

## Cleaning in Excel

- First task was to clean Genres that had *N/A* values. Compiled list of games that needed a genre then added genre into sheets.
- Once added, perform a **VLOOKUP** on the **Final Games** sheet against the **Games Missing Genre** sheet to update games with *Genre = N/A*.
- Once this was completed performed **Split Text to Columns** feature to split games with multiple Genres into their own column

## Cleaning in SQL

- Added Columns to separate the **Release Date** column. This will create a column for the year, month & date.
  - -- add new columns for year dates
  - ALTER TABLE ['List-of-best-selling-videogames\$']
  - ADD year\_release int
  - 
  - -- add new columns for moth & day dates
  - ALTER TABLE ['List-of-best-selling-videogames\$']
  - ADD month\_release int, day\_release int
  - 
  - -- add data into year new columns
  - UPDATE ['List-of-best-selling-videogames\$']
  - set year\_release = YEAR([Release date])
  - 
  - -- add data into month new columns
  - UPDATE ['List-of-best-selling-videogames\$']
  - set month\_release = MONTH([Release date])
  - 
  - -- add data into month new columns
  - UPDATE ['List-of-best-selling-videogames\$']
  - set day\_release = DAY([Release date])
  -
- Updated the table to replace commas with spaces or / in order to avoid import issues with CSV files.
  - --Replace comma with space for publisher
  - UPDATE ['List-of-best-selling-videogames\$']
  - set Publisher = REPLACE(Publisher,',',' ')
  -

- --Replace comma with space for Developer
- UPDATE ['List-of-best-selling-videogames\$']
- set Developer = REPLACE(Developer,',','/')
  -
- --Replace comma with space for game
- UPDATE ['List-of-best-selling-videogames\$']
- set Game = REPLACE(Game,',',' ')
  -
- --Replace comma with space for genre
- UPDATE ['List-of-best-selling-videogames\$']
- set Genre = REPLACE(Genre,',',' ')
  -
- Added the value N/A for null values in the **Genre** column
  - -- Add a N/A for null genere
  - UPDATE ['List-of-best-selling-videogames\$']
  - set Genre = 'N/A'
  - WHERE
  - Genre IS NULL
- Also Find the N/A for **Genre** column
  - SELECT
  - Game,
  - Genre
  - FROM
  - ['List-of-best-selling-videogames\$']
  - WHERE
  - Genre = 'N/A'
  -