Instituto Superior Técnico

Information Visualization

Final Report

Authors:

André Filipe Cruz Fonseca – 84977

Gonçalo Filipe Lourenço Tomás – 76491

Samuel Simão Canada Gomes – 76415

1 INTRODUCTION

The objective of the visualization we created is to show an overview of the critics done to several computer games during the last two decades. To help this analysis, we considered multiple variables such as game genres, ESRB ratings, publishers and developers.

An analysis of videogame reviews is very important for both users and videogame companies in order to evaluate the progress of videogames as well as to help predict future trends. Moreover, being a relatively recent industry, there are yet very few analytical studies in videogames which makes this work an important contribute.

In our visualization, we analyzed data from a website named "metacritic" that gathers game scores from both users (user score) and critics ("metascore"). With such data, we supported the following tasks:

1. Compare user scores with "metascores" in several genres.
2. Relate maturity rating with user score/"metascore" and number of reviews during the last 20 years.
3. Relate the game genres with the maturity rating.
4. Find dependencies between "publishers" and "developers", and analyze features of each "publisher" and their "developers".
5. Compare developers with the respective scores of their games.

Each task was created with the objective to answer some related questions:

Task 1: Is there a relation between game genre and score? Are there relevant differences between user and critical score?

Task 2: Is there a relationship between "metascore", user score and maturity rating? Which maturity rating has a higher number of reviews? Will these values change over time?

Task 3: Is there a relation between game genre and maturity rating?

Task 4: Which publishers have a larger developer force? Which publishers have better developer consistency (distribution of average "metascore" for each developer)?

Task 5: Which developers have a better critical reception?

2 RELATED WORK

As previously discussed, videogame analytics are an understudied field and, as such, there are not many studies about it; at least publicly... Most major videogame companies have a marketing team that is usually responsible for collecting and analyzing such data. Their results, however, are not publicly announced.

One of the most important analytical report about videogames is the yearly report of "Essential facts about the

computer and video game industry" [1]. This document was created by the Entertainment Software Association (ESA) with the cooperation of multiple game publisher and its objective is to summarize multiple yearly statistics in a single document.

Another analysis we found was the recent CGD presentation on videogame data called "Awesome Video Game Data"[2]. This presentation is more focused in the distribution of the multiple player bases as well as videogame revenues.

Our work, however, is different from the examples presented in this section. Our focus is on critical reception of videogames and how they are distributed across the multiple genres, publishers and developers that comprise the videogame spectrum.

3 DATA

The dataset found is based on the website metacritic.com and is a compilation of videogame data including the "metascore": an average critical score obtained from multiple websites and magazines.

This dataset was found on a GitHub repository [3] created by Dmitry Velikiy, who extracted the data from the metacritic website and compiled it in a csv file.

In order to filter and alter the data, we first needed it to be parsed. However, we had some problems doing so. The culprit was the game "Steins; Gate". In its title, the character ";" is used. This is a reserved character in csv to separate columns. This entry was manually deleted to avoid these parsing problems.

In its original state, the data included the following attributes for each game: title, metascore, critics_reviews_count, user_score, user_reviews_count, maturity_rating (ESRB rating), publisher, developer, genre, genre_tags, release_date, platform and link (website link).

After analysing the importance of each attribute to answer our questions, we decided to remove some of them such as platform (even though studying platform differences is interesting, unfortunately, the data found is exclusively of PC games which makes it redundant to the analysis), genre_tags (which had multiple entry for each game and also had a big variety between them making them difficult to analyse) and links (which added no information to the dataset).

After removing these attributes, we processed multiple NA[4] entries in the dataset. These entries were mostly found in the maturity_rating and user_score/user_reviews_count categories.

And last but not least, we found that a new, unexpected KA[5] maturity_rating appeared in the list of ratings. After some search we found out that this was an old rating category that was later replaced by the now in vigour E[6] rating. Therefore, we changed all entries with a KA rating to an E rating. Also some games in the dataset hadn't yet been rated and these entries were also removed.

Having filtered the data, we then divided the data in 5 parts (one for each task) and further manipulated it for the purpose of each task, because we didn´t need all attributes for all tasks. In addition to this, for each task we calculated some derived

---

1 http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf

2 http://www.gdcvault.com/play/1021828/Awesome-Video-Game-Data

3 https://github.com/codenotfound/metacritic-analysis

4 NA – Not attributed

5 KA – Kids to Adults

6 E - Everyone

measures to obtain better and interesting results.

Task 1: To compare user scores with "metascores" in several genres, we selected user_score, user_reviews_count, metascore, critics_reviews_count, genre attributes. Then we calculated the average userscore, average metascore and the total number of reviews (both user and critic) for each genre.

Task 2: In this task the attributes chosen were maturity_rating, critics_reviews_count, metascore, user_reviews_count, user_score and release_date (which was used to retrieve the years). It was also calculated the total number of reviews (user and critic) and average score (user and metascore) for each maturity rating, to answer all questions associated to the task and perhaps discover some trends along the analysed timespan.

Task 3: For this task we selected maturity_rating and genre attributes and we calculated the number of games for each pair maturity rating/genre to find relations between the pairs.

Task 4: In this task we selected the publisher, developer, user_score, metascore, critics_reviews_count and user_reviews_count. Then we split this task in two branches. In the first branch it was calculated the number of distinct developers for each publisher. In the second branch we calculated the average metascore and standard deviation of metascore for each publisher and developer and then with those values we calculated the developer consistency (average of metascore of the developer) and standard deviation (calculated with the standard deviation of metascore). Then the two branches were merged and we chose to only choose to include publishers with developer force bigger than two.

Task 5: Was selected the publisher, developer and metascore attributes and calculated the average metascore for each developer in each publisher to discover which developers have a better critical reception.

4 VISUALIZATION

4.1 OVERALL DESCRIPTION

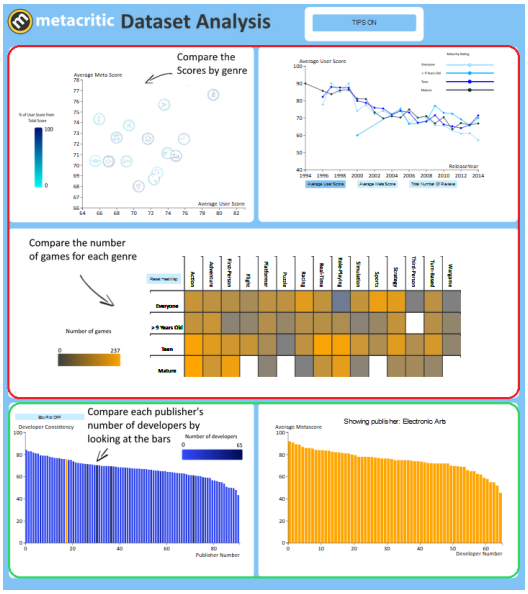

Image 1 – Visualization Layout divided into part 1 (red) and part 2 (green) and all 5 idioms

Our visualization is mainly divided in two parts:

The first part concerns game genres and maturity ratings and is composed by three graphs. One for each question. The first graph (scatterplot) is related with game genres and the second graph (line chart) with maturity rating. Clicking elements in both highlights them in the third graph (heat map).

The second part concerns publishers and developers. The first bar chart sorts all publishers by the number of developers they have and by selecting a publisher the average "metascore" of every developer for the selected publisher is displayed in the second bar chart.

The layout also includes hints to help understand how each visualization works and how they interact. They can be toggled on and off in the top right of the layout.

As for the individual idioms:

Idiom 1 – Scatterplot: The relationship between the game genre (represented by a glyph) and both critical (x axis) and user (y axis) score can be analysed. Also the number of games in each genre is represented by the colour of the glyph. A scale to the left allows the user to compare the relative colour values. Highlighting a glyph shows the genre name and point position.

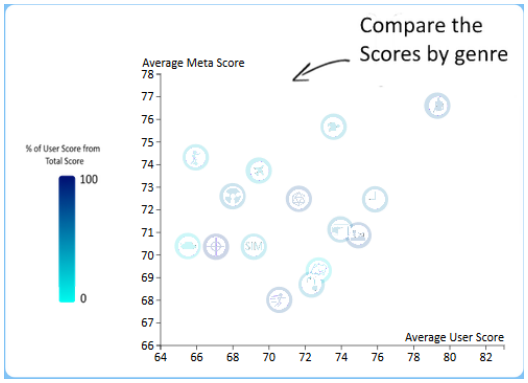

Image 2 – Idiom 1: Genre Scatterplot

Idiom 2 – Line Chart: The evolution of user reviews, critical reviews and number of reviews (y axis) can be observed through the last 20 years (x axis). A group of buttons placed below the idiom can be used to switch between the 3 possible attributes. Each maturity rating is represented by a line and more detailed information about each maturity rating can be obtained by hovering the corresponding line.

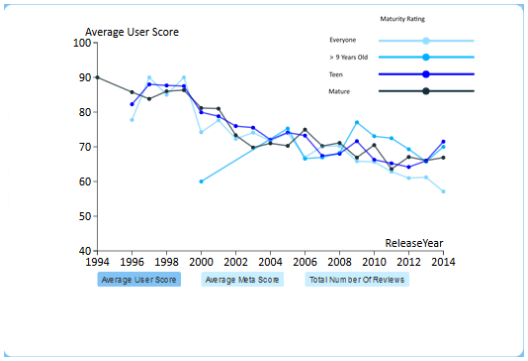

Image 3 – Idiom 2: Maturity Rating Line chart

Idiom 3 – Heat Map: The number of games for each pair genre/maturity rating is displayed in each cell by a different color. A scale to the left allows the user to compare the relative colour values. Absolute numerical values are shown by hovering over a cell and can be locked into showing the value by pressing it. Entire columns and rows can also be locked by interacting with idioms 1 and 2. To reset the locked cells, the button “Reset heat map” must be pressed.

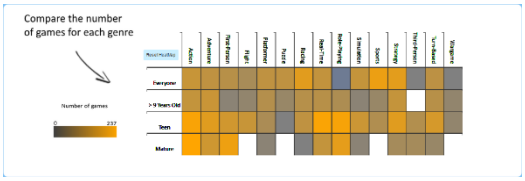

Image 4 – Idiom 3: Genre and Maturity Rating Heat Map

Idiom 4 – Bar Chart and Box Plot: In this idiom, every bar represents a publisher and its length (the y axis) is associated with its developer consistency (average score of all developers that work for that publisher). Also the developer force (number of developers) of each publisher is represented by the colour of the bar. A scale at the top allows the user to compare the relative colour values. Instead of the medium values, a box plot can be used by pressing “TODO”. Hovering a bar (or box plot) shows the name and the numerical value of developer force for the corresponding publisher. Clicking the bar changes Idiom 4’s data to show the developers of the chosen publisher.

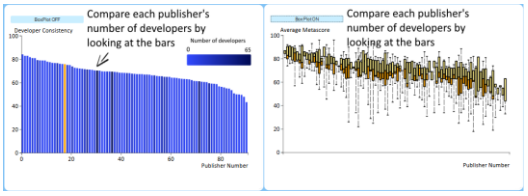

Image 5 – Idiom 4: Publisher’s Bar Chart and Box Plot

Idiom 5 – Bar Chart: In this idiom, every bar represents a developer and its

length (the y axis) is associated with the average metascore of that developer's games. Clicking over a bar shows the name and the numerical value of average metascore for the corresponding developer.

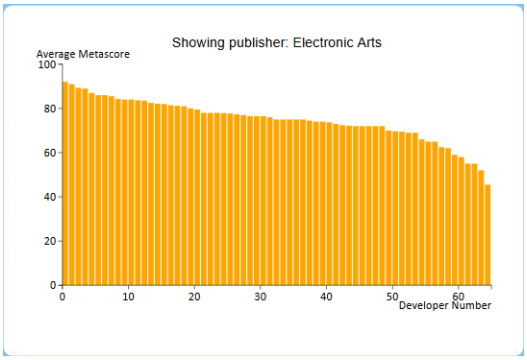

Image 6 – Idiom 5: Developer's Bar Chart

4.2 RATIONALE

Idiom 1 – We selected a scatterplot, because it's the best idiom to compare two quantitative attributes (the score). Also, it allows us to represent the percentage of user reviews when compared to the total number of reviews, by varying between two color hues (divergent order). This idiom was not changed since its initial inception.

Idiom 2 – We selected a line chart, because it's one of the best idioms to represent a time based attribute, mapped to discrete points (years), allowing us to observe trends resulting from the variation of the quantitative attribute average metascore score through time (for each maturity rating / number of reviews). During development, we added an option to show user reviews and, as such, the button used to change from score and number of reviews was replaced by three buttons, one for each attribute with the currently active button highlighted.

Idiom 3 - We used a heat map matrix, because it's a good way to represent two key attributes (genre and maturity rating) and provides a compact summary of the number of games attribute using a 2D matrix alignment. This idiom was not changed since its initial inception.

Idiom 4 – The chosen representation for this data was a bar chart because it is one of the best idioms to compare the categorical attribute publisher to the quantitative attribute developer consistency. We also planned on using the color of each bar to represent the developer force. Initially we thought in adding a filter that uses the attribute developer force to change the number of displayed publishers. However, since there are hundreds of publishers, each bar on the screen would be no more than one or two pixels wide making it impossible to analyze the data effectively. Instead, since there are many indie studios (auto financed developers) which are not the target case study for this idiom, we decided to remove all publishers with only one or two developers. This move effectively reduces the amount of publishers to approximately a third of the initial dataset.

Idiom 5 – This idiom was originally conceived as a grouped bar chart in which a group consisted of all the developers for a given publisher. Since there are more than 1000 developers, this approach would never work. There are too many developers to use a bar to represent each one. Instead, we only show the developers for a given publisher at a time as selected in idiom 4. We added at the last minute a box plot in idiom 4 just so we could answer more questions about the selected developers in this idiom 5.

Besides all these changes, we also added interaction methods to all idioms in order to expose more specific information about selected attributes.

4.3 DEMONSTRATE THE POTENTIAL

One of the question we asked was: "Which maturity rating has a higher number of reviews? Will these values change over time?" from task 2. To answer this

question the only necessary step is to press the “Total number of reviews” button.

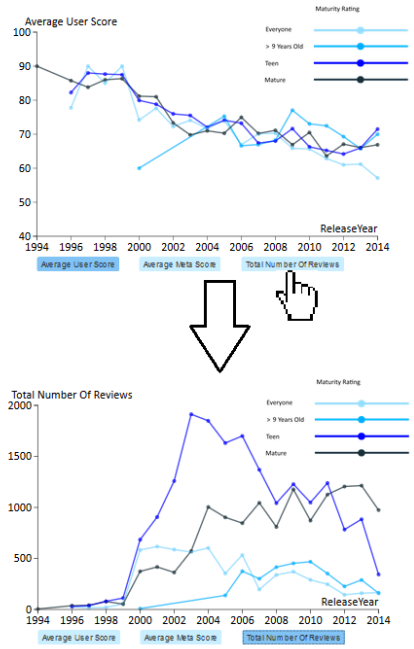

Image 7 – Answering the questions from task 2: “Which maturity rating has a higher number of reviews? Will these values change over time?”

As shown in the graph, this question can be easily answered. In fact, the maturity ratings with the higher number of games is (at least until 2010) the “Teens” category. Moreover, it shows us an interesting pattern. While games targeted for children have a large presence during the first years of analysis, their numbers drop during the latter years. Mature games do quite the opposite and actually have the highest number of reviews from 2010 forward. Also, games for teenagers rise in popularity until around the middle of the analyses time only to decrease again. The conclusion is obvious. Our data shows a relevant increase in the average age of gamers. Now most gamers are adults who prefer the more complex and shocking experiences of mature games.

Also in task 3 we asked the question “Is there a relation between game genre and maturity rating?” Answering this question is even easier. Just by looking at the data, it is possible to conclude that a big percentage of games are action games and that they are more prevalent in both “Mature” and “Teen” ratings. Also “Teen” ratings have a big number of real-time and role-playing games. With a lot less games in total, the “Everyone” ratings have mostly racing, sports and strategy games.

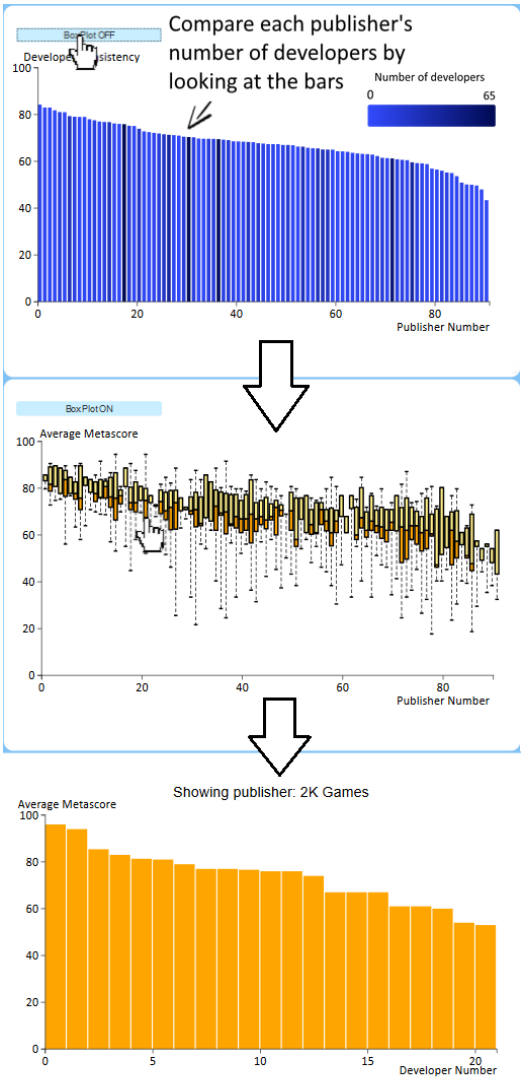

Image 8– Answering the questions from task 5: “Which developers have a better critical reception?”

And lastly, in the fifth task, we asked the question “Which developers have a better critical reception?”. To answer this question we must first change idiom 4 to box plot mode. After that, we select the developers with the highest top whiskers

and then look at the first developer on idiom 5. With some comparison we can conclude interesting facts. One of which was a developer like “Bioware” who both worked for “LucasArts” and then “Electronic Arts” showing a relative equal product quality (93 to 89). As such, at least in this case, the change of publisher didn’t affect the product quality.

5 IMPLEMENTATION DETAILS

We divided the visualization in two parts: one in which we coded the idioms for the tasks 1, 2 and 3 and the other in which we coded the idioms for the tasks 4 and 5.

In the first part, we used 3 “svg” elements which were connected to 3 elements in html file.

We used an external “css” to style the axes and for the buttons we used the library bootstrap.

For the animations we used the command:

“(…).transition().duration(x)”.

5.1 TASK 1

For the Scatterplot we appended one image and circle for each point of the dataset and we superimposed them to create the composed glyphs. We appended additional text elements to the pop-ups. The file path for every image is the same except for the final few characters. These were saved in an array to be accessed in the creation of the “svg” elements of the type “svg:img”.

On the ”mouseover” function of the image elements, we increased the opacity attribute of the element in focus to 1, and reset it to 0.2 in the “mouseout“ function.

5.1 TASK 2

For the “linechart”, we appended elements of the type line and circle, and additional elements of the type text to the pop-ups.

On the “mouseover” we reduced the opacity attribute of the elements not in focus, and reset them to 1 in the “mouseout” function (*lineSelection(i)*).

When pressing the buttons we changed the circle and line positions and redrew the axes to accommodate our several aspects of analysis.

5.1 TASK 3

In the heat map we appended one “rect” element per point in the dataset and adapted the fill attribute of that “rect” to match a colour scale for the number of games.

For the highlighting, we reduced the opacity of the “rects” selected, either by mouse over or by selecting an element of the upper charts (linked highlighting).

For both the scatterplot and the line chart, by clicking the different elements, we implemented linked highlighting. To do this we accessing the “rects” of the heat map and performing a change in their opacity (function *alterGenreInHeatmap(i)* and *alterRatingInHeatmap(i)*).

5.1 TASK 4 & 5

In the bar charts we follow the same scheme of the first labs, but we use two “svg” elements to accommodate each idiom.

In order to calculate the dataset corresponding to the developers of each publisher, we filtered all the developers to find the ones associated with a particular publisher (function *calcCluster()*).

We altered the colour of each “rect” element according to a colour scale defined by the number of the developers of each publisher.

On the “mouseover” function of the “rect” elements of the charts we set the stroke attribute (colour of the outline of the “rect”) to change, and switched the state of the text in the bottom to be visible with the corresponding text.

On the “mouseout” function of the “rect” elements in the bar charts, we reconfigured their attributes to reset.

In the boxplot, we used another svg (svg_box_plot) which wrapped the svg elements necessary to draw the chart.

Two scales were created for the correct positioning of each chart and its elements.

Those elements were built for each entry in the dataset and included one dotted line (by changing the attribute “stroke-dasharray” of the element) element for the main bars, two line elements for the whiskers and two “rect” elements for the center box.

A horizontal and vertical axis were included just like in the bar charts.

6 CONCLUSION AND FUTURE WORK

Our work in this visualization was an enriching one in terms of skill development. A good visualization must try to show the biggest amount of information to the user with the least cognitive load possible. In order to achieve the best compromised solutions a good knowledge of the human brain and perceptions is required. As such, during this work a lot was learned about the multiple human properties human beings are sensible to: color, rotations, shapes, dimensional properties such as length, areas, volumes, depth, etc. We also learned to use new tools. The first was “Pentaho Data Integration” which we used to filter, sort, derive and convert the original data. We also learned how to use “d3.js”, a JavaScript library for producing dynamic, interactive data visualizations in web browsers.

If we had more time to work on this project, we would probably spend more time planning the sketches before implementing the idioms. The most time consuming part of the project was the implementation of the idioms in D3. As such, when we found out that some implementations had to be changed that costed us a lot more time than it would have if we noticed the problems in the prototyping phase.

If we had 3000 dollars and 1 more month to work on this project, we would use the money to finance our own statistical data and gather information not only from “metacritic” but also multiple websites and magazines. With this new data we could also expand the analysis for not only the PC platform but also both home and portable consoles, even mobile market if possible. We could then observe the platform contrasts in terms of maturity ratings, genres and, of course, review scores.