

TAG-ORIENTED MARKET ANALYSIS OF ONLINE MARKETPLACES USING VISUAL ANALYTICS

A visual analytics system for market analysis of the desktop
game online marketplace “Steam”

Valerio Di Stefano - 1898728

“Engineering in Computer Science” Master’s Degree - **Visual Analytics** (2023/2024)

Sapienza University of Rome

INDEX

- 1. INTRODUCTION**
- 2. DATA**
- 3. VISUALIZATIONS & ANALYTICS**
 - 3.1. TREEMAPS
 - 3.2. HISTOGRAMS
 - 3.3. PARALLEL COORDINATES VIEW
 - 3.4. TAGS RANKING
 - 3.5. GENERAL TAGS INFO
 - 3.6. TAGS SIMILARITY SCATTERPLOTS & CHORD DIAGRAM
 - 3.7. TAGS SCATTERPLOT / BUBBLE CHART
- 4. RELATED WORKS**
 - 4.8. FOOD & BEVERAGE CASE STUDY
 - 4.9. VISUAL ANALYTICS FOR COMPETITIVE INTELLIGENCE
 - 4.10. OTHER RELATED WORKS
- 5. USE CASES & INSIGHTS**
 - 5.8. INTENDED USERS AND USE CASES
 - 5.9. USE CASES & INSIGHTS EXAMPLES
- 6. CONCLUSIONS AND FUTURE WORKS**
- 1.8. GENERAL GAMES INFO
- 1.9. GAME RESULTS
- 1.10. TAGS TIME SERIES

INTRODUCTION

MARKET ANALYSIS vs MARKET RESEARCH

Market Analysis

The activity of gathering information about conditions that affect a marketplace.

Studies the attractiveness of a market (degree to which a market offers opportunities to an organization).

Market Research

The activity of gathering information about consumers' needs and preferences.

Aimed at identifying the needs of the market, the market size and the competition.

VISUAL ANALYTICS & MARKET ANALYSIS

- Visualizations allow **more users** to **understand data and insights**
 - Business data is often relegated to specific data analysts and data analytics division in companies
 - Visualizations make complex data more accessible
- Visualization allow to identify **markets/customers trends and patterns**
- **Interactive data exploration** allows to dive deeper into **specific aspects of the market**

ONLINE MARKETPLACES

Digital hubs where buyers and sellers meet to trade goods and services

Opposed to businesses' **physical retail locations and stores**.

Globalization and digital transformation led to a wide spread use of e-commerce platforms, online stores, digital goods/services marketplaces.

Covid-19 pandemic accelerated the digital transformation and led to a **shift in consumer behaviors**, from physical to online.

THE LITERATURE

Various publications on companies/industries business data visualizations.

Few research publications on visual analytics solutions for market analysis and market research (centered around specific markets and industries).

Almost no research publications on visual analytics solutions for market analysis and market research in broad consumer online marketplaces.

Existing solutions focus on **traditional businesses** and **financial markets**.

ABSENT LITERATURE - POSSIBLE REASONS

1) Lack of publicly available data

- Businesses don't share their data to maintain their competitive advantage
- Lack of data hinders research developments and system implementations

2) Integrated business intelligence tools

- Online marketplace platforms let their business partners use integrated tools
- Integrated tools focus on “internal business intelligence” (no market overview)
- External solutions development is discouraged by lack of data and APIs

3) Business vs Open Research

- Existing solutions are often B2B products rather than open research attempts

STEAM'S VIDEO GAME MARKETPLACE

- **Biggest** online distribution platform in the desktop gaming industry
- Over **85.000** games and software products
- Over **130 million** monthly active users (as of the end of 2023)
- Owned by **Valve**
 - Private american company
 - \$13 billion in revenue in 2022
 - 44% desktop gaming market share (as of the end of 2022)
- Online marketplace for digital goods (video games and software)
- Presents similar characteristic to online/physical goods/services marketplaces
 - Diverse products/services offerings
 - Vendors variety
 - Search and filter options
 - Global accessibility
 - User reviews and ratings
- No sharing of financial and business data

DATA

STEAM GAMES DATASET (KAGGLE + PYTHON)

INITIAL GAMES DATASET

- **Kaggle Dataset:** “Steam Games Dataset” (by Martin Bustos)
- **84.368 unique data items**
- **39 attributes** each
 - 29 attributes have simple data values
 - 10 attributes have complex data structures as values (lists of values, list of key-value pairs, list of objects)
- **No business data attached**
 - No game sales & revenue data

FINAL GAMES DATASET

- **Python** used for processing data
- Filtered data items ➔ **73.078 game items**
- Edited attributes ➔ **17 attributes**
 - Removed 12 non-relevant attributes
 - Aggregated 4 structured attributes
 - Kept 11 simple data attributes
 - Obtained 6 final complex data attributes (including **list of tags**)
 - + “**Estimated copies sold**” (Boxleiter Method)
 - + “**Estimated revenue**” (Fixed Price Discount)

GAME TAGS (TAG-ORIENTED SYSTEM)

Every game item can have an associated **list of tags**

- **Max 20 tags** per game item, taken from a **pool of 448 tags**
- Game publisher defines initial list, Steam platform users can suggest changes
- Tags in list are **ranked by relevance** to the associated game (publisher + users votes)

New “tags dataset” (computed by iterating over all games’ list of tags)

- **423 tags** included (out of 448 total)
- Computed **name**, **number of games** and **tag category** (“Genre”, “Sub-Genre”, “Feature”)
- Computed **t-SNE and MDS coordinates** from **tags similarity** value
(for each pair $\langle x,y \rangle$ of tags, number of games with both x and y in games’ tags list)

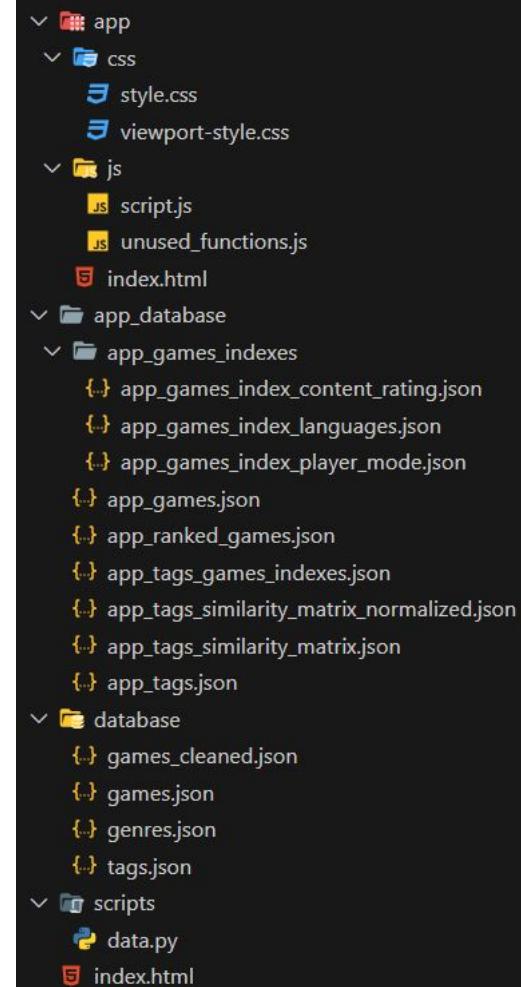
SYSTEM'S FILE STRUCTURE

“dataset” folder ➔ Initial raw data

- Kaggle: “games.json” + “games_cleaned.json”
- Processed (with Python): “tags.json” + “genres.json”

“app_database” folder ➔ Data used by the system

- Main Datasets: **app_games.json** + **app_tags.json** (by category)
- Other Datasets:
 - **app_tags_similarity_matrix.json** (+ normalized version)
 - **app_ranked_games.json** (game IDs sorted by attributes)
 - **app_tags_games_indexes.json** (tags' associated game IDs)
- Indexed Datasets: **app_games_indexes** (sub-folder)

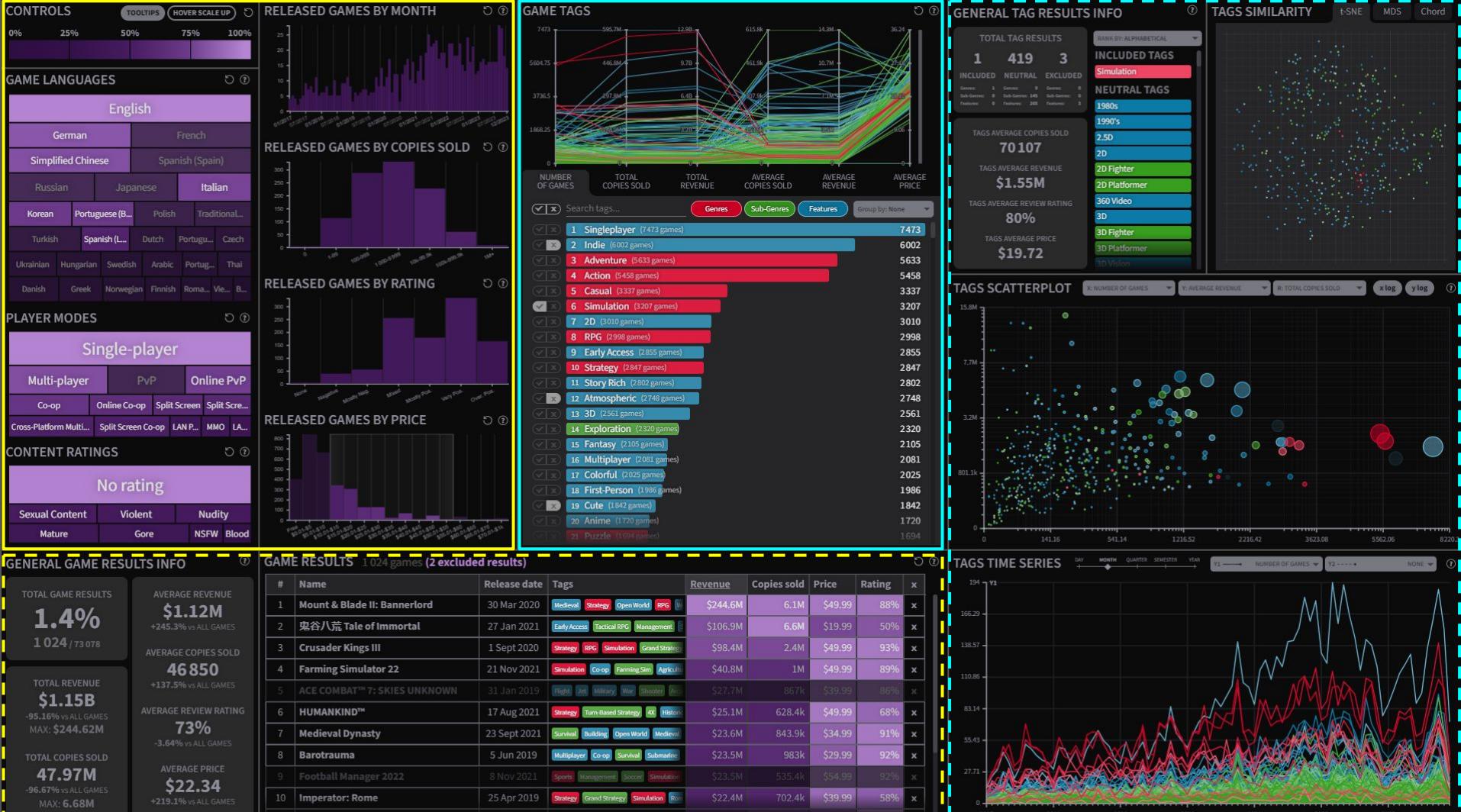


VISUALIZATIONS & ANALYTICS



VISUALIZATIONS & SYSTEM FEATURES

- **10 interactive, coordinated and customizable visualizations**
(treemaps, histograms, scatterplots, bubble charts, rankings, parallel coordinate views, chord diagrams, tabular visualizations, line charts, numerical visualizations)
- System designed for **4K resolutions** ➔ Single sections **scaling to full size**
- **Detailed infos & instructions** for each visualization (associated "?" button)
- **QoL Features:** general controls, hover tooltips, legends, color consistency
- Interleaved or separate **games and/or tags exploration**



TAG-ORIENTED SYSTEM

System allows for both:

- **Exploration of game items** (based on user defined filters, global and local)
- **Exploration of tags** (based on user defined filters, global and local)

Tag-oriented market analysis

- **Analysis of data about single tags** (rather than single product or product types)
- Allows to **analyze business data for dishomogeneous products and services**
- Allows to **compare data about overlapping subsets of items**
- Allows to **discover additional global insights** (w.r.t. single items analysis)

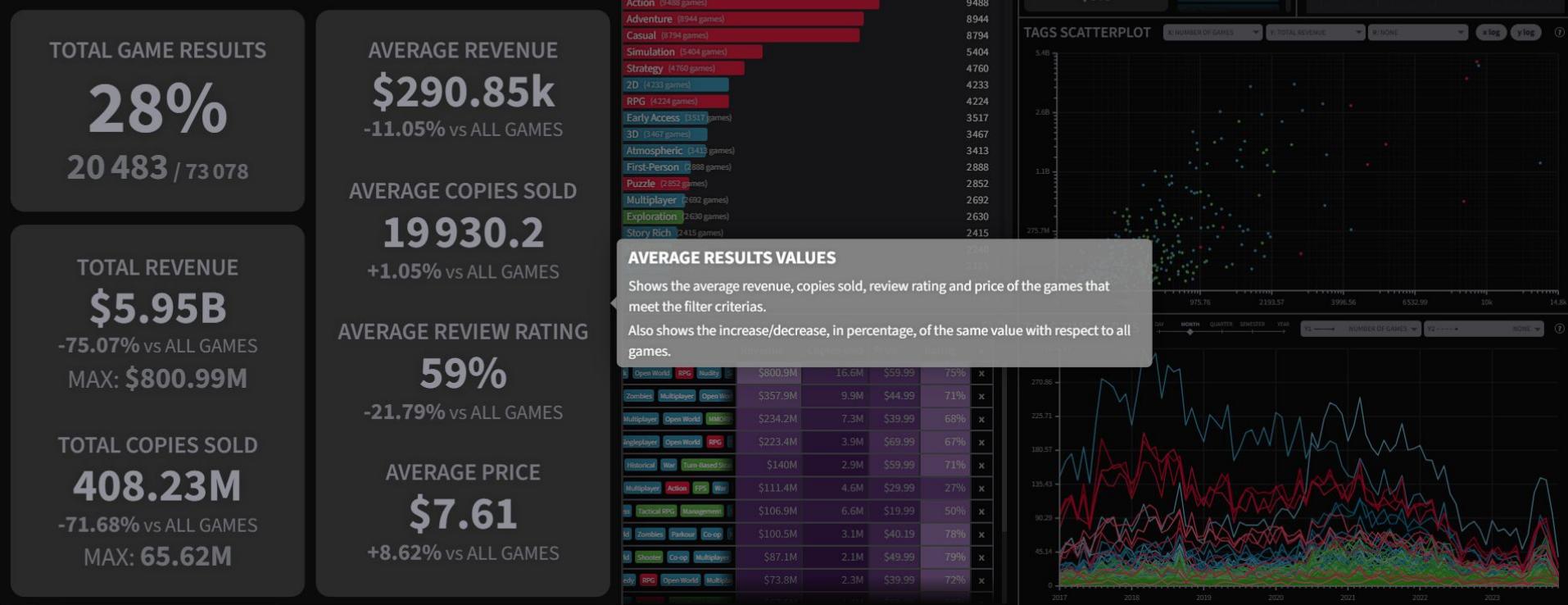


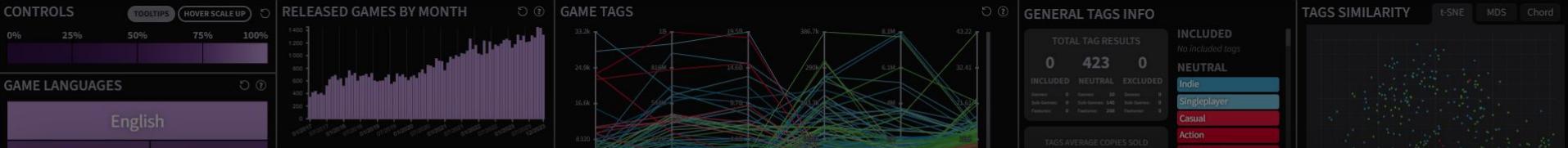






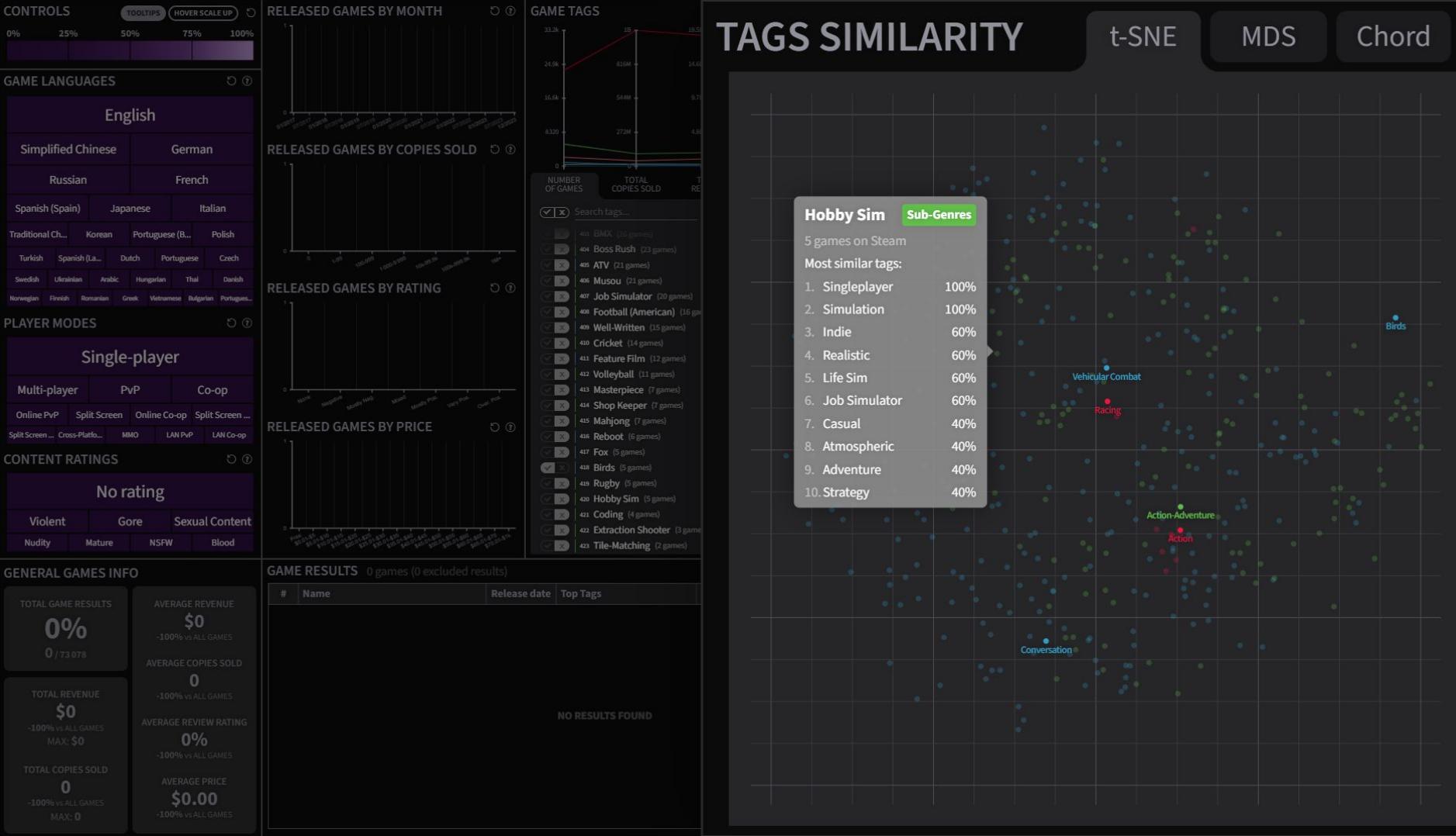
GENERAL GAMES INFO

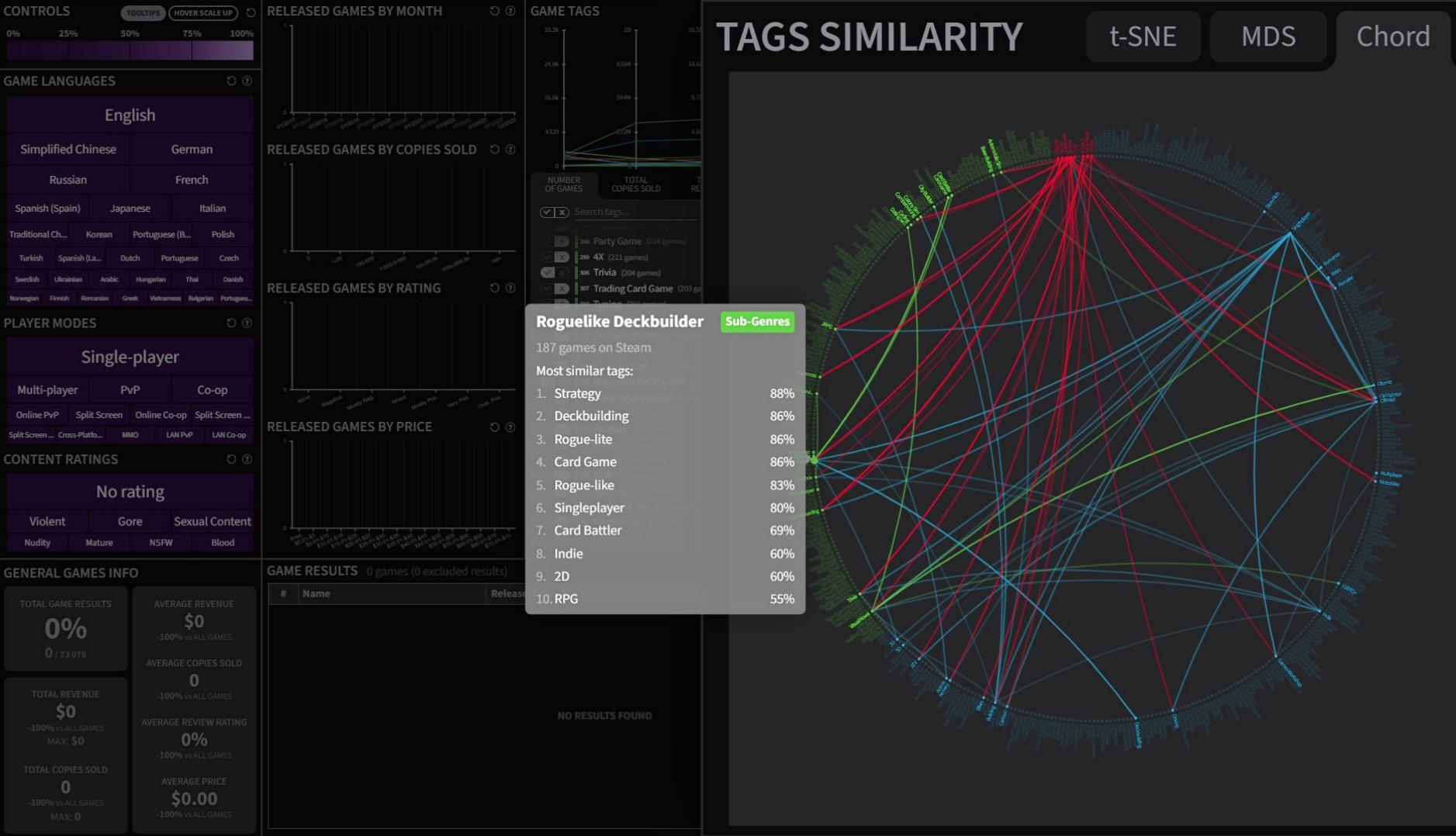




GAME RESULTS 73 078 games (0 excluded results)

#	Name	Release date	Top Tags	Revenue	Copies sold	Price	Rating	x
1	Rust	8 Feb 2018	Survival Crafting Multiplayer Open World Open World Survival Craft	\$831.1M	25.9M	\$39.99	87%	x
2	Cyberpunk 2077	9 Dec 2020	Cyberpunk Open World RPG Nudity Singleplayer Sci-fi Futuristic	\$800.9M	16.6M	\$59.99	75%	x
3	ELDEN RING	24 Feb 2022	Souls-like Relaxing Dark Fantasy RPG Difficult Open World Action	\$786.3M	16.3M	\$59.99	90%	x
4	ARK: Survival Evolved	27 Aug 2017	Open World Survival Craft Survival Open World Multiplayer Dinosaur	\$430.1M	17.9M	\$29.99	82%	x
5	DayZ	13 Dec 2018	Survival Zombies Multiplayer Open World Action PVP Shooter	\$357.9M	9.9M	\$44.99	71%	x
6	Sekiro™: Shadows Die Twice - GOTY Edition	21 Mar 2019	Souls-like Difficult Action Singleplayer Ninja Stealth Adventure	\$290M	6M	\$59.99	95%	x
7	Monster Hunter: World	9 Aug 2018	Co-op Multiplayer Action Open World RPG Third Person Character	\$275.4M	11.4M	\$29.99	86%	x
8	Mount & Blade II: Bannerlord	30 Mar 2020	Medieval Strategy Open World RPG War Sandbox Multiplayer	\$244.6M	6.1M	\$49.99	88%	x
9	Sea of Thieves	3 Jun 2020	Adventure Multiplayer Open World Pirates Co-op Action Online	\$238.7M	7.4M	\$39.99	90%	x
10	New World	28 Sept 2021	Massively Multiplayer Open World MMORPG RPG Adventure Action	\$234.2M	7.3M	\$39.99	68%	x
11	Starfield	5 Sept 2023	Space Singleplayer Open World RPG Sci-fi Exploration First-Person	\$223.4M	3.9M	\$69.99	67%	x
12	Red Dead Redemption 2	5 Dec 2019	Open World Adventure Story Rich Western Action Multiplayer	\$218.3M	9.1M	\$29.99	89%	x
13	Forza Horizon 4	9 Mar 2021	Racing Open World Driving Multiplayer Online Co-Op Automobile	\$211.3M	4.4M	\$59.99	89%	x
14	The Forest	30 Apr 2018	Open World Survival Craft Survival Open World Horror Crafting Adventure	\$198.3M	12.4M	\$19.99	95%	x
15	Halo: The Master Chief Collection	3 Dec 2019	Great Soundtrack FPS Classic Multiplayer Story Rich Sci-fi Comedy	\$195.9M	6.1M	\$39.99	93%	x
16	Far Cry® 5	26 Mar 2018	Open World Co-op Action Multiplayer FPS Shooter Adventure	\$193.3M	4M	\$59.99	80%	x
17	Assassin's Creed® Odyssey	5 Oct 2018	Open World RPG Singleplayer Assassin Action Adventure History	\$190.9M	3.9M	\$59.99	89%	x
18	Valheim	2 Feb 2021	Open World Survival Craft Survival Online Co-Op Open World Multiplayer	\$180.8M	11.3M	\$19.99	95%	x
19	Total War: WARHAMMER II	28 Sept 2017	Strategy Fantasy Turn-Based Strategy RTS Grand Strategy War	\$177.7M	3.7M	\$59.99	93%	x
20	Divinity: Original Sin 2 - Definitive Edition	14 Sept 2017	Tactical RPG RPG Turn-Based Strategy Exploration Story Rich Casual	\$170M	4.7M	\$44.99	96%	x





TAGS SCATTERPLOT

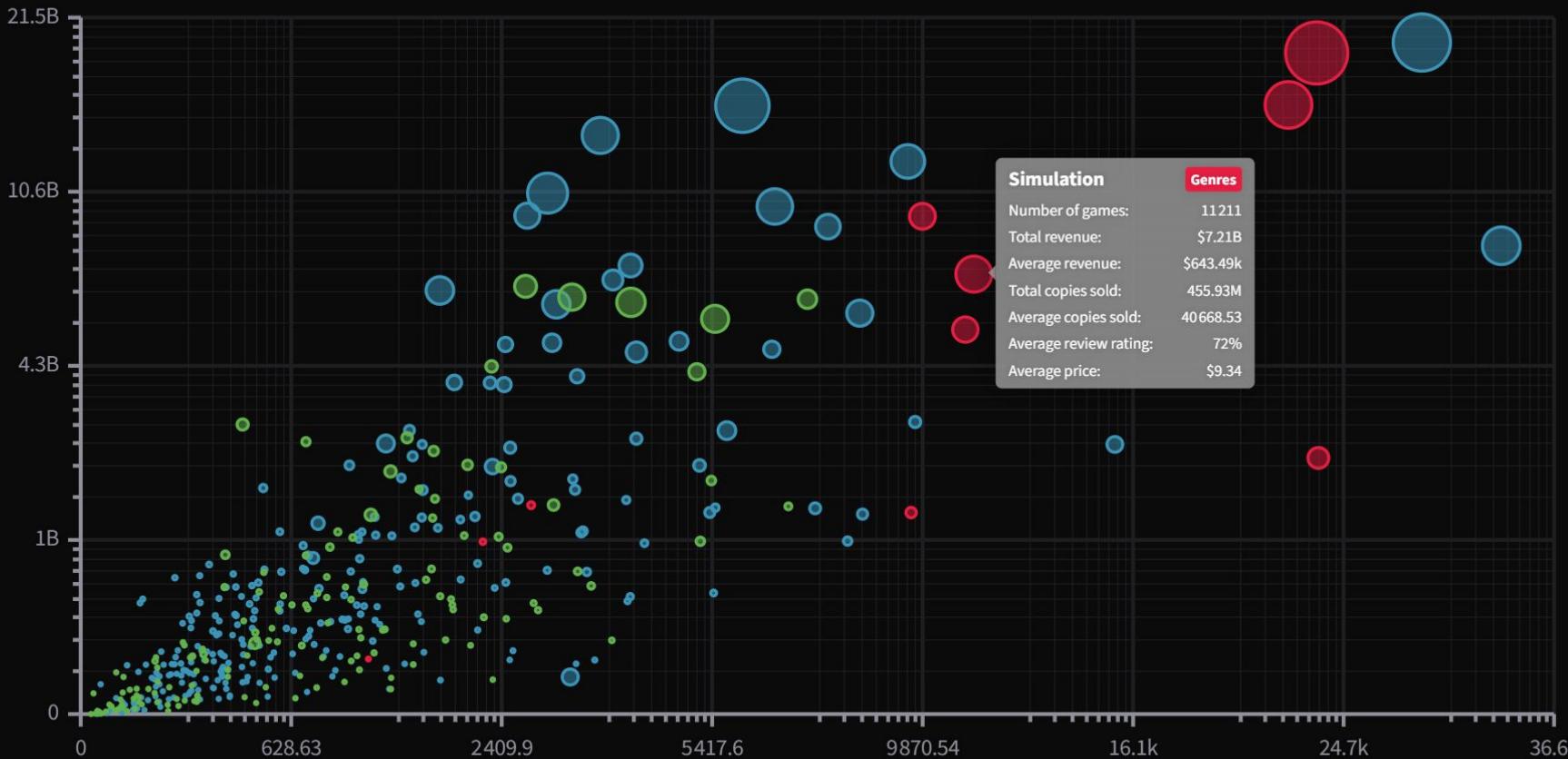
X: NUMBER OF GAMES

Y: TOTAL REVENUE

R: TOTAL COPIES SOLD

x log y log

?



The chart displays the total revenue for the game 'New World' across different platforms. It also shows the total revenue for other games like 'Sea of Thieves', 'Starfield', and 'Grand Theft Auto V'. The chart highlights the massive success of 'New World' with over \$234 million in revenue.



RELATED WORKS

TagReel: A Visualization of Tag Relations among User Interests in the Social Tagging System

Joohee Bae

Integrated Design Laboratory, Department of Media
 Graduate School of Ajou University
 Suwon, Korea
 jooya0211@ajou.ac.kr

Kyungwon Lee

Division of Digital Media
 Ajou University
 Suwon, Korea
 kwlee@ajou.ac.kr

Abstract— Social tagging systems provide users with the ability to share information and extend their field of knowledge. The purpose of this paper is to explore the tag relations of user interest in these systems to examine the semantic relationships of tag in terms of user interests. To do this, a classifying method was used to characterize words into seven classes. Experiments were limited to the five areas of interests: music, photography, games, books and video. Samples were taken from 50 English-language web pages in the social bookmarking service Delicious. The top 7 tags were extracted from each web page that was tagged as music, photography, games, books and videos. Via the results of the tag analysis of this study, relationship between tags and user interests were demonstrated. TagReel employed a radial visualization method, which has the advantage of giving an overview of the whole cluster of connection among tags. TagReel can represent the semantic concept of connection between tags and between interests. It can contribute to raising the cognitive power and awareness of tag usage in various categories and show the relative weight of tags.

Keywords:visualization; social tagging; tag; resource; user; social bookmarking

I. INTRODUCTION

Social tagging has received a considerable amount of interest in recent years due to its advantages in sharing information. Through this system, users create a collection of items of personal interest and assign individual keywords to each of the resources in that collection [1].

Individual keywords, referred to as tags, relate to different types of content or to the work of a user. In contrast to categorization, tagging represents a new approach to organizing information. Nonhierarchical classification allows data to be freely gathered, allows easy access, and has the ability to move directly to other types of content. A study of tagging by Riddle (2005) found that tagging helps to increase personal recall and that it has beneficial social effects while allowing for serendipity and novelty [2]. Tagging can freely represent a user's own ideas as keywords and can allow users to manage information individually. Additionally, tagging groups related URLs together [3] and the same tags also cross over diverse fields of resources [4]. Thus, it has the potential to connect a variety of types of content. Users

are used to describing resources of interest by means of one or more freely chosen keywords and exploring every user and tags with the purpose of discovering various resources of interest [5]. Accordingly, tags are considered as user interests and collecting a new resource also thought as user activities of interests.

The first part of this paper investigates the structure of social tagging system and illustrates relationships among users, tags, and resources. After ascertaining the major patterns, the second section identifies the classification of tags and conducts a qualitative study of the distribution of tag types. In the last section, the visualization of tags in the social tagging system is demonstrated and summarizes the overall discussion.

II. RELATED WORK

A. Social Tagging

Social tagging refers to the tags people use to collect the various keywords to organize a subject. Tagging makes it possible to annotate freely various resources such as images, videos and web pages. Although there is no rule to making words, users become used to sharing and organizing in an unauthorized way in the social tagging system. The importance of social tagging is that it is a reliable way to perceive participatory behavior. This is because users tag proper words to the documents even when no one is looking, for the purpose of sharing with other users and for their own retrieval.

Weinberger (2005) insists that tagging causes people who belong to the social community to form their interests and to develop a similarity of view [6]. Thus, if users have the same tags, it represents that they are sharing deep commonalities with each other. The purpose of tagging is to manage their own data as well as to share and understand other users' thoughts and ideas.

Therefore, tagging is regarded as a method of communication. For these reasons, websites based on the social tagging system, such as Flickr, Delicious, Last.fm and YouTube, have advantages for observing personal interests as well as public interests, simultaneously.

“TAG-REEL” VISUALIZATION

Key Insights

- 2009 paper by Bae et al. [3]
- Radial chart to visualize tag relationships
- Circle of points groped by color & radial distance (hierarchical data structure)
- Static visualizations → Difficult data exploration

Differences with Proposed Solution

- Interactive visualization
 - Allows for global + local data exploration
 - Coordinated view customization
- Tags only grouped by color
 (classification by category, no hierarchy)

Visual Analytics for Market Research: A Case Study of F&B Industry

1st Choong Kah Mann
Asia Pacific University
Kuala Lumpur, Malaysia

2nd Preethi Subramanian
Asia Pacific University
Kuala Lumpur, Malaysia
dr.preethi@apu.edu.my

Abstract—The rapid globalization as well as the technological advancement has led to the world economies to be increasingly interconnected. The international market economics and the lowered trade barriers have also prompted businesses to venture beyond their country of origin in order to obtain greater market share and sales revenue. Numerous such failed global expansion projects can be attributed to the lack of understanding of the international market. Therefore, it is important for companies to perform thorough analysis in order to identify a target region and direct the analytical resources on the target region before the expansion and such an analysis becomes highly possible with the abundance of data. This paper focuses on applying visual analytics for the purpose of market research in the food industry from various points of view. Visual data analytics has been deployed to ascertain the region that holds the most potential. Analysis also includes the performance of the major competitors within the identified region to draw insights and conclusion on the market.

Keywords—Data analytics, Globalization, Market research, Target region, Visual analytics.

I. INTRODUCTION

Food and beverage (F&B) industry has witnessed phenomenal growth due to various factors such as the rapidly growing population, technological advancement and supply chain logistics and F&B has become one of the most important components of the world economics. According to the report gathered by Statista in 2019, the food and beverage industry is expected to achieve approximately 76, 647 million dollars this year and predicted to rise up to 105, 898 million dollars 4 years later. Given the lucrative market, it is expected that the market is getting increasingly competitive over the years. In order to ensure that the business remains competitive and sustains a larger market share, international expansion is a vital strategy used by several companies within the food and beverage industry [1].

The force of globalization has also contributed to the international expansion of various companies where the import and exportation of goods can be facilitated by cargo ships, planes, and the internet [2]. Additionally, the lowered trade barriers have positively contributed to the integration of international economies [3][4]. Considering the vast opportunities and greater sales, many companies are joining the global expansion band wagon. Although many have been successful in gaining market share in international market, there are also many who have failed. Factors such as the lack of understanding of consumer preferences [5], logistics and supply chain issues [6] and the failure to gather the needs of the local consumer market [7] have caused countless companies to fail in their global expansion and suffer significant monetary loss. Based on these findings, it is important for the company to understand the target market before decision making regarding international expansion. Specifically, in-depth market research of the target location

becomes essential. It is also important to remember that similar strategies may not work across international borders as culture, likes and preferences vary. This is also attributed to the differences in the dining habits, consumer preferences, and the failure to anticipate the strategies used by the other competitors that are leading players in the local market.

This paper focuses on international expansion through a data-driven approach where visual analytics is employed for strategic decision making. The scope is within the confines of the fast food industry. Because of the company's branding and persistence to not serve pre-fried food, the fast food chain gradually dominated the fast food market in America.

II. RELATED WORKS

Last year, the company had become one of the largest fast food chains in America by sales alone as well as surpassing its' much larger competitors such as McDonalds, Subway, and KFC [8][9]. The company's great success in America served as an encouragement to expand globally. Without much research, the company had opened up its' first international branch in the UK. However, the formula of success was no longer effective in the UK market. Consequently, the company's first international branch in the UK is closed down in just 6 months of operation [10]. Even though public relation can address some concerns, it is essential that the company must formulate a more effective expansion plan by first understanding the local market. Based on the article by Yoder, Visich, and Rustambekov (2015) [11], the evaluation of the market size and competitive analysis on a few selected locations are important before any decision for global expansion is made. The ROI of the expansion can be determined from the market while the competitive analysis will help determine the subsequent plans for the expansion. In other words, it is essential to identify profitable locations and analyse the competitors that are already in the target market. This section has thus far explained the domain and the problem focus. The rest of the paper addresses the dataset, analytical tools, and the analytical methods used to conduct the market research. Following the discussion on the methods, the results of the analysis is reported with suitable recommendations.

III. PROPOSED METHOD

Two main objectives have been formulated for better international expansion plan:

- To analyze the market potential of each region outside of America through population size and demographics analysis.
- To identify the locations and density of competitors in the fast food industry in selected countries through visual analytics.

FOOD & BEVERAGE CASE STUDY

Key Insights

- 2022 paper by Mann et al. [4]
- Importance of market research for businesses
- Benefits of visual analytics for market research
- Weak use of visual analytics (static infographics)
- Focuses on traditional business practices
- Market analysis focused on competitor analysis

Differences with Proposed Solution

- Stronger use of visual analytics solutions (interactive, customizable & coordinated views)
- Focuses on online marketplaces (Steam platform)
- Employs a tag-oriented market analysis approach

Article

Visualising Business Data: A Survey

Richard C. Roberts *  and Robert S. Laramee 

Visual and Interactive Computing Group, Swansea University, Swansea SA2 8PP, UK;
R.S.Laramee@swansea.ac.uk

* Correspondence: richardroberts1992@gmail.com; Tel.: +44-1792-602-609

Received: 25 September 2018; Accepted: 9 November 2018; Published: 17 November 2018



Abstract: A rapidly increasing number of businesses rely on visualisation solutions for their data management challenges. This demand stems from an industry-wide shift towards data-driven approaches to decision making and problem-solving. However, there is an overwhelming mass of heterogeneous data collected as a result. The analysis of these data become a critical and challenging part of the business process. Employing visual analysis increases data comprehension thus enabling a wider range of users to interpret the underlying behaviour, as opposed to skilled but expensive data analysts. Widening the reach to an audience with a broader range of backgrounds creates new opportunities for decision making, problem-solving, trend identification, and creative thinking. In this survey, we identify trends in business visualisation and visual analytic literature where visualisation is used to address data challenges and identify areas in which industries use visual design to develop their understanding of the business environment. Our novel classification of literature includes the topics of businesses intelligence, business ecosystem, customer-centric. This survey provides a valuable overview and insight into the business visualisation literature with a novel classification that highlights both mature and less developed research directions.

Keywords: business; customer; ecosystem; visualisation; survey

1. Introduction and Motivation

As businesses make the transition to digital solutions, they become overwhelmed by the volume of data they collect. The continued evolution of improved hardware propagates a cycle of collecting larger quantities of data at a lower cost. Despite the investment made to collect data, it still only accounts for a fraction of the process required for useful output. Interpretation of data is vital to unlock the potential value held within and to make the most informed decisions. Companies often employ teams of data analysts to achieve this. However, this can be very costly. In addition to the cost, it also limits the number of people who may understand or access the analysis. Employing visualisation approaches enables employees with a wide range of backgrounds to view and understand it [1]. Opening the analysis to a wider audience encourages ideas and provokes discussion about the nature of the behaviour under investigation [2]. This broadening of the audience highlights a unique benefit that data visualisation and visual analytics can offer.

Visualisation and visual analytics have the capability to overcome the challenges associated with large datasets ([3] p. 56) and multidimensional relationships [4]. In business, the holistic nature of “big picture” approaches is valuable [5], providing a complete overview of a scenario or situation. Pairing this requirement of large-scale data analytics with the capabilities of data visualisation produces fruitful and thought-provoking output.

The body of business visualisation literature is growing rapidly. During the IEEE VIS 2014 conference, a workshop entitled “Business” focused on the conversion of business data into meaningful visual insight which aids in better decision making [6]. The workshop was so popular that a second

BUSINESS DATA VISUALIZATIONS

- 2018 survey by Roberts et al. [5]
- Analyzes various business-oriented visual analytics solutions grouped in 4 categories (“Business Intelligence”, “Business Ecosystem”, “Customer Centric”, “Financial Visualizations”)
- Few visual analytics solutions for market analysis
- No market analysis in online marketplaces
 - 1990s: First online marketplaces appear
 - 2000s: Globalization & WWW Internet access
 - 2010s: Boom of e-commerce and online shops

Type	Source	Business Intelligence		Business Ecosystem		Customer Centric	
		Internal Intelligence	External Intelligence	Business Ecosystem	Customer Behaviour	Customer Feedback	
Primary Data	Intentional, Active, Digital Collection	Otsuka et al. [20]			Yaeli et al. [21] Nagaoka et al. [22]		
	Intentional, Active, Research Study Data	Burkhard [23] Sedlmair et al. [24] Kandel et al. [18] Aigner [25] Lafon et al. [26]	Bresciani and Eppler [27] Bertschi [28] Keahey [29]	Merino et al. [30] Basole et al. [31]	Dou et al. [32]		Brodbeck and Girardin [33]
Hybrid Data	WebScrape		Ramesh et al. [34]	Lu et al. [35]	Shi et al. [36] Sijtsma et al. [37]		Chen et al. [38] Ziegler et al. [39] Oelke et al. [40] Wu et al. [41] Hao et al. [42] Saitoh [43] Fayoumi et al. [44] Haleem et al. [45] Saga and Yagi [46]
		Wright [47] Vliegen et al. [48] Bai et al. [49] Nicholas et al. [50] Roberts et al. [51] Kumar and Belwal [52] Roberts et al. [53]	Ferreira et al. [54]	Wattenberg [55] Wu and Phillips [56] Basole et al. [57] Basole et al. [58] Deligiannidis and Noyes [59] Basole et al. [60] Iyer and Basole [61] Schotter et al. [62] Basole et al. [63]	Woo et al. [64] Hanafizadeh and Mirzazadeh [65] Kameoka et al. [66] Wu et al. [67] Sathiyanarayanan et al. [68]		Kang et al. [69]
Secondary Data	Business Process	Du et al. [70] Broeksema et al. [71] Ghooshchi et al. [72] Bachhofner et al. [73] Lea et al. [74]	Hao et al. [19] Hao et al. [75]	Basole [76] Basole and Bellamy [77]			
	Business By-product	Gresh and Kelton [78] Eick [79] Keim et al. [80]	Liu et al [81]	Oflacques et al. [82] Ko et al. [83]	Rodden [84] Nair et al. [85]		

MarketAnalyzer: An Interactive Visual Analytics System for Analyzing Competitive Advantage Using Point of Sale Data

S. Ko^{†1}, R. Maciejewski^{‡2}, Y. Jang^{§3}, and D. S. Ebert^{¶1}

¹Purdue University

²Arizona State University

³Sejong University

Abstract

Competitive intelligence is a systematic approach for gathering, analyzing, and managing information to make informed business decisions. Many companies use competitive intelligence to identify risks and opportunities within markets. Point of sale data that retailers share with vendors is of critical importance in developing competitive intelligence. However, existing tools do not easily enable the analysis of such large and complex data; therefore, new approaches are needed in order to facilitate better analysis and decision making. In this paper, we present MarketAnalyzer, an interactive visual analytics system designed to allow vendors to increase their competitive intelligence. MarketAnalyzer utilizes pixel-based matrices to present sales data, trends, and market share growths of products of the entire market within a single display. These matrices are augmented by advanced underlying analytical methods to enable the quick evaluation of growth and risk within market sectors. Furthermore, our system enables the aggregation of point of sale data in geographical views that provide analysts with the ability to explore the impact of regional demographics and trends. Additionally, overview and detailed information is provided through a series of coordinated multiple views. In order to demonstrate the effectiveness of our system, we provide two use-case scenarios as well as feedback from market analysts.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Pixel-based visualization, Linked Views, Geospatial, Temporal, Multi-variate, Business Intelligence, Competitive Intelligence, Market Analysis—

1. Introduction

The underlying goal of a business is to increase (or at least maintain) its current market share and to maximize its profits within the market. In order to pursue this goal, analysts must constantly explore and analyze market share data changes that are relevant to their current business sector. Their goal is to forecast changes in the market as a means of controlling and expanding the company's current market share. This exploration, analysis, and prediction of the market share is termed *competitive intelligence* (CI) [Kah98]. Companies use CI to compare themselves to other companies, to identify market risks and opportunities and to evaluate the potential impact of new sales strategies.

In order to generate intelligence reports, many companies extract information from a variety of sources using various methods of data collection and analysis (e.g., networking with company rivals, examining security filings, patent application analysis). One key data source is point of sale data that retailers share with vendors. This point of sale data is temporal, multivariate, and spatial in nature; therefore, it is well suited for analysis in a visual analytics environment. However, it is difficult to find systems that manage the characteristics of point of sale data effectively. In this paper, we present MarketAnalyzer, a visual analytics system for exploring, comparing, analyzing, and predicting trends of point of sale data. We have worked directly with analysts to provide proper and accurate analysis of their point of sale data (e.g., 288 stores with 36 different products) to increase their understanding and improve their market insight. We use an enhanced pixel-based visualization approach [KK94, Kei00] in MarketAnalyzer to efficiently utilize limited screen space for the large store and product

[†] ko@purdue.edu

[‡] rmacieje@asu.edu

[§] jangy@sejong.edu: corresponding author

[¶] eberd@purdue.edu

“MARKET-ANALYZER” SYSTEM

Key Insights

- 2012 paper by Ko et al. [2]
- Focuses on “Competitive Intelligence”
(competitor analysis to forecast market changes)
- Visualizations to compare sales, trends, and growth of a main company and a selected secondary company over a certain time interval
- 3 entities: “products”, “stores” and “companies”

Differences with Proposed Solution

- Focuses on market analysis for online marketplaces
- 1 entity above all: “products” (tag-oriented analysis)

Commercial Visual Analytics Systems – Advances in the Big Data Analytics Field

Michael Behrisch, Dirk Streeb, Florian Stoffel, Daniel Seebacher, Brian Matejek
Stefan Hagen Weber, Sebastian Mittelstädt, Hanspeter Pfister, Daniel Keim

Abstract—Five years after the first state-of-the-art report on Commercial Visual Analytics Systems we present a reevaluation of the Big Data Analytics field. We build on the success of the 2012 survey, which was influential even beyond the boundaries of the InfoVis and Visual Analytics (VA) community. While the field has matured significantly since the original survey, we find that innovation and research-driven development are increasingly sacrificed to satisfy a wide range of user groups. We evaluate new product versions on established evaluation criteria, such as available features, performance, and usability, to extend on and assure comparability with the previous survey. We also investigate previously unavailable products to paint a more complete picture of the commercial VA landscape. Furthermore, we introduce novel measures, like suitability for specific user groups and the ability to handle complex data types, and undertake a new case study to highlight innovative features. We explore the achievements in the commercial sector in addressing VA challenges and propose novel developments that should be on systems' roadmaps in the coming years.

Index Terms—System Comparison, Commercial Landscape, Visual Analytics Research, Advances, Development Roadmap.

1 INTRODUCTION

In 1890, Herman Hollerith revolutionized the world of data analysis with a creative and innovative idea: he used punch cards to collect and analyze the US census data. Using punch cards saved two years and five million dollars over the manual tabulation techniques used in the previous census while enabling more thorough analysis of the data [1]. We currently face an analogous development in the Big Data Analysis field, where commercial Visual Analytics (VA) systems allow a faceted confirmatory or a data-driven exploratory analysis of large amounts of data in significantly less time than years ago. Today, the success of many businesses relies on efficient and effective analysis of massive quantities of data.

Bertin [2] and Tukey [3] consider the possible levels of data, information, and analysis. They summarize data analysis into three levels: presentation, confirmatory, and exploratory analysis. Over the last decade, a significant amount of research explores presentation and confirmatory analysis in the commercial VA field. Specifically, dashboarding systems enable users to gain quick insights with faceted filtering functionality. Confirmatory analysis scenarios are supported either by focusing on simple visual interactive Overview + Detail displays or by incorporating increasingly automatic analysis techniques into coordinated view systems. We claim that although current commercial VA systems have been developed with the aim to support the exploration of large quantities of data, they currently do not sufficiently support

exploratory analysis scenarios. In particular, we see a scarcity of supportive environments where the domain expert and the machine work in an interplay towards formulating and validating hypotheses. This dearth is due to several reasons: (1) often users are left alone in finding a starting point in their analysis; (2) the communication of non-trivial hypotheses is challenging; (3) automatic algorithms for validating interesting findings are not scalable or even implemented in the systems. This survey counterbalances the efforts of the community against the needs and requirements imposed by the Big Data Era. Further, we ask which steps should be taken in the future by examining past directions to allow for exploratory data analysis in Big Data scenarios.

We revise and update the 2012 state-of-the-art report on commercial VA systems following the original methodology and rationale of Zhang et al. [4]. We build our comparative market overview on an encompassing list of 46 relevant commercial VA systems.¹ These chosen systems reflect current market shares [5]–[7] and encompass the broad product categories within the field: e.g., data discovery, visual software, Business Intelligence (BI), innovative, and niche products.

Our survey is structured along two primary dimensions. In a *user/task* oriented view we claim that three user groups with potentially overlapping skill sets are interested in commercial VA systems: (1) Upper management, e.g., CEOs, who make critical business decisions based on prepared presentations; (2) Domain experts who have extensive domain knowledge and can formulate hypotheses; (3) Data analysts and engineers who do not necessarily know the data in advance but have the challenge of finding a needle in a large amount of complex data, potentially at high velocity.

The second dimension structuring this survey relates to the functional capabilities of commercial VA products. Therefore, we approached all 41 vendors (five offered two candidate products)

COMMERCIAL V.A. SYSTEMS

- 2018 survey by Behrisch et al. [1]
- Describes the evolution of **commercial visual analytics solutions** (uses 10 case studies)
- Analysis and evaluation of V.A. solutions features, performance, and usability from users' POV.
- Users classified in “**upper management**”, “**domain experts**”, and “**data analysts/engineers**”
- Not centered around business data visualizations
- Used to guide development of proposed solution
 - ➡ Describes methodologies for evaluating commercial visual analytics systems
 - ➡ Focused on data analysts user group criterias: extensibility, interactivity, and data handling

¹ The complete list is on our website <http://commercialtools.dbvis.de/>

USE CASES & INSIGHTS

INTENDED USERS & USE CASES (1)

“AAA” GAME STUDIOS

- In-house game development
- Large number of **employees**
- Various company/business **divisions**
- **Multi-platform self-publishing**
- Priorities:
 - Conduct in-depth **market analysis**
 - **Analyze competitors**
 - Assess product-market fit
 - Visualizing **trends and growths**

“INDIE” GAME STUDIOS

- In-house game development
- Small number of **employees**
- Unstructured companies
- **Single platform self/external publishing**
- Priorities:
 - Effective **resource allocation**
 - Ensure studios’ **financial sustainability**
 - Estimate **sales and revenue**
 - **Funding** (publishing deal/crowdfunding)

INTENDED USERS & USE CASES (2)

GAME PUBLISHERS

- Companies **funding** game studios
- Cover studios' development costs
- At game launch:
 - Recoup initial funding
 - Get a percentage fee on revenue
- Priorities:
 - Analyze game's "**marketability**"
 - **Find customers**
 - Ensure games' **financial success**

MORE USERS & USE CASES

- Discover **genre/sub-genre "niches"** (low product supply, high demand)
- Evaluate recent **market viability** of a product **being developed**
- Identify **trends** to predict gaming market direction (for **new products development**)
- Help with **product, price, placement and promotion** (4 "Ps" of marketing)

USE CASES & INSIGHTS

System designed to be “**versatile**” (for different types of users & use cases)

- **In-depth filters** allow to cover a **wide range of user needs**
- Visualizations are **highly customizable** (w.r.t. plotted data & views layout)
- Exploration of **both games and tags/genres** (locally and globally)

More use cases are possible, and more types of users can discover insights

Multiple insights can be discovered by interacting with the system

- **Insights examples** provided in project report and discussed during demo

CONCLUSIONS & FUTURE WORKS

CONCLUSIONS & FUTURE WORKS

Main Work's Topic: Visual analytics system for a “tag-oriented” market analysis of online marketplaces centered around the “Steam” desktop gaming distribution platform

Possible research directions and **future works**:

- **Translate** the visual analytics tag-oriented market analysis methodology to different online marketplaces
- Use “official” business data (system used estimated business data)
- **Extend** the tag-oriented methodology to tags with a hierarchical structure (system used tags with no defined hierarchical structure)

END OF THE PRESENTATION

DEMO?

