, the visualisation system is categorised by targeted country and attacking country. Users can then combine these two for traversal. For example, users can explore the system B that only the bombing incidents of the selected attacking countries on a specific date (date/ month/ year) would be visible on the chart with a timeline.

Implementation approach

Users can control the opacity of the visual elements through an interactive device since visual elements are possible to be overlapped to combat over-plotting when down-scaling. Over-plotting potentially cause loss of information and can occur when an excessive amount of data items is rendered within a limited resolution [3]. Overlapped elements can be more visually noticeable to the users by adjusting the opacity. This system with bar charts and scatterplot utilises several selections including interval selection and point selection for data exploration through brushing and linking. Tooltip is also added for additional contextual information of each datapoint.

Design comparison

Colour encoding

- **Visualisation System A** uses colour to distinguish between different data categories effectively. For Countries of Service, it uses bold primary colours (red, green, blue). This choice is informed by cultural associations, for example, the colour represented in the country's flag. For the bombing altitude plot, the colormap 'viridis' is used with lower altitudes being purple, while higher are yellow. This aligns with convention where darker=low and brighter=high values. A continuous colour scale is used because it's more appropriate for numeric values.
- **Visualisation System B** uses the categorical colour scheme 'category20' provided by the Altair throughout the whole system to represent attacking country and targeted country [Screenshot B1]. It is decided not to manually map colours into different countries since some countries may have similar colour representation based on their flags. It may defeat the purpose of utilising colour as a retinal variable to differentiate data values.
- **Visualisation System C** uses colour to differentiate categorical data, such as model of plane and country. Furthermore, System C also uses colour to draw a contrast in quantitative data, such as the heat map of number of attacks per country in WW1. The two main colours used are yellow and purple, which are contrasting colours. Yellow for a high number of attacks and purple for a low number. Additionally, green is used for a moderate amount. This use of colour intends to highlight data that stands out.

Chart types

- **Visualisation System A** uses a mixture of bar charts and scatter plots on top of maps formed from geoJSON files with historically accurate country border. Each point on the scatter plot represents one mission, with coordinates determining its position. This was a natural choice as we were trying to represent special data. It gives extra context about which countries were bombed and effectively shows areas with higher concentration of bombing missions, and therefore we can roughly estimate where the frontlines were for a given year. The bar charts were chosen because they are the standard choice for categorical comparisons.
- **Visualisation System B** utilises stacked bar chart, heatmap, scatterplot to visualise data. Stacked bar chart is used to visualise the proportion of attacking countries' bombings by targeted countries. Heatmap is used to showcase the attacking countries and to enable more diverse interaction through brushing. Scatterplot is used to visualise the frequency of incidents throughout WW1 and at a specific time unit.
- **Visualisation System C** uses a stacked bar chart, scatterplot, bar chart, and heatmap to visualise data. The visualisations achieve a similar goal to System B, however, different attributes are highlighted, such as model of plane. Overall, system B and C are visually similar, yet slightly different regarding the content of the data.

Interaction methods

- **Visualisation System A** allows users to select areas of the map they want to zoom in on to better differentiate data points. Hover annotations are implemented to view extra data for specific data points. This interaction technique allows exploring the data at a granular level without overwhelming the main view. It's a good choice for providing details-on-demand. All 4 of the visualisation systems are linked by the year slider, so they can all update when the year is changed to show a generalised selection.
- **Visualisation System B** allows user interaction mainly through brushing, interval selection, and point selections. Users can interact with the system by brushing a customisable area on the chart [Screenshot B2], brushing through interval, and clicking on specific data item(s). Any selection made

is linked across the system to achieve linking and filtering dynamically. Interaction device such as slider is provided to adjust opacity of the chart objects [Screenshot B6].

- **Visualisation System C** allows user interaction in all 4 visualisations. Point selection is used in all graphs to let the user focus on specific data, such as the model of plane in visualisations 1 and 2. For visualisations 3 and 4, brushing is used, as the user can select a country which they want to focus on, e.g. France, and the relevant information will be more visible while the rest of the visualisation will be more transparent.

Visual encoding

- **Visualisation System A** Uses opacity for visual encoding. The points in the scatter plots have an alpha (transparency) value of 0.5. This allows some visibility of overlapping points, indicating areas of higher density.
- **Visualisation System B** utilises size and opacity for visual encoding. Size is used to achieve quantitative perception for users to estimate the quantitative difference between chart objects since size is a superior visual channel to represent orderability [2]. Opacity is also used to provide better visibility of overlapped chart objects when a country operated bombing missions to multiple countries on the same day [Screenshot B3].
- **Visualisation System C** also uses opacity to highlight or downplay certain data. Visualisations 1, 3, and 4, make categorical data more transparent if it has not been selected. However, visualisation 2, the scatter plot, makes non selected data completely transparent. This was done due to the large amount of data that overlaps with each other in this graph.

Tooltip design

- **Visualisation System A** Tooltips are triggered by hovering over a point on the map. They display additional data, namely Mission Date, Target Location, Altitude, and Target Type. These are the most relevant details to provide to fulfil the visualisation tasks without overwhelming the user. The tooltips don't have a border and they are not opaque, so that they stay visually lightweight.
- **Visualisation System B** uses the tooltip feature provided by Altair is used to display additional contextual information for visualisation system B. A pop-up window containing information about attacking countries, targeted country, and count of incidents is shown when a user hovers on a chart object [Screenshot B4]. Mission IDs of individual incident are not shown since this system focuses on the frequency of bombing incidents over time rather than distribution of each individual incident.
- **Visualisation System C** also uses the tooltip feature in visualisations 1, 2, and 4 to allow the user to gain more detailed information by placing the cursor over the specific data item they are interested in.

Layout and Composition

- **Visualisation System A** uses a grid of 2x2, with the scatter plot maps on the top row and the bar charts on the bottom row. Placing the maps side-by-side allows the user to quickly compare similar zoomed in areas. In addition, the user could select different areas of the map on each plot to compare the density of mission in different places. The bar charts are placed below the maps to create a visual hierarchy where the user is likely to start with maps for geographical context, then move down to the bar charts for more detailed comparisons.
- **Visualisation system B** employs dashboard layout and hierarchical layout for enhanced readability. The legend, attacking country by targeted country chart, and the proportion chart are aligned horizontally due to identical attribute on the Y-axis. The charts showing incidents in the same day, same month, and same year are displayed vertical to show hierarchical order according to the time units [Screenshot B7].
- **Visualisation System C** uses a legend for visualisations 1, 2, and 4. This allows the user to see the categorical data in visualisations 1 and 2 and quantitative data in visualisation 4, which is a heat map.

In visualisation 3 and 4, the x-axis displays the different countries for both graphs, this is done so that the user can easily understand the data. This also aids in brushing.

User evaluation comparison

The evaluation methodology aimed to gather comprehensive feedback from participants regarding their experience with the system. To achieve this, we employed a structured approach that involved three main components:

- **Positive Feedback:** Users were asked to highlight aspects of the system that they found favourable or beneficial. This included features, functionalities, or overall experiences that they enjoyed or found particularly useful. By prompting users to express positive feedback, we aimed to identify strengths and successes of the system that could be highlighted and leveraged in future iterations.
- **Areas for Improvement:** Users were requested to identify areas where they believed the system could be enhanced or optimised. This encompassed any shortcomings, limitations, or challenges they encountered while using the system. By soliciting feedback on areas for improvement, we aimed to pinpoint specific aspects of the system that required attention or refinement, guiding future development efforts.
- **Displeasing Aspects:** Users were encouraged to share any aspects of the system that they found displeasing, unfavourable, or unsatisfactory. This could include features that did not meet their expectations, usability issues, or any elements of the system that detracted from their overall experience. By allowing users to express negative feedback, we aimed to uncover pain points and areas of dissatisfaction that needed to be addressed to enhance user satisfaction and system usability.

The collected reviews were segregated into distinct categories based on their content. Within each category, the analysis involved identifying common themes, patterns, or recurring issues mentioned by users. Frequency of specific issues or recommendations was analysed to identify trends or prioritise areas for improvement. Manual analysis enabled the team to capture the diverse perspectives and experiences expressed by users in their reviews.

System A

Based on the user evaluations provided for System A, it is evident that users found the interactivity of the system to be the most commendable aspect. There are mentions that the ability to view data by selecting different years increases interactivity. The users appreciate being able to manipulate the data and explore various aspects of it based on their preferences. Users praised the map visualisation for excellently representing Allied Bombing in Europe during World War 1. There are consistent suggestions for improvement, particularly in enhancing interactivity, providing contextual information, and clarifying symbols used on the map. Recommendations include adding more interactive features, such as directional indicators and airport markings, and providing descriptions to explain the meaning of symbols and arrows used in the visualisation.

System B

The users found that the visualisation effectively communicated information about countries attacked, number of bombs, and time of attack events. This clarity allowed users to quickly grasp key insights without confusion. Users noted that the visualisation facilitated comparison of data by displaying views side by side. This feature enabled users to compare the number of bombings in different countries and understand the distribution of attacks more easily. There are suggestions for improvement regarding user experience, with recommendations to integrate more interactive components and ensure clarity in visual design choices. Feedback also highlighted areas where certain design elements may hinder usability, such as scatter plot points potentially covering each other in concentrated data areas

System C

Users found several strengths in System C, ranging from its user interface design to its ability to effectively visualise complex data. They noted that System C did an excellent job of showing complex connections in the data, making it the most impressive part of the system. This indicates the system successfully translates complex World War I dataset into visual representations that users can easily grasp, aiding in understanding complex patterns and correlations. There are also consistent suggestions for improvement, particularly in terms of interactivity and clarity. Suggestions for improvement include enhancing interactivity to facilitate data exploration, refining colour schemes for better differentiation, and reorganising charts to clarify relationships between variables.

Future work

Based on the evaluations and suggestions received, our future work will aim at enhancing user engagement and comprehension through strategic improvements in interactivity, information clarity, and data narrative. Emphasising on information visualisation concepts, we plan to implement the following enhancements:

1. **Dynamic Interaction:** To address the feedback for more dynamic data exploration across all systems, we will incorporate interactive filtering and search functionalities. This includes brush able timelines for temporal data navigation, clickable legends for selective visibility, and dynamic querying mechanisms to enable users to construct personalised views of the dataset. Utilising concepts such as focus, and context and dynamic queries will significantly improve user engagement and data discoverability.
2. **Information Layering and Contextualisation:** For the map visualisation system, adding directional indicators and marking key historical sites will provide users with essential context. Implementing layered information, where users can toggle between different levels of data detail (e.g., strategic bombing routes, historical battlefronts), aligns with the layered encoding concept. This approach enhances data comprehension without overwhelming the user.
3. **Visual Clarity and Data Density:** To overcome challenges related to visual clarity, particularly in distinguishing data points and understanding symbols, we will refine our visual encoding strategies. This includes employing a more distinct and intuitive colour palette, adjusting data point sizes for better visibility, and utilising tooltips for immediate context provision. Moreover, employing semantic zooming can offer users varying levels of detail and clarity at different zoom levels, improving navigation and understanding of complex datasets.
4. **Enhanced Documentation and Definitions:** Recognising the need for clearer definitions and annotations, especially for users unfamiliar with WWI context, we will integrate comprehensive tooltips and an interactive glossary. This information scaffolding supports learning and exploration, making the visualization systems more accessible to a broader audience.
5. **Reorganization for Analytical Insight:** Feedback on the relationship between bomber performances and their mission frequency highlights the necessity for a more analytical and comparative visualisation approach. We will restructure the presentation to enable side-by-side comparisons and temporal progression tracking, employing coordinated views and linked highlighting to facilitate deeper analytical insights into aircraft development and strategic evolution during WWI.

Implementing these suggestions will significantly enhance our visualisation systems, making them not only more interactive and informative but also more intuitive and accessible for diverse user groups. These improvements, grounded in core information visualisation concepts, aim to elevate the user experience from simple data viewing to engaging in a meaningful exploration and understanding of WWI bombing missions.

References

- [1] Luyi Bai, Xinyi Duan, and Bin Qin. 2022. Adaptive query relaxation and top-k result sorting of fuzzy spatiotemporal data based on XML. *International journal of intelligent systems* 37, 3 (2022), 2502–2520. <https://doi.org/10.1002/int.22781>
- [2] David H. S. Chung, Daniel Archambault, Rita Borgo, Darren J. Edwards, Robert S. Laramée, and Min Chen. 2016. How Ordered Is It? On the Perceptual Orderability of Visual Channels. *Computer graphics forum* 35, 3 (2016), 131–140. <https://doi.org/10.1111/cgf.12889>
- [3] James Walker, Rita Borgo, and Mark W. Jones. 2016. TimeNotes: A Study on Effective Chart Visualization and Interaction Techniques for Time-Series Data. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (January 2016), 549–558. <https://doi.org/10.1109/TVCG.2015.2467751>

Image Reference

Unkown Photographer (1915) *Frankreich, Bombardierung Calais [Photograph]*. German Federal Archive. <https://www.bild.bundesarchiv.de/dba/de/search/?query=Bild+146-2008-0051>

Work Log

Cameron Watson 2392496w:

- System C code
- Design Comparison
- Demo Video

Yuman Li (Amanda) 2842084l:

- The data
- Future work
- Test System

WingChau Lau (Terry) 2872109l:

- System B code
- Generalised selection
- Design comparison
- Demo Video

Raphael Taylor 2940447t:

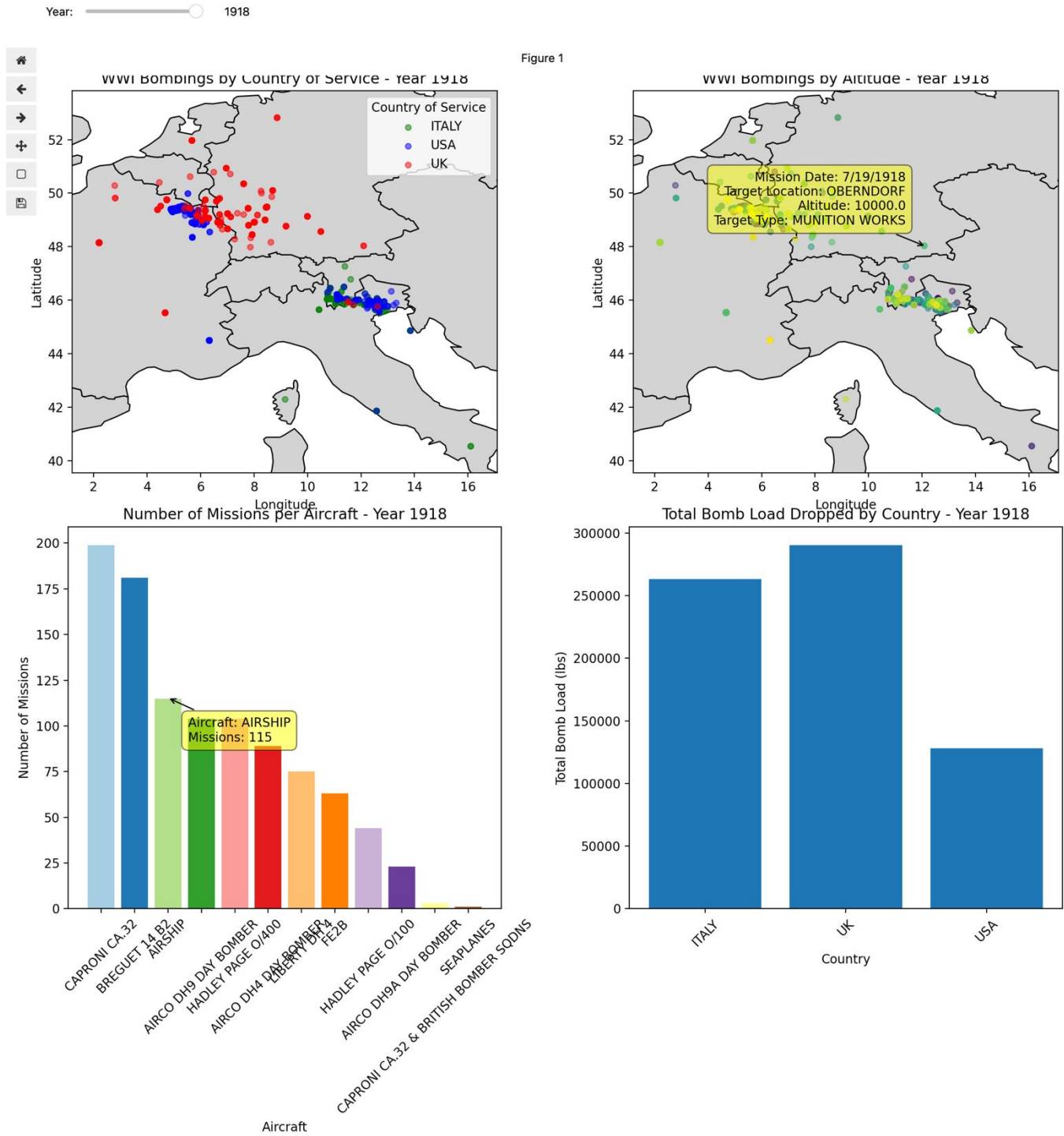
- System A code
- Design Comparison
- Demo Video

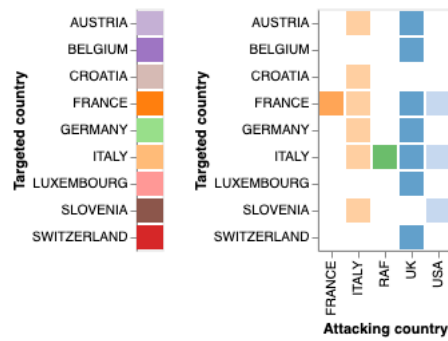
Harsh Suvarna 2892925s:

- Visualisation Tasks
- User Evaluation

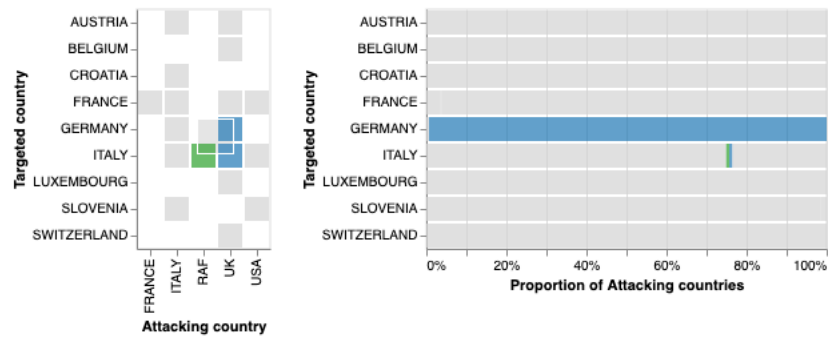
Appendix

SYSTEM A Screenshots

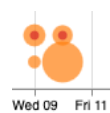




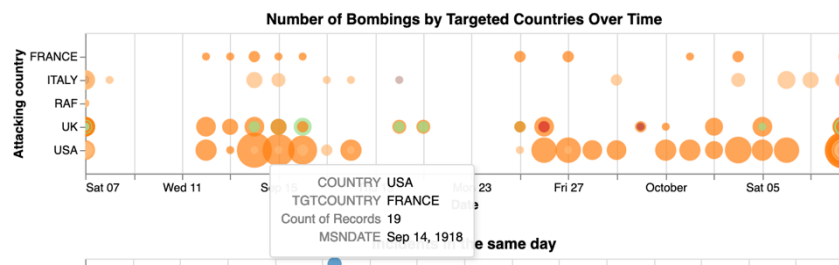
Screenshot B1



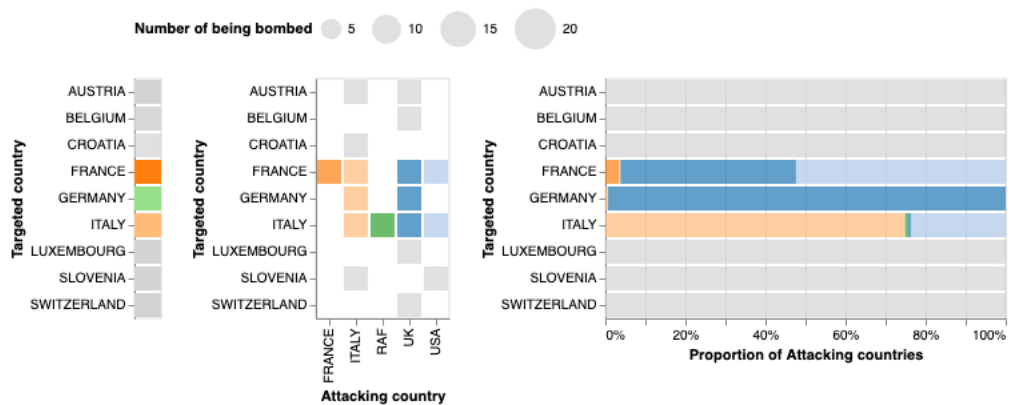
Screenshot B2



Screenshot B3



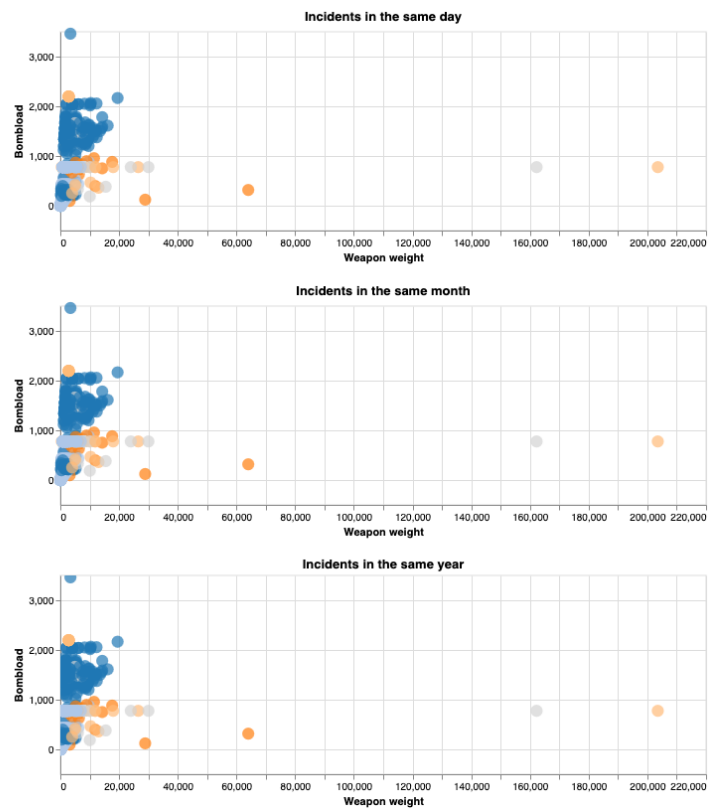
Screenshot B4



Screenshot B5

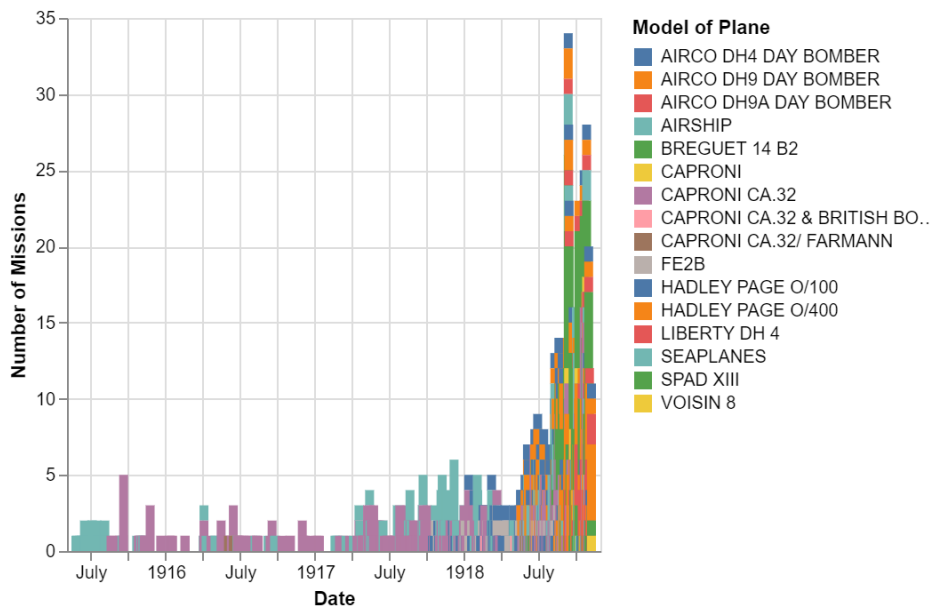
opacity:  0.7

Screenshot B6

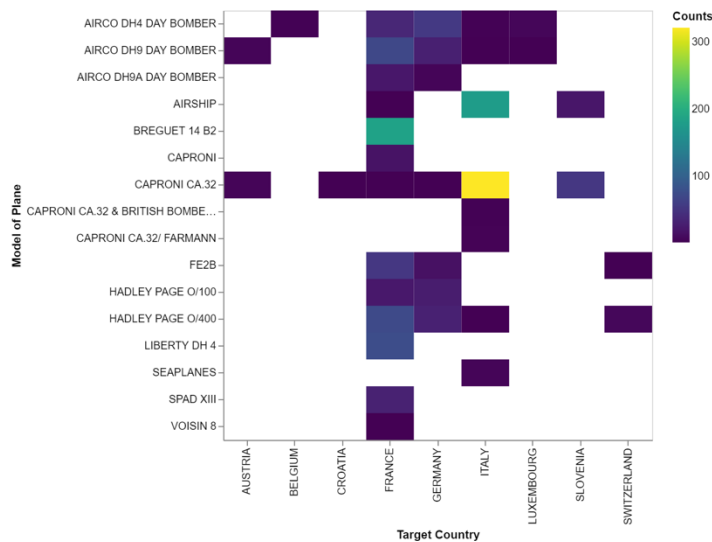
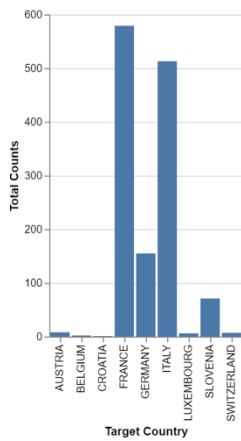
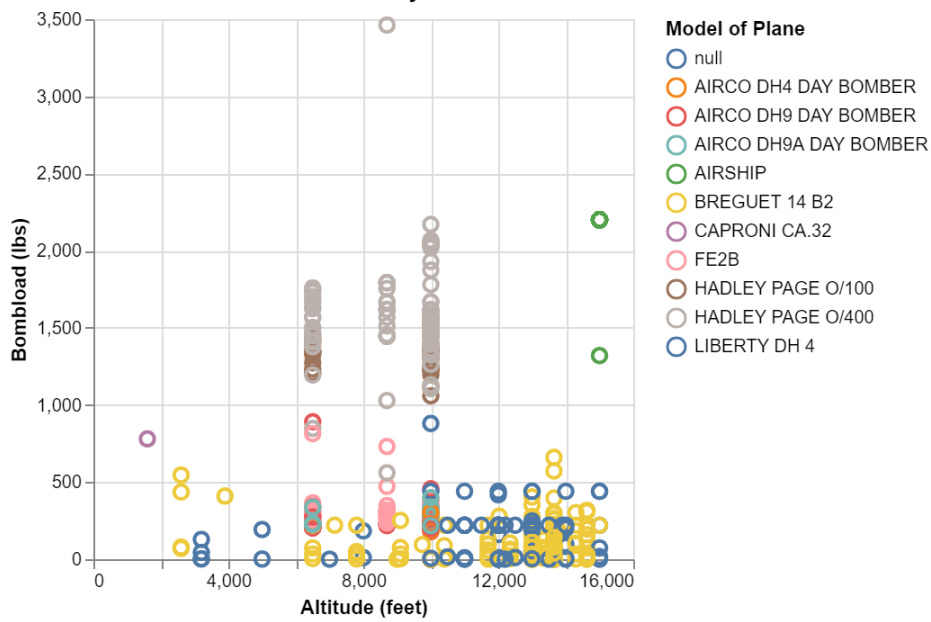


Screenshot B7

SYSTEM C Screenshots



Altitude vs. Bombload by Aircraft Model



Test Results

a.

1. This system effectively utilises JupyterLab UI elements and incorporates Vega visualisation. Its interface is well-designed, and the visualisation tools are adept at representing complex data in an easily interpretable format. The use of colors and fonts enhances readability and user interaction. However, the system could improve by integrating more interactive elements to allow for dynamic data exploration. The code implementation and documentation aspects are thorough, supporting the ease of use **【InfoVis_System.html】**.

2. The map visualisation excellently represents Allied Bombing in Europe during World War 1. The use of Leaflet for mapping and the addition of detailed polyline routes add significant value to the visual representation of historical data. This system excels in its ability to display geographical data accurately and intuitively. However, the user interface might benefit from more interactive features, like filters or search options, to navigate through different years or locations more efficiently **【Map_Visualisation_copy_4.html】**.

3. This system offers a comprehensive and detailed visualization of bombing missions during World War 1, effectively using bar and circle marks in its charts. The data representation is clear, providing insightful information about various aspects of the missions. The visualizations successfully convey the scale and impact of the operations. However, there might be room for enhancing the user experience by integrating more interactive components for data exploration and comparison **【WW1_BOMBING_SYS.html】**.

In summary, while all three systems demonstrate strong capabilities in visualizing complex datasets, there is potential for further enhancement in interactivity and user engagement.

b.

For Map_visualisation_copy:

The visualization might be enhanced by introducing directional indicators, such as arrows, on the black or blue lines representing the allied bombing routes, to show their start and end points. Additionally, marking key airports on the map could significantly improve comprehension by providing context to the depicted routes.

In the case of InfoVis_Note:

This visualization system showcases three key aspects. Firstly, a histogram details the relationship between the dates and the number of bombing missions during the first war, with an interactive feature allowing users to highlight all missions involving the same model of plane by clicking on a column segment. The second aspect illustrates the relationship between fighter altitude and payload capacity. Lastly, it maps out the connection between target countries and the number of attacks. Recommendations include expanding the first visualization for better clarity, and enhancing the second with interactive elements, such as enabling users to click on aircraft types to highlight related data, moving beyond static color representations for differentiation.

Regarding WW1_BOMBING_SYS:

The option for users to adjust opacity might seem like a relinquishment of design responsibility. Effective visual system design entails the designer making informed decisions about optimal transparency levels to present to the users, rather than delegating this choice to them. Ensuring the visual clarity and effectiveness of the display should be a priority, achieved through deliberate design choices rather than user adjustments.

c.

1. The map display of System 1 is very detailed, and you can view the data by selecting different years, which increases interactivity. The blue lines and arrows indicate the direction and scope of the bombing activities, but it does not provide enough information to explain the symbols on the map. Such as the specific meaning represented by the size, width or color of the arrow. If users use it directly, they will have doubts about obtaining information. It is recommended to add function descriptions.
2. System 2 displays the countries attacked, the number of bombs, and the time of attack events very clearly. By placing the views side by side, you can quickly compare the number of bombings in different countries and the proportion of countries attacked, and their distribution is clear at a glance. The disadvantage is that in the scatter plot in the third row, larger scatter points may cover other scatter points, especially in areas where the data is concentrated, which may cause some data to be hidden.
3. The first picture of System 3 clearly shows the changing trend of bombing missions in different years through time series. The disadvantage is that the large number of aircraft models results in a wide range of colors in the histogram, and the stacking of similar colors makes it difficult to distinguish the information. The heat map in the third picture simultaneously displays the number of bombings and the proportion of participating countries, clearly indicating the main target countries and the frequency of attacks by different countries. These charts summarize a large amount of historical data on World War I aircraft bombing missions. They are unique in design but need improvement in data density and visual identification.

d.

1. The map visualization clearly shows the bombing effects for allies (Germany, Austria-Hungary) in World War I. As the bombing mission need bombers start from Airport to the target site. This graph uses clear straight lines to show all mission that succeed in bombing. The graph can clearly demonstrate the course of war in WWI in a special vision. What's more, the develop of aircraft in war can be also speculated through this chart.

However, if this map is shown to someone who lacks of knowledge of the WWI, it may cause some misunderstanding. As the missionary base and targets are not clear in this chart. It also lacks of other form to support the title, like the battlefront of each year, total bombers used during one mission, and basic border of all countries that in war as it may show to people who is not living in Europe.

2. The chart illustrates the countries that suffer from strategical bombing in WWI. It provides a clear view of the bombing strategy of those countries that involved in WWI. It also shows the time options, frequency options and target options of those main countries in WWI.

But what needs to be point out is that some items lacks of clear definition, like RAF, is it refer to Royal Air Force? Also, the target of charts shown is hard to recognize.

3. This chart shows the mission executed by different models of bombers during WWI. From this chart a clear view of aircraft development in WWI is shown through those charts. It also claims different types of plane's performance.

What I think can be improve is that those charts can be re-organized to show a clearer relationship between bomber's performances and their mission frequency.