

TRABAJO PRÁCTICO - GESTIÓN DE DATOS

2º CUATRIMESTRE 2022



Integrantes: Juan Cruz Rey Alvarez, Juan Ignacio Schuhmann, Nicolas De Rosso, Rodrigo Gutson

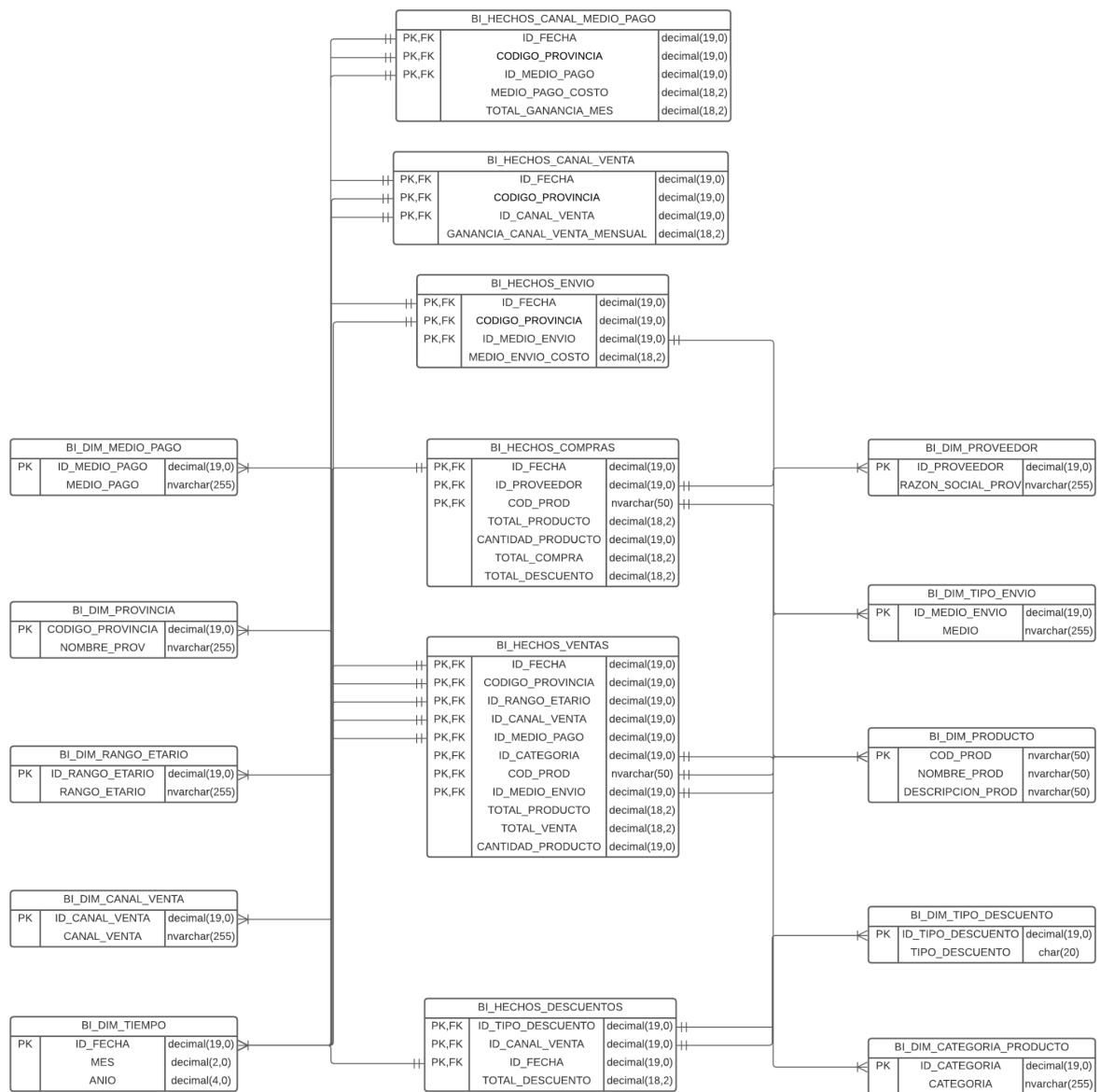
DER

3

DECISIONES:

4

DER



DECISIONES:

Se modelan las dimensiones necesarias (pedidas por el enunciado) para poder realizar las vistas solicitadas. Además, agregamos la siguiente dimensión de proveedor para poder cumplir con la vista “Aumento promedio de precios de cada proveedor anual.”, ya que distinguimos por razón social.

En cuanto a tabla de hechos, creamos las que consideramos necesarias para tener los datos accesibles de forma rápida para las vistas y evitar tener que hacer muchos joins (lo que degradaba significativamente la performance), lo que nos sucedió en un comienzo, por eso lo cambiamos al modelo actual (antes teníamos todo en la tabla de hechos ventas, ahora tenemos: de compras, de medios de envío, de canales de venta, de ventas, de medios de pago y de descuentos)

Creamos unas funciones auxiliares que nos sirvieron para obtener ciertos datos de las tablas más fácilmente, y otras para el proceso de migración del modelo normalizado al modelo BI

Para cargar la tabla hechos_compras usamos un cursor ya que queríamos sumar la cantidad y el total por producto para cada fecha si ya existía. Se podría haber resuelto sin cursores (con group by por ejemplo) pero la performance de esta procedure no era un problema (al correrla individualmente tardaba un segundo*), por lo que decidimos no modificarla.

En la de hechos_ventas nos sucedió algo similar. Además esta se encarga de, por cada venta que lee, insertar sus descuentos según su fecha en hechos_descuento, discriminándolos por su tipo. Como en la venta ya teníamos esa información (además del canal de venta y otros), nos pareció más sencillo que lo levante directamente desde ahí. Por este mismo motivo inserta en hechos_envio, hechos_canal_venta, y hechos_medio_pago, si no existían previamente.

Para el rango etario asumimos que los intervalos son mutuamente excluyentes (recordando que el corchete incluye al valor y el paréntesis no):

- <25: [0,25)
- 25 - 35: [25-35)
- 35 – 55: [35-55]
- >55: (55-infinito) (caso donde cae si no cumple ninguno anterior)

Luego como aclaración extra, en donde se piden porcentaje se casteo a varchar para poder agregarle el símbolo ‘%’

A su vez, cuando se pedían condiciones como el top 5, top 3, top N, por ciertos criterios (más específicamente en los casos donde usamos ROW_NUMBER() OVER (PARTITION BY...)) pusimos la condición del N en el select de la vista. Realizamos una pregunta en el foro de si esto era correcto o debía estar dentro de

la vista, pero no recibimos respuesta hasta la modificación de este documento así que lo dejamos en el SELECT.