INF 551- FOUNDATIONS OF DATA MANGEMENT

FINAL REPORT

JEUGE PARAISO

Group Members:

Pratishtha Singh **8957278830** **pratisht@usc.edu**

Sharad Narayan Sharma **4011180766** **sharadna@usc.edu**

Project Idea:

- Gamers often don't know which games to buy, since there is a large variety to pick from.
- This variety arises due to different consoles, gaming genres, game producer, popularity, etc. We wanted to make an application which gives a one-stop solution to the above problems.
- Our web application will give a gamer insight into the best-selling games of the last 40 years, based on different criteria such as sales, genre, publisher, console availability, game versions, etc. This will help him/her to buy a game after a quick one-stop analysis.

Technology Stack Used:

- Database - Firebase
- Server-side scripting - Flask
- Front-end - HTML, CSS, JavaScript

Dataset: *"Video Game Sales"*

https://www.kaggle.com/gregorut/videogamesales

Shape of the Data Set: 16,600 rows x 11 columns

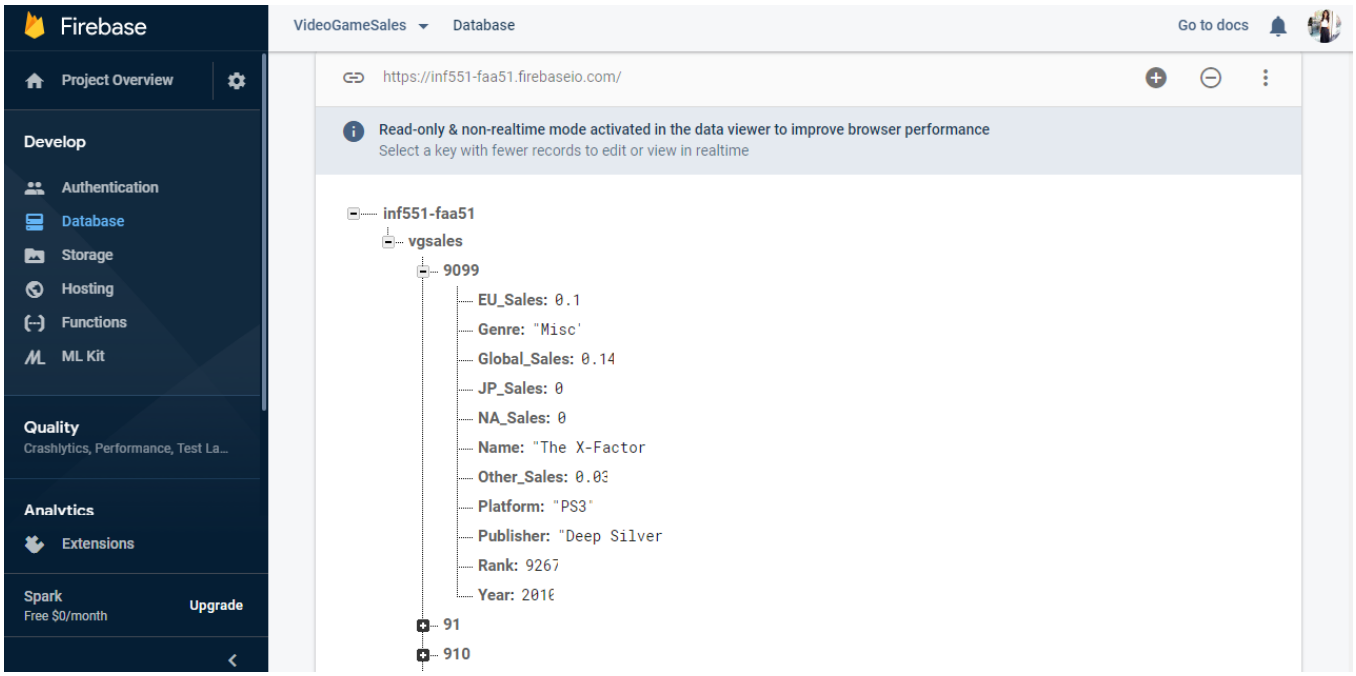
Example Data:

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
2	Super Mario	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
3	Mario Kart W	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
4	Wii Sports R	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33
5	Pokemon Re	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1	31.37
6	Tetris	GB	1989	Puzzle	Nintendo	23.2	2.26	4.22	0.58	30.26
7	New Super M	DS	2006	Platform	Nintendo	11.38	9.23	6.5	2.9	30.01
8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.2	2.93	2.85	29.02
9	New Super M	Wii	2009	Platform	Nintendo	14.59	7.06	4.7	2.26	28.62
10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31
11	Nintendogs	DS	2005	Simulation	Nintendo	9.07	11	1.93	2.75	24.76

Reason for choosing the Dataset:

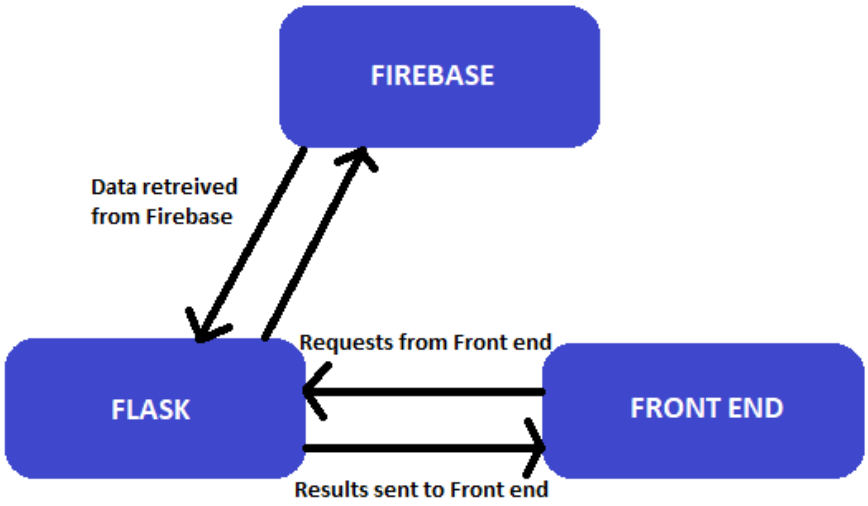
- The dataset has both numerical and non-numerical attributes.
- Size of the dataset makes it ideal for this application as it is more than 1k and less than 100k.
- It is an interesting dataset as it encompasses all those attributes needed for a gamer to buy or analyze a game.

Firebase Database:



Data Flow:

- On button clicks on the front end, the request is send to the server side handler (Flask). Based on these requests, it extracts live data from Firebase every time a gamer chooses a particular sorting or filtering option.
- We perform manipulations on this retrieved data using Python on the back-end.
- The required dataset is sent back to the front end from Flask.



Working Components:

It is a cloud-based Web application with the following features and implementation.

Jeuge Paraiso

What's your genre? Genre Filter

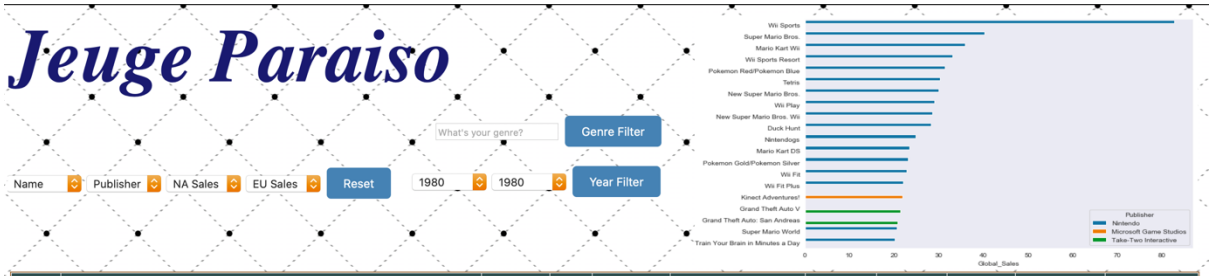
Name Publisher NA Sales EU Sales 1980 1980 Year Filter

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37
6	Tetris	GB	1989	Puzzle	Nintendo	23.20	2.26	4.22	0.58	30.26
7	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.50	2.90	30.01
8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.20	2.93	2.85	29.02
9	New Super Mario Bros. Wii	Wii	2009	Platform	Nintendo	14.59	7.06	4.70	2.26	28.62
10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31
11	Nintendogs	DS	2005	Simulation	Nintendo	9.07	11.00	1.93	2.75	24.76
12	Mario Kart DS	DS	2005	Racing	Nintendo	9.81	7.57	4.13	1.92	23.42
13	Pokemon Gold/Pokemon Silver	GB	1999	Role-Playing	Nintendo	9.00	6.18	7.20	0.71	23.10
14	Wii Fit	Wii	2007	Sports	Nintendo	8.94	8.03	3.60	2.15	22.72

Original landing page

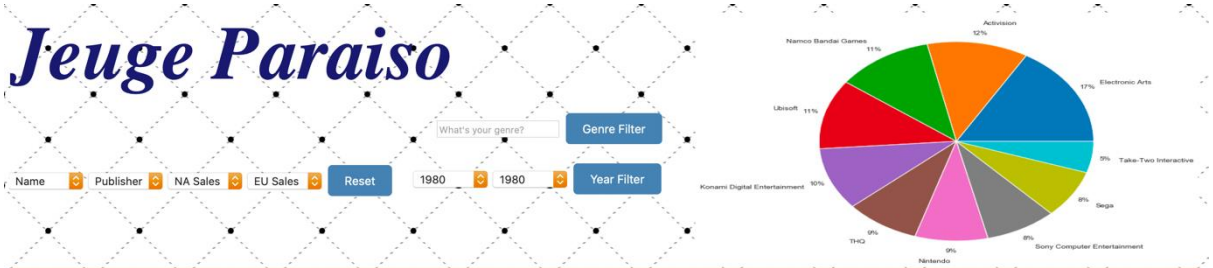
(i) **Sorting:**

- for non-numerical data on attributes of ‘Name’ and ‘Publisher’.
- for numerical data on attribute of ‘NA_Sales’ and ‘Global_Sales’.



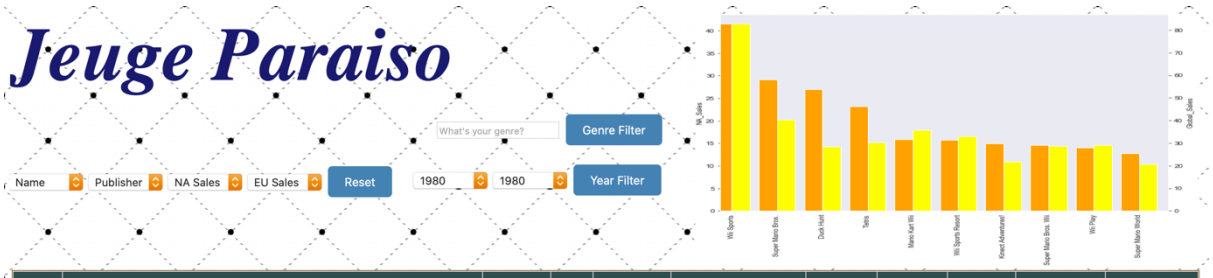
Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
12919	Zyuden Sentai Kyoryuger: Game de Gaburincho!!	3DS	2013	Action	Namco Bandai Games	0.00	0.00	0.05	0.00	0.05
15422	Zwei!!	PSP	2008	Role-Playing	Falcom Corporation	0.00	0.00	0.02	0.00	0.02
6279	Zumba Fitness: World Party	Wii	2013	Misc	Majesco Entertainment	0.14	0.11	0.00	0.03	0.27
6977	Zumba Fitness: World Party	XOne	2013	Misc	Majesco Entertainment	0.17	0.05	0.00	0.02	0.23
7925	Zumba Fitness Rush	X360	2012	Sports	505 Games	0.00	0.16	0.00	0.02	0.19
11968	Zumba Fitness Core	X360	2012	Misc	505 Games	0.00	0.06	0.00	0.01	0.07
11765	Zumba Fitness Core	Wii	2012	Misc	505 Games	0.00	0.07	0.00	0.01	0.07
502	Zumba Fitness 2	Wii	2011	Sports	Majesco Entertainment	1.54	1.07	0.00	0.28	2.88
118	Zumba Fitness	Wii	2010	Sports	505 Games	3.50	2.64	0.00	0.67	6.81
3433	Zumba Fitness	PS3	2010	Sports	505 Games	0.28	0.21	0.00	0.09	0.59

Name Descending Sorting



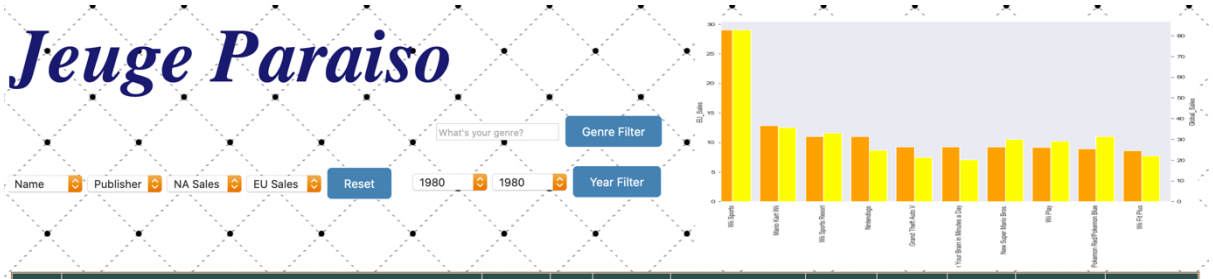
Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
11021	Family GameShow	Wii	2009	Misc	Zushi Games	0.08	0.00	0.00	0.01	0.09
14913	Monster Mayhem: Build and Battle	DS	2009	Simulation	Zushi Games	0.02	0.00	0.00	0.00	0.02
14345	Dodge Racing: Charger vs Challenger	DS	2009	Racing	Zushi Games	0.03	0.00	0.00	0.00	0.03
5211	North American Hunting Extravaganza	Wii	2008	Sports	Zushi Games	0.33	0.00	0.00	0.03	0.36
5473	Chicken Blaster	Wii	2009	Shooter	Zushi Games	0.31	0.00	0.00	0.02	0.33
14274	Jig-A-Pix: Wild World	DS	2009	Puzzle	Zushi Games	0.03	0.01	0.00	0.00	0.03
5919	Reel Fishing: Angler's Dream	Wii	2009	Sports	Zushi Games	0.26	0.01	0.00	0.02	0.30
15468	Garfield Gets Real	DS	2009	Action	Zushi Games	0.00	0.02	0.00	0.00	0.02
15429	Defendin' De Penguin	DS	2008	Strategy	Zushi Games	0.02	0.00	0.00	0.00	0.02
12886	Defendin' De Penguin	Wii	2008	Strategy	Zushi Games	0.05	0.00	0.00	0.00	0.05
13132	Sea Monsters: A Prehistoric Adventure	Wii	2007	Adventure	Zushi Games	0.04	0.00	0.00	0.00	0.05
15110	Indianapolis 500 Evolution	X360	2009	Racing	Zushi Games	0.02	0.00	0.00	0.00	0.02
9840	Dodge Racing: Charger vs Challenger	Wii	2009	Racing	Zushi Games	0.11	0.00	0.00	0.01	0.12

Publisher Descending Sorting



Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31
6	Tetris	GB	1989	Puzzle	Nintendo	23.20	2.26	4.22	0.58	30.26
3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
16	Kinect Adventures!	X360	2010	Misc	Microsoft Game Studios	14.97	4.94	0.24	1.67	21.82
9	New Super Mario Bros. Wii	Wii	2009	Platform	Nintendo	14.59	7.06	4.70	2.26	28.62
8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.20	2.93	2.85	29.02
19	Super Mario World	SNES	1990	Platform	Nintendo	12.78	3.75	3.54	0.55	20.61
7	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.50	2.90	30.01
5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37

NA Sales Descending Sorting



Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
11	Nintendogs	DS	2005	Simulation	Nintendo	9.07	11.00	1.93	2.75	24.76
17	Grand Theft Auto V	PS3	2013	Action	Take-Two Interactive	7.01	9.27	0.97	4.14	21.40
20	Brain Age: Train Your Brain in Minutes a Day	DS	2005	Misc	Nintendo	4.75	9.26	4.16	2.05	20.22
7	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.50	2.90	30.01
8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.20	2.93	2.85	29.02
5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37
15	Wii Fit Plus	Wii	2009	Sports	Nintendo	9.09	8.59	2.53	1.79	22.00
14	Wii Fit	Wii	2007	Sports	Nintendo	8.94	8.03	3.60	2.15	22.72
12	Mario Kart DS	DS	2005	Racing	Nintendo	9.81	7.57	4.13	1.92	23.42

EU Sales Descending Sorting

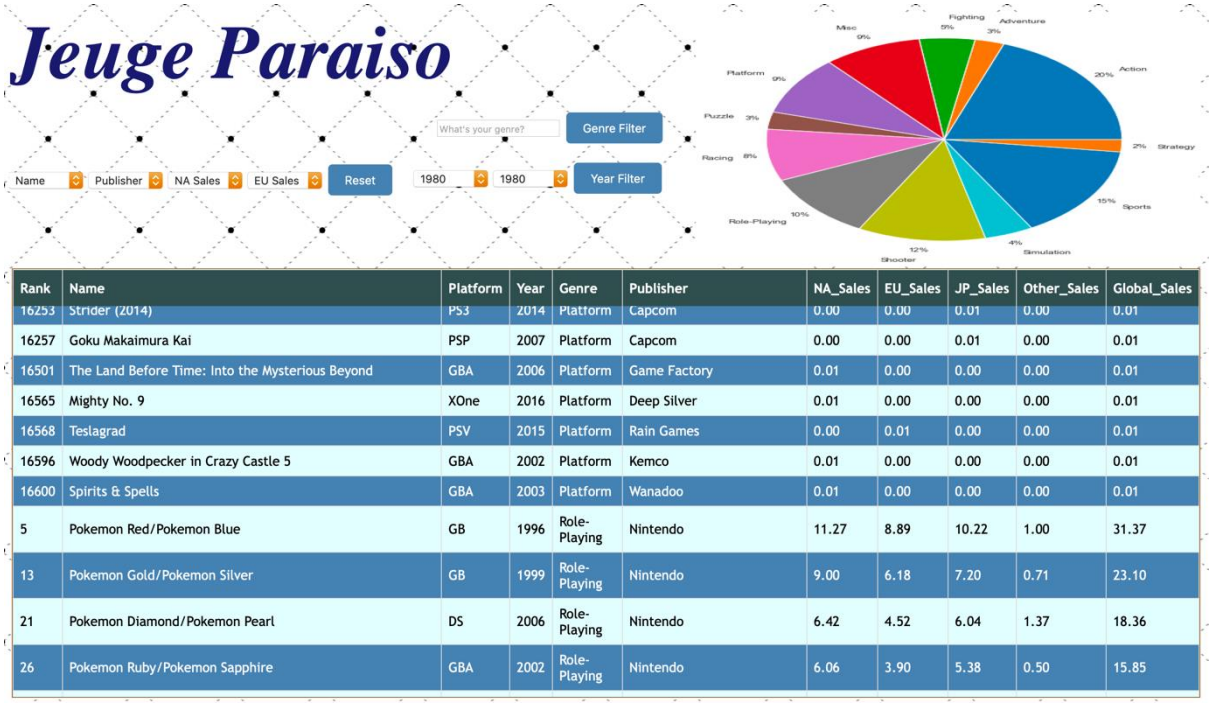
Feature Implementation:

- We've created four sorting buttons, each with the ascending and descending options.
- As explained before, flask gets the data from firebase after the request is sent from the front end.

- We convert this data into a dataframe, and then sort it based on the request received from the gamer.
- The sorted dataframe is displayed as a HTML table on the front end.
- Since we're using Flask, we've used 'render template()' to return four different templates on the basis of the sorting option selected. This is because each option is associated with a different visualization as well. (discussed later in this document).

(ii) **Filter function**

- for non-numerical data on the attribute 'Genre'.
- for numerical data on the attribute 'Year'.



Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
16253	Strider (2014)	PS3	2014	Platform	Capcom	0.00	0.00	0.01	0.00	0.01
16257	Goku Makaimura Kai	PSP	2007	Platform	Capcom	0.00	0.00	0.01	0.00	0.01
16501	The Land Before Time: Into the Mysterious Beyond	GBA	2006	Platform	Game Factory	0.01	0.00	0.00	0.00	0.01
16565	Mighty No. 9	XOne	2016	Platform	Deep Silver	0.01	0.00	0.00	0.00	0.01
16568	Teslagrad	PSV	2015	Platform	Rain Games	0.00	0.01	0.00	0.00	0.01
16596	Woody Woodpecker in Crazy Castle 5	GBA	2002	Platform	Kemco	0.01	0.00	0.00	0.00	0.01
16600	Spirits & Spells	GBA	2003	Platform	Wanadoo	0.01	0.00	0.00	0.00	0.01
5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37
13	Pokemon Gold/Pokemon Silver	GB	1999	Role-Playing	Nintendo	9.00	6.18	7.20	0.71	23.10
21	Pokemon Diamond/Pokemon Pearl	DS	2006	Role-Playing	Nintendo	6.42	4.52	6.04	1.37	18.36
26	Pokemon Ruby/Pokemon Sapphire	GBA	2002	Role-Playing	Nintendo	6.06	3.90	5.38	0.50	15.85

Genre Filter on syllable 'pl'

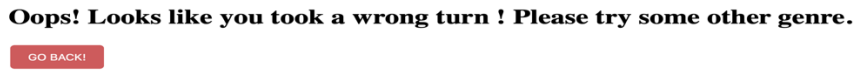
Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31
6298	Beamrider	Atari 2600	1984	Racing	Activision	0.26	0.01	0.00	0.00	0.27
2957	Mappy	NES	1984	Platform	Namco Bandai Games	0.00	0.00	0.69	0.00	0.69
2515	Clu Clu Land	NES	1984	Puzzle	Nintendo	0.42	0.10	0.28	0.02	0.82
1640	Pac-Man	NES	1984	Puzzle	Namco Bandai Games	0.27	0.08	0.85	0.02	1.22
1562	Hogan's Alley	NES	1984	Shooter	Nintendo	0.68	0.16	0.41	0.02	1.27
1326	4 Nin uchi Mahjong	NES	1984	Misc	Nintendo	0.00	0.00	1.45	0.00	1.45
1859	Lode Runner	NES	1984	Puzzle	Hudson Soft	0.00	0.00	1.10	0.00	1.10
1235	F1 Race	NES	1984	Racing	Nintendo	0.00	0.00	1.52	0.00	1.52
940	Pinball	NES	1984	Action	Nintendo	0.80	0.19	0.83	0.03	1.85
756	Tennis	NES	1984	Sports	Nintendo	0.48	0.11	1.56	0.02	2.17
299	Golf	NES	1984	Sports	Nintendo	1.22	0.28	2.46	0.04	4.01

Year Filter on 1984 - 1993

Feature Implementation:

GENRE:

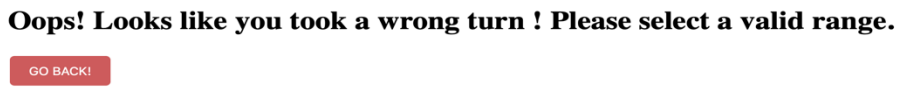
- The gamer can choose to enter any syllable of the gaming genre he/she is interested in. This can be entered in any case as well.
- On the back-end, our code contains a list of all the possible genres present in the database.
- The entered syllable is converted into lower case and compared with the list of genres(all in lower case). If the syllable matches any of the words in the list, those matched words are appended to another list.
- The above case insensitive comparison logic helps the gamer to input any kind of text in any case.
- Now, on the data retrieved from firebase, we pass items in the newly appended list to ‘Genre’.
- The results of the above operations are appended to an empty dataframe and the final dataframe result is displayed as an HTML table on the front end.
- Since we’re using Flask, we’ve used ‘render template()’ to return a different template when genre filtering is selected . This is because each option is associated with a different visualization as well. (discussed later in this document).
- **Error Handling** : If the user enters a genre or syllable that does belong to any genre in the database, it is taken to a page (with the help of render_template( ))where it is prompted to enter the correct genre.



Error Page for ‘Genre’

YEAR:

- The starting and end years are taken from a POST form method of two separate drop-down boxes.
- These two years along with the years between them are then stored in a list.
- Now, on the data retrieved from firebase, we pass the items in the newly appended list to ‘Year’.
- The results of the above operations are appended to an empty dataframe and the final dataframe result is displayed as an HTML table on the front end.
- Since we’re using Flask, we’ve used ‘render template()’ to return a different template when genre filtering is selected . This is because each option is associated with a different visualization as well. (discussed later in this document).
- **Error Handling** : If the user enters a starting year which is greater than the ending year, it is taken to a page (with the help of render_template())where it is prompted to select the correct range of years.



Error Page for ‘Year’

(iii) **Visual Representation:**

- Non-dynamic graphs and charts for sorting and filtering functions: (the above attached screenshots have the respective visualizations)
- On clicking of ASC/DES of **Name** button : Bar Chart of the top 20 selling games of the last 40 years, along with their respective publishers.
- On clicking of ASC/DES of **Publisher** button : Pie-chart showing publishers' contribution to the global sales of games in the last 40 years.
- On clicking of ASC/DES of **NA Sales** button : Double bar chart of the top 10 games with the highest sales in North America, and comparison of the NA Sale figures to the global sales of these games.
- On clicking of ASC/DES of **EU Sales** button : Double bar chart of the top 10 games with the highest sales in Europe, and comparison of the EU Sale figures to the global sales of these games.
- On clicking of **Year Filter** button : Line graph of global sales of all the top games over the entire year range from 1980 - 2020
- On clicking of **Genre Filter** button : Pie-chart showing genres' contribution to the global sales of games in the last 40 years.

Group Responsibility:

Pratishtha Singh:

- Loading of the data on Firebase
- Implementing filtering function on the dataset

Sharad Narayan Sharma:

- Implementation of Sorting function
- Creation of Visualization graphs.

NOTE : We've re-submitted the final presentation PPT as we made some changes to our final project.