From I Do to I Don't: What Tears Relationships Apart

Maria Campos
95629
maria.d.campos@tecnico.ulisboa.pt

Sofia Morgado
95675
sofia.morgado@tecnico.ulisboa.pt

ABSTRACT

The topic of divorce has always generated relevant discussions around the world, its legal grounds vary from country to country and divorce rates in Western countries are reported in news outlets to be at an all-time high.

It is also important to understand that divorce is a multifaceted issue with many implications and consequences that extend beyond the parties involved, such as children.

With this, there are interesting questions an individual can ask. Does the theory of compatibility between zodiac signs hold true against the divorce rates? Do bigger or lower age gaps in marriages result in longer lasting marriages or not? How does the change of season affect the number of divorces? Christmas time is known to be quite a chaotic time for families after all.

INTRODUCTION

When thinking about what we wanted to do for our project, the idea was to be able to visualize the theory of compatibility between zodiac signs came to mind since it is a topic that is met with a lot of skepticism by some, while others firmly believe its significance and influence. This led to the rising of other questions such as whether larger age gaps result in longer marriages or not, and if there is a particular season in which more divorces happen.

After contemplating and discussing which conclusions we would like our visualization to reach we proposed the following questions:

1. How has the divorce rate changed in the past 15 years in Xalapa? What is the distribution of divorce dates across months and seasons?
2. What is the age distribution for divorcing partners, for each gender in Xalapa?
3. Is there a relation between zodiac sign compatibility and divorce rates in Xalapa?
4. How does the divorce rate in Xalapa compare to divorce rates all over the world?
5. Does a lower or higher combination of income and education result in more divorces?

It is important to highlight that our dataset specifically pertains to data from the region of Xalapa in Mexico and our questions are all tailored to address the specific trends and tendencies specific to Xalapa.

RELATED WORK

During research, we found a notebook on Kaggle from a user that used the Xalapa dataset to assess compatibility between signs [7]. The results were displayed in bar charts. They present the combinations of zodiacs with the most and the least divorces but then they do not relate it with the compatibility values from the compatibility matrix they use. So, no conclusion is reached about higher or lower compatibilities being related with frequency of divorces.

Moreover, this project focuses on gaining insights directly from the data and not with providing an interactable visualization that lets the user perform exploratory analysis. And besides, it only deals with the zodiac compatibility and does not consider other factors and interesting questions regarding divorce, such as age gap, education level and income.

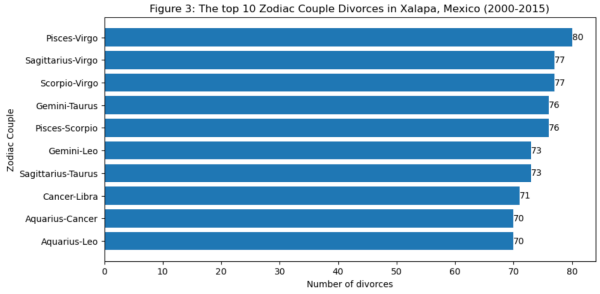

Figure 1. Results from the investigation on Zodiac Compatibility.

The inspiration for the Chord Diagram came from a visual representation for a Chinese Zodiac matchmaking guide which is available in Len De Groot's online portfolio [6] and can be seen in Figure 2. In this visualization, the user inputs their birthday and the compatibilities of their sign with others will be displayed, color coded by how good the compatibility is. Even though, this visualization provided the structure and even interaction of the Chord Diagram, it didn't quite delve into using real people to assess the theory of compatibility, which is what we wanted to do.

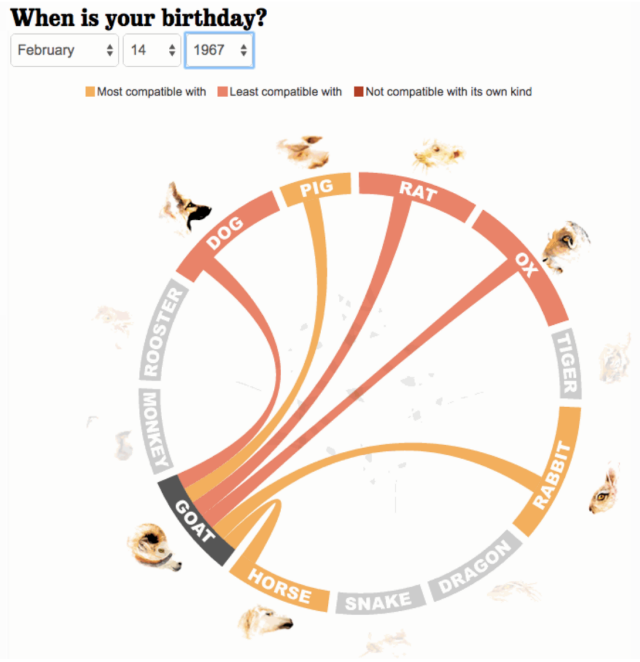

Figure 2. Year of the Goat: Chinese Zodiac Matchmaking Guide.

The idea to compare divorces in Xalapa to divorces worldwide and the Year Slider came from the visualizations available at [8]. As can be seen in Figure 3, the Year Slider is a double-range slider which can be interacted with freely to choose the desired year interval and the corresponding data will be displayed.

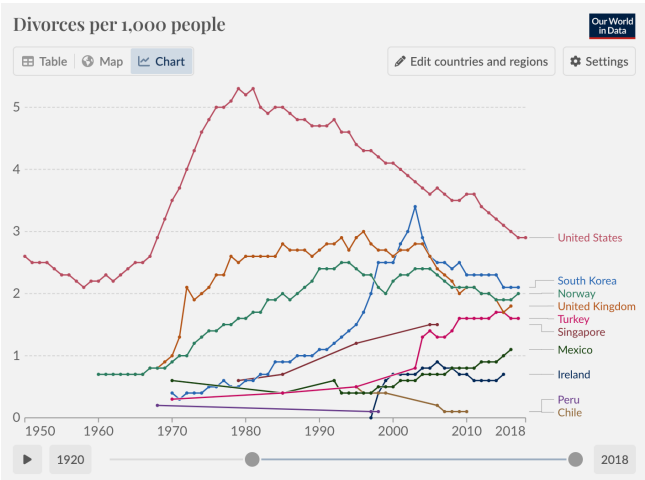

Figure 3. Divorces per 1,000 people Visualization.

THE DATA

To be able to display compatibilities between zodiac signs, we needed a dataset that contained the birth dates for each couple, so that we could derived the zodiac sign.

All datasets we used in this project can be found available online at Kaggle [1], except for the third dataset which can be found at [8].

The first dataset we found was named *"divorces_2000-2015_translated.csv"*, containing records of approximately 4,900 divorces over a fifteen-year period (2000-2015) in the city of Xalapa, Mexico. It contained information of each divorced couple like birth dates, marriage dates, ages, employment status, birth municipalities, etc. The positive aspect of this dataset was the inclusion of birth dates, which are usually not a disclosed due to privacy reasons.

The second dataset is called *"comp_matrix.csv"* and has records containing the compatibility rates for each combination of zodiac sign which were obtained from [4].

The third dataset is called *"data_world.csv"* and contains divorce rates from various countries in the world from 2000 to 2015. The dataset was slightly modified to include divorce rates from Xalapa as well, which were calculated using the first dataset and total population records from Xalapa population available at the National Institute of Statistics and Geography of Mexico [5]. We did think we would find data for the population of Xalapa for each year of 2000 to 2015, however we only found census data from a five-to-five-year interval. Since we could not find any other reliable sources for this data and we preferred working with official data, we compromised to use the same population for each year depending to which interval it belonged. The same calculation for divorce rates worldwide was used for calculating the Xalapa divorce rates. The formula is the following:

$$Divorce\ Rate = \left(\frac{Number\ of\ Divorces}{Total\ Population}\right) x\ 1000$$

To prepare the data for our visualization, we discarded, from the first dataset, missing information and information that was not relevant for our work such as birth municipality of the couple. To do this, we used both the Pandas Python Library [2] and Pentaho Data Integration [3] to process and clean the dataset. These tools also helped us to derive the age gaps of each couple, the zodiac signs, compatibilities, and the season in which divorce happened. We also translated from Spanish to English the columns containing levels of education and employment status. This process resulted in a refined dataset with 2069 records.

VISUALIZATION

Overall Description

Layout

As can be seen in Figure 4, our dashboard is divided in:

The filters for age range, education level and monthly income were placed on the right of the visualization.

The visualizations are the focus of the dashboard. We have a region for each one, them being the Cleveland Dot Plot, the Chord Diagram, the Bar Chart, and the Sunburst Chart.

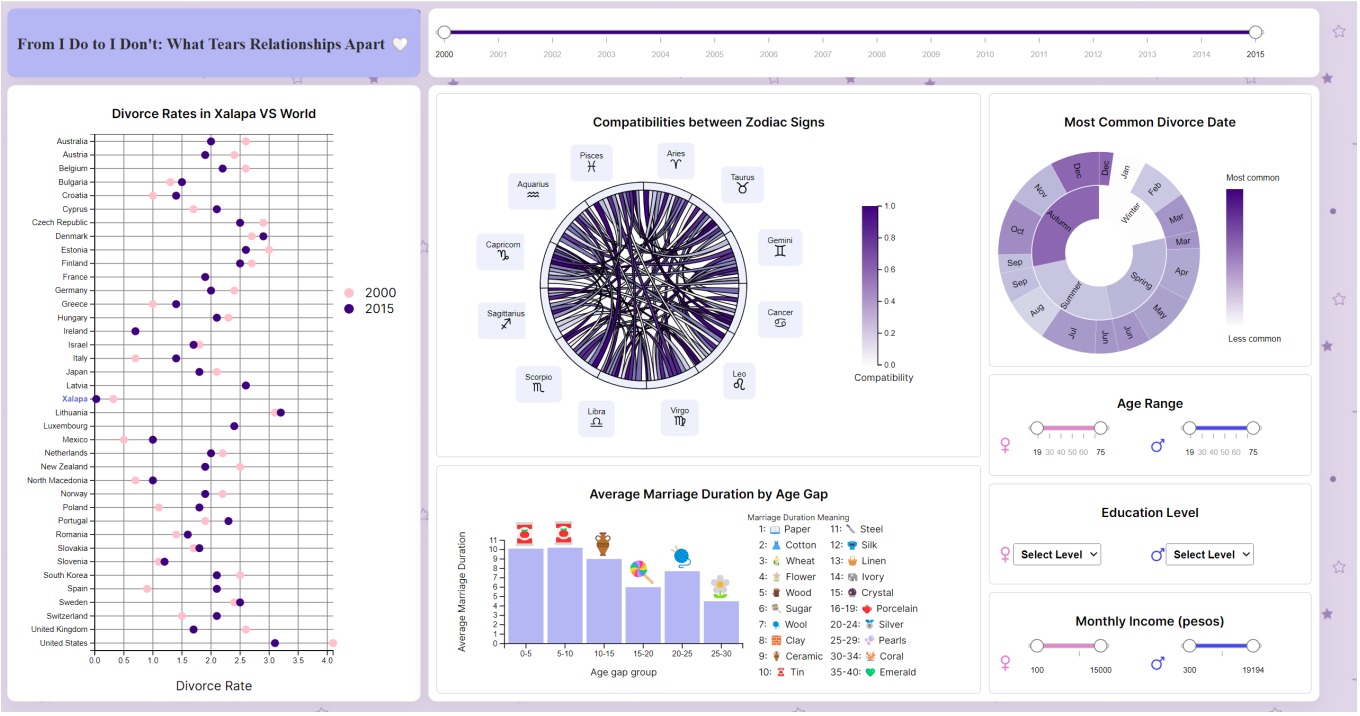

Figure 4. Dashboard Overview.

Age Range Slider

The Age Range Slider is a dual slider in which the user can filter the age for each partner.

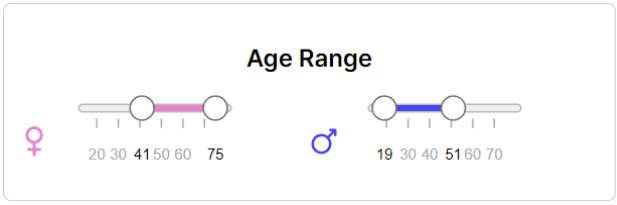

Figure 5. Age Range Slider, woman's age filtered between 41 and 75 years, and man's age filtered between 19 and 51 years old.

Level of Education Filter

The Level of Education Filter is a dropdown checkbox in which the user can filter the various levels of educations for each partner – Without Education, Elementary School, Middle School, High School, College and Other.

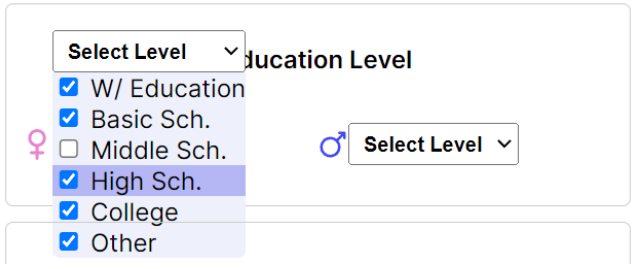

Figure 6. Education Level, with level of education for woman filtered to show results without Middle School level.

Monthly Income Slider

The Monthly Income slider is a dual-range slider in which the user can filter the various income intervals for each partner.

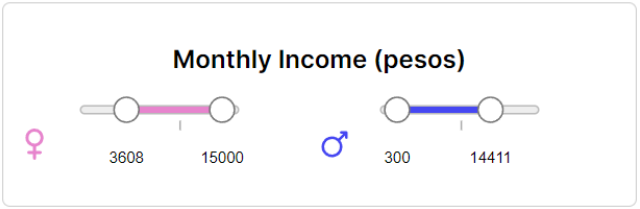

Figure 7. Monthly income (pesos), woman monthly income filtered between 3608 and 15000 and man filtered between 300 and 14411.

Year Slider

The dual-range slider allows the user to define a specific timeframe they wish to explore, providing higher granularity to examine changes over time. For instance, a user could set the interval to post-2008, during the economic crisis, to examine whether divorces increased or decreased.

Figure 8. Year Slider, filtered between 2004 and 2013.

Cleveland Dot Plot

The Cleveland Dot Plot represents the variations in divorce rates over the years in Xalapa and around the world. The user sees the divorce rates in the x-axis and the various countries and Xalapa in the y-axis. A color will be encoding

the minimum year (2000) and maximum year (2015). The user can change the minimum and maximum year with the year slider.

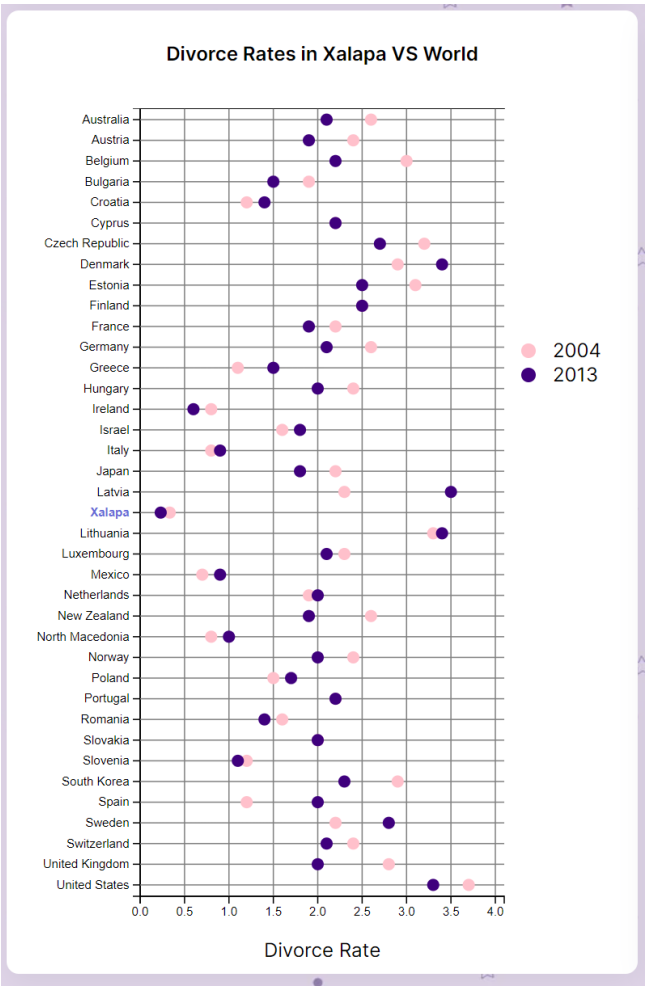

Figure 9. Cleveland Dot Plot, divorces rates filtered between 2004 and 2013 by the Year Slider.

Chord Diagram

The Chord Diagram represents the various degrees of compatibility between the zodiac signs, the color of each link represents how compatible the couple is, a gradient on the left of the diagram illustrates the varying colors and the values they encode (darker the purple means higher compatibility). Each zodiac has their corresponding symbol next to it to facilitate recognition. The opacity of each link encodes how many divorces in the dataset happened with those signs, a higher opacity means more divorces happened. The user can hover the mouse over the name of each zodiac sign and highlight the links to facilitate checking compatibilities for a specific sign, moreover the user can hover each link and, after highlighting that link, a tooltip will appear and specify which combination of signs that link is for, their compatibility value and how many divorces happened between them. The user can also filter divorces with that zodiac if the user clicks on the zodiac name button.

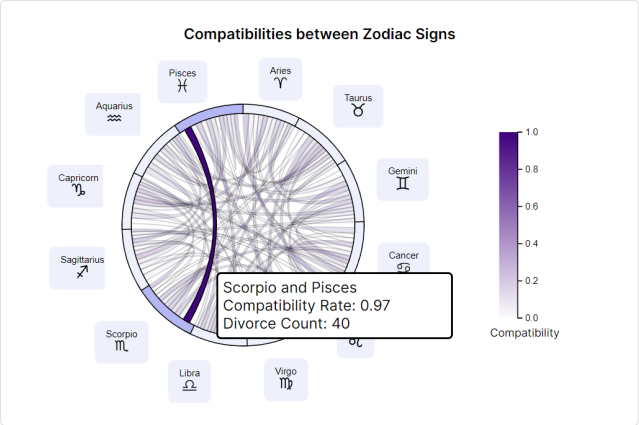

Figure 10. Chord Diagram with hover on Scorpio and Pisces link.

Bar Chart

The Bar Chart represents the various age gaps between couples and their respective marriage durations. The x-axis shows the age gaps and the y-axis the marriage duration between those couples. On top of each bar, there is an icon representing the gift the couple would receive had they stayed married. The user can hover the mouse over a bar representing each age gap range and a tooltip will appear showcasing the minimum marriage duration for that range and the average duration of marriages. If the user clicks, he can also filter the data for couples with that age gap.

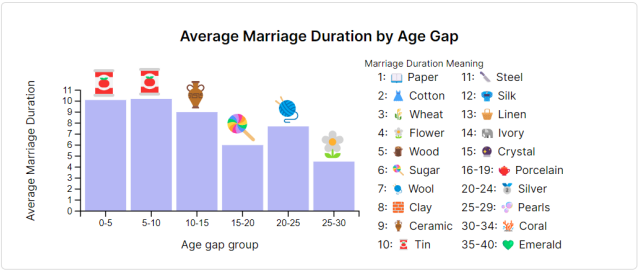

Figure 11. Bar Chart.

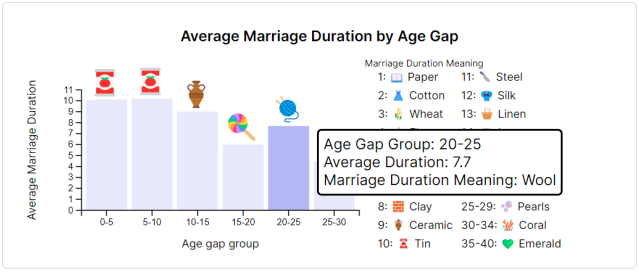

Figure 12. Bar Chart hovered on Age Gap Group 20-25.

Sunburst

The Sunburst represents the divorce frequency based on a hierarchy. Firstly, the year is divided in the four seasons and then divided into months. The user can click on the season and will be shown the divorce distribution for the months and days of that season. The user can click on the months and will be shown the divorce distribution during

the days of those months. All hierarchies are color coded based on a color scale that assigns darker shades of purple to higher divorce frequencies and lighter shades to lower divorce frequencies to facilitate pattern finding. The user can hover the mouse over each season, month or day and a tooltip will appear showcasing the divorce frequency for that component of the sunburst.

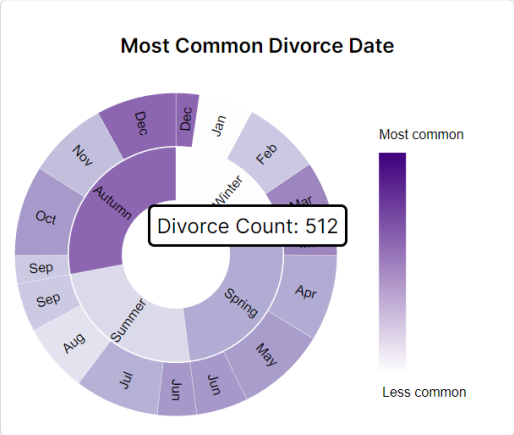

Figure 13. Sunburst hovered on Autumn.

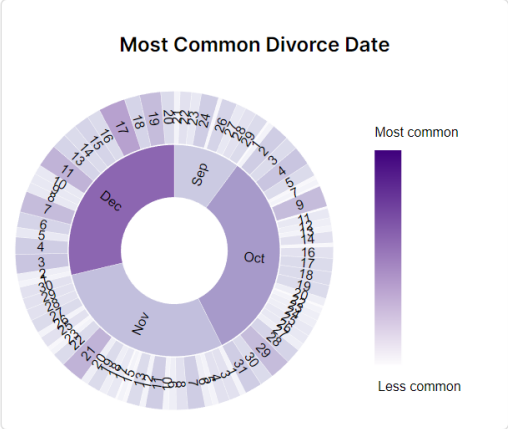

Figure 14. Sunburst, after clicking on Autumn and showing the months and days on this season.

Rationale

Our original sketch proposes a visualization mostly like the final product. However, there are notable changes. One of them is the replacement of the calendar heat map by the sunburst. The change was due to the low internal interaction the calendar heatmap provided. By changing to the sunburst, the user would be shown a general view first and the interaction would progressively show more specific data.

The decision to use the Chord Diagram to illustrate the compatibilities between zodiac signs came from the fact that this type of graph is excellent at depicting relationships between entities. There was also the fact that when researching zodiac sigs, we came across numerous representations where they are arranged in a circular layout as can be seen in figure 16 and we could take advantage of presenting the user with something he can recognize. The custom visualization is also depicted here, extending the Chord Diagram to encode the divorce frequencies using the thickness of the links. The higher the divorce count, the higher the thickness of the link. This idea ended up being discarded and another custom visualization was introduced, which is addressed in a later section.

The labels with each zodiac sign name became buttons since the drawing does not consider the 144 possible combinations of zodiacs. The buttons were needed to facilitate the viewer to analyze and filter each zodiac sign's data.

We opted for a purple color palette. This choice was motivated by the spiritual symbolism often associated with this color, relation to the introspective dimension of the zodiac signs.

To further aid in conclusion reaching, exploration and interactivity, the user has available the Year Filter that updates all idioms, affecting the visualization on a global scale. And then, the Age Range Filter, the Education Level Filter, and the Monthly Income Filter update all other idioms except the Cleveland Dot Plot, since that idiom is more focused worldwide while the others are more focused on the characteristics of the couples.

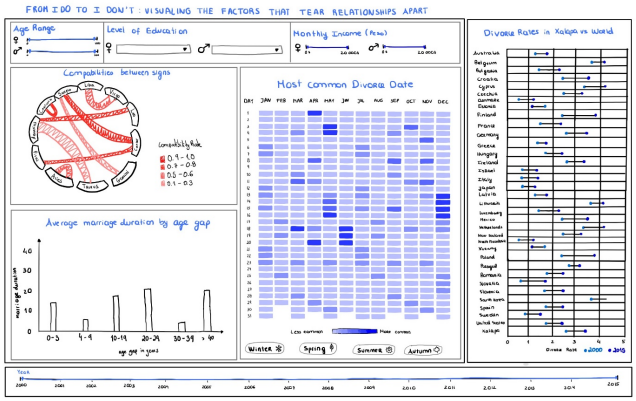

Figure 15. Original sketch of our visualization.

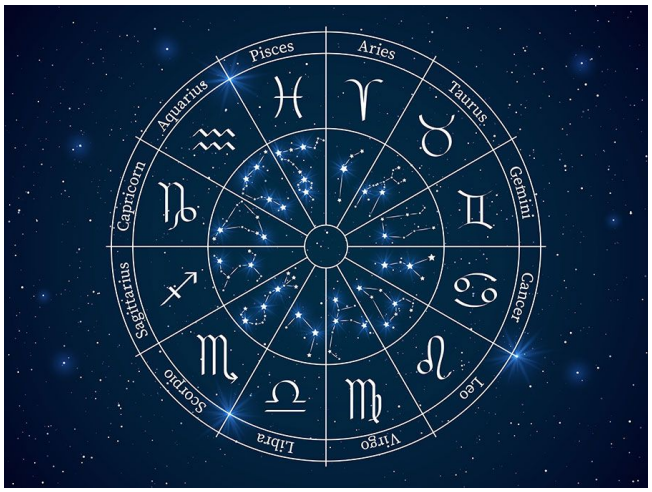

Figure 16. One of the many circular representations of zodiac signs.

Custom Visualization

For our custom visualization, we tackled the Chord Diagram and the Bar Chart. On the Chord Diagram, we changed the thickness of the arc to vary according to the divorce count. On the Bar Chart, we added icons on top of each bar that, according to the marriage duration, would represent the gift that the couple would receive, had they stayed married. This adds more meaning to this visualization, and a more quick and fun way for the user to compare bars between age gap groups.

The Potential

For this subsection, we will be presenting how to answer some of our questions with our visualization.

1: How has the divorce rate changed in the past 15 years in Xalapa? What is the distribution of divorce dates across months and seasons?

We can answer this question by looking at the Sunburst chart and finding the higher concentration of divorces through the colors of the chart by season and month. We can see that most of divorces happens in Autumn, specifically in December, before Christmas, and surprisingly, October is the second most common month, followed by May. Winter is the least common season, and January is the least common month, followed by August.

2: What is the age distribution for divorcing partners, for each gender in Xalapa?

To answer this question, we must look at the bar chart and compare the marriage duration to the age when divorced by each gender.

3: Is there a relation between zodiac sign compatibility and divorce rates in Xalapa?

We can answer this question by looking at the chord graph and see if signs with higher compatibility (darker color) have thicker lines (divorce count) and if couples with lower compatibility (lighter color) have thinner lines. If we want to check the values themselves, we can hover the mouse in each arc.

4: How does the divorce rate in Xalapa compare to divorce rates all over the world?

We can answer this question by looking at the Cleveland Dot Plot, by looking at the Xalapa divorce rates to the other countries. We find that, contrary to popular belief, for most of the countries, the divorces rates have declined.

5: Does a lower or higher combination of income and education result in more divorces?

To answer this question, we can filter by level of education and level of income and see how the data changes in the different idioms.

For example, lets compare Couples with College education and higher income against couples with lower education than college and lower income for both partners.

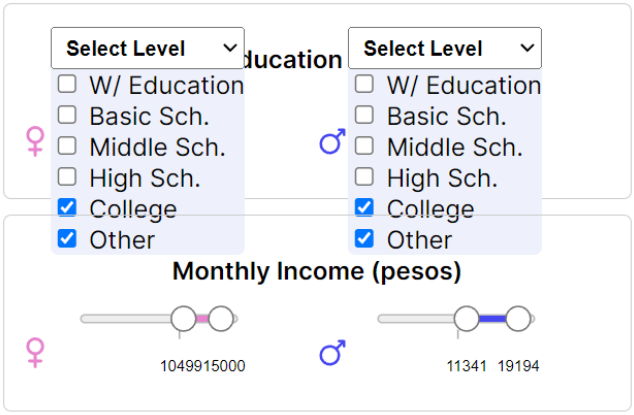

Figure 17. Filtered Education Levels and Monthly Income.

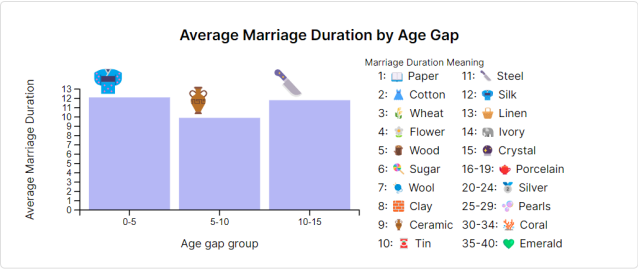

Figure 18. Resulting Bar Chart when applied the filters on figure 17.

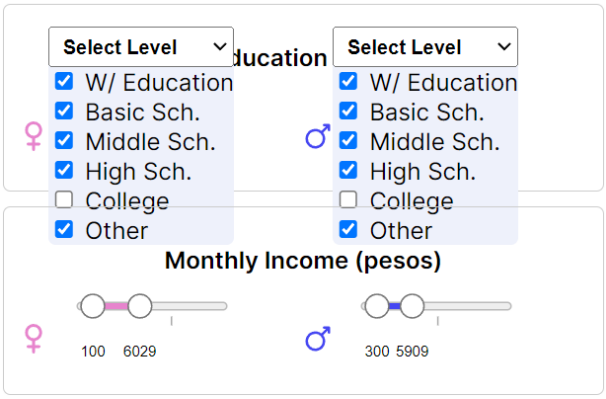

Figure 19. Filtered Education Levels and Monthly Income.

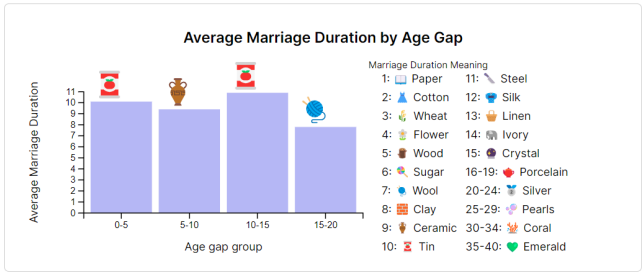

Figure 20. Resulting Bar Chart when applied the filters on figure 19.

As expected, the first group of couples tend to have higher marriage durations, but the second group of couples have openness for larger age gap between couples.

IMPLEMENTATION DETAILS

Implementation

One of the challenges we faced in our project was to transform the labels in the Chord Graph into interactive buttons. This allows for a more intuitive selection mechanism and better interactivity, while still maintaining the information over which zodiac sign the user is looking at. To achieve this effect, we had to add button behavior to the zodiac labels, manage mouse click and mouse hover interaction and update the visualization in a manner that filtered by the correct zodiac but also, if the user wished to go back, filtered back to the original graph.

Another challenge was encoding the thickness on the chord diagram based on the compatibilities, we had to track the compatibilities between zodiacs in a matrix and then calculate the maximum divorce count for each time the chord diagram was updated in order to divide the counts for that value in order to normalize the values, if this was not done the links would have inconsistent sizing due to the difference between the start angles and end angles of each link being different between them.

We created a state object to track the values for all the filters in our dashboard. This object gets updated whenever someone interacts with a filter in the visualization, triggering a function that re-filters the original data based on the current filter settings. The idiom charts then get updated with the new filtered data using D3's event handling.

To make the data more engaging, we annotated each marriage duration with a relevant emoji. The emojis corresponded to the traditional gift associated with each anniversary number. For example, the 1st anniversary (paper) was marked with a 📖 book emoji, the 2nd (cotton) with a 👗 dress emoji, and so on up to the 40th (emerald) marked with an 💚 green heart emoji. The emoji annotations were handled similarly to other chart elements, binding them to the data dynamically using D3 and styling them with CSS.

Problems

As can be seen in Figure 2, the idea was to have a calendar heat map to showcase the higher divorce frequencies by day, month, and season. However, this approach did not provide any internal interaction, therefore we switched to a sunburst graph that can be interacted with and show the user a general view first and upon interaction show progressively more specific data.

One of the challenges we faced had to do with the custom visualization. Originally, we had proposed a custom chord diagram which would encode the compatibilities between the signs and the frequency of divorces between that combination of signs in the thickness of the links in the chord graph. However, we had made a great oversight. The links in the chord diagram have angles, which humans are typically poor at perceiving. So, what happened was that links with equal thickness (meaning equal divorce frequency) but with different angles would not look similar. Even so, we kept this because we tested with some users who were able to perceive the difference but did not end up choosing this as our custom visualizations. Further testing with users would need to be done to assess the effectiveness of this way of encoding divorce count.

CONCLUSIONS AND FUTURE WORK

The development of this visualization has highlighted the importance of considering scalability problems when handling large amounts of data. This relates especially with an important problem with the visualization to tackle in the future. The initial view of the chord diagram can be overwhelming to users and be the source of some confusion. To improve the readability of the initial diagram, edge bundling should be performed on the arcs of the chord diagram.

We would have also liked to add a chart more complex than the bar chart to showcase the age gaps and marriage duration, like a radar chart. A radar chart could enhance the

visualization by allowing viewers to easily discern patterns and relations between these two variables.

An additional idea we would have liked to add was, when the user clicked on a button for a specific zodiac sign, the arcs of the chord diagram to morph into the arcs on the other clicked on sign.

Moreover, we must be aware that there is no universal way to calculate compatibility between zodiac signs. There are other formulas online that consider gender and even moon signs (based on the precise location of the moon in the sky by the time one was born) and ascendant signs (based on the zodiac sign that was rising in the eastern horizon when one was born) – the ones we used in the project were the widely known sun signs. It could be interesting to explore other calculations for compatibility.

It could also be interesting to explore the Chinese zodiac sign theory of compatibility. By, once again, taking advantage of the dataset having records of birthdates, we can use this system that assigns each year an animal sign and element. By cross-referencing these with the Western zodiacs, we can discover potential correlations with divorce rates and marriage dynamics.

Finally, we could also introduce some storytelling techniques such as introducing the city of Xalapa to the user or even introduce each zodiac as a character with a personality composed of the common traits of that zodiac so the user can relate to them and even get to know them better.

REFERENCES

1. Kaggle. 2023. (2023). https://kaggle.com
2. Pandas. (2023). Python Data Analysis Library. Retrieved from https://pandas.pydata.org/
3. Pentaho Corporation. (2023). Pentaho Data Integration. Retrieved from https://www.hitachivantara.com
4. Numerology Sign. (n.d). Zodiac Compatibility. Numerology Sign. https://numerologysign.com/zodiac-compatibility/
5. Instituto Nacional de Estadística y Geografía. (2023). Areas Geograficas. https://en.www.inegi.org.mx/
6. Len de Groot. (2015). Year of the Goat: Chinese Zodiac Matchmaking Guide. Len de Groot. http://lendegroot.com/
7. Investigation on Zodiac Compatibility. (n.d). Kaggle. https://www.kaggle.com/investigation-on-zodiac-compatibility
8. Marriages and Divorces Visualizations. (n.d) Our World In Data. https://ourworldindata.org/marriages-and-divorces