Wine Consumption Production

Yngve Skorge
University of Vienna

Nikola Golubovic
University of Vienna

Viktoria Lazarova
University of Vienna

ABSTRACT

This paper will concentrate on both, the wine consumption and production in the world and the distribution of different types of grapes in state of California. By using the tableau application our team provided, we managed to gain insights into where investments in the wine business might harvest the most success in the future. Furthermore we want to find out what type of grape was grown where in CA in relation to the acreage in each county. Our target was to create a useful tool to visualize large amount of wine data in a manner that the user still finds it easy to work with the provided interface.

Index Terms: Visualization—Consumption—Production—Visualization techniques Ratio—Visualization—Visualization design and evaluation methods—Acreage of Grapes

1 MOTIVATION

For visualizing large amounts of wine data in a useful way we first needed to find a database that allowed us to work with consistent data. The University of Vienna supplied us with the website http://www.wineinstitute.org for our research, but we had to manually put together snippets of wine related data to build a functional database. After creating a consistent database we needed to decide which user we can and want to address to then design a visualization tool that fits that specific type of user. We decided to focus on a user group that is particularly interested in finding good investment opportunities in the wine producing business (internationally and in CA). By creating a software that would enable those user types to explore wine statistics in a way that gives them productive intuition where to invest in what type of wine, we would succeed our goals.

2 RELATED WORK

During the research for our project we found one visualization tool similar to our problem, that visualizes the world coffee consumption (Fig 1). We found the idea of combining the Coffee Consumption (kg per person per year) with the worlds map very interesting (Fig 0) and we also tried to implement this idea in our project[1].
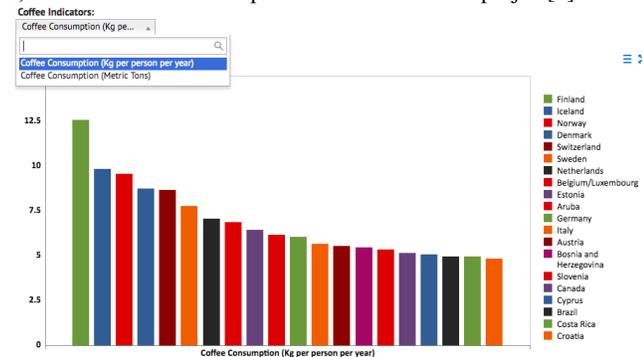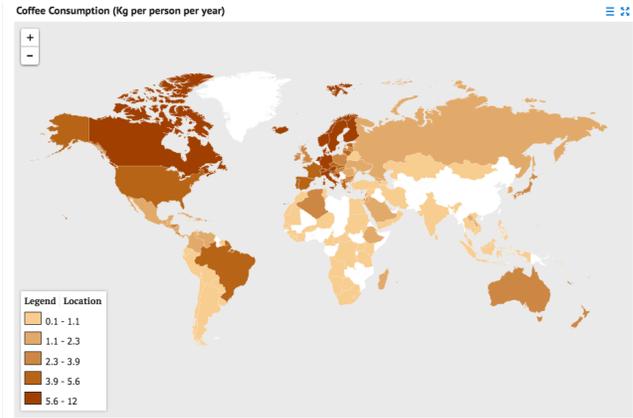

Word Coffee Consumption Bar Chart

Figure 1: Word Coffee Consumption Map

Another similar project we also found was from our colleges from previous years. Their tool was about Wine and Vine Statistics, implemented in d3. Their data source was completely different and we didnt find their implementation connected to our use case, also we didnt find it very interactive, which as a very important point for us.(Fig. 2)

We also wanted to compare our project with our first ideas that we had. We considered it could be interesting to show how the visualization has changes as we learned about Tableau and about whole process of showing the data visually. Our first Low-Fi Prototypes had two dashboards. One should represent production and consumption of wine data (figure 10), and second Types of Grape data (figure 11). On the first one, it would be three sheets, one bar graph, and two line charts representing Consumption and production over the years. We thought it would be nice to show consumption and production data on one dashboard, and data about types of grape on the second. No question that we were wrong, and that would not give results that we wanted. Only thing we left like planed was the map of California from second dashboard, third sheet. Thanks to feedback we received, we removed the bubble chart which stood for representing types of grape and which type was most made. We learned that quantity shouldn't be displayed with bubble chart, because users wouldn't recognize the difference between some values. Instead of that, we used bar chart.

3 APPROACH

- We decided to create two dashboards for our visualization story since our data didnt really allow us to combine them in one dashboard efficiently.

- Because of the feedback we received from our supervisors we narrowed down our use cases from three to one main use case - Barbara Queen (age 27), who is a potential wine producer in California and would also like to expand her wine sales internationally. Her main needs are to understand the market in CA, especially what grapes are most wanted, or produced and how many acres of which grape type are grown. Furthermore, she needs to know what countries produce the least and consume
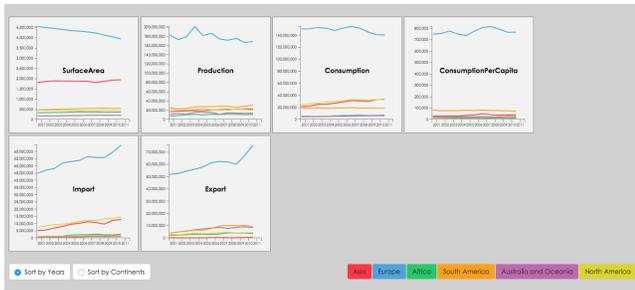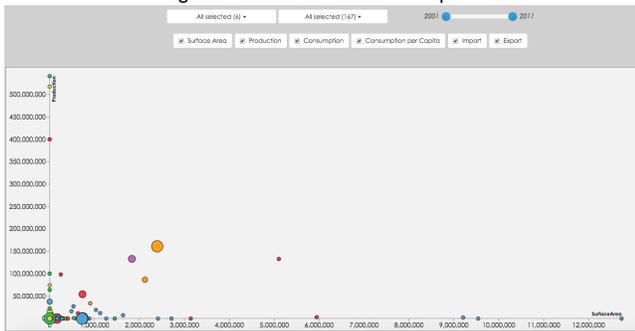
Figure 2: Word Wine Consumption

Figure 3: Word Wine Consumption Plot

the most wine so she can consider exporting her good to those countries.

- In order to realize her needs, we created two dashboards, one for the types of grape in California and one with world wide consumption and production.

- With the first dashboard (Figure 1), CA Type of Grapes, we wanted Barbara to have the opportunity to find what type is mostly produced and in which county and also how the acres of each type are distributed throughout each county in California. We created widgets for each type of grape, so when Barbara clicks on them (by choosing one or multiple), she can see in which county the specific type grows most successfully and how many acres of that type are used to produce wine. Vice versa she can also click on each county individually and see how many distinct types of grape are harvested there and the total acres for that county.

- We wanted to make the visualizations as interactive as possible so we implemented techniques such as:

  – widgets
  – filters
  – zoom
  – tool tips
  – aggregation

  for this view to optimize usability.

- The second dashboard shows the production and consumption of wine worldwide. With this visualization we wanted Barbara to see where the lowest production and the highest consumption occurs internationally. We created a consumption/production view so that she can select countries on the map and get a direct comparison between consumption and production for that specific country.
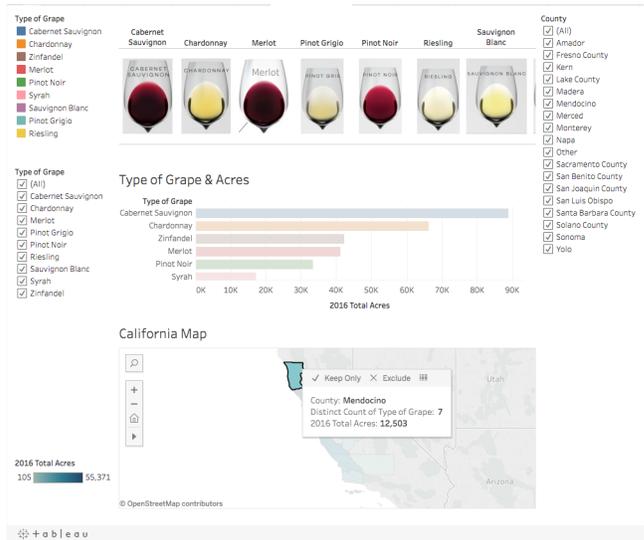

Figure 4: Types of Grape Dashboard

- To make it more interesting, we added several filters, so Barbara can specify the range she is interested in. Once she has chosen her range, she can see certain countries on the Consumption vs. Production chart and click on a country, so that it will be referenced on the map.

- When she hovers over that country on the map, she will see the most important data displayed via tooltip. Then the ratio between the consumption and the production(consumption/production) of the wine worldwide is calculated - if the ratio is high, this means, the production is low and consumption is high. We also made a general top 10 list of countries in which the ratio between production and consumption is the highest. Those would be the countries Barbara could invest in. Furthermore, we added hyperlinks for, her so she can compare data she finds online with our data.

4 IMPLEMENTATION

- By using the fully established platform Tableau we were given a huge range of diverse implementation tools. But as mentioned above many of our data sources were not directly relating to each other, so even with Tableau we needed to join different data sheets on at least one key attribute before moving on to the visualization. In our case we joined 2 sheets (Wine Production Consumption worldwide) on their almost identical attribute Country Name. As said, our projects consists of two separate dashboard, showing completely different areas on a very different ways, and in this section we would like to introduce you with the implementation of these.

- The first dashboard (see Figure 5) consists of three different sheets. We named it "Types of Grape", because its focus is exactly on different types of grape, where were they made and how much (in acres).

  – In first sheet we made the widgets, which are on the top of the dashboard. Here you can see the pictures of glasses with wine in a row, which represent so called "widgets". To make it happen, we had to import our images in as graphics(?). Then we assigned every picture to one type of grape, and use it as filter in our dashboard.

We edited the pictures in open-source software called Inkscape Vector Graphic Editor (OS: Ubuntu), where we were able to crop pictures, gave them certain resolution, so the every picture was the same size, and at the end we added the names of the corresponding type of grape on the top of the pictures.

– The second sheet shows us how many acres of specific type of grape are made. We used one bar graph - on X axis are values, on Y axis types of the grapes. If one click on one of the bars, not only this sheet will change, but the other sheets as well. As filters, Sheet 1 and this sheet work the same way, one can choose certain type of grape, but additionally sheet 2 shows how much acres were made. The sheet 2 is shown in the middle of the dashboard.

– On the last sheet (California Map), we showed the map of California, and put the longitudes and latitudes of every county where certain type of grape is made. The best way to show our data per county, in our opinion, were the Tool tip. By hovering over the county you can see how many acres were produced and number of types of grape, and of course, the name of the county.

• The second dashboard called World map dashboard (see figure 6) is based on our second part of a database. We imported 2 excel sheets, one for World Wine Production and second for World Wine Consumption. Only way to connect them was with one common column - Country name. Again, it consists of three different sheets - one map and two bar graphs. Only third sheet, Top Ratio, isn't interactive with two others, because we wanted that user has always an overview on the country where is best to invest. How and what exactly show us this dashboard you can read in following paragraphs, which detailed describe our sheets:

– First Sheet is a map. We put the coordinates in axis and got the map the way we wanted. It is connected with sheet 2, so it would show same countries which are displayed in sheet 2. At first, it will show only top 10 countries by Consumption. That can be changed by (re)setting the range of Consumption on the filter on the right side. Also, map can be zoomed and if one hover over the country with the mouse, he can see tooltip with most important information like name, consumption, production and ratio (see Figure 7). Of course, if you click on it, other sheets will be changed. This sheet stands at the top of the dashboard.

– Below left, we have our second sheet. It is a bar graph, which shows us the consumption and production in acres. On x-axis we have country name. For each country there are two bars, one for consumption and one for production. As said in description of first sheet, at the start, it shows only 10 top countries by consumption. By changing the range the the sheet changes with it. For better overview, we grouped bars in yellow and blue color. Yellow stands for production and blue consumption.
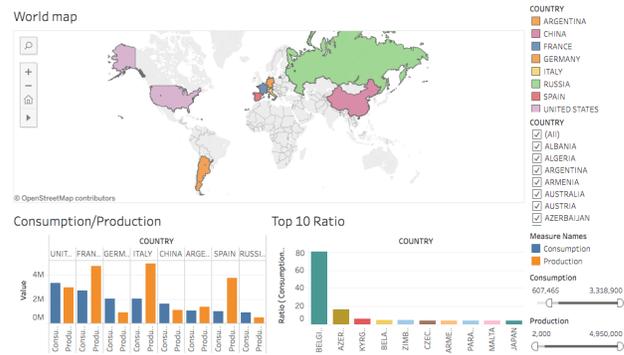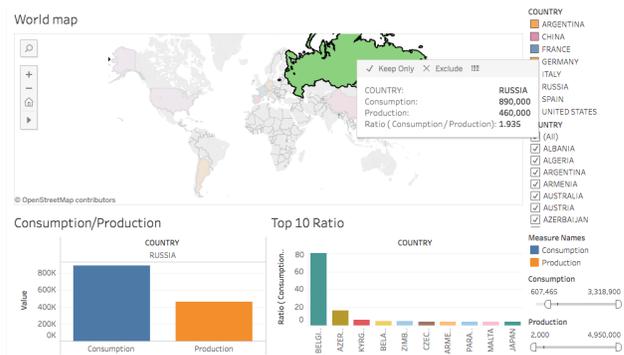

Figure 5: World map

Figure 6: World Wide Consumption and Production Map Russia selected

– Third Sheet is one down right. It shows us the top 10 countries on ratio scale. We wanted to put something on our own here, and show it on the dashboard. So we thought of new formula, which would help our visualization and showing our data on one interesting way. As we discussed the best countries to invest in are the ones that have high consumption and low production, we created bar graph which presents exactly that. As said in Approach, by dividing our number of consumption and production, we get the ratio and clear result where is best to invest.

5 RESULTS

The Tableau application we created enabled us to explore static data dynamically in an interactive environment. It allowed us to make discoveries that were hard to find in raw data otherwise. One of those discoveries was through our self created Top 10 Ratio Sheet in the Worldwide Consumption/Production dashboard. By dividing the consumption and the production attribute of each country with each other we created a dividend that gives us disclosure about where the ration between consumption and production is optimal(Figure 6).

In our dashboard we narrowed down the top 10 ratios for our users to have a direct overview of what countries have the best preconditions for wine investments. The winner was clearly Belgium with an astonishing result of 80.67 ratio, a country in which the consumption is most definitely much higher then the
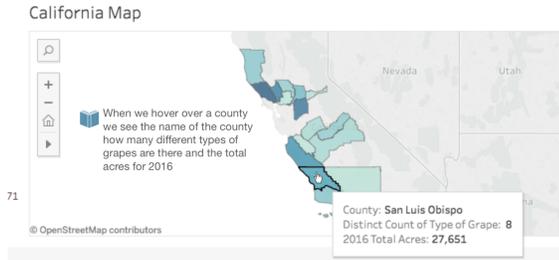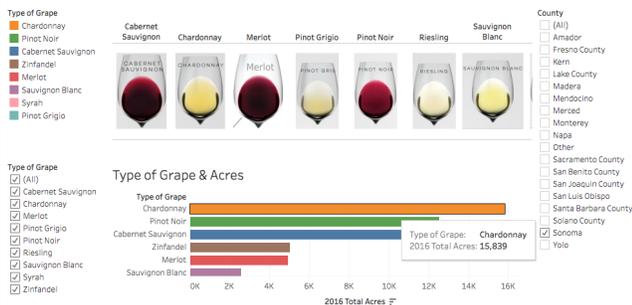
Figure 7: CA Maps with Tool Tip

Figure 8: Types Of Grape Somona

Figure 9: Production/Consumption Prototype

Figure 10: Types Of Grape Prototype

production (around 80 times). But since geographical location play a big role in starting business our application user might also not want to invest in one of the top 3 ratio countries. But adding filters for both consumption and production our user can check both factors individually for each country on the maps.In Figure 7 you can see that one can select a country such as Russia and comfortably see that this country has higher consumption then production.

With our Types of Grape dashboard there are unique results for each type of grape respectively. An example would be to select Type of Grape Cabernet Sauvignon and you will quickly experience that this grape has the highest growing rate with 89.027 acres in 2016. An example would be to select the Type of Grape Cabernet Sauvignon and you will quickly experience that this grape is most prevalent grape in CA.

But in reverse you can click on a selected county and you can find out useful information such as: in what county was the highest acreage in 2016, in this case it was Sonoma with 55,371 acres and 8 distinct types of grapes. Interesting here is, that in San Joaquin County there are 9 distinct types of grapes but the acreage is only 44,656 acres. In comparison to San Joaquin Count, Sonoma has less distinct types of grapes but a higher total acreage. When you look closer at Somona you will see that the grape Chardonnay, Pinot Noir and once again Cabernet Sauvignon is widely spread and has a high acreage(Figure 9). By using the tool tip (Fig.8) you can also access quick results: County, Distinct Count of Type of Grape and 2016 Total Acres.

6 DISCUSSION

One of the weaknesses in our approach was that because of the inconsistency of the data we had to implement two dashboards (Figure 5 Figure 6) for our implementation instead of one. In the beginning there were two excel sheets with data - one for worldwide wine consumption and one for worldwide wine consumption. In order to use this information properly in our visualization tool,as
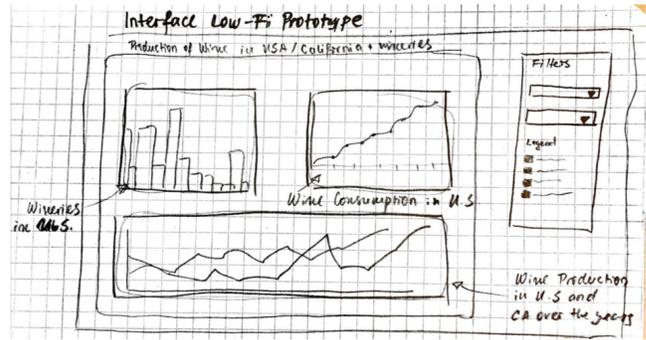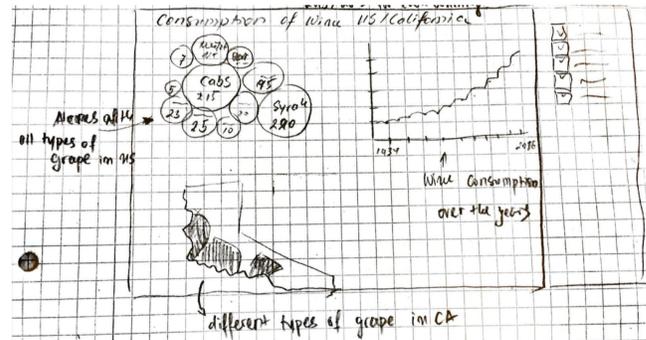
mentioned, we had to join the two tables sets. The problem there was to find a common attribute to join the two sheets, because in almost every column there was some data missing. At the end we managed to join the data by country name, which is not the best solution, but be could achieve what we wanted.

Another downside is that the data we use for our tool is static, which means that it cannot be updated. Then we found that sometimes we during making the Type of Grapes Dashboard, Tableau would not recognize some of the counties, although in our database they were written the same as in the Table Database the maps did not seem to pick it up. Eventually we had to manually insert the Latitude and Longitude for the corresponding county.

Regardless the problems we faced with the data sets we found on wineinstitute.org, we managed to complete our goal and create a tool that is interactive and easy to use. The widgets we implemented, representing each wine sort in U.S., make filtering data easier and more fun.

Another strength, separately from the project implementation, was the team work throughout our project. As a team we managed to separate the tasks equally, each and everyone of the team members were very collaborative.

Some of the lessons we learned during this project is that teamwork is very important. As a team we managed to separate the tasks equally and each and everyone of the team members was very compatible to works with, which helped us to get things easily done and without misunderstanding each other. We also figured out that whilst working in a team of 3 people, everybody has different ideas of the way they want their visualization should look like. That is why it is imperative to each create one's own prototype(e.g. Figure 8 9) and then to combine the best of all of them.
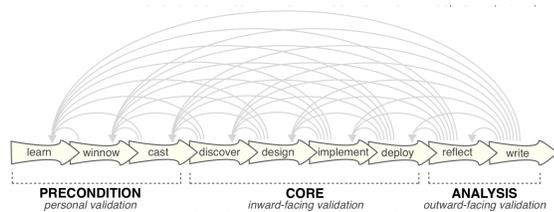
Figure 11: Nine-stage design study methodology framework classified into three top-level categories.

E.g in Figure 10 and 11 we created our first Prototype of how our dashboard should look like, when you look at the current version of both dashboards you can see how much has changed and been added. Also through reading Michael Sedlmair, Miriah Meyer, Tamara Munzner paper[3] we found it useful to use Nine-stage design study methodology framework which are classified into three top-level categories: Precondition, Core Analysis. This dynamic iterative process helped us go through the stages in Figure 12 over and over until we had a application that we were happy with.

One more thing we learned was, that there are many things people can do with the tool Tableau and also the struggles one can face while using it - like not being very compatible with macOS, which created some problems during the implementation. Also that once published on Tableau online the application slows down immensely.

7 A CLEAR SEPARATION OF TASKS BETWEEN THE GROUP MEMBERS FOR THIS MILESTONE

For this milestone we separated the sections of the paperwork so everyone of us can contribute in this task. Yngve Skorge wrote *Motivation* and *Approach*, parts of the *Implementation*, *Results* as well as finishing some details on the tool. Nikola Golubovic was responsible for the *Implementation* section, *Approach*, the Images through the paper and for updating the project page. Viktoriya Lazarova did the section *Related work* and *Discussion* work.

8 REFERENCES

[1] https://knoema.com/ajnnqlf/coffee-consumption-across-the-world?indicators=Coffee

[2] http://homepage.univie.ac.at/a0906960/plots.html

[3] Design study methodology: Reflections from the trenches and the stacks, Design study methodology: Reflections from the trenches and the stacks, Michael Sedlmair, Miriah Meyer, Tamara Munzner