

# BIG DATA

PROYECTO FINAL

Jonathan Perez Sedova



# Informe para el medio de comunicación – Proyecto de Business Intelligence

## Introducción

En respuesta a la solicitud del medio de comunicación, hemos desarrollado un conjunto de informes interactivos en PowerBI que analizan la evolución global de la pandemia de COVID-19 y su relación con el rendimiento comercial de nuestra empresa. Los datos sanitarios proceden de OurWorldinData (archivo cases\_deaths.csv), mientras que la información de ventas y clientes se ha obtenido del dataset "Proyecto 3: Segmentación de Clientes en ecommerce" disponible en Kaggle. Tras limpiar y combinar ambas fuentes, se han construido cuatro páginas de análisis que presentamos a continuación junto con las principales conclusiones.

## Resumen mundial de la pandemia

La primera página muestra indicadores clave de la pandemia: **103 millones de casos acumulados y 1 millón de muertes** a nivel global. El gráfico temporal de **casos nuevos y muertes nuevas** evidencia un pico pronunciado a comienzos de 2021, seguido de sucesivas oleadas que van disminuyendo hacia finales de 2021. El mapa de burbujas destaca a países como **Estados Unidos, China e India** como los más afectados, y una gráfica de barras sitúa a estos tres junto a **Francia, Alemania, Brasil y Japón** entre los diez primeros en número de casos totales.

En conjunto, los datos muestran la magnitud de la pandemia y ponen de relieve que la carga de casos se concentra en unas pocas naciones de gran población y actividad económica.



## Análisis de ventas

La segunda página se centra en el rendimiento comercial. Aquí se visualizan **ventas totales de 9,75 millones de euros** y un **ticket medio de 1,88 euros** por factura en el periodo comprendido entre el 1º de enero de 2020 y el 31 de diciembre de 2021. El gráfico de líneas compara **ventas diarias** con **ventas acumuladas**, evidenciando un crecimiento sostenido a lo largo del periodo con picos puntuales en septiembre y diciembre de 2021.

El análisis mensual revela que **septiembre, octubre y diciembre** son los meses con mayor facturación, mientras que el desglose por países muestra una concentración abrumadora en el **Reino Unido**: este mercado genera la mayor parte de las ventas, muy por encima de **Países Bajos, EIRE, Alemania y Francia**. Esta concentración puede atribuirse a que la empresa tiene su base operativa en el Reino Unido, lo cual se refleja en la base de clientes y en la estrategia comercial.

## Análisis de clientes

La tercera página estudia la base de clientes. Durante el periodo analizado la empresa atendió a **4 373 clientes únicos** y vendió **5 millones de unidades**. El **ticket medio mensual** oscila entre 1,6 y 2,0 euros, con un máximo en junio de 2021. Un gráfico de barras identifica a los **diez clientes con mayor volumen de ventas**, destacando que unos pocos ID concentran una proporción significativa de la facturación. Además, se observa que **el 90 % de las unidades se venden en el Reino Unido**, mientras que el resto de países aportan volúmenes marginales.

También se incluyó un mapa para visualizar el número de clientes por país: la mayor densidad se sitúa en **Europa occidental**, con pocos clientes en otros continentes. Estos hallazgos sugieren que la empresa debería reforzar la fidelización de sus grandes clientes y valorar la expansión a nuevos mercados



## Relación entre COVID-19 y ventas

La cuarta página se centra en cómo se mueven las ventas cuando cambian los **casos nuevos** en los países donde opera la empresa. Para resumir la relación, se muestran dos indicadores: **Ratio Ventas/Casos (0,16)** y **Ventas por 1.000 casos (159,50)**. Estos valores deben interpretarse como una referencia comparativa dentro del periodo y la geografía analizada, ya que dependen del nivel de casos y de dónde se concentran las ventas.

En el gráfico **Casos nuevos vs Ventas totales** se aprecia un patrón visual bastante claro en varios tramos del año. En el **primer tramo (enero-febrero)**, cuando se observan picos frecuentes de casos nuevos, las ventas tienden a **frenarse** y muestran oscilaciones sin un crecimiento sostenido. En el **segundo tramo (abril-mayo)** este efecto se ve aún más: los casos nuevos se mantienen altos durante varias semanas y las ventas quedan **casi planas**, con menos repuntes, lo que sugiere un periodo de menor actividad comercial.

Después, en el **tercer tramo (junio-julio)**, cuando los casos nuevos bajan de forma más marcada, las ventas dejan de caer y pasan a una fase de **estabilización**. En el **cuarto tramo (septiembre-noviembre)**, con niveles de casos más controlados y sin picos tan constantes, las ventas muestran un **crecimiento más claro** y se observan máximos de la serie. Finalmente, en el **quinto tramo (final de noviembre-diciembre)** vuelve a aparecer una subida fuerte de casos; en ese momento las ventas mantienen un nivel alto, pero se nota una **ligera pérdida de impulso** (más ruido y pequeños retrocesos) respecto al tramo anterior.

La comparación de variaciones mensuales refuerza esta lectura: los cambios en ventas no siguen exactamente el calendario de máximos de contagios, pero sí se aprecia que los periodos con picos de casos tienden a asociarse con ventas más “contenidas”, mientras que los periodos de menor presión sanitaria permiten una recuperación gradual.

## Conclusiones generales

En conjunto, los informes muestran que la evolución de ventas y clientes tiene un componente propio del negocio, pero en la página de relación se observa una señal coherente: cuando aumentan los casos nuevos durante varias semanas, las ventas tienden a frenarse, y cuando los casos bajan o se estabilizan, las ventas se recuperan. Por tanto, la relación parece más compatible con un efecto indirecto (incertidumbre, restricciones, logística y hábitos de compra) que con una coincidencia aleatoria, aunque con estos datos no se puede afirmar causalidad.