



# VENTAS DE VIDEOJUEGOS

# UN BREVE REPASO SOBRE NUESTRO DATASET

- El dataset a usar fue obtenido de la página VGChartz.com mediante un proceso de raspado web o “web scraping”.
- El conjunto de datos representan ventas.
- Los juegos móviles e independientes no se incluyen.
- Los juegos que tienen descargas gratuitas no cuentan como ventas.
- Las ventas son divididas por regiones: norteamérica, europa, japón, otra regiones y global.
- Estos datos son acumulados hasta 2016, por lo que los juegos más nuevos tienen menos ventas.





# OBJETIVOS DE ANÁLISIS PLANTEADOS

01

**Predicción:** ¿Cuáles serán las ventas globales de un videojuego específico en función de sus características, como el nombre, la plataforma, el año de lanzamiento, el género y el editor?

02

**Agrupación:** A partir de un algoritmo de agrupación, se necesita generar n cantidad de grupos según las características del dataset para luego caracterizar cada grupo en base a sus atributos.



# Objetivo de la presentación



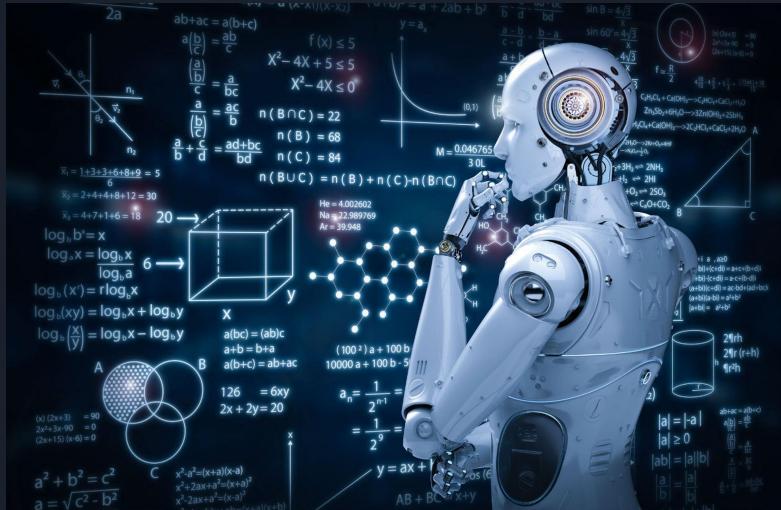
El desarrollo de la presentación se centrará en proporcionar una visión general y accesible sobre cómo abordamos los objetivos planteados, sin adentrarnos en detalles técnicos. Presentaremos cómo utilizamos algoritmos de regresión para predecir las ventas y algoritmos de agrupación que nos permitirán comprender mejor la diversidad de juegos y adaptar estrategias de marketing en consecuencia.



# 01: REGRESIÓN

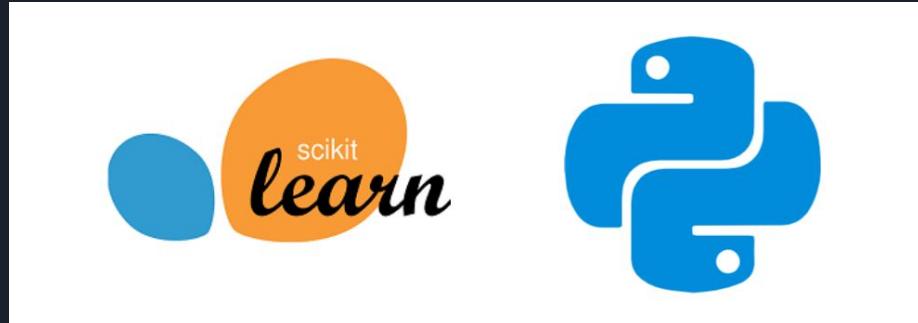
# PASO 1: ENTRENAMIENTO DE MODELOS

- Random Forest
  - Regresión lineal
  - Regresión Ridge
  - SVR.

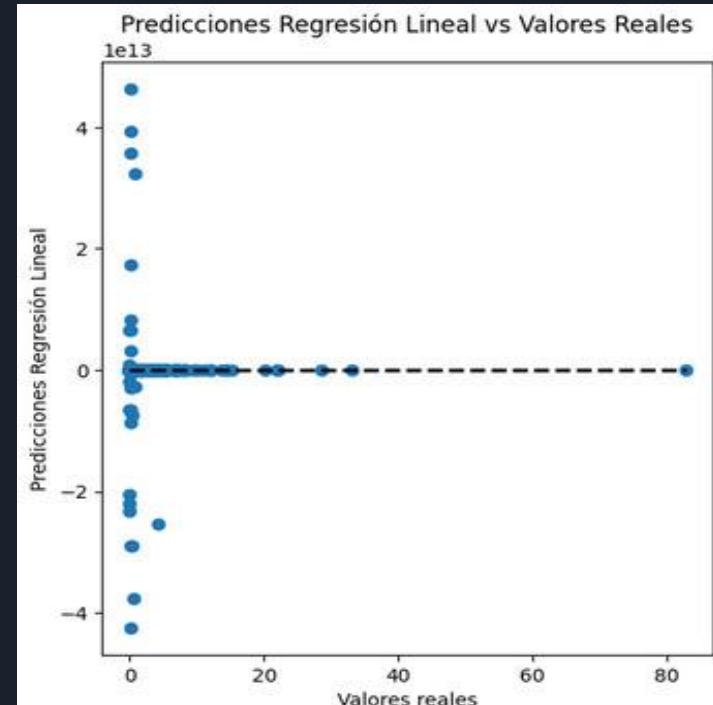
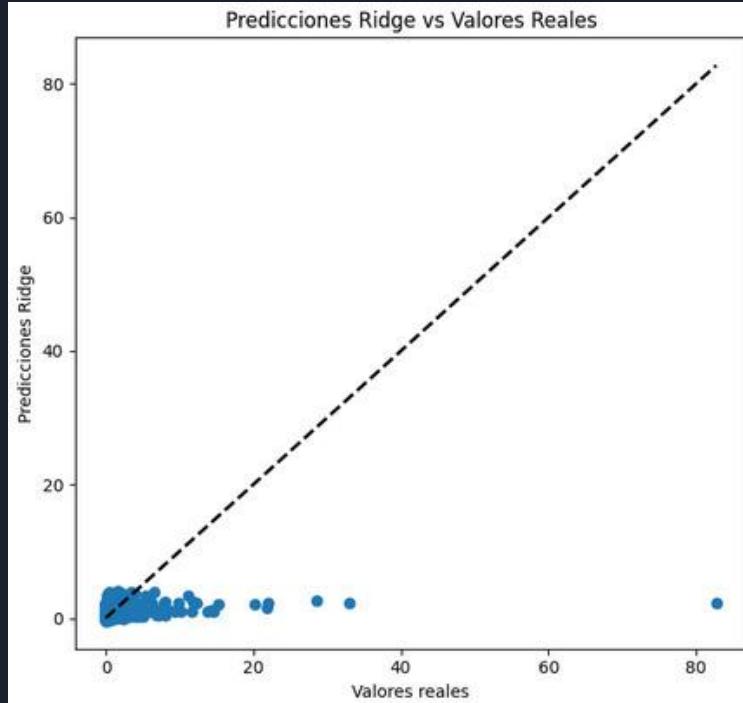


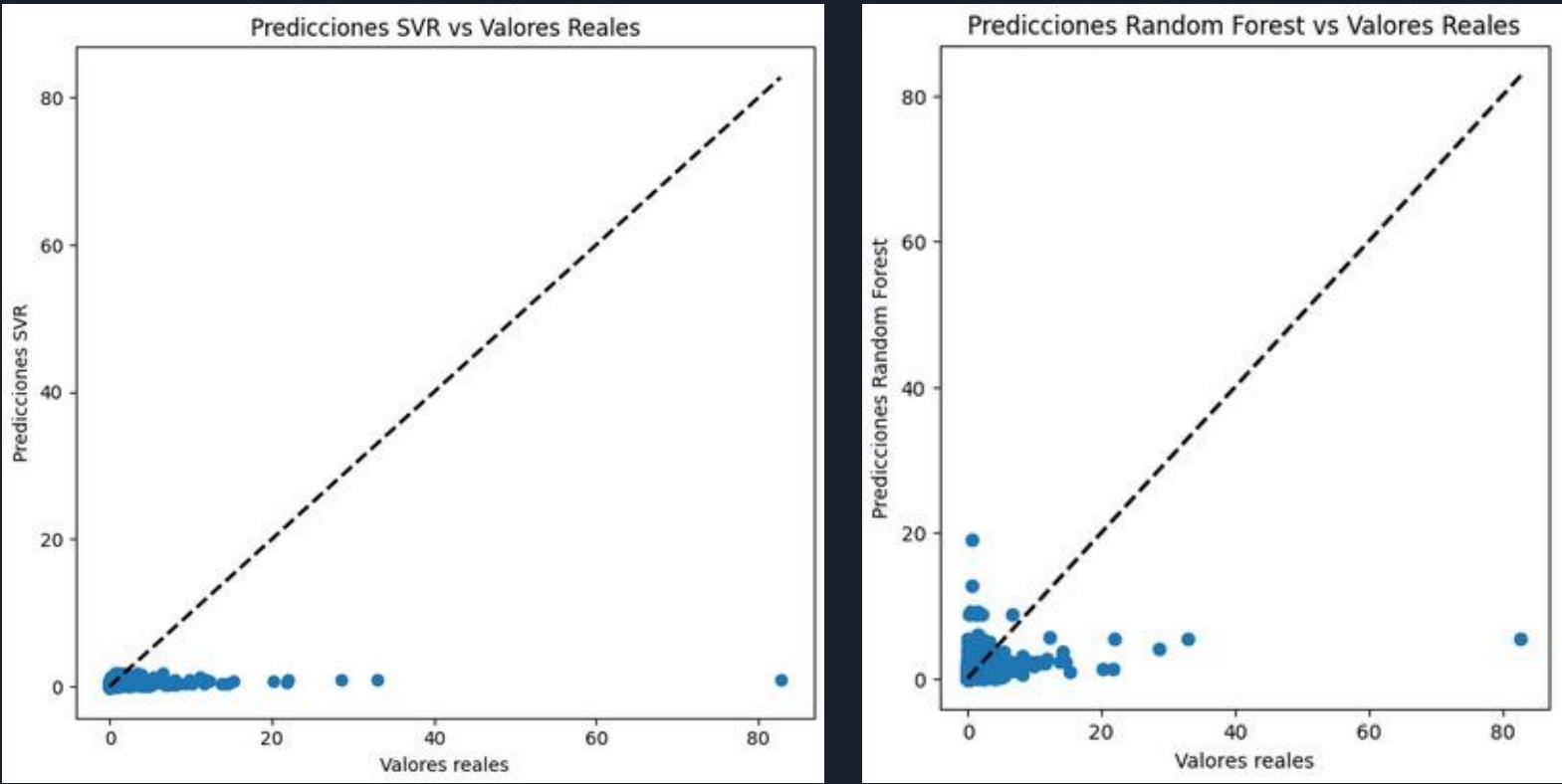
## PASO 2: EVALUACIÓN Y ERROR

- Rendimiento de cada uno comparando sus errores
- Gráfica comparativa permite visualizar sus rendimientos y compararlos.



# COMPARACIÓN DE ERRORES

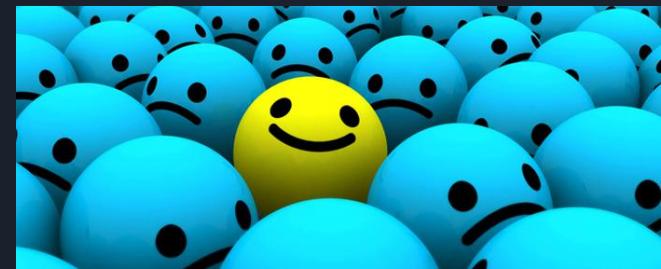




# CONCLUSIONES SOBRE ERRORES

- El mejor algoritmo para predecir es el de Ridge debido a que su error ha demostrado ser bajo lo cual sugiere que es capaz de generalizar “bien” a datos nuevos.
- La Máquina de Soporte Vectorial para Regresión es un poco mayor que el error de Ridge, aún proporciona predicciones bastante precisas.
- El error de Regresión Lineal es extremadamente alto, lo que sugiere que el modelo puede no ser adecuado para nuestro conjunto de datos.
- El modelo de Bosques Aleatorios tiene un error cercano al de Ridge muestra una capacidad sólida para predecir las ventas globales con precisión

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}}$$



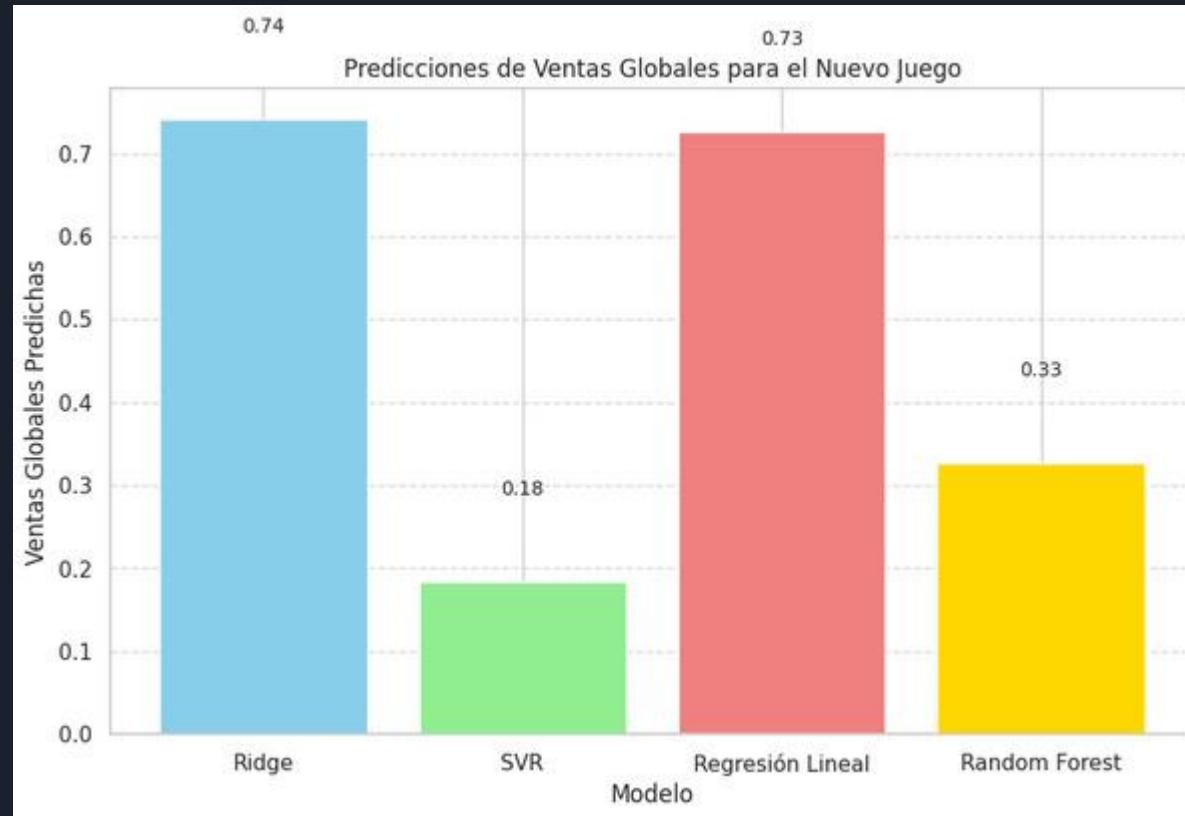


# PASO 3: GENERALIZAR EN DATOS NUEVOS

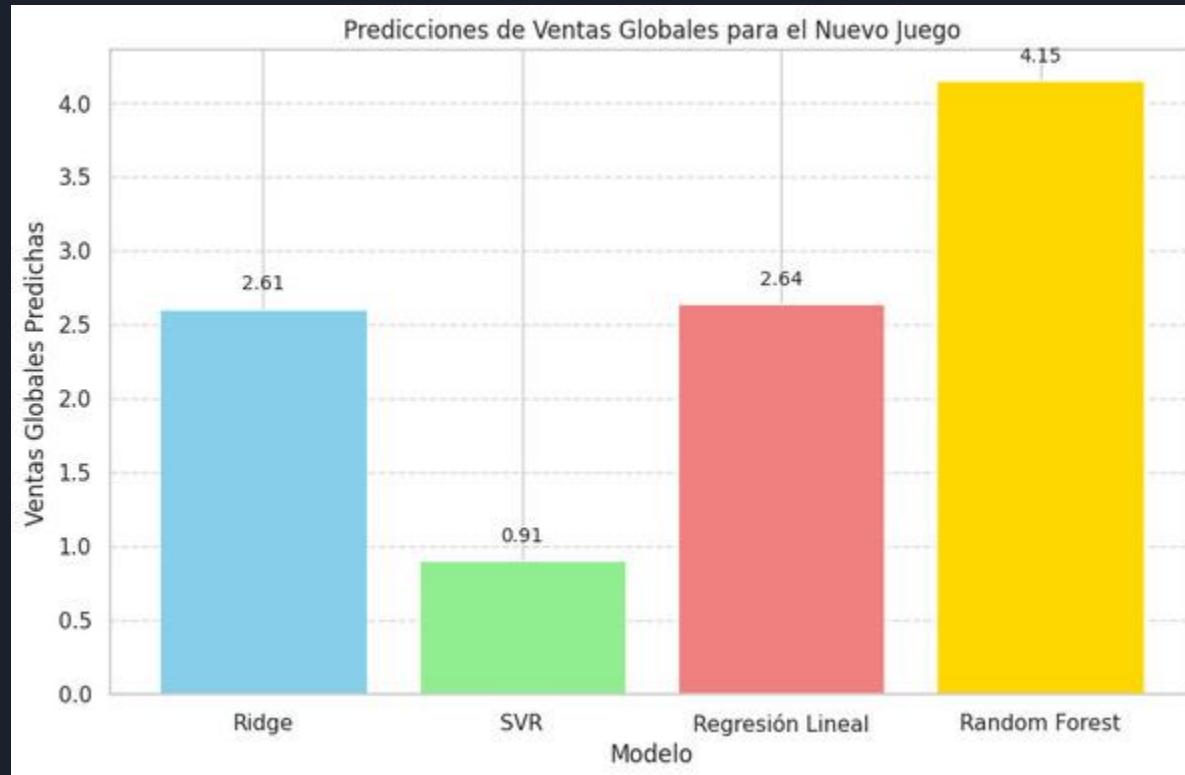
Se realizaron algunas predicciones utilizando los algoritmos para determinar las ventas que podrían tener futuros juegos entorno a sus características

- Predicción 1: ¿Cuáles serán las ventas para un juego de la empresa “Take Two interactive” de género “Action” para la plataforma PC ?
- Predicción 2: ¿Cuáles serán las ventas para un juego de la empresa “Nintendo” de género “Platform” para la consola Wii ?
- Predicción 3: ¿Cuáles serán las ventas para un juego de la empresa “Ubisoft” de género “Sports” para la plataforma PC?

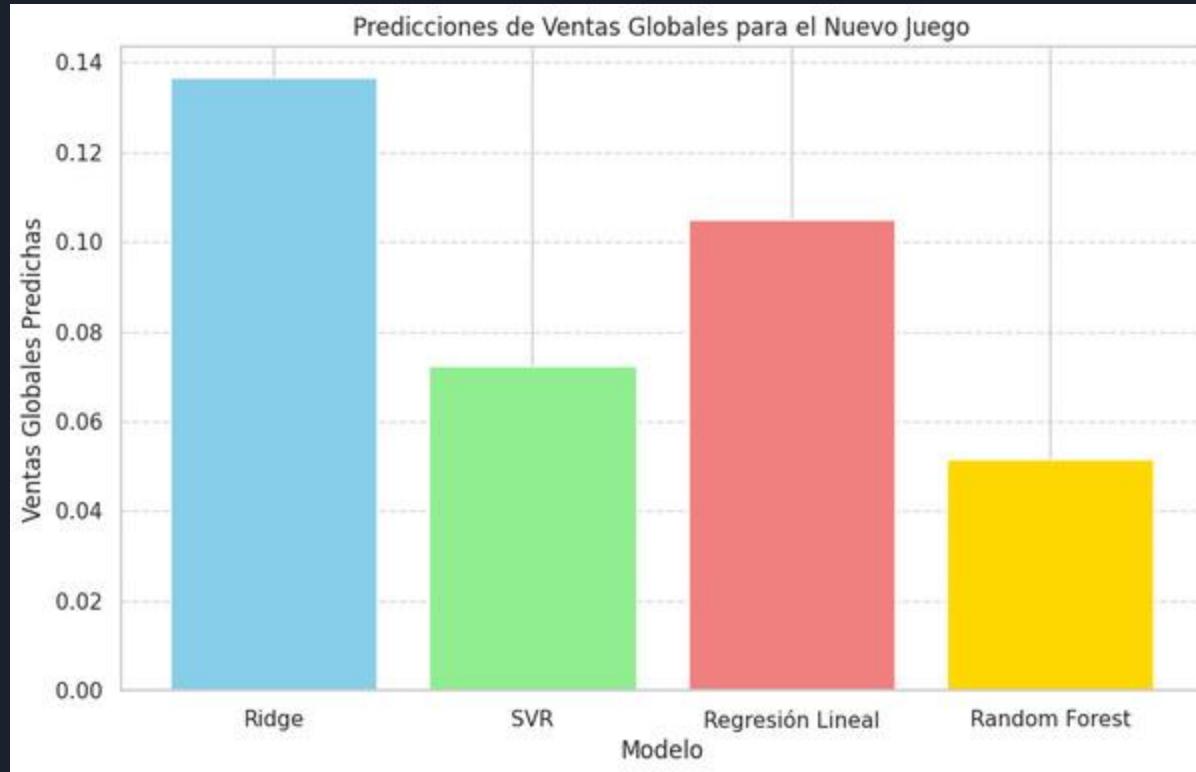
Resultado primera predicción:



## Resultado segunda predicción:

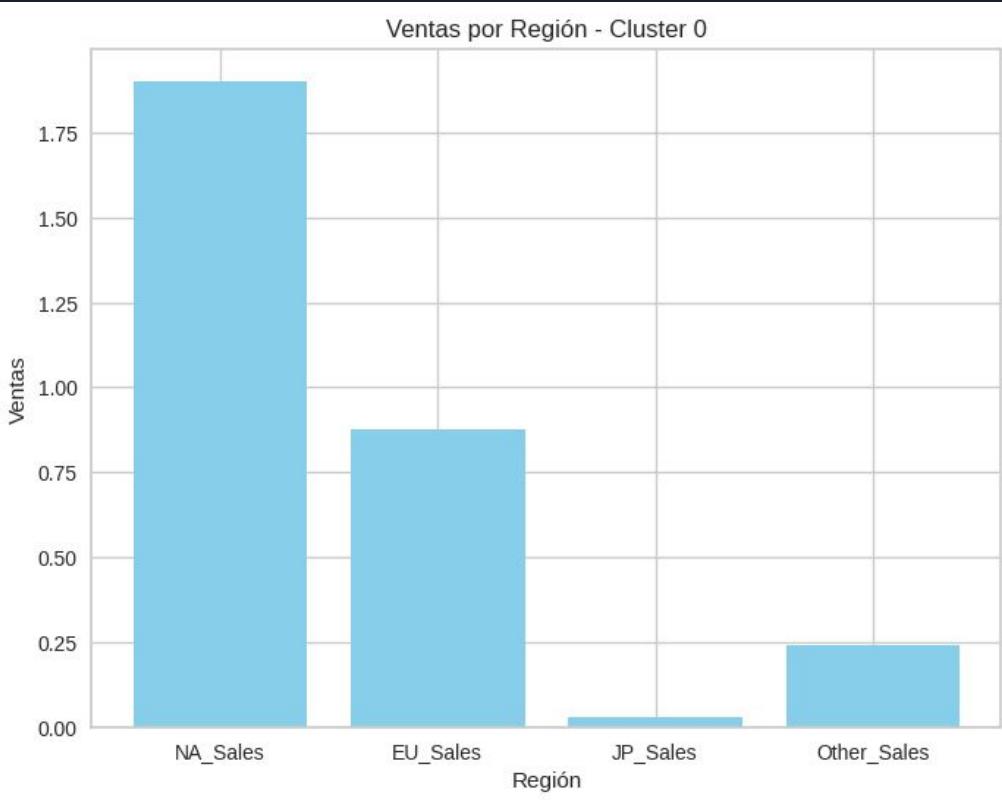


## Resultado tercera predicción:

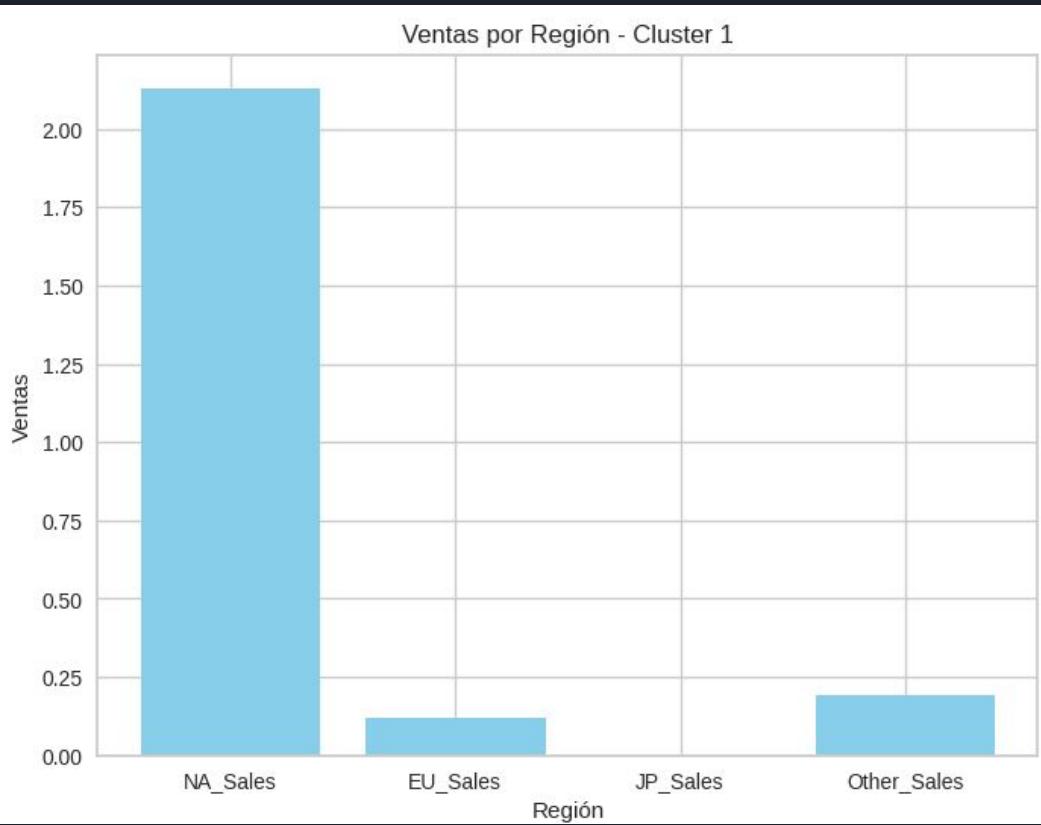




## 02: AGRUPACIÓN

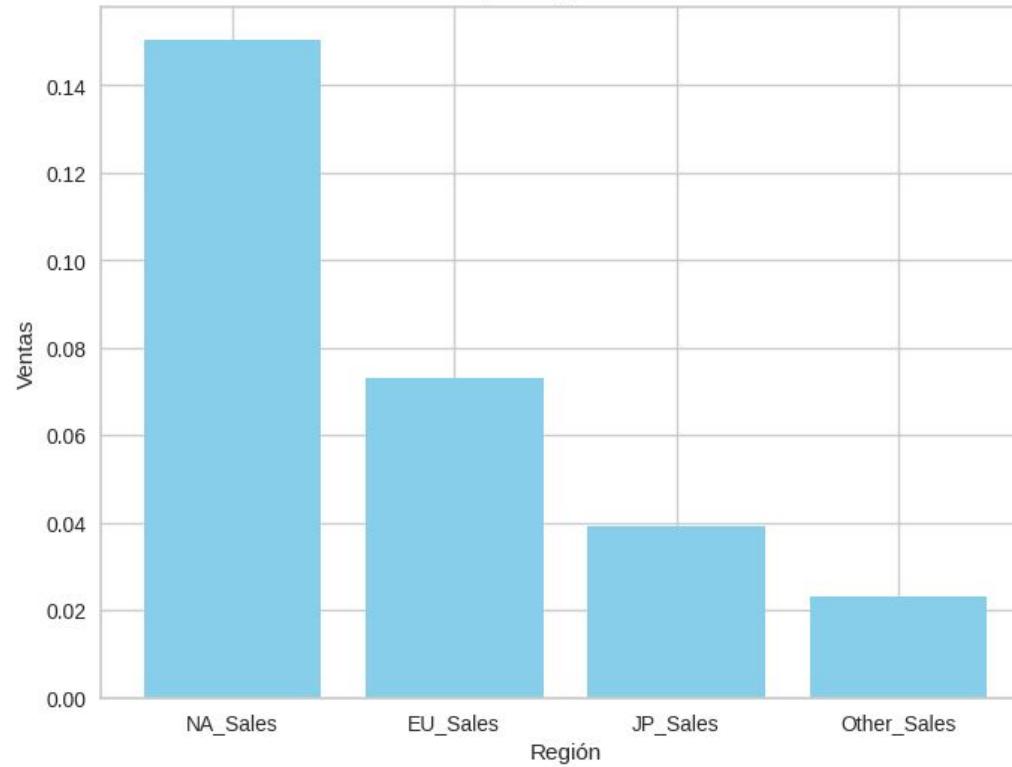


Género: Shooter

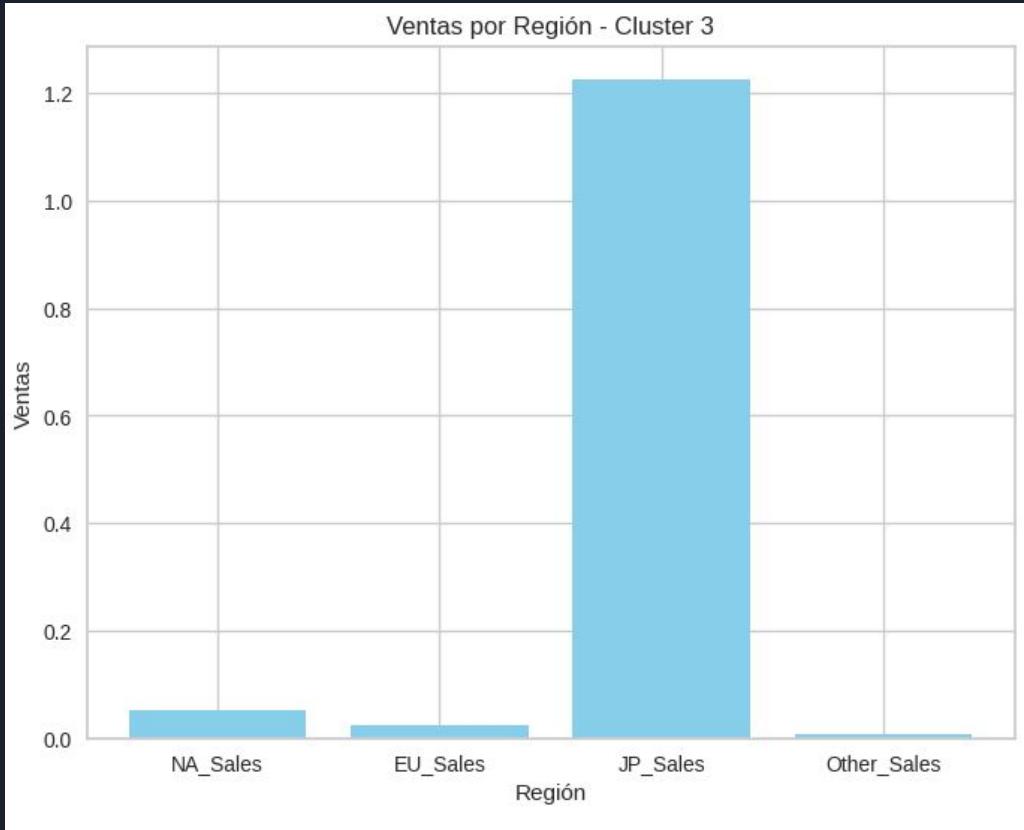


Género: Sports

Ventas por Región - Cluster 2



Género: Puzzle



Género: Role-Playing

Gracias

