

A dark blue vertical bar on the left side of the slide. A blue arrow points to the right from this bar, containing the date.

2-11-2022

A4 - EDA basat en Arbre de decisió

M05 – Projectes Big Data

Contenido

Videogame Sales Dataset	2
Per què?	2
Hipòtesi	2
Resolució	2
Primera hipòtesi	2
Segona hipòtesi	3
Tercera hipòtesi	3
Arbre de decisió	4
Millor decisió	4
Segona decisió	4
Tercera decisió	4

Videogame Sales Dataset

El Dataset escollit forma una recopilació de les ventes de videojocs a nivell mundial des de l'any 1980 fins al 2016 ordenat segons el nombre de ventes, consta dels següents camps:

Rànquing, Nom, Any, Gènere, Editor, Nombre de ventes a Amèrica, Europa, Japó, Altres i Globals.

Per què?

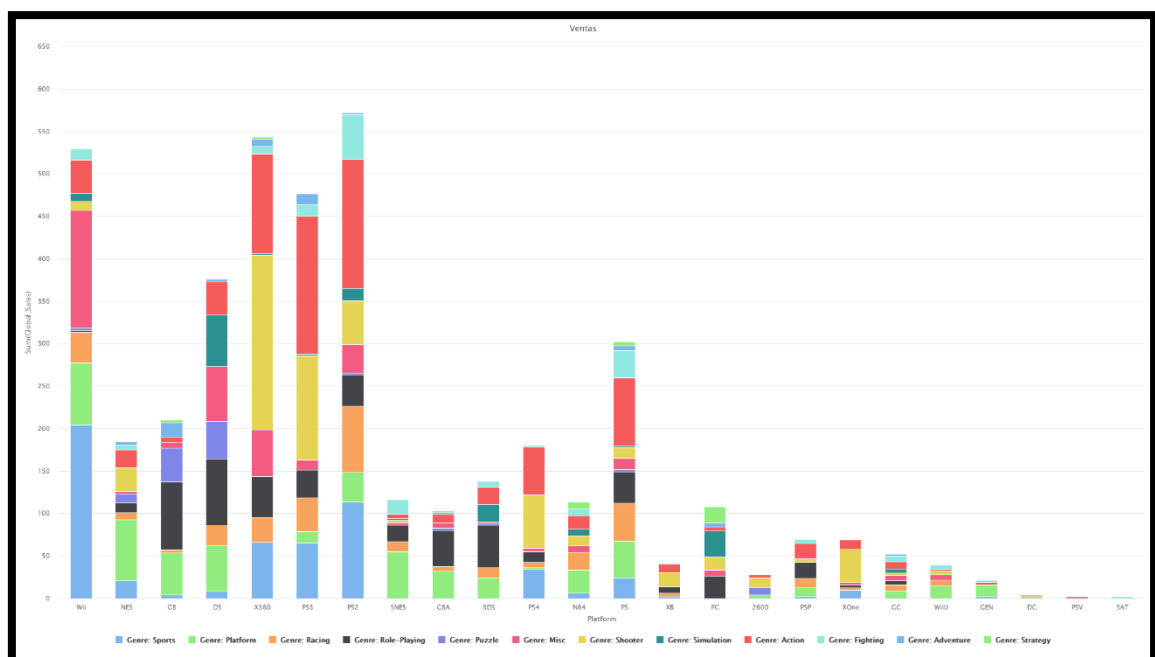
He escollit aquest Dataset ja que es un camp que m'interessa i domino, i crec que està ben estructurat per a poder fer un bon anàlisi de dades i respondre les preguntes que tinc a continuació.

Hipòtesi

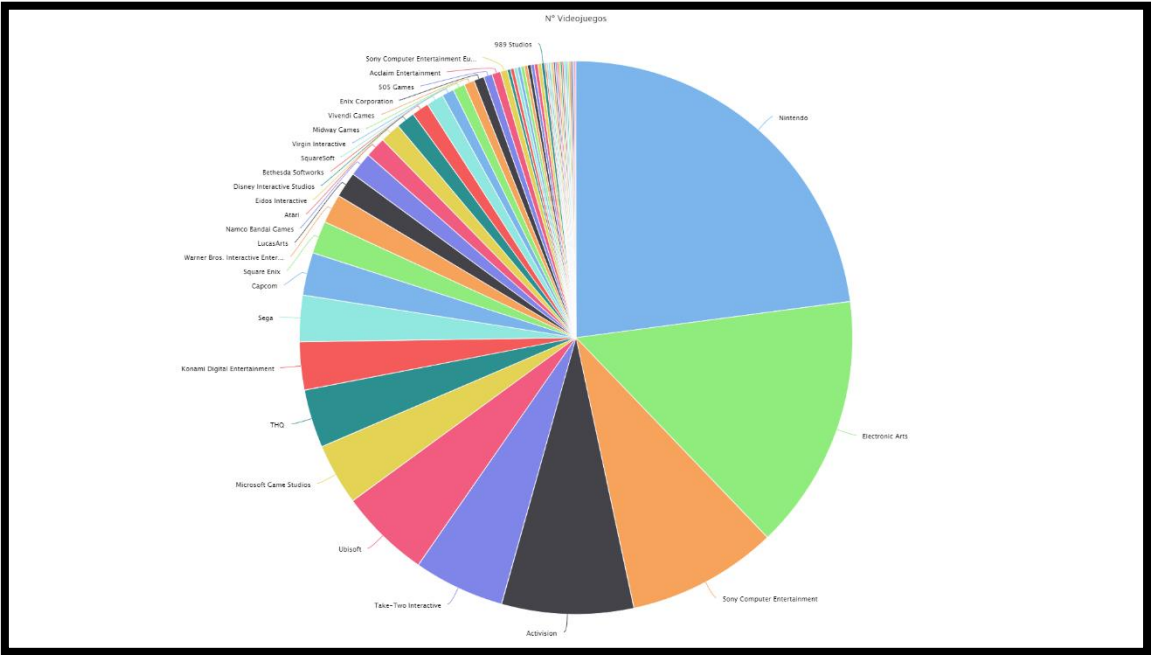
1. Quina consola té el major nombre de ventes i quin gènere tenen?
2. Quin editor ha venut més videojocs?
3. Quina és la progressió de ventes a les diferents regions?

Resolució

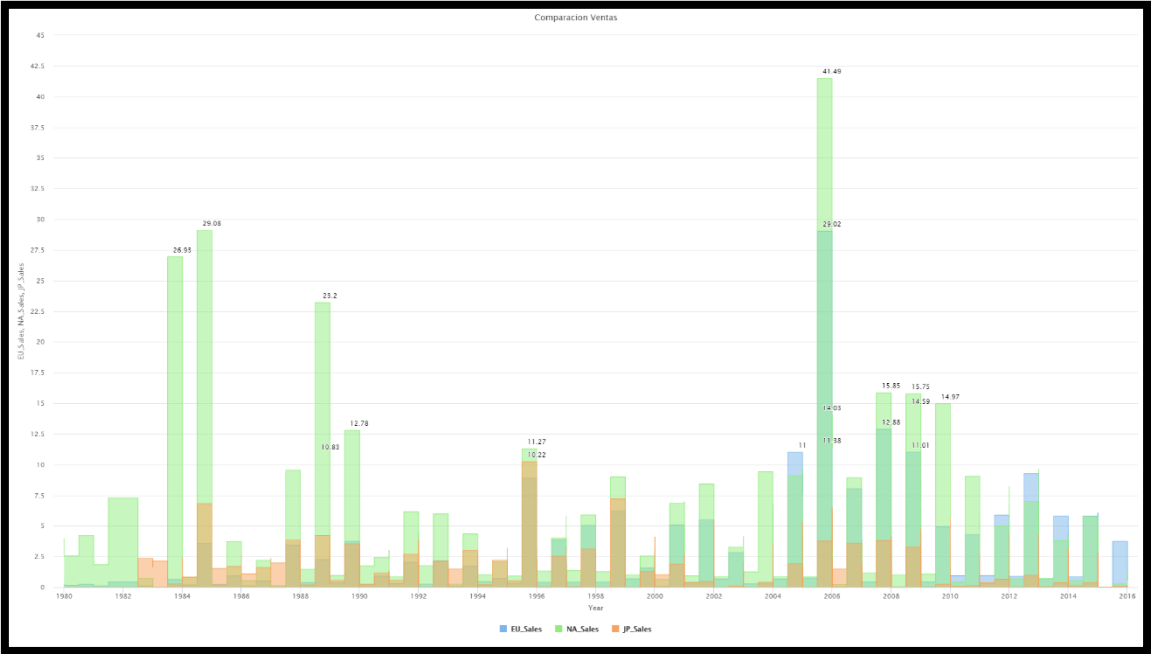
Primera hipòtesi



Segona hipòtesi



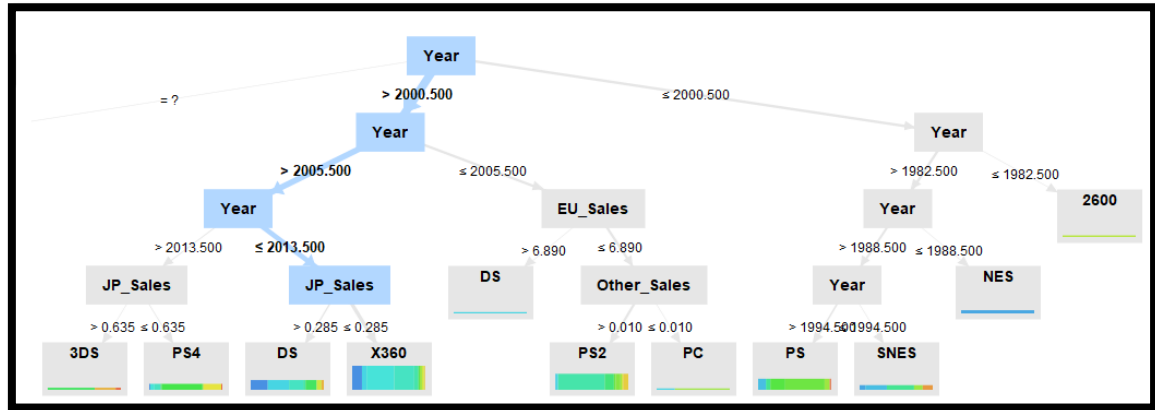
Tercera hipòtesi



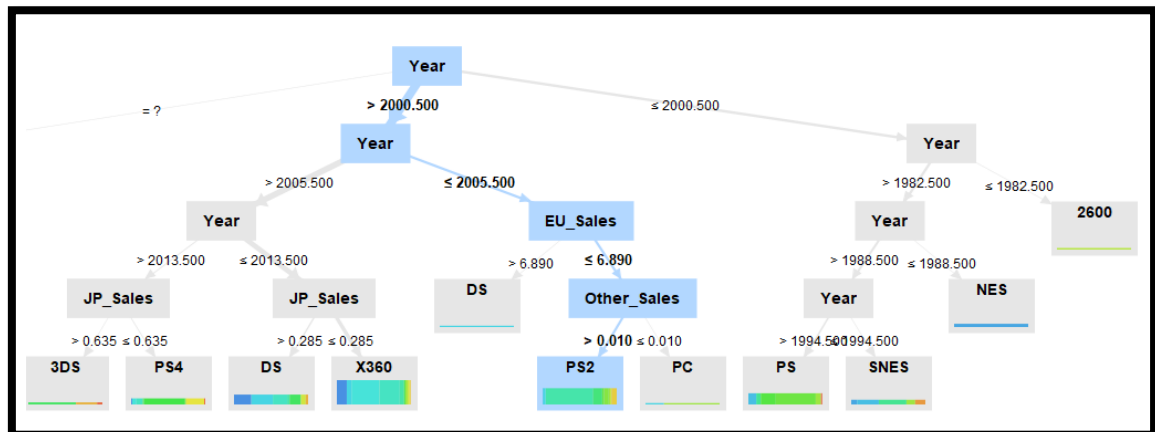
Arbre de decisió

Per a poder fer una bona predicció de ventes segons la plataforma, ens haurem de fixar sobretot en l'any, ja que segons aquest, les ventes a una plataforma poden ser desastroses degut al poc ús d'aquesta, i també la regió on es produeixin aquestes ventes.

Millor decisió



Segona decisió



Tercera decisió

