

EQUIPO: Big Dreams

INTEGRANTES: Linck – Brun – Solana

## **ACTIVIDAD 1:** Ciudad Inteligente (Smart city)

- Caso de Uso: Gestión del tráfico urbano y seguridad pública.
- Descripción Ampliada: Las ciudades modernas están usando Big Data para mejorar la vida de sus ciudadanos. Un ejemplo es la gestión del tráfico y la seguridad:
  - Gestión del Tráfico: Sensores en las calles y cámaras de tráfico recopilan datos de flujo vehicular en tiempo real. Esta información se procesa para ajustar los semáforos de forma dinámica y re-enrutar el tráfico en caso de congestión o accidentes.
  - Seguridad Pública: Las cámaras y otros sensores recopilan datos para analizar patrones de movimiento y comportamiento. Utilizando análisis predictivos, las fuerzas de seguridad pueden identificar áreas de alto riesgo para desplegar recursos y prevenir delitos antes de que ocurran. Los datos de llamadas de emergencia, redes sociales y reportes ciudadanos también se integran en este análisis.

### **CONSIGNAS:**

#### 3. Investigación y Discusión:

- En tu grupo, analicen en profundidad el caso de estudio asignado.
- Utilicen los siguientes puntos como guía para su análisis y tomen notas de sus respuestas:

#### **1. Las 5 V's del Big Data:**

■ **Volumen:** *¿Qué tipo de datos masivos se generan o manejan? ¿De qué escala hablamos (terabytes, petabytes)?*

#### **RESPUESTA:**

Entendemos que los datos que se manejan son masivos por el tamaño de la ciudad, principalmente porque se analizaría cada zona, calle y espacio de la misma. Estamos hablando de una escala de Petabytes.

Tipos:

- Imágenes y video de cámaras de tráfico y seguridad.
- Registros de sensores en la calle
- Datos GPS de vehículos

- Reportes de los ciudadanos (apps, redes sociales, llamadas de emergencia)

■ **Velocidad:** *¿Los datos se procesan en tiempo real o por lotes? ¿Por qué esa velocidad es crucial para el éxito de la empresa?*

**RESPUESTA:**

- Los datos se procesan en tiempo real ya que está conectado con Waze y otras aplicaciones de tráfico
- Es crucial esa velocidad porque necesitan reaccionar rápido a los movimientos en el tráfico para poder regularlo y accionar en el momento

■ **Variedad:** *¿Qué tipos de datos se utilizan (estructurados, no estructurados, semi-estructurados)?*

**RESPUESTA:**

Utiliza datos todo tipo de datos

- Datos estructurados: conteos de vehículos, velocidad promedio, registros de incidentes.
- Datos no estructurados: videos, audios de llamadas al 911, fotos.
- Semi estructurados: sensores de velocidad, datos GPS

■ **Veracidad:** *¿Qué desafíos de calidad y confiabilidad de datos podrían enfrentar?*

**RESPUESTA:**

- Los datos pueden estar algo desfasados en el tiempo
- Los sensores y cámaras pueden estar dañados o pueden tener fallas
- Llamadas de emergencia falsas

■ **Valor:** *¿Cuál es el beneficio de negocio (ganancias, eficiencia, satisfacción del cliente) que se obtiene del Big Data en este caso?*

**RESPUESTA:**

- Se obtiene una mejor gestión de los elementos para regular el tráfico.
- Ayuda a que haya menos embotellamientos y eficientiza el tráfico
- Hace que los caminos tomados por los conductores sean más seguros y rápidos
- Proyección de estados de tráfico en el futuro

○ 2. Almacenamiento:

■ ¿Dónde se almacenarán estos datos? ¿Creen que sería un sistema de archivos distribuido como HDFS, un Data Lake o una base de datos más tradicional?

**RESPUESTA:**

Entendemos que almacena los datos en un sistema de Archivos Distribuido HDFS, porque los mismos manejan almacenamiento de grandes volúmenes está diseñado para almacenar archivos de gran tamaño. Principalmente en el caso de Gestión del Tráfico.

También podría ser un sistema de archivos Data lake ya que tiene que almacenar archivos de video que no están estructurados. Principalmente en el caso de seguridad Pública.

■ ¿Qué desafíos de escalabilidad y costo enfrentarían al almacenar estos datos?

**RESPUESTA:**

Costo de almacenar un gran volumen de datos. Con respecto a la escalabilidad, debería poder seguir subiendo información para poder encontrar patrones de comportamiento

○ 3. Procesamiento y Análisis:

■ ¿Qué tipo de procesamiento se necesita (por lotes o en streaming)?

**RESPUESTA:**

El procesamiento que se necesita en este caso es en tiempo real. Porque las respuestas que da Smart city son instantáneas.

■ ¿Qué herramientas de análisis serían las más adecuadas (ej. SQL, Python, machine learning)?

**RESPUESTA:**

Machine learning para poder analizar los datos en tiempo real y podemos analizar patrones complejos y hacer predicciones. Python para poder analizar los videos y fotos de las cámaras

○ 4. Gobernanza y Seguridad:

■ ¿Qué datos sensibles o personales podrían estar manejando? (ej. datos personales de clientes, historial de navegación)?

**RESPUESTA:**

- Podrían estar manejando datos personales sobre los viajes recorridos por cada usuario
- También un registro de las patentes de los autos en circulación por ciertas zonas

■ ¿Qué desafíos de seguridad y privacidad tendrían que considerar para proteger la información?

**RESPUESTA:**

- Garantizar que los datos sean sólo accesibles para el ministerio de seguridad y los agentes relacionados con la seguridad de la ciudad y no para cualquier persona.
- Asegurar que los datos sean confidenciales