Propuesta de proyecto

**Engagement Inteligente: Optimización de Contenidos en Redes Sociales**

Integrantes del equipo:

1. *Barbara Sanchez – Data Analyst BI*
2. *Ducelvis Arcia – Data Analyst BI*

Caso de negocio

| **Industria y contexto del caso de negocio a desarrollar:** | *Proyecto de análisis en marketing digital para mejorar el rendimiento de publicaciones en redes sociales.* |
| --- | --- |
| **Descripción del Problema:** | Falta de claridad sobre qué factores impulsan el engagement y variabilidad en los resultados de los posteos. El análisis se realizará con herramientas profesionales en la nube (Azure SQL + Power BI), para replicar un entorno corporativo real. |
| **Objetivo principal del proyecto:** | Identificar patrones de alto rendimiento y predecir resultados de futuras publicaciones. |
| **Resultados esperados & Impacto:** | * Mejoras en KPIs de engagement. * Dashboard para toma de decisiones. * Modelo predictivo de rendimiento. |

Fuentes de datos

| **Fuentes de Datos:** | *Se trabajará con el dataset “Social Media Engagement Dataset", obtenido desde:*  [*https://www.kaggle.com/datasets/subashmaster0411/social-media-engagement-dataset*](https://www.kaggle.com/datasets/subashmaster0411/social-media-engagement-dataset)  *Este dataset tiene información sobre:*  *fecha\_posteo, hora\_posteo, tipo\_contenido, formato, likes, comentarios, shares, impresiones* |
| --- | --- |
| **Descripción de los Datos:** | 1. *Archivo CSV. “Engagement.csv”:*  * *Total de 28 Columnas* * *Registros \** * *5.18 MB*   *El dataset se descargó y se trabajó en Python para limpieza inicial ETL y EDA.*  *Posteriormente, fue cargado a una base de datos en la nube (Azure SQL), donde se integró con una tabla calendario generada desde Python.*  *Finalmente, los datos fueron utilizados en Power BI para crear un Dashboard interactivo, con visualizaciones clave para toma de decisiones.* |

Plan o estrategia de análisis

| **Requerimientos del caso:** | *Como analistas, se requiere identificar los factores que generan mayor engagement, con base en la estructura del dataset. Se aplicarán métricas de interacción y modelos de predicción para anticipar el rendimiento de futuras publicaciones. Esto permitirá sugerir estrategias óptimas de contenido.*  *Para ello es necesario abordar el siguiente plan de acción*   1. *Limpieza y estructura del dataset(EDA)* 2. *Análisis exploratorio de datos* 3. *Análisis de métricas clave( cálculo de engagement rate, interacciones por impresión, y rendimiento por categoría).* 4. *Los datos se cargaron en una base de datos en la nube (Azure SQL) y se trabajaron con consultas SQL en Azure Data Studio, replicando un entorno de trabajo real. También se integró una tabla calendario desde Python para facilitar análisis temporales.Consultas Avanzadas y Vistas necesario de la base de datos* 5. *Definición de KPIs y evaluación continua (establecimiento de métricas de éxito y monitoreo con nuevos datos para ajustar estrategias).* 6. *Visualización de datos* |
| --- | --- |
| **Tareas y tecnologías** | 1. *Planificación y Gestión del Proyecto:*     1. *Definición de roadmap y tareas clave*    2. *Organización y seguimiento en Trello.*    3. *Uso de repositorio GitHub para control de versiones.*    4. *Evaluación de desvíos y ajuste de cronograma si es necesario.* 2. *Tratamiento del Dataset Python :*    1. *Limpieza de datos con pandas y validaciones con NumPy.*    2. *Revisión de integridad y coherencia del dataset.*    3. *Preparación estructurada para análisis exploratorio y modelado.* 3. *Exploración y Detección de Patrones con Python:*    1. *Visualización de métricas clave con matplotlib y seaborn.*    2. *Análisis por subgrupos: formatos, tipos de publicaciones, días y horarios.*    3. *Identificación de variables influyentes en el rendimiento.* 4. *Creacion y coneccion del data set en Azure SQL :*   *De ser necesario generar consultas y vistas claves que aporten al desarrollo del análisis*    *5. Visualización Final:*  *a. Diseño del dashboard en Power BI enfocado en insights accionables.  b. Inclusión de filtros por categoría, tipo y rendimiento esperado.* |