VISUEL - A Web Dynamic Dashboard for Data Visualization

VISUEL - Un Tablero Dinámico Web para la Visualización de Datos

Antonella S. Antonini[1], María Luján Ganuza[1], and Silvia M. Castro[1]

[1] *VyGLab Research Laboratory, DCIC, Universidad Nacional del Sur and Institute for Computer Science and Engineering (CONICET-UNS), Bahía Blanca, Argentina*

{antonella.antonini, mlg, smc}@cs.uns.edu.ar

Abstract

Data visualization aims to explore and analyze data quickly, interactively, and intuitively using visual representations. Faced with the constant growth of data in terms of volume and diversity, visualization techniques must confront the challenge of dealing with increasingly large datasets in terms of representation, interaction, and performance. Therefore, these techniques must be able to effectively convey the characteristics of the information space and inspire discovery. In this article, a web dynamic dashboard for data visualization called VISUEL is presented. VISUEL supports multiple coordinated views, integrating visualization techniques such as scatterplots, parallel coordinates, and box plots, and interactive schematic maps to represent information enriched with spatial references. VISUEL is fully interactive, supporting traditional interactions like filtering, selection, brushing and linking, and zooming, among others. It also allows the user to configure the visual representation of their data, by selecting the color and shape of the representations. The usefulness of this tool is illustrated using real-life dataset related to the wine industry in Argentina. Important aspects of the case study are discovered through the construction and analysis of multiple views.

Keywords: Data Visualization, Visual Analysis, Visualization Dashboard, Visualization Tool.

Resumen

La visualización de datos tiene como objetivo explorar y analizar los datos de forma rápida, interactiva e intuitiva mediante representaciones visuales. Ante el constante crecimiento de los datos en términos de volumen y diversidad, las técnicas de visualización deben afrontar el desafío de lidiar con conjuntos de datos cada vez más grandes en términos de representación, interacción y desempeño. Por lo tanto, estas técnicas deben ser capaces de transmitir de manera efectiva las características del espacio de información e inspirar el descubrimiento. En este artículo, se presenta un tablero web dinámico para la visualización de datos llamado VISUEL. VISUEL admite múltiples vistas coordinadas, integrando técnicas de visualización como diagramas de dispersión, coordenadas paralelas, diagramas de caja, y mapas esquemáticos interactivos para representar información enriquecida con referencias espaciales. VISUEL es totalmente interactivo y admite interacciones tradicionales como filtrado, selección, brushing and linking, y zoom, entre otras. También permite al usuario configurar la representación visual de sus datos, seleccionando el color y la forma de las representaciones. La utilidad de esta herramienta se ilustra utilizando datos reales relacionados con la industria del vino en Argentina. Se descubren aspectos importantes del caso de estudio mediante la construcción y el análisis de múltiples vistas.

Palabras claves: Visualización de Datos, Análisis Visual, Tablero de Visualización, Herramienta de Visualización.

1 Introduction

Visualization tools have become very popular and useful, allowing the user to explore their datasets and to amplify cognition through the interactive analysis of visual representations of their data [1]. As data is produced at an incredible rate, visualization tools face the challenge of dealing with ever-larger datasets. Frequently, a single view of the dataset may not be enough to present the potentially interesting relationships between the data. To address this problem, the multiple coordinated views technique can be applied, easing the recognition of hidden relationships from observed data [2].

The popularization of the Internet infers that a natural evolution for these visualization tools is the migration to a web platform [3, 4]. With a web application, the users can access the same web tool from any computer connected to the Internet with minimal specifications. In general, web applications are easy to maintain, and can be quickly updated and published without the need to install new software on each computer. However, the development of web applications presents various challenges, like better interface components, customizable interfaces, support for multiple users, support for large datasets, better management of user interactions, and faster responses to user interactions, among others [5].

In this article, a web dynamic dashboard for data visualization called VISUEL, is presented. VISUEL incorporates traditional data visualization techniques such as scatterplots, ternary plots, box plots, parallel coordinates, among others. In addition, VISUEL includes interactive schematic maps for visualizing spatial data and a series of interaction techniques that aid in the interactive exploration of datasets.

The following sections briefly describe the necessary concepts of data visualization (section 2) and the related work (section 3). Section 4 describes VISUEL, including its main characteristics, design, applied technologies, and examples. Then, an example scenario of multidimensional data analysis using VISUEL is described in section 5. Finally, in section 6, the final considerations and the directions for future work are discussed.

2 Data Visualization

Data visualization techniques allow the creation of visual and interactive representations of datasets. Data visualization tools are computer tools that, based on visualization techniques, allow the user to visually manipulate and reorganize data in an efficient way to support their analysis. Using visualization tools, the user could, among other tasks, perform general exploration of the dataset, focus on a subset of the data, reduce the dataset according to a previously defined criterion, and deduce relationships between data items [6, 7].

Given the large amount and variety of data and the various types of users, it is difficult to find a single way to present information in an accessible and useful way to meet the needs of all visualization consumers. In this context, the concept of Dashboard arises as a set of visual representations of the data that allows users to actively participate in the process of analyzing the dataset. A Dashboard is a tool that provides an interactive and centralized means to visualize, analyze and extract relevant information from different datasets in an interactive, intuitive, and visual way [8, 9]. A Dashboard can use two or more different views to support the process of exploring a dataset. One approach to support multiple views is to apply the traditional technique of multiple coordinated views (MCV) [2]. With MCV, the user can simultaneously build several different visual representations for the same dataset to better analyze the data. Coordination ensures that changes made in one view are propagated to all other views, keeping the representations consistent. Coordinated data interaction mechanisms should facilitate the discovery of non-trivial data relationships.

In conclusion, a good visualization tool should provide, at a minimum, techniques, and interactions that allow the user to easily and efficiently manipulate visual representations of the data to better understand the characteristics and relationships of the dataset. Additionally, the MCV technique of the same dataset can greatly facilitate user analysis.

3 Related Work

Over the last few decades, multiple tools have been proposed to explore and analyze datasets. While some of them help domain-specific data analysis (e.g. biology data analysis [10], visualizations in meteorology [11], and visualization of geological data [12, 13, 14], etc.), other tools focus on general-purpose data exploration [15, 16, 17]. The currently available tools and some examples are briefly listed below.

In 2005, North and Shneiderman [18] proposed a classification for information display systems based on MCV. In their description, the systems are classified according to their level of flexibility in data, views, and coordination. Based on this classification and the tools surveyed for our work, this taxonomy was extended, classifying the systems according to their level of flexibility in:

–Accessibility: it refers to whether users have access to the applications easily and for free.

–Data upload: it refers to whether users can load their own datasets with different characteristics, structures and from different application domains.

–View selection: it refers to whether the application provides different views, and whether these views can be configured and linked.

–Interactions: it refers to whether users can configure different views provided by the application for exploratory analysis as appropriate for their data.

Next, an analysis of the compiled tools framed in this classification is presented. The results of this analysis are summarized in Table 1.

Some tools offer a free version of their product, but with limitations such as Zoho Reports [16], which presents a free alternative available for 10 days for different operating systems, or SAS Visual Analytics [19], which also offers a free data analysis product, but limited to certain operating systems. Microsoft Power BI [20], Sisense [21], IBM Watson Analytics [22], and QlikView [23] are some of the non-free platforms used mostly by companies for the analysis of their data.

Studies show that users are slower when using web applications than when using desktop applications due to the limited interaction mechanisms provided by web browsers and the lack of delimitation between web browsers and web applications. However, without doubt, desktop applications have many more drawbacks [5]. Desktop applications (e.g. [24]) must be developed on multiple platforms to support all users, must be downloaded and installed to be used, and are often more problematic to manage and maintain. Numerous web applications have been implemented in recent years to overcome these problems. Users can access the same web tool from any computer connected to the Internet with minimal specifications. ChartBlocks [25] is an online tool where data import is more complex than in other applications that have automated modules or extensions for specific data sources. However, it does not require coding and creates visualizations from spreadsheets, databases, and live feeds. Another example of an online tool is Infogram [26], an infographics software that offers several interactive charts and numerous maps to help the user visualize data in an enjoyable way.

Tableau [15] is a desktop application for visual analytics for business, allowing the creation of charts, graphs, maps, and many other graphics. In this case, there is also the option of a cloud-hosted service, which allows the user to view reports online and in the mobile app. However, in its public version, it does not allow to keep the data analysis private.

Another aspect of data analysis applications to consider is their level of flexibility in data loading.

Applications such as Google Analytics [17], offer a wide variety of charts and support multiple browsers. However, importing the user's dataset can be a complex task and, beyond the tutorials and forums available, have limited support. Other tools, such as Zoho Reports[16], are available for different operating systems, but their free version limits the size of the dataset.

Several applications allow the user to upload his own dataset freely, but not to choose the views to display. These applications are static dashboards, where the user can load his own data and configure some parameters to display them but the result is a static representation.

Other applications offer flexibility in both data loading and choice of views to display. Snap-Together [27, 18, 28], SpotFire [29], Voyager 2 [30], SAS Visual Analytics [19], Sisense [21], and Tableau Software [15] are some examples of these applications.

Regarding interaction, applications like Microsoft Power BI [20] and IBM Watson Analytics [31] allow the users to view and modify data in terms of attribute names and values, and data types. Other applications are primarily focused on exploring information (selecting, navigating, and querying) [27, 28, 18], while business tools such as QlikView [23], Alteryx [32], and academic tools such as Voyager 2 [30], DataSite [33], NorthStar [34], and ForeSight [35], allow users to choose from a set of visual representations (like heat maps, pie charts, line charts, etc.) and, in some cases, visual attributes (like color, texture, etc.).

Modern visualization tools like Keshif [36], NorthStar [34], Voder [37], and Tableau [15] allow users to interactively hover, and zoom in and out over visualizations. Tools like Keshif [36], VisTrees [38], VisFlow [39], and Tableau [15] also offer interactions such as interactive filtering.

Some systems consider different alternatives for the selection of data items and coordination of views. NorthStar [34] links the two dimensions and creates a scatterplot that shows the correlations between the two dimensions. On the other hand, MyBrush [40] focuses entirely on the brushing and linking dimensions. It allows, interactively configuring different components of the brushing and linking operation, such as source, link, and target.

PanoramicData [41], Tableau [15], iVisDesigner [42], Voder [37], DataSite [33], IBM Watson Analytics [31], among others, allow users to compose different visualizations techniques such as scatterplots, pie charts, and histograms, to perform bivariate data analysis. Some tools such as Tableau [15], IBM Watson Analytics [22], Alteryx [32] also allow users to perform join operations on multiple related tables in the same database.

This article introduces VISUEL, a free, easy-to-access web-based system that enables users to design dynamic data visualizations dashboards and explore various datasets interactively. It is a visualization tool based on MCV that integrates various visualization techniques into a user-friendly, efficient and pleasant visual environment. The system supports visual analysis interactions, such as brushing and linking, and visualization customizations. Generated visualizations can be exported to various formats: .PNG, .JPG, .PDF and SVG.

4 VISUEL

4.1 Availability and Design

VISUEL is a coordinated multi-view data visualization tool developed in JavaScript language[1], using React.js[2] for the user interface, and Highcharts[3] and Plotly[4] to display the data. It is a web visualization tool, so no installation process is required. It can be accessed for free from any browser (such as Google Chrome, Internet Explorer, etc.) on the website *http://vyglab.cs.uns.edu.ar/visuel*.

The user interface of the application is designed as a combination of two sections: the logging section and the visualization section. The visualization section (see Figure 1) includes a data upload area, a data table area that allows the user to view and manipulate the loaded data, and a visualization area.

4.2 Data Visualization Techniques

In the visualization area (see Figure 1), the user can display their data using different visualization tech-

[1]https://www.javascript.com/
[2]http://reactjs.org/
[3]https://www.highcharts.com/
[4]https://plot.ly

Table 1: Summary of investigated exploratory data analysis tools. Note: '●' represents the tool 'supports' the operation, and '◐' represents the tool 'support with restrictions' the operation.

Name of tool	Accessibility	Data upload	View selection	Interactions
Zoho Reports [16]	◐	◐	●	●
SAS Visual Analytics [19]	◐	●	●	●
Microsoft Power BI [20]	◐	●	●	●
Sisense [21]	◐	●	●	●
IBM Watson Analytics [31]	◐	●	●	●
QlikView [23]	◐	●	●	●
ChartBlocks [25]	●	◐	◐	◐
Infogram [26]	●	●	◐	◐
Tableau Software [15]	◐	●	●	●
Google Analytics [17]	●	◐	●	●
Snap-Together [27, 28, 18]	◐	●	●	●
SpotFire [29]	◐	●	●	◐
Voyager 2 [30]	●	●	◐	◐
Alteryx [32]	◐	●	●	●
DataSite [33]	◐	●	◐	●
NorthStar [34]	◐	●	●	●
ForeSight [35]	◐	●	●	●
Keshif [36]	◐	◐	◐	●
Voder [37]	◐	◐	◐	◐
VisTrees [38]	◐	●	◐	●
VisFlow [39]	●	◐	◐	◐
MyBrush [40]	●	◐	◐	●
PanoramicData [41]	●	●	◐	●
iVisDesigner [42]	◐	●	◐	◐

niques and integrate those views into a dashboard. The integrated views are correlated. Next, the visualization techniques supported by VISUEL are briefly described.

From now on, for consistency, it will be considered that a dataset is made up of a number n of data items d, and a number x of dimensions dim.

Scatterplot. This technique is based on Cartesian axes to display the values of the two dimensions of a dataset. Each axis represents one dimension of the data. The data are displayed as a collection of graphical elements (points, in most cases). The position of each graphical element is defined by the values of the dimensions associated with each axis (see view D in Figure 1). It is one of the most widely used methods, allowing visual identification of patterns and extreme values [43, 44].

Histogram. It is a visualization technique that shows the frequency distribution of data according to one dimension in a dataset. A histogram groups a range of results into columns along the horizontal axis. The vertical axis represents the frequency of occurrences in the data for each column. In our implementation, the user must select the dimension to trace. If the selected dimension is numeric, the data range is determined and then the horizontal axis is divided into C (number of classes or bars chosen by the user) equal segments (see view F in Figure 1). Finally, the distribution of the classes for each specific range is calculated. If the dimension is categorical, each column refers to a particular category and its height corresponds to the frequency of occurrences of the category.

Ternary Plot. This graph is used to represent trivariate data in which the three dimensions represent the proportions of a whole. The dimensions are plotted on the sides of an equilateral triangle. Each must have values between 0 and some maximum value (default 1) and the sum of the three dimensions must equal that maximum (see view G in Figure 1). Ternary plots are used in many fields, such as earth science[45, 46, 47, 14], agriculture[48], biology [49], among others.

Box plot. A boxplot [50] is a graph that summarizes a dataset measured on an interval scale. It is used to show the shape of the distribution, its central value, and variability. The graph consists of the most extreme values in the dataset (maximum and minimum values), the lower and upper quartiles[5], and the median. Box plots are especially useful for showing comparisons of two or more sets of data items(see view K in Figure 1).

Parallel Coordinates. It is a technique designed in

[5]The lower quartile is the median of all values to the left of the median of the whole set of data. The upper quartile is the median of all values to the right of the median of the whole set of data.

Figure 1: The visualization area in VISUEL. In this area, the user can display their data using different visualization techniques and integrate those views into a dashboard. In this analysis session the user configured a (a) Choropleth Map, a (b) Line Chart, (c) Bar Chart, a (d) Scatterplot, (e) Radial Coordinates, a (f) Histogram, a (g) Ternary Plot, a (h) Pie Chart, a (i) Glyph Map, a (j) Bubble Map, a (k) Box plot, and (l) Parallel Coordinates. At the top left is the data uploading area, and at the top right is a link to the data table area.

the 1970's that has been applied to a wide range of multidimensional datasets [51, 52]. In this method, each dimension of the dataset is associated with an axis A and the x axes are arranged as evenly spaced parallel lines. A data item is represented as a set of connected points, one on each axis. Then, a data item d_i is represented by a polygonal that intersects each axis at the point corresponding to the value of dimension associated with that axis for d_i. In VISUEL, when generating the visualization of parallel coordinates, the user can interactively select the dimensions that he/she wants to add/remove in the chart (see view L in Figure 1). Data points are automatically generated as polylines on the N axes.

Radial Coordinates. Radial and Parallel Coordinates are mathematically equivalent techniques [53, 54, 55]. The difference lies in the geometric layout (see view E in Figure 1). As in parallel coordinates, each dimension of the dataset is associated to an axis A, and the data items are represented by polygonal lines, but in Radial Coordinates diagram the axes radiate at equal angles from a common origin.

Choropleth Map. This graph is widely used for the representation of geospatial data. This kind of map allows the user to display quantitative data using graphical properties such as hue, saturation, or luminance, in predetermined geographic units, such as countries, provinces, states, departments, and regions (see view A in Figure 1). Users can quickly colorize the map simply by selecting a dimension from the drop-down list.

Bubble Map. It is one of the most widely used tools for geospatial data analysis. It is based on a map on which circles representing data items are superimposed. The location of the circles is defined from two dimensions selected by the user, corresponding to the latitude and longitude of the data item. The size of the circle represents a numerical value in the geographic area (see view J in Figure 1).

Line Chart. It is a type of chart that displays information as a series of data points connected by straight line segments. It is similar to a scatterplot except that the measurement points are ordered (typically by their x-axis value) and joined with straight line segments (see view B in Figure 1).

Bar Chart. The Bar chart displays data by a series of bars, each of which represents a particular category. The height of each bar is proportional to the value it represents. The horizontal axis of the graph shows the categories being compared and the vertical axis represents a measured value (see view C in Figure 1).

Pie Chart. This chart is a circle divided into sectors. Each sector represents a numerical value and its arc length is proportional to the quantity it represents. A pie chart shows the relation of the parts to the whole of a categorical or nominal variable (see view H in Figure 1).

Heatmap. A Heat map is a graphical representation of data in which individual values contained in a matrix are represented in different color ranges according to their magnitude (see Figure 8). The color variation can be by hue, saturation, and/or intensity.

Glyph Map. Glyph Maps use a small glyph or icon, namely bar chart, pie chart (see view I in Figure 1) or heat map, to represent multivariate data at each geographic location.

4.3 Data input format

To make this tool easily applicable to a wide range of problems, VISUEL admits CSV files (comma-separated list of values) as input data format, a standard format that is commonly used and exported through spreadsheets and similar applications. The input file should have the format depicted in Figure 2. Each row in the table represents a data item *d* described by a set of dimensions *dim*. The first row of the spreadsheet is reserved for the dimension's labels, that VISUEL will use to identify the dimensions. The following rows are reserved for the data items, that should be loaded in consecutive rows. All data items in the file must have the same number of dimensions (bigger than zero), and it is not required that the data items or dimensions follow a specific order. The first cell of unused rows or columns should be empty.

To start the analysis process, the user must load a file from his own computer with data previously compiled in this format (*.csv). VISUEL works with two types of dimensions: numeric (integers or real) and categorical (variables that contain a finite number of different categories or groups). When loading a dataset, the system infers the data types of each column (dimension) and displays the whole dataset in the data table panel. For example, the input file of Figure 2 contains three data items and seven dimensions. VISUEL will infer that the dimensions with labels *dimension4* and *dimension7* are categorical and that the rest are numerical. The user can modify the data types inferred by the application by clicking on the label of the corresponding column.

It is recommended to process the database before starting to visualize the data. During this stage, the removal of useless dimensions or the cleaning of erroneous data, among other tasks, can be performed. If necessary, restructuring the data table can be another important procedure, as the structure of the dataset will influence how well certain visualization techniques will be supported.

Finally, it should be noted that to configure the visualization views it may be necessary for your data source to include certain types of information. For example, the Bubble Map requires the user to include in their dataset the geographic dimensions (latitude and longitude) and the numeric dimension that will be associated with the size of the bubble. The Choropleth Maps of Argentina and the World, require that one dimension be defined with the names of the provinces or countries respectively, and to indicate the numeric dimension that will be associated with the color. In Glyph Maps, one dimension with the names of the provinces must also be specified and one or more numeric dimensions must be selected.

4.4 Interactions

VISUEL supports a rich set of interactions that allow the user to explore their datasets and drill down to the regions of interest. This subsection briefly describes each of them:

Selection/Brushing and Linking. Selection is one of the best-known and most widely used interactions in the visual exploration process, as it enables to isolate subsets of data and then highlight, eliminate or analyze them in-depth. While brushing allows data items within a chart region to be highlighted, linking connects the highlighting between two or more charts so that all linked charts highlight the same data items. VISUEL supports these interactions on most views. In Parallel Coordinates, the user can select a subset of data by drawing a line on axes that cross the polylines (see view A in Figure 5). In other charts such as Scatterplots and Triplots, the user can select data items using a rectangular brush. Figure 3 shows an example of Brushing and Linking in VISUEL, where a subset of the data is brushed in the Scatterplot (red rectangle in view *A*), and the selected points are highlighted in view *B* (Bubble Map).

Filtering. Users often need to reduce the size of datasets to explore and analyze certain subsets in depth. One way to do this is by filtering. Dimension filtering on radial and parallel coordinate charts, for example, hides some of the dimensions to reduce clutter while preserving the core information of the dataset. From a drop-down box, the user can manually select the dimensions to be represented in the chart (see view A in Figure 4).

Navigation. Navigation allows scrolling through visualizations and focusing on areas of interest. A common navigation technique is zooming in/out, present in most visualization tools, it allows to selectively enlarge the data visualization and to scroll through the surrounding data. In addition, techniques such as Boxplot implement a semantic zoom that allows additional information to be revealed

	A	B	C	D	E	F	G	H
1	dimension1	dimension2	dimension3	dimension4	dimension5	dimension6	dimension7	
2	23	5.7	0.005	blue	0	3456	true	
3	45	8.9	0.56	red	1	3424	false	
4	234	0.9	0.05	blue	1	532	true	
5								

Figure 2: Example of a data input file with three data items and seven dimensions. The dimensions with labels *dimension1*, *dimension2*, *dimension3*, *dimension5*, and *dimension6* are numerical and the dimensions with labels *dimension4* and *dimension7* are categorical.

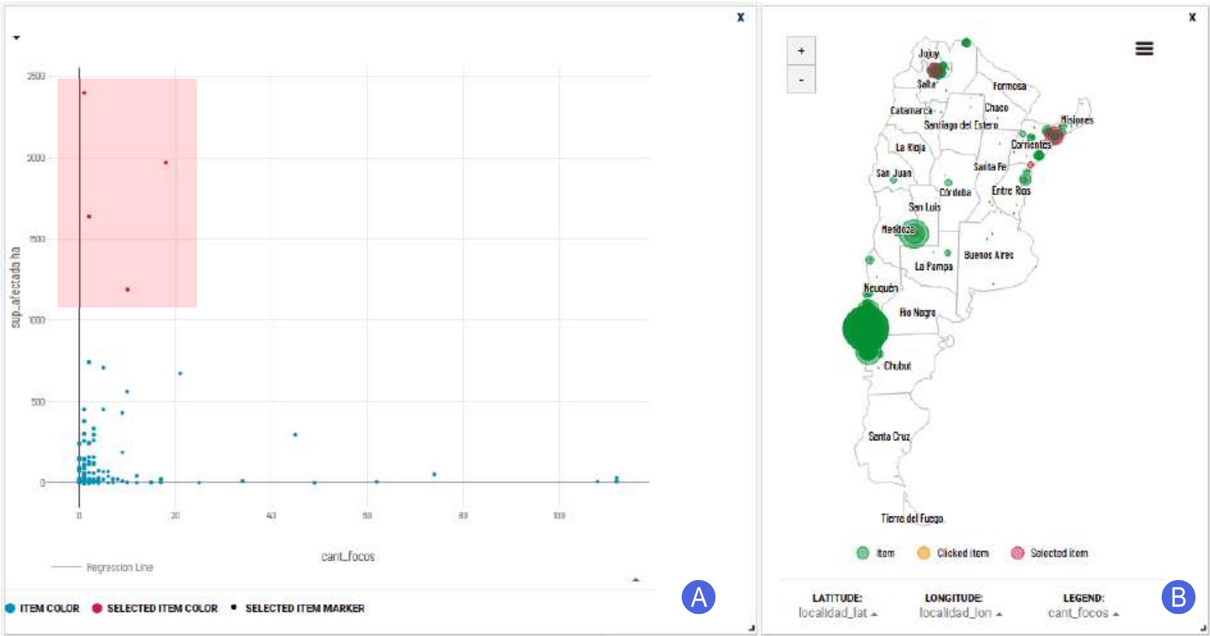

Figure 3: Brushing and Linking in VISUEL. A subset of the data is brushed in the Scatterplot, and the selected items (red points in view A) are highlighted in the Bubble Map (view B).

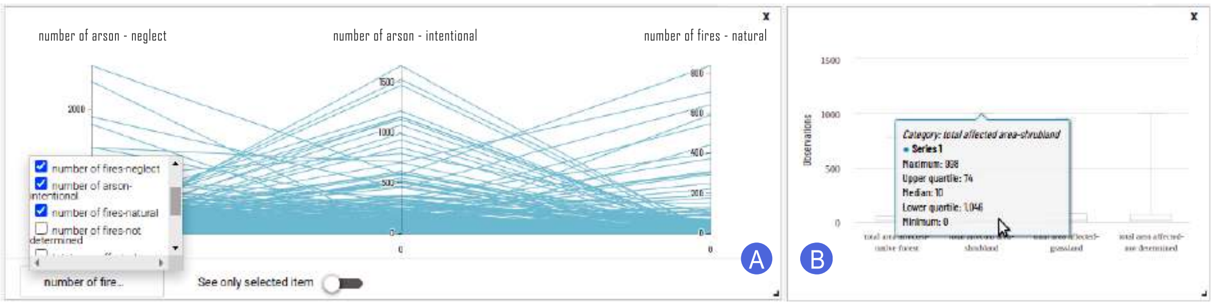

Figure 4: Filtering and Semantic Zoom in VISUEL. In view A the user can manually select the dimensions to be plotted in the parallel coordinates using a drop-down box (filtering). In view B, the user can ask for additional information to be included on the views using a boxplot box (semantic zoom).

on user request (see view B in Figure 4). The user can select a Boxplot box on the screen to obtain additional information (the median, approximate quartiles, and the lowest and highest data points in the box).

Visual Encoding. VISUEL provides the user with the necessary tools to customize the visualizations according to their needs, allowing the modification of the visual appearance of the data items, such as size, shape, and color. View D in Figure 5 shows the result of modifying the visual appearance (in particular, the shape) of the data items. From a drop-down box, the user can manually select the graphical representation of the data in the view (see view C in Figure 5). In addition, it is possible to change the visualization technique applied to display the data. In the Glyph map, for example, the user can choose between bar charts or pie charts to visualize his dataset on the map.

Axes reordering. The order of axes in parallel coordinates has a great influence on the visual exploration of datasets. VISUEL allows users to manually select and rearrange the PCP axes to find a good distribution. View B in Figure 5 shows the result of reordering the axes with labels "number of fires-intentional" and "number of fires-neglect" from view A.

5 VISUEL in use. Vitiviniculture in Argentina

The Argentine wine sector has stood out since the 1990s, for the sustained growth in its production and exports. Argentina is the fifth world producer of wine and the main exporter of grape must worldwide. Both in Cuyo[6], the main wine-producing region in South America, as in the rest of the country, viticulture is a source of employment and one of the factors that highlight Argentina's economic importance in the world.

Given the regional and national importance of this activity in Argentina, the evolution of production, consumption, and marketing, and the regional and international growth of the agribusiness are analyzed.

5.1 Dataset

The National Institute of Viticulture (INV)[7] is a public law institution, linked to the Ministry of Economy of the Argentine Republic, whose objective is to control the genuineness of wine products and control the production, distribution, and marketing of ethyl alcohol and methanol. The INV publishes annual research and statistical reports on the commercialization of wines in the domestic and foreign markets, imports of wines and musts per year, cultivated areas, and various statistics relevant to the wine industry in Argentina. The data used for the analysis were extracted from these reports. Different tables were prepared to configure the visualizations. This procedure resulted in four datasets used in this case study.

The first dataset corresponds to the distribution of vineyards and wineries, by territorial unit in Argentina, and has 526 records. The first row of the spreadsheet is reserved for the labels. The first two columns contain the location (*latitude* and *longitude*) of the department or district, and the remaining two contain the *number of vineyards* and the *number of wineries* respectively.

The second dataset consists of 254 records and contains the data corresponding to the cultivation, production, consumption, and commercialization of wine by province from 2010 to 2020. The first column contains the *name of the province* and the following columns have the following data: the *year*, the *area cultivated* in hectares in the province, the *number of vineyards*, the *production of grapes* in quintals, the *number of quintals of grapes destined for wine production*, the *number of quintals of grapes destined for fresh consumption and raisins*, the *production of wine* in hectoliters, and the *participation in the internal and external market*.

The third dataset corresponds to wine exports in hectoliters by country of destination per year. The first column contains the *name of the destination country*, and the following columns contain the *volume of wine exported per year (2013-2020)*, in hectoliters. The dataset contains 48 records.

The last dataset has 20 records corresponding to total wine production and marketing per year. The first column indicates the year, and the next four columns indicate wine production in hectoliters, the participation of Argentina in the internal and external market, and the volume of imported wine (Hl).

5.2 Distribution of vineyards and wineries

Considering the importance of wine production in Argentina, the first characteristic that can be analyzed is the distribution of vineyards and wineries across the country.

As the dataset contains geospatial references, the analysis can start by configuring the bubble map view. For this, the latitude and longitude from the dataset are selected, distributing the bubbles across the map according to the locations of the territorial units that have wine production. Besides, the size of the bubbles can be associated with any quantitive dimension of the dataset. If the number of wineries is selected, the size of each bubble reflects the number of wineries located in a territorial unit (see Figure 6A). Alternatively, as shown in Figure 6B, the number of vineyards can be associated with the size of the bubbles.

Figure 6 shows the results of these configurations. The territorial distribution of the bubbles reveals a significant concentration of vineyards and wineries in the provinces of Mendoza and San Juan, as expected, due to the fact that these are the provinces that traditionally concentrated wine production and elaboration.

As a next step, the distribution of the cultivated vid area across the country and its evolution over the last ten years can be analyzed. For this, the Choropleth Map provided by VISUEL is employed. For the configuration of the Choropleth Map, the dimension "Province" from the dataset is chosen and the quantitative dimension "Cultivated area " is associated with the color of each province.

Figure 7 shows the distributions of the cultivated vid areas across Argentina in 2000 (A), 2010 (B), and 2020 (C). Figures 7B and C reflect a slight drop in the last 10 years in the provinces of Mendoza and

[6]The Cuyo region is made up of the provinces of Mendoza, San Juan, and San Luis.

[7]https://www.argentina.gob.ar/inv

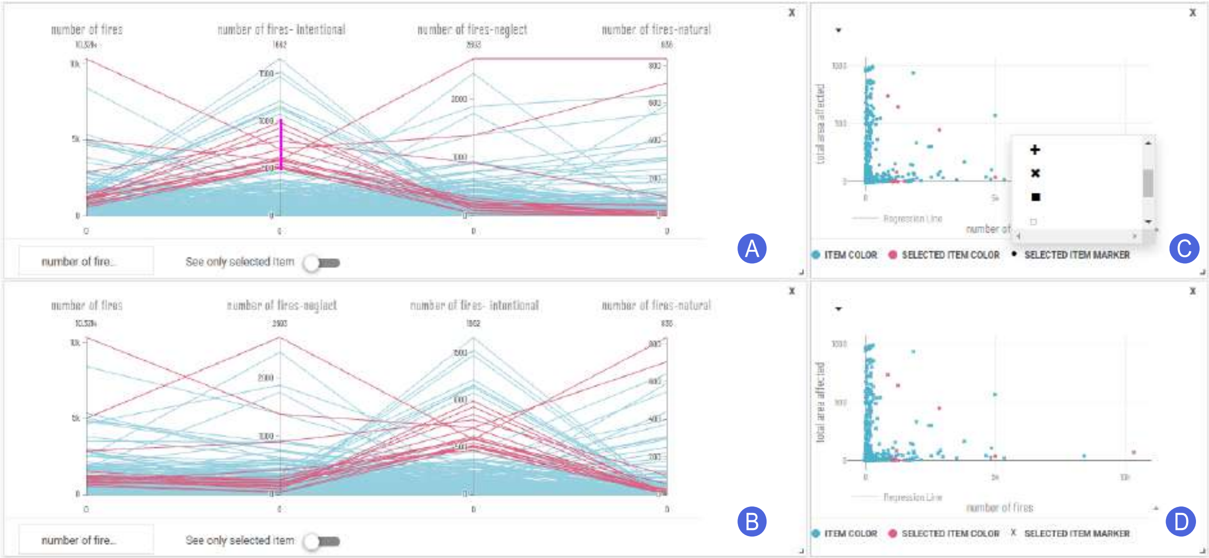

Figure 5: Selection, axes reordering, and visual encoding in VISUEL. In view A a subset of data in parallel coordinates is selected. In view B, the subset of data selected in A is highlighted and axes are reordered. In view C, the user can manually select the graphical representation of the data using a drop-down box (visual encoding). Finally, in view D, the results of the modifications performed in C are shown.

San Juan. However, the total area is still significantly greater than the area registered in 2000 (Figure 7A). In the last 20 years, there was a notable increase in the cultivated vid areas in provinces like Mendoza, Misiones, Neuquén, La Pampa, and Chubut.

5.3 Wine Trends in Argentina

In Argentina, 9,430 hectoliters of wine are consumed per year, which places the country in seventh place in the world for per capita consumption. According to the data published by the INV, red wines lead the consumption trends in most provinces with 71% of the total, followed by white with 20% and rosé with 9%.

The analysis of wine consumption trends in Argentina starts by analyzing the area cultivated with vine (Ha) for the production of red, white, and rosé wines in each province in 2013-2019. To do this, the Argentina heatmap view was set up selecting the dimension “Province” from the dataset and associating the dimension “Year” with the columns, the dimension “Color” with the rows, and the dimension “Cultivated area” with the heatmaps cells.

Figure 8 shows the result of the settled configuration. In most glyphs, it can be observed how the cultivated area for red wine production predominates over that for white and rosé wine. In addition, it is appreciated how it fluctuated over the years, particularly in the provinces of Santiago del Estero in 2017 and Misiones in 2013.

The next step is to analyze the wine consumption by wine color, configuring the Argentina histogram map view. To do this, the “Province” dimension is selected and the consumption of red, rosé, and white wine is associated with each bar of the histogram. Figure 9 shows the consumption preferences in Argentina in 2017 (A) and 2018 (B). It can be deduced that there is a significant increase in wine consumption in the country in general, and in particular in the provinces of Catamarca, Salta, La Rioja and Entre Ríos. The increase in the consumption of white wines in certain provinces such as Buenos Aires, Jujuy, Catamarca, and La Rioja, and of rosé wines in La Pampa can be also highlighted. On the other hand, there was a decrease in the consumption of white wines in La Pampa and rosé wines in Buenos Aires and Jujuy.

5.4 The role of Argentina in the international wine market

Regarding foreign trade, Argentina is among the top 10 wine exporters, distributing its products to more than 130 countries around the world [8]. Due to the role that Argentina occupies in the foreign wine market, the next goal is to carry out a year-on-year comparison of wine exports by country of destination, to analyze how the foreign market in our country has changed over the years. To this end, the Choropleth Map view can be configured by selecting the “Destination country” dimension from the dataset, and associating, in the first place, the exports corresponding to the year 2013 with the color of the countries. In the same way, different Choropleth Map views to visualize exports from 2013 to 2020 can be configured.

Figure 10 shows the results of these configurations and exposes how the foreign market changed for Ar-

[8] https://enolife.com.ar/es/de-junio-2019-a-junio-2020-argentina-paso-de-10mo-a-8vo-exportador-mundial-de-vino/

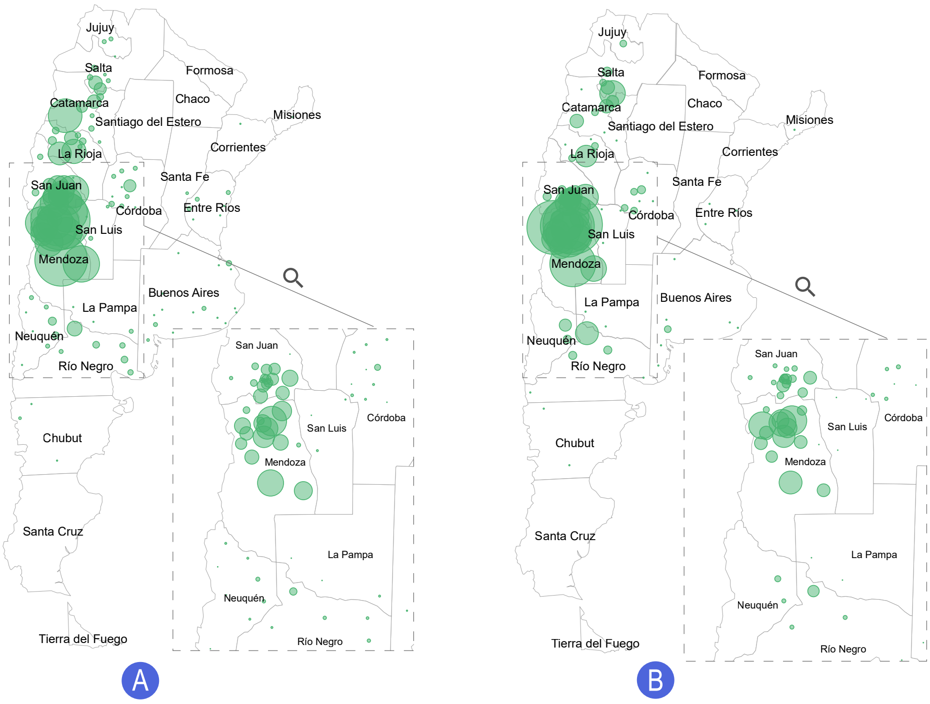

Figure 6: Bubble Maps in VISUEL showing the distribution of (A) vineyards and (B) wineries across Argentina.

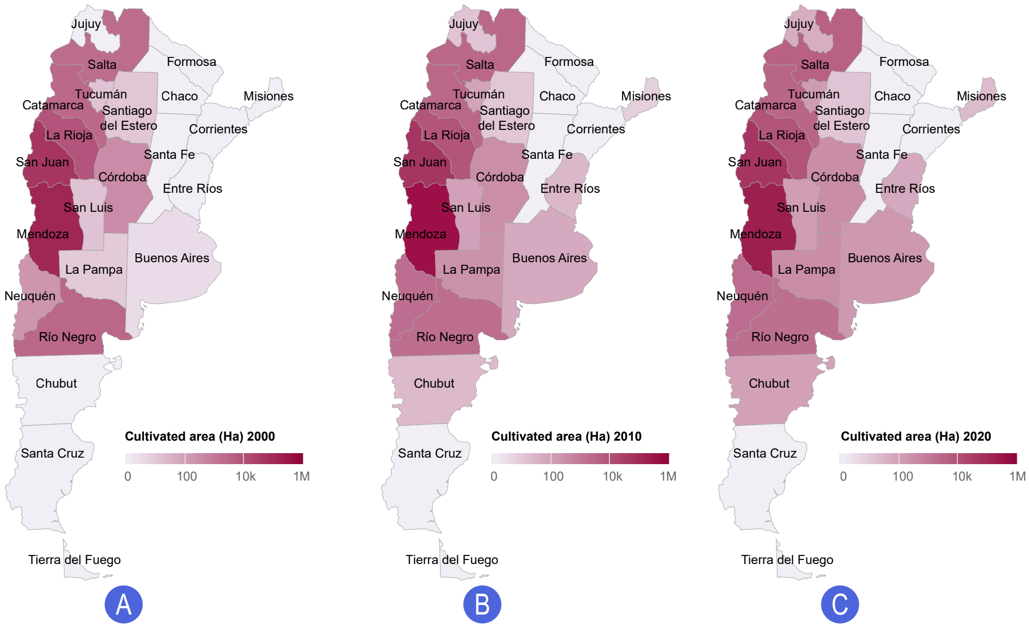

Figure 7: Choropleth Maps in VISUEL showing the distributions of the cultivated vid areas across Argentina in (A) 2000, (B) 2010, and (C) 2020.

gentina in terms of export volume and destination countries. In this figure, it can be seen how the volume exported to the United States progressively decreases over the years, however, it remains as the first market. On the other hand, it can be seen that the United Kingdom, Canada, China, and Brazil are also large importers of our win products. Finally, in 2020 there is an increase in the volume of wine exports, compared to 2019, which allowed the country to climb several positions in the world ranking of wine exporting countries.

5.5 The unforgettable Mendoza harvest of 2016

The effects of climate change undoubtedly affected Mendoza in 2015 A summer that started out hot and then turned rainy and cool, significantly reduced the

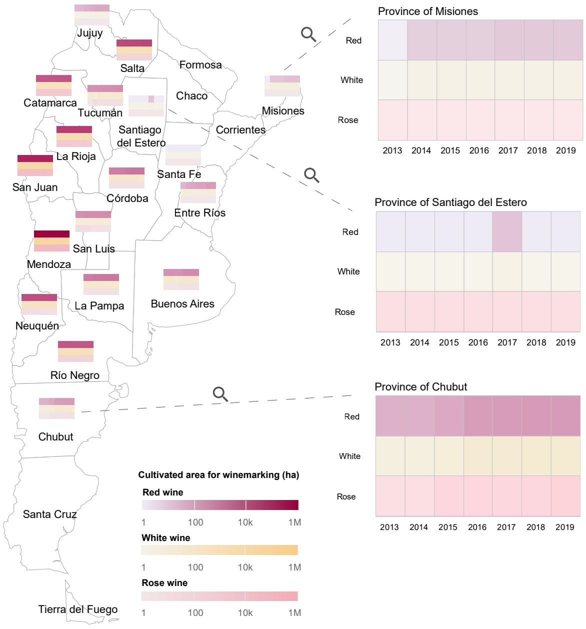

Figure 8: Heatmap in VISUEL showing the area cultivated with vine (Ha) for the production of red, white and rosé wine in the 2013-2019 period

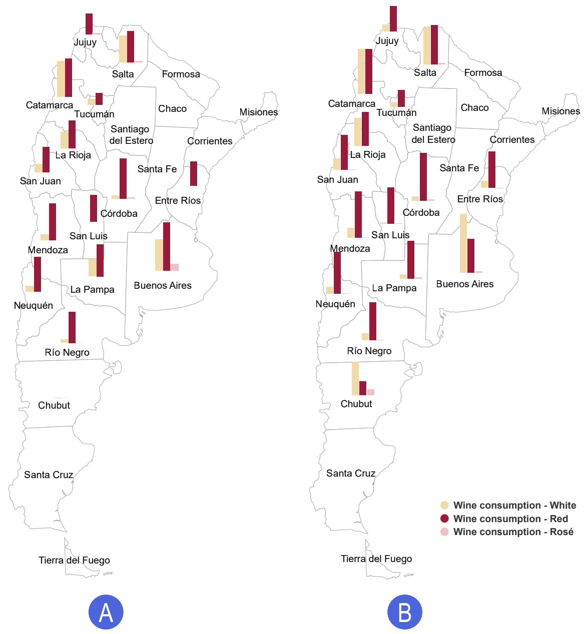

Figure 9: Glyph maps in VISUEL showing the preferences of Argentines in wine consumption in (A) 2017 and (B) 2018.

yield of several vineyards in the area. However, conditions leveled off in February and the producers managed to end a favorable season with wines of good quality and concentration. In this context, the 2016 harvest was not a good one, especially in Mendoza, which had to endure various inclement weather conditions, such as frost, hail, and very high humidity, and diseases, such as grapevine moth and peronospora.

To continue the analysis, an interactive dashboard to observe what happened in the Argentine grape harvests from 2010 to 2020 was created. A Choropleth Map view of Argentina was configured to show the distribution of cultivated vine areas. The dimension "Province" from the dataset was selected and the quantitative dimension "Cultivated area" was associated with color. Next, two-time series charts were set up. First, the Bar chart was used to visualize the data corresponding to wine production and processing. For this, the "Year" dimension was selected to determine the interval of the time series and link the dimensions "Total grape production (qq)", "Destination of grapes-Wine(qq)", "Destination of grapes-Fresh consumption (qq)" and "Destination of grapes-Raisins (qq)" with the bars of the chart. Also, the line chart to show the information corresponding to the grapevine crop was set up, associating the dimension "Cultivated area (Ha)" with the line of the chart.

Figure 11 shows the result of these settings. To see in detail what happened in the Mendoza vintages, VISUEL allows filtering this information by selecting the province in the Choropleth Map (see view A in Figure 11). All dashboard views will be automatically updated at the same time showing the data of interest. By observing the views and interacting with them it can be seen, for example, that the cultivated area with vines in Mendoza in 2016 (approximately 158,000 ha), was slightly lower than the previous year (see view C in Figure 11). However, the harvest of that year (approximately 10 million quintals) had a sharp drop of almost 40% of its production with respect to 2015, as expected (see view B in Figure 11). Alternatively, as shown in Figure 12, the province of San Juan can be selected on the Choropleth Map finding that in 2016, San Juan had a slightly higher harvest than the previous year (see view B in Figure 12).

5.6 About wine importation

After analyzing Argentina's role in the wine market, it may be difficult to understand why a country that is among the world's leading exporters would need to import wine or must from other countries. To understand this situation, the relationship between wine production, domestic and foreign trade, and wine imports can be analyzed. For this matter, the Line chart view is configured, selecting the Year dimension to determine the interval of the time series, and associating the "Wine production(Hl)", "Participation in MI(Hl)", "Participation in ME(Hl)", and "Wine import(Hl)" dimensions with the lines of the chart.

Figure 13 shows the result of the above configuration. Looking at it in detail, it can be seen that a considerable drop in wine production (red line) is followed by an increase in imports in the subsequent year (green line). Therefore, it can be deduced that faced with the adverse situation of multiple abnormally scarce harvests, particularly that of 2015 and 2016, the wine

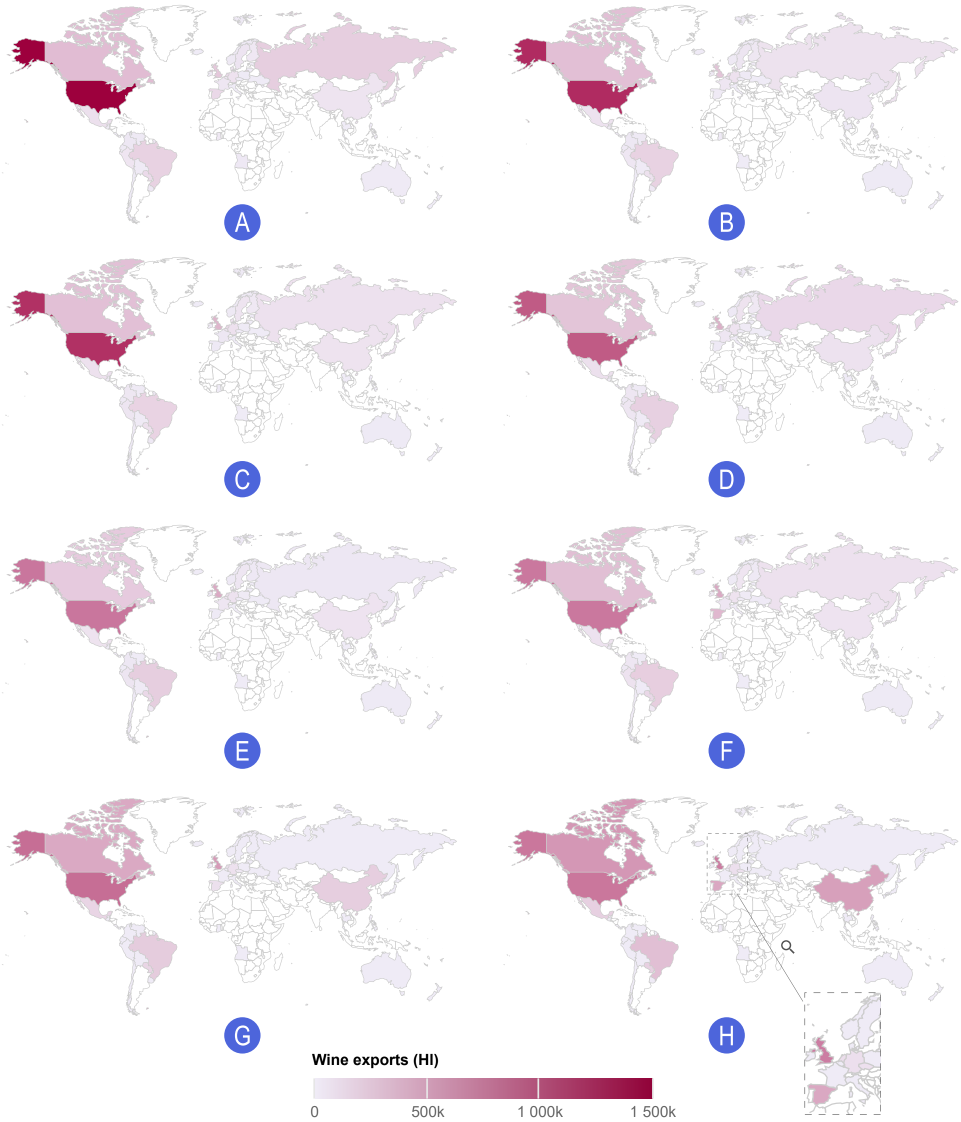

Figure 10: Choropleth maps in VISUEL showing wine exports in hectoliters by country of destination in (A) 2013 - (H) 2020.

stock of Argentine wineries was significantly reduced, forcing the importation of wine from other countries to ensure the supply of the domestic and foreign markets.

6 Conclusions and Future Work

For many years, researchers have made a great effort to improve existing visualization techniques and have proposed multiple valuable software tools that help users in exploratory data analysis.

This paper introduces VISUEL, a public data visualization web dashboard that integrates multiple visualization techniques: Parallel Coordinates, Scatterplots, Triplots, Box plots, Histograms, among others, and basic interactions that aid in exploratory data analysis such as zooming, brushing and linking, and filtering. It also allows the analysis of geo-referenced data as it includes several types of map visualization: Bubble Maps, Choropleth Maps, and Glyph Maps.

Our system was designed to be a flexible tool for the interactive generation of customized visualization environments. VISUEL allows users to quickly and dynamically configure, resize and combine multiple views, and coordinate them to build dashboards without programming, according to their specific data and task needs.

While VISUEL already provides very comprehensive functionality, there are still several ways to extend it. We highlight a number of directions for future work:

- Explore alternatives to increase the processing capacity of the application. Currently, VISUEL can handle relatively large databases (approximately 10,000 data items with 10-15 dimensions each.), however, as the number of data increases, the browser's capabilities decrease.

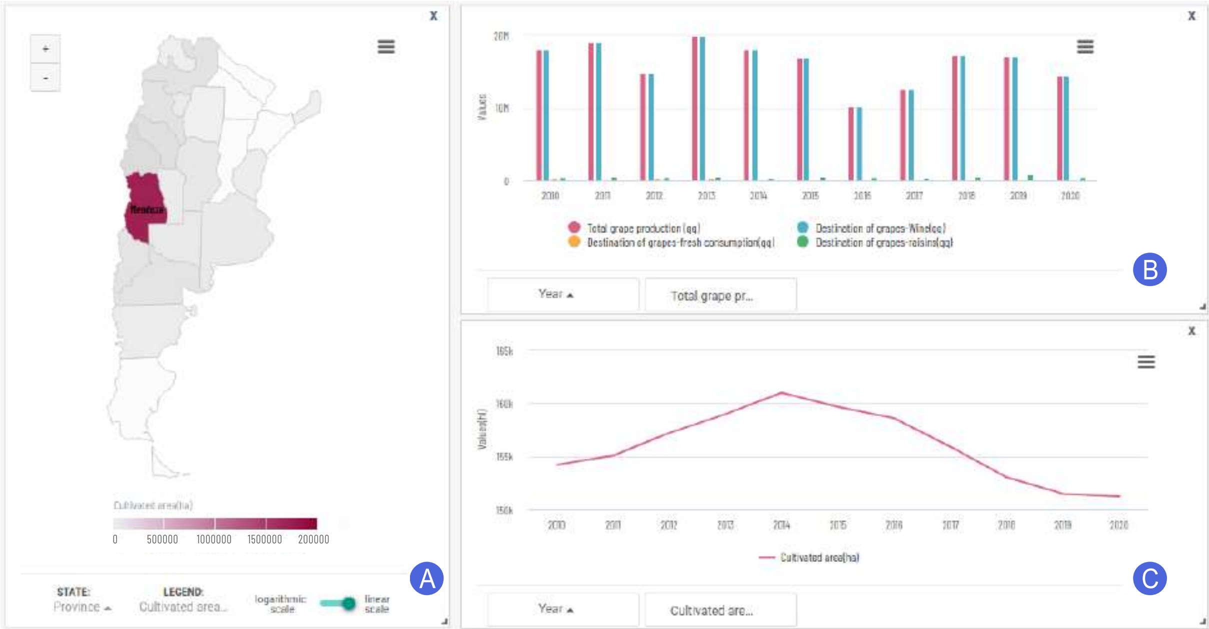

Figure 11: Screenshot of a dashboard created with VISUEL showing data related to the harvests from 2010 to 2020 in the province of Mendoza with different views of the data: Choropleth Maps (view A), Bar Chart (view B), and Line Chart (view C).

- Focus the Usability of the tool on inexperienced users. Provide the user with a mechanism that suggests the appropriate techniques and corresponding configuration options based on the data uploaded by the user.

- Allow users to save sessions, including useful settings and activity histories.

Competing interests

The authors have declared that no competing interests exist.

Authors' contribution

ASA, MLG, and SMC carried out the conception of this work and the design of the web application. ASA implemented the web application and performed the state of the art. ASA, MLG, and SMC formulated and analyzed the case study. All authors worked on the general writing of the article and the reviews provided and approved the final manuscript.

References

[1] S. T. Card, J. D. Mackinlay, and B. Scheiderman, *Readings in Information Visualization, using vision to think*. San Francisco: Morgan Kaufmann, 1999.

[2] M. Scherr, "Multiple and coordinated views in information visualization," *Trends in information visualization*, vol. 38, pp. 1–33, 2008.

[3] S. Wozny, "Web based data visualization solutions in quality assurance." Available at: `https://eestec.net/media/it-team/wozny15.pdf`, september 2015. FORSCHUNGSPRAXIS.

[4] F. Mwalongo, M. Krone, G. Reina, and T. Ertl, "State-of-the-art report in web-based visualization," *Computer graphics forum*, vol. 35, no. 3, pp. 553–575, 2016.

[5] P. Pop, "Comparing web applications with desktop applications: An empirical study," *Linköping University, Linköping*, 2002.

[6] T. Munzner, *Visualization analysis and design*. CRC press, 2014.

[7] C. Tominski and H. Schumann, *Interactive Visual Data Analysis*. CRC Press, 2020.

[8] S. Few, *Information dashboard design: The effective visual communication of data*, vol. 2. O'Reilly Sebastopol, CA, 2006.

[9] S. Wexler, J. Shaffer, and A. Cotgreave, *The big book of dashboards: visualizing your data using real-world business scenarios*. John Wiley & Sons, 2017.

[10] M. Lawrence, E.-K. Lee, D. Cook, H. Hofmann, and E. Wurtele, "explorase: Exploratory data analysis of systems biology data," in *Fourth International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV'06)*, pp. 14–20, IEEE, 2006.

[11] M. Rautenhaus, M. Böttinger, S. Siemen, R. Hoffman, R. M. Kirby, M. Mirzargar, N. Röber, and R. Westermann, "Visualization in meteorology—a survey of techniques and tools for data analysis tasks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 12, pp. 3268–3296, 2018.

[12] M. L. Ganuza, S. M. Castro, G. Ferracutti, E. A. Bjerg, and S. R. Martig, "Spinelviz: An interactive 3d application for visualizing spinel group minerals," *Computers & Geosciences*, vol. 48, no. 0, pp. 50–56, 2012.

[13] M. L. Ganuza, G. Ferracutti, M. F. Gargiulo, S. M. Castro, E. Bjerg, E. Gröller, and K. Matković, "The

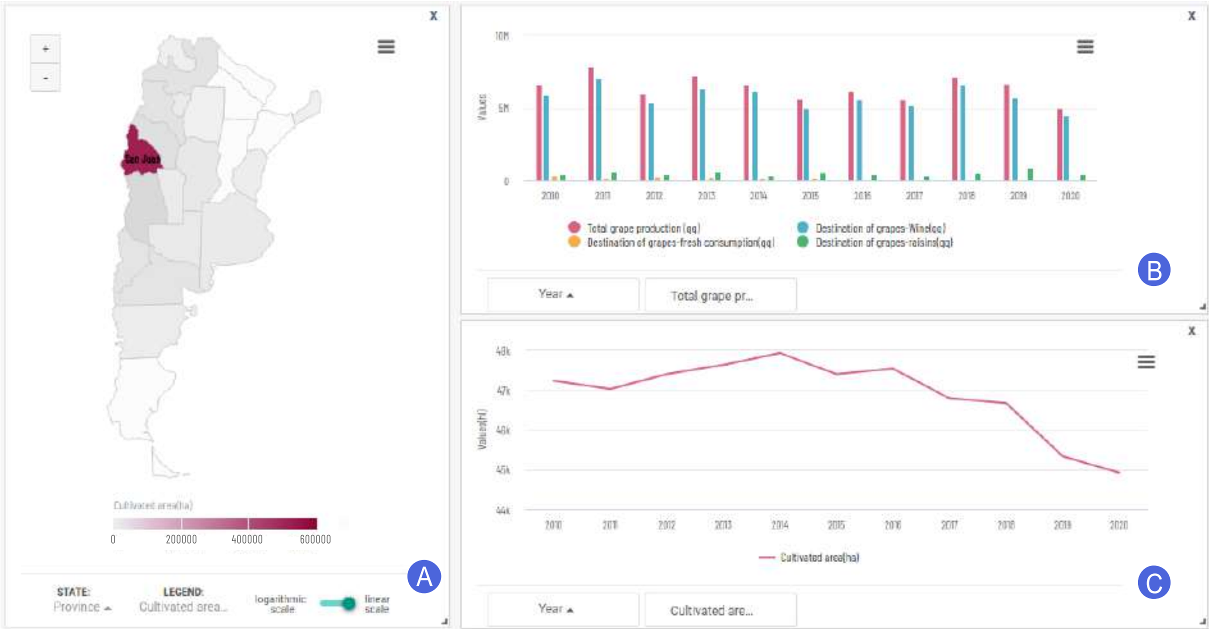

Figure 12: Screenshot of a dashboard created with VISUEL showing data related to the harvests from 2010 to 2020 in the province of San Juan with different views of the data: Choropleth Maps (view A), Bar Chart (view B), and Line Chart (view C).

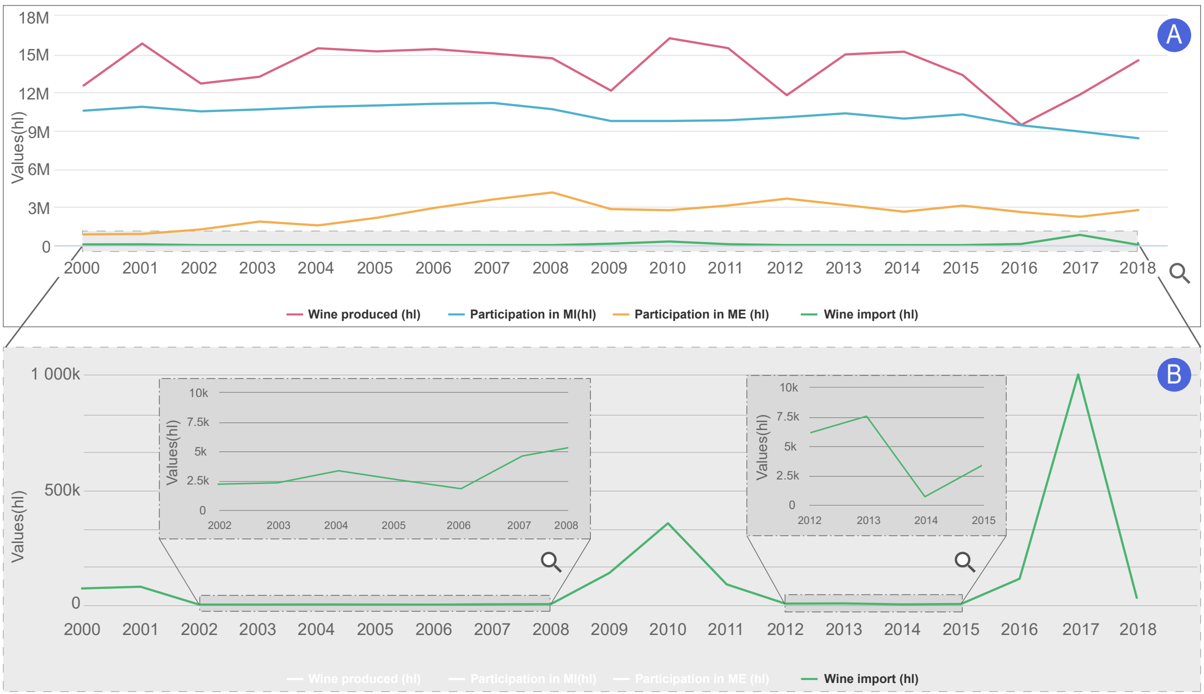

Figure 13: (A) Line chart in VISUEL showing the relationship between exports, imports, and the production (Hl) of wine in Argentina in 2000-2018. (B) A magnification of the imports line is shown at the bottom for clarification.

spinel explorer—interactive visual analysis of spinel group minerals," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1913–1922, 2014.

[14] A. S. Antonini, M. L. Ganuza, G. Ferracutti, M. F. Gargiulo, K. Matković, E. Gröller, E. A. Bjerg, and S. M. Castro, "Spinel web: an interactive web application for visualizing the chemical composition of spinel group minerals," *Earth Science Informatics*, vol. 14, no. 1, pp. 521–528, 2021.

[15] "Business dashboards: Data at a glance." Available at: `https://www.tableau.com/solutions/topic/business-dashboards/`. Accessed on 2021-03-25.

[16] "Zoho reviews — technologyadvice." Available at: http://technologyadvice.com/products/zoho-reviews/. Accessed on 2021-03-25.

[17] "Google analytics." Available at: https://analytics.google.com/analytics/web/. Accessed on 2021-04-12.

[18] C. North and B. Shneiderman, "Snap-together visualization: A user interface for coordinating visualizations via relational schemata (2000)," tech. rep., HCIL, Inst.for Systems Research, Dpt. of Computer Science, University of Maryland, 2005.

[19] "Sas visual analytics — sas." Available at: https://www.sas.com/en_us/software/visual-analytics.html/. Accessed on 2021-03-25.

[20] "Microsoft power bi." Available at: https://powerbi.microsoft.com/. Accessed on 2021-04-16.

[21] "Sisense reviews — technologyadvice." Available at: http://technologyadvice.com/products/sisense-review/. Accessed on 2021-03-25.

[22] F. Anderson, "Getting started tutorial for ibm watson analytics," *IBM Corp, New York*, 2012.

[23] "Qlik view." Available at: https://www.qlik.com/us/. Accessed on 2021-04-16.

[24] "Scimago graphica." Available at: http://graphica.app/. Accessed on 2021-09-16.

[25] "Chartblocks." Available at: http://www.chartblocks.com/en/. Accessed on 2021-04-19.

[26] "Infogram." Available at: https://infogr.am/. Accessed on 2021-04-19.

[27] C. North and B. Shneiderman, "Snap-together visualization: Coordinating multiple views to explore information," tech. rep., HCIL, Inst.for Systems Research, Dpt. of Computer Science, University of Maryland, 1999.

[28] C. North and B. Shneiderman, "Snap-together visualization: can users construct and operate coordinated visualizations?," *International Journal of Human-Computer Studies*, vol. 53, no. 5, pp. 715–739, 2000.

[29] C. Ahlberg, "Spotfire: an information exploration environment," *ACM SIGMOD Record*, vol. 25, no. 4, pp. 25–29, 1996.

[30] K. Wongsuphasawat, Z. Qu, D. Moritz, R. Chang, F. Ouk, A. Anand, J. Mackinlay, B. Howe, and J. Heer, "Voyager 2: Augmenting visual analysis with partial view specifications," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2648–2659, 2017.

[31] R. High, "The era of cognitive systems: An inside look at ibm watson and how it works," *IBM Corporation, Redbooks*, vol. 1, p. 16, 2012.

[32] R. L. Sallam, J. Tapadinhas, J. Parenteau, D. Yuen, and B. Hostmann, "Magic quadrant for business intelligence and analytics platforms," *Gartner RAS core research notes. Gartner, Stamford, CT*, 2014.

[33] Z. Cui, S. K. Badam, M. A. Yalçin, and N. Elmqvist, "Datasite: Proactive visual data exploration with computation of insight-based recommendations," *Information Visualization*, vol. 18, no. 2, pp. 251–267, 2019.

[34] T. Kraska, "Northstar: An interactive data science system," *Proceedings of the VLDB Endowment*, vol. 11, no. 12, pp. 2150–2164, 2018.

[35] Ç. Demiralp, P. J. Haas, S. Parthasarathy, and T. Pedapati, "Foresight: Rapid data exploration through guideposts," *arXiv preprint arXiv:1709.10513*, 2017.

[36] M. A. Yalçın, N. Elmqvist, and B. B. Bederson, "Keshif: Rapid and expressive tabular data exploration for novices," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 8, pp. 2339–2352, 2017.

[37] A. Srinivasan, S. M. Drucker, A. Endert, and J. Stasko, "Augmenting visualizations with interactive data facts to facilitate interpretation and communication," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 672–681, 2018.

[38] M. El-Hindi, Z. Zhao, C. Binnig, and T. Kraska, "Vistrees: fast indexes for interactive data exploration," in *Proceedings of the Workshop on Human-In-the-Loop Data Analytics*, pp. 1–6, 2016.

[39] B. Yu and C. T. Silva, "Visflow-web-based visualization framework for tabular data with a subset flow model," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 251–260, 2016.

[40] P. Koytek, C. Perin, J. Vermeulen, E. André, and S. Carpendale, "Mybrush: Brushing and linking with personal agency," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 605–615, 2017.

[41] E. Zgraggen, R. Zeleznik, and S. M. Drucker, "Panoramicdata: Data analysis through pen & touch," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2112–2121, 2014.

[42] D. Ren, T. Höllerer, and X. Yuan, "ivisdesigner: Expressive interactive design of information visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2092–2101, 2014.

[43] W. S. Cleveland and R. McGill, "The many faces of a scatterplot," *Journal of the American Statistical Association*, vol. 79, no. 388, pp. 807–822, 1984.

[44] R. A. Becker and W. S. Cleveland, "Brushing scatterplots," *Technometrics*, vol. 29, no. 2, pp. 127–142, 1987.

[45] D. I. Benn and C. K. Ballantyne, "The description and representation of particle shape," *Earth Surface Processes and Landforms*, vol. 18, no. 7, pp. 665–672, 1993.

[46] G. B. Sidder, "Petro. calc. plot, microsoft excel macros to aid petrologic interpretation," *Computers & Geosciences*, vol. 20, no. 6, pp. 1041–1061, 1994.

[47] M. L. Ganuza, G. Ferracutti, F. Gargiulo, S. Castro, E. A. Bjerg, E. Gröller, and K. Matković, "Interactive visual categorization of spinel-group minerals," in *Proceedings of the 33rd Spring Conference on Computer Graphics*, pp. 1–11, 2017.

[48] D. Gozdowski, D. Sas, J. Rozbicki, W. Madry, J. Golba, M. Piechocinski, L. Kurzynska, M. Studnicki, and A. Derejko, "Visualizing diversity of yield determination by its components for winter wheat cultivar with ternary plot," in *Colloquium Biometricum*, vol. 41, -, 2011.

[49] M. Kozak *et al.*, "Visualizing adaptation of genotypes with a ternary plot," *Chil. J. Agric. Res*, vol. 70, no. 4, pp. 596–603, 2010.

[50] J. W. Tukey *et al.*, *Exploratory data analysis*, vol. 2. Reading, Mass., 1977.

[51] A. Inselberg and B. Dimsdale, "Parallel coordinates: a tool for visualizing multi-dimensional geometry," in *Proceedings of the First IEEE Conference on Visualization: Visualization90*, pp. 361–378, IEEE, 1990.

[52] A. Inselberg, "A survey of parallel coordinates," in *Mathematical Visualization*, pp. 167–179, Springer, 1998.

[53] B. Kovalerchuk, "Visualization of multidimensional data with collocated paired coordinates and general line coordinates," in *Visualization and Data Analysis 2014*, vol. 9017, p. 90170I, International Society for Optics and Photonics, 2014.

[54] B. Kovalerchuk, *Visual knowledge discovery and machine learning*, vol. 144. Springer, 2018.

[55] B. Kovalerchuk and V. Grishin, "Adjustable general line coordinates for visual knowledge discovery in nd data," *Information Visualization*, vol. 18, no. 1, pp. 3–32, 2019.

Citation: A. S. Antonini, M. L. Ganuza, S. M. Castro. *VISUEL - A Web Dynamic Dashboard for Data Visualization*. Journal of Computer Science & Technology, vol. 22, no. 1, pp. 42–57, 2022.
DOI: 10.24215/16666038.22.e03
Received: November 9, 2022 **Accepted:** January 24, 2022.