Video game sales data cleaning report

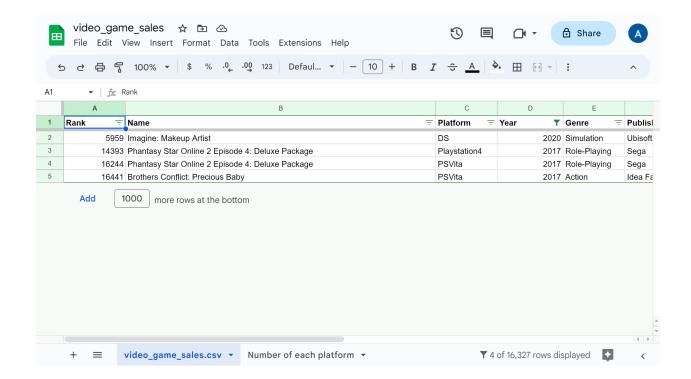
Platform column

Some of the platform names were vague so I changed them be more recognizable (Ctrl+H, find and replace):

- 2600 -> Atari2600
- DC -> Dreamcast
- GC -> Gamecube
- GB -> Gameboy
- GBA -> GameboyAdvance
- GG -> GameGear
- GEN -> Genesis
- NG -> NeoGeo
- PS -> Playstation
- PS2 -> Playstation2
- PS3 -> Playstation3
- PS4 -> Playstation4
- PSV -> PsVita
- SAT -> Saturn
- SCD -> SegaCD
- TG16 -> TurboGrafx16
- WS -> WonderSwan
- XB -> Xbox
- X360 -> Xbox360
- XOne -> XboxOne

Year Column

- When sorting the year column I discovered 271 N/A values. Since this only accounted for 1% of the dataset, I deleted those rows.
- The dataset only featured games released up to 2016, so I manually changed the 4 values that were past 2016 after looking up the correct release date.



Publisher column

There were 35 N/A values. The publishers either couldn't be found or were the sole publisher in the dataset so it was deemed insignificant so all 35 rows were dropped.(Conditional formatting > Format Rules > Text is exactly N/A)

General

- Data > Data cleanup > Trim whitespace
- Changed all the sales columns names to "_Copies(millions)" for clarity