ASSIGNMENT 1:

Video Game Sales

https://www.kaggle.com/datasets/archissave/video-game-sales?resource=download

This data sets contains the Name, system released on, year of release, publisher, genre, and the sales split across NA, EU, and JP.

I'm interested in this topic because I love games and I've been gaming since I was a kid. My goal with this data set is to map sales trends within series (Call of Duty, Battlefield, Super Mario, etc.) as well as overall yearly sales trends in order to see how consumer tastes might shift from year to year.

For example, did the sales of one game within a series increase because it was more well-received or because the general sales of the year increased.

I really like this data set since this kind of information is relevant to me too since I'm also a senior in the DSA major. I would be interested to see if you could focus on how salary rises between experience levels. I also like the idea of looking closer at salary differences between in-person and remote since remote has become a lot more common particularly for programming jobs since covid.

ASSIGNMENT 2:

\*\*\*DISCLAIMER: Outputs might look weird due to display constraints. Such as line breaks in the middle of a record.\*\*\*

\*\*Question 1\*\*

```

warda@LAPTOP-LI9T1JR8 MINGW64 ~

$ pwd

/c/Users/warda

```

\*\*Question 2\*\*

```

warda@LAPTOP-LI9T1JR8 MINGW64 ~

$ cd desktop/'Advanced Data Analytics'

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics

$ mkdir 'assignment 2 folder'

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics

$ ls

'assignment 2 folder'/ marks.txt vgsales.csv

```

\*\*Question 3\*\*

```

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics

$ cp vgsales.csv 'assignment 2 folder'

```

\*\*Question 4\*\*

```

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics

$ cd 'assignment 2 folder'

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ ls

vgsales.csv

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ pwd

/c/Users/warda/desktop/Advanced Data Analytics/assignment 2 folder

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ head vgsales.csv

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

5,Pokemon Red/Pokemon Blue,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 Mario Bros.,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 Mario Bros. Wii,Wii,2009,Platform,Nintendo,14.59,7.06,4.7,2.26,28.62

```

\*\*Question 5\*\*

```

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ head -1000 vgsales.csv > first1000vgsales.csv

```

One message not enough

\*\*Question 6\*\*

```

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ grep 'Call of Duty' vgsales.csv > filteredvgsales.csv

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ head filteredvgsales.csv

30,Call of Duty: Modern Warfare 3,X360,2011,Shooter,Activision,9.03,4.28,0.13,1.32,14.76

32,Call of Duty: Black Ops,X360,2010,Shooter,Activision,9.67,3.73,0.11,1.13,14.64

34,Call of Duty: Black Ops 3,PS4,2015,Shooter,Activision,5.77,5.81,0.35,2.31,14.24

35,Call of Duty: Black Ops II,PS3,2012,Shooter,Activision,4.99,5.88,0.65,2.52,14.03

36,Call of Duty: Black Ops II,X360,2012,Shooter,Activision,8.25,4.3,0.07,1.12,13.73

37,Call of Duty: Modern Warfare 2,X360,2009,Shooter,Activision,8.52,3.63,0.08,1.29,13.51

38,Call of Duty: Modern Warfare 3,PS3,2011,Shooter,Activision,5.54,5.82,0.49,1.62,13.46

41,Call of Duty: Black Ops,PS3,2010,Shooter,Activision,5.98,4.44,0.48,1.83,12.73

56,Call of Duty: Modern Warfare 2,PS3,2009,Shooter,Activision,4.99,3.69,0.38,1.63,10.69

62,Call of Duty: Ghosts,X360,2013,Shooter,Activision,6.72,2.63,0.04,0.82,10.21

```

I filitered my entire data set to just show Call of Duty which is a massive franchise and I might want to do some analysis with this series

\*\*Question 7\*\*

```

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ grep -E [0-9]{4} vgsales.csv | sed -e 's/201[0-9]/2010/g' -e 's/202[0-9]/2020/g' -e 's/199[0-9]/1990/g' -e 's/198[0-9]/1980/g' -e 's/200[0-9]/2000/g' > replacedvgsales.csv

warda@LAPTOP-LI9T1JR8 MINGW64 ~/desktop/Advanced Data Analytics/assignment 2 folder

$ head replacedvgsales.csv

1,Wii Sports,Wii,2000,Sports,Nintendo,41.49,29.02,3.77,8.46,82.74

2,Super Mario Bros.,NES,1980,Platform,Nintendo,29.08,3.58,6.81,0.77,40.24

3,Mario Kart Wii,Wii,2000,Racing,Nintendo,15.85,12.88,3.79,3.31,35.82

4,Wii Sports Resort,Wii,2000,Sports,Nintendo,15.75,11.01,3.28,2.96,33

5,Pokemon Red/Pokemon Blue,GB,1990,Role-Playing,Nintendo,11.27,8.89,10.22,1,31.37

6,Tetris,GB,1980,Puzzle,Nintendo,23.2,2.26,4.22,0.58,30.26

7,New Super Mario Bros.,DS,2000,Platform,Nintendo,11.38,9.23,6.5,2.9,30.01

8,Wii Play,Wii,2000,Misc,Nintendo,14.03,9.2,2.93,2.85,29.02

9,New Super Mario Bros. Wii,Wii,2000,Platform,Nintendo,14.59,7.06,4.7,2.26,28.62

10,Duck Hunt,NES,1980,Shooter,Nintendo,26.93,0.63,0.28,0.47,28.31

```

I grouped all entries by decade which can allow for a better visual on shifts in preference of games between decades. I chose decades because it prevents outliers where maybe one year released way more games in one genre/series and groups them into a larger time period.

\*\*Question for R\*\*

```

> setwd('C:/Users/warda/Desktop/Advanced Data Analytics/Assignment 2 folder')

> getwd()

> my\_data <- read.csv('filteredvgsales.csv', header = FALSE)

> head(my\_data)

V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11

1 30 Call of Duty: Modern Warfare 3 X360 2011 Shooter Activision 9.03 4.28 0.13 1.32 14.76

2 32 Call of Duty: Black Ops X360 2010 Shooter Activision 9.67 3.73 0.11 1.13 14.64

3 34 Call of Duty: Black Ops 3 PS4 2015 Shooter Activision 5.77 5.81 0.35 2.31 14.24

4 35 Call of Duty: Black Ops II PS3 2012 Shooter Activision 4.99 5.88 0.65 2.52 14.03

5 36 Call of Duty: Black Ops II X360 2012 Shooter Activision 8.25 4.30 0.07 1.12 13.73

6 37 Call of Duty: Modern Warfare 2 X360 2009 Shooter Activision 8.52 3.63 0.08 1.29 13.51

```

ASSIGNMENT 3:

Attached is my rpub link as well as my raw RMB file. I do most of my explanation within the document, so I will leave this mostly empty.

<http://rpubs.com/Tecnikle/1340228>

ASSGINMENT 4:

Question 1: Attached image showing DataLemur completion

Question 2 Loading: Attached image

Question 2 First Query:

```

duckdb> SELECT name, platform, year FROM vgsales

...> WHERE year BETWEEN '2000' AND '2010'

...> limit 5;

┌───────────────────────┬──────────┬──────┐

│ Name ┆ Platform ┆ Year │

╞═══════════════════════╪══════════╪══════╡

│ Wii Sports ┆ Wii ┆ 2006 │

│ Mario Kart Wii ┆ Wii ┆ 2008 │

│ Wii Sports Resort ┆ Wii ┆ 2009 │

│ New Super Mario Bros. ┆ DS ┆ 2006 │

│ Wii Play ┆ Wii ┆ 2006 │

└───────────────────────┴──────────┴──────┘

```

Question 2 Second Query:

```

duckdb> SELECT platform, round(sum(global\_sales),2) AS sales FROM vgsales

...> GROUP BY platform

...> ORDER BY sales DESC;

┌──────────┬─────────┐

│ Platform ┆ sales │

╞══════════╪═════════╡

│ PS2 ┆ 1255.64 │

│ X360 ┆ 979.96 │

│ PS3 ┆ 957.84 │

│ Wii ┆ 926.71 │

│ DS ┆ 822.49 │

│ PS ┆ 730.66 │

│ GBA ┆ 318.5 │

└──────────┴─────────┘

```

A screenshot of a black screen

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

Question 2 Third Query:

```

duckdb> SELECT name, platform, year FROM

...> (SELECT name, platform, year, global\_sales FROM vgsales

...> WHERE global\_sales > 10)

...> WHERE year BETWEEN '1900' AND '2000';

┌─────────────────────────────────────────┬──────────┬──────┐

│ Name ┆ Platform ┆ Year │

╞═════════════════════════════════════════╪══════════╪══════╡

│ Super Mario Bros. ┆ NES ┆ 1985 │

│ Pokemon Red/Pokemon Blue ┆ GB ┆ 1996 │

│ Tetris ┆ GB ┆ 1989 │

│ Duck Hunt ┆ NES ┆ 1984 │

│ Pokemon Gold/Pokemon Silver ┆ GB ┆ 1999 │

│ Super Mario World ┆ SNES ┆ 1990 │

│ Super Mario Land ┆ GB ┆ 1989 │

│ Super Mario Bros. 3 ┆ NES ┆ 1988 │

│ Pokémon Yellow: Special Pikachu Edition ┆ GB ┆ 1998 │

│ Super Mario 64 ┆ N64 ┆ 1996 │

│ Super Mario Land 2: 6 Golden Coins ┆ GB ┆ 1992 │

│ Gran Turismo ┆ PS ┆ 1997 │

│ Super Mario All-Stars ┆ SNES ┆ 1993 │

└─────────────────────────────────────────┴──────────┴──────┘

```

Question 3:

I honestly prefer using R because it is the most similar to coding I'm comfortable with. Command line had a very large learning curve for me cause it felt like nothing I knew really transfered and SQL feels too rigid and structured to handle every potential niche case I may encounter. R allows me to handle niche cases a little bit more flexibly compared to SQL but also not being as confusing as command line

ASSIGNMENT 5:

Part 1 Query 1:

```

https://pokeapi.co/api/v2/pokemon/whimsicott

```

https://pokeapi.co

This query uses the pokemon api provided on the public apis link to call all the information regarding whimsicott which is my favorite pokemon

Part 1 Query 2:

```

https://thereportoftheweekapi.com/api/v1/reports/?min\_rating=7

```

https://www.thereportoftheweek.com

This query searches the report of the week api to find all reports he posted with a minimum rating of 7.

Part 1 Query 3:

```

https://api.chess.com/pub/leaderboards

```

https://www.chess.com/home

This query returns the top 50 leaderboard players on the website chess.com

Part 2:

Github Repo Link: https://github.com/Tecnikle/Adv-Data-Video-Game-Sales-Analysis

Github pages Link: <https://tecnikle.github.io/Adv-Data-Video-Game-Sales-Analysis>

Part 3 Query 1:

```

Brandon@DESKTOP-KSDFHRO MINGW64 ~

$ curl 'https://tecnikle.github.io/Adv-Data-Video-Game-Sales-Analysis/vgsales.json' | jq '.[0:10]'

```

This returns the first 10 items in my json

Part 3 Query 2:

```

Brandon@DESKTOP-KSDFHRO MINGW64 ~

$ curl 'https://tecnikle.github.io/Adv-Data-Video-Game-Sales-Analysis/vgsales.json' | jq '.[] | select(.Global\_Sales >= 20)'

```

This returns all items that had a global sales above 20 million

Part 3 Query 3:

```

Brandon@DESKTOP-KSDFHRO MINGW64 ~

$ curl 'https://tecnikle.github.io/Adv-Data-Video-Game-Sales-Analysis/vgsales.json' | jq '.[] | select(.Publisher == "Activision")' | jq -s '.[0:5]'

```

This returns the first 5 items in the JSON that were published by Activision

A screenshot of a computer screen

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

ASSIGNMENT 6:

https://public.tableau.com/app/profile/brandon.faulkenberry/viz/VGsalesDashboard/Sheet1?publish=yes

https://public.tableau.com/app/profile/brandon.faulkenberry/viz/VGsalesDashboard/Sheet2

My first Worksheet is a pie chart with an interactive year filter. What it shows is the top 10 publishers and how big of a share of the pie they hold. We can learn how large a publisher is and kind of find when they fall out of the top 10 as well as who takes their place. It can also help to show dominance, for example nintendo almost always holds a majority of the pie.

My second worksheet is a line graph with no filter. It shows the top 10 publishers of all time and their sales through the years. We can learn what years publishers published a game that had a large amount of sales.