**Ejercicio 1** (puntaje: 1 por pregunta)

Desarrollar los siguientes conceptos:

1. Describa que es una tabla de hecho y qué elementos se pueden encontrar en dicha tabla.

*En un modelo estrella, una* ***tabla de hechos*** *es aquella que contiene los datos medibles o cuantificables de un negocio. Generalmente, almacena métricas numéricas como ventas, ingresos, costos, etc.*

*En una tabla de hechos podemos encontrar los siguientes elementos:*

* *Medidas: Son los valores que queremos analizar, como el total de ventas o la cantidad de productos vendidos.*
* *Claves Foráneas: Referencias a las dimensiones, que permiten contextualizar las medidas con datos adicionales (como cliente, fecha, producto).*
* *Agregaciones: Datos resumidos que facilitan cálculos sobre grandes volúmenes de información.*

1. Indicar las principales diferencias entre medidas y dimensiones (defina cada una y mencione algunas diferencias).

*Medidas: Son los valores numéricos que se quieren analizar y sobre los que se pueden hacer cálculos (suma, promedio, etc.). Ejemplo: cantidad de productos vendidos, monto total de una transacción.*

*Dimensiones: Son atributos que permiten categorizar o filtrar las medidas. Ejemplo: fecha, producto, cliente.*

*Diferencias principales:*

* *Las medidas son generalmente numéricas y almacenadas en la tabla de hechos.*
* *Las dimensiones son descripciones de los datos y suelen encontrarse en tablas separadas con información más detallada.*
* *Las dimensiones permiten realizar análisis y segmentaciones en los datos almacenados en la tabla de hechos*

1. ¿Cuál es la clave de la dimensión tiempo?

*La dimensión tiempo suele tener una clave primaria que identifica cada período de tiempo único. Esta clave puede ser un identificador numérico o una fecha formateada. Ejemplo: Un identificador autogenerado (Ejemplo: id\_fecha = 20250609 para representar el 9 de junio de 2025).*

1. ¿Por qué no puede existir una cardinalidad de muchos a muchos entre una tabla Fact y una tabla de Dimensiones?

*La relación entre una tabla de hechos y una tabla de dimensiones no puede ser de muchos a muchos porque esto complicaría la integridad del modelo estrella y dificultaría la consulta eficiente de datos.* ***Agregación de datos:*** *Si la relación fuera de muchos a muchos, sería difícil realizar cálculos como sumas, promedios y otros tipos de agregaciones sin generar inconsistencias o duplicaciones en los valores.*

**Ejercicio 3** (puntaje: 6)

La empresa AgroTrack S.A. se especializa en la gestión de producción agrícola inteligente mediante el uso de sensores loT y plataformas de análisis en la nube. Su servicio permite a productores rurales monitorear en tiempo real variables como humedad del suelo, temperatura ambiental, uso de maquinaria y rendimiento de cultivos.

Con el crecimiento de su base de clientes y la incorporación de nuevos dispositivos, AgroTrack desea mejorar su capacidad de análisis de datos para optimizar la toma de decisiones agronómicas, identificar patrones de producción y prever condiciones que impacten en los cultivos. La empresa recopila información proveniente de sensores instalados en los campos, que registran mediciones periódicas relacionadas al clima, al suelo y al uso de maquinaria.

Entre los atributos relevantes a considerar se encuentran:

Para las maquinarias: tipo de maquinaria, modelo, año de fabricación, consumo promedio, capacidad operativa, frecuencia de uso.

Para los sensores: tipo de sensor, variable medida (temperatura, humedad, etc.), precisión, unidad de medida, intervalo de muestreo y estado de funcionamiento.

Para los clientes: nombre del productor, tipo de cliente (individual o empresa), ubicación, antigüedad como cliente, nivel de tecnificación del campo, tipo de cultivo habitual, contacto principal, y canal de comunicación preferido.

Para los cultivos: tipo de cultivo, variedad, ciclo de crecimiento estimado, requerimientos hídricos, y tolerancia a condiciones climáticas.

Además, mantiene registros detallados de las campañas agrícolas de sus clientes, incluyendo cultivos sembrados, fechas de siembra y cosecha, rendimiento por hectárea y condiciones meteorológicas durante el ciclo de cultivo. que registran mediciones periódicas relacionadas al clima, al suelo y al uso de maquinaria. Además, mantiene registros detallados de las campañas agrícolas de sus clientes, incluyendo cultivos sembrados, fechas de siembra y cosecha, rendimiento por hectárea y condiciones meteorológicas durante el ciclo de cultivo.

La dirección de AgroTrack ha solicitado al área de análisis de datos el diseño de un modelo que permita realizar consultas relacionadas con el rendimiento promedio por cultivo y por región, la influencia de condiciones climáticas en los resultados de producción, el comportamiento del suelo en diferentes parcelas, la eficiencia del uso de maquinaria, y la comparación entre campañas agrícolas de diferentes años y zonas.

También se desea responder a preguntas como:

¿Cuáles son las condiciones óptimas de humedad y temperatura que se repiten en campañas exitosas?

¿Qué regiones muestran mayor variabilidad en el rendimiento?

¿Qué maquinaria presenta mayor frecuencia de uso?

¿Cómo varían las mediciones de los sensores a lo largo del tiempo?

¿Hay zonas donde las condiciones climáticas han cambiado significativamente en los últimos años?

¿Cuáles son las condiciones óptimas de humedad y temperatura que se repiten en campañas exitosas?

¿Qué regiones muestran mayor variabilidad en el rendimiento?

¿qué maquinaria presenta mayor frecuencia de uso?

Una regla de negocio importante a considerar es que un cliente puede poseer múltiples lotes agrícolas, y que en cada campaña no necesariamente se siembran todas las hectáreas disponibles en un lote. Esta variabilidad debe ser contemplada en el análisis para obtener indicadores más precisos sobre el uso de la tierra y el rendimiento.

Restricciones y consideraciones adicionales:

* Identificar las claves primarias y foráneas de cada tabla.
* Definir las relaciones entre la tabla de hechos y las dimensiones.
* Especificar las jerarquías dentro de las dimensiones, si es aplicable
* Asegurarse de que el modelo sea flexible y escalable para permitir futuras incorporaciones de datos y análisis más complejos.

Objetivo:

El objetivo final es crear un diseño de modelo dimensional que facilite la comprensión y el análisis de los siniestros reportados en la compañía de seguros, permitiendo a los usuarios realizar consultas eficientes y obtener información significativa para la toma de decisiones informadas.

TUIA-2025