

UNIVERSIDAD RAFAEL LANDÍVAR

FACULTAD DE INGENIERÍA

ANÁLISIS DE DATOS

SECCIÓN 1 VESPERTINA

ING. JAIRO SALAZAR

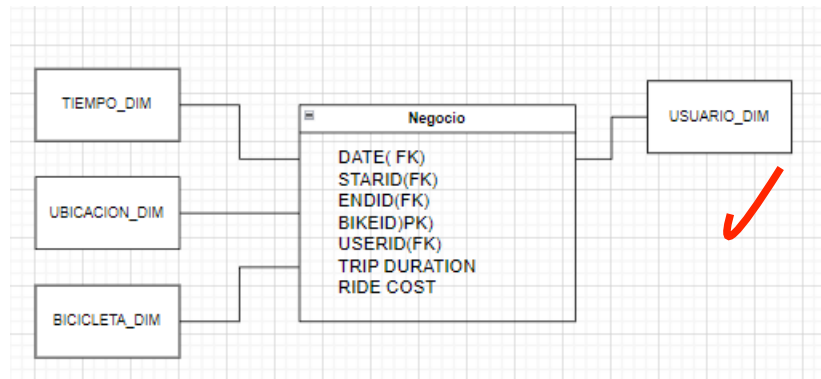
EXAMEN PARCIAL 1

Julio Anthony Engels Ruiz Coto 1284719

GUATEMALA DE LA ASUNCIÓN, MARZO 6 DE 2023

CAMPUS CENTRAL

1. Efectué un diseño dimensional con los datos provistos. Identifique plenamente las dimensiones y la tabla de hechos, describiendo cada uno de los campos que pertenecen a cada entidad (35 puntos)



En la dimension de TIEMPO_DIM = iría start time, stop time, tripduration, start year , start month, start day, start hour. En la dimension UBICACION_DIM = start station ID, stop station id , start station latitude, start station longitude, stop station latitude, stop station longitude, en la dimension de BICICLETA_DIM = Bike id, en la dimension de usuario USUARIO_DIM = user type , birth year, gender.

2. Si usted pudiera enriquecer el diseño agregando datos, campos o otras fuentes de datos, de manera general y sin conocer mucho del negocio, describa que otras cosas le agregaría al diseño. (10 puntos)

Si yo pudiese enriquecer el diseño yo como analista agregaría la información sobre la demografía del usuario con estos datos me permitirían es analizar la demanda de bicicletas entre diferentes grupos de usuarios por ejemplo la edad, la ocupación, el ingreso. También agregaría la información sobre el clima y las condiciones meteorológicas en las que se experimenta en cada viaje con esto me permitiría analizar como me afecta el clima en la demanda de bicicletas.

3. Dejando aparte el tema de los costos y presupuestos de poder implementar un sistema de data warehouse, y tomando en cuenta los objetivos generales de tener un sistema DW/BI vistos en clase, describa que otros argumentos podría proveer para poder apoyar la implementación de un sistema de DW/BI. (15 puntos)

Mejorar la calidad de los datos, como así un sistema de DW/BI puede integrar los datos de diferentes fuentes y limpiarlos para asegurarse de que estos sean precisos y consistentes. Una mayor eficiencia un sistema de DW/BI tiene la posibilidad de automatizar procesos y reducir dicho tiempo y la búsqueda necesaria para obtener dicha información y analizar datos. También facilidades en toma de decisiones, el sistema DW/BI da información procesada y resumida en la que esta permite a los gerentes y tomadores de decisiones tener una visión un poco mas clara del negocio y tomar decisiones informadas y respaldadas por los datos.

4. Como puede observar, no en todos los casos el año de nacimiento del cliente se ha registrado correctamente, asumiendo que el se mejora el app por medio de la cual se contratan los servicios de uso de las bicicletas, obligando ahora a actualizar el año de nacimiento de cada usuario, tomando en cuenta el concepto de las slow changing dimensiones vistos en clase, cual de los tipos de SCD utilizaría. JUSTIFIQUE SU RESPUESTA (20 puntos).

Me inclino mas por utilizar un SCD de tipo 2, esta permite mantener un historial de los cambios en los datos, se pueden mantener un trazo completo de los cambios en el año de nacimiento de los usuarios lo que puede ser útil para los análisis posteriores.

en este caso para mi en mi opinion aplicaria mas la 0.

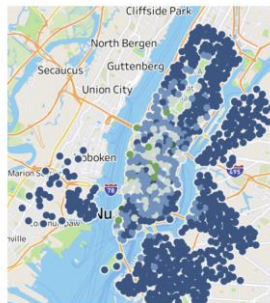
✓/✓

5. El director del programa de NYC Bike, desea conocer cuales son las estaciones de bicicletas con mas demanda, de una manera grafica, El equipo de visualización de datos ha trabajado en tres propuestas para que usted elija, tomando en cuenta lo visto en clase, cual de estas propuestas elegiría usted y porque. JUSTIFIQUE SU RESPUESTA. (visualizaciones en la siguiente hoja). (10 puntos)

Yo como analista de datos me iría por la propuesta 3 ya que se presenta un mapa de calor en la que describe la demanda de bicicletas en diferentes, estaciones de la ciudad , esta es una visualización clara y fácil de interpretar para todo tipo de personas y le permite identificar rápidamente las estaciones con mayor demanda por todo nueva york.

Para mi esta es la mas cargada visualmente y menos recomendable

Propuesta 3



6. Se le presentan en la hoja no. 5 del presente documento, una serie de visualizaciones de datos del set de datos mencionado. Tomando en cuenta los conceptos vistos en clase de Everyday data Viz y Visual Discovery, clasifique cada uno de las visualizaciones presentadas en alguna de estas dos clasificaciones. Justifique cada elección. (10 puntos)

Visualización 1 : visual Discovery

Visualizacion 2: Everyday data viz

✓

Visualizcion 3: visual Discovery

Visualización 4: Everyday data viz