

Optimización de Redes de Transporte Urbano

Optimización del transporte urbano mediante análisis de datos en MySQL para mejorar eficiencia y experiencia del usuario.

Integrantes:

- Escobar Arana Dario
Felipe
- Heredia Villamar Kimberly
Elizabeth
- Quinde Montecé Josué
Lenin
- Vera Cajape Neyvi
Mageles
- Villareal Mora Sebastian
Emiliano

Fundamentos de Datos para la Optimización del Transporte

La gestión de flotas y rutas de transporte urbano genera muchos datos, que para este proyecto clasificamos en cuatro tipos principales:

Tipos de datos en transporte urbano:

- **Demanda:** pasajeros y tarifas.
- **Costos:** gastos por ruta.
- **Operación/Planificación:** horarios y viajes.
- **Maestro:** rutas, paradas y conexiones.



DEFINICIÓN Y ENFOQUE

Delimitación y Lógica del Análisis de Datos



Delimitación Temporal y Espacial

- Análisis inicial basado en tiempos *programados*.
- Enfoque en rutas, viajes y tramos específicos.
- Horas pico definidas como 06:00–09:00 y 17:00–20:00 (parametrizable).

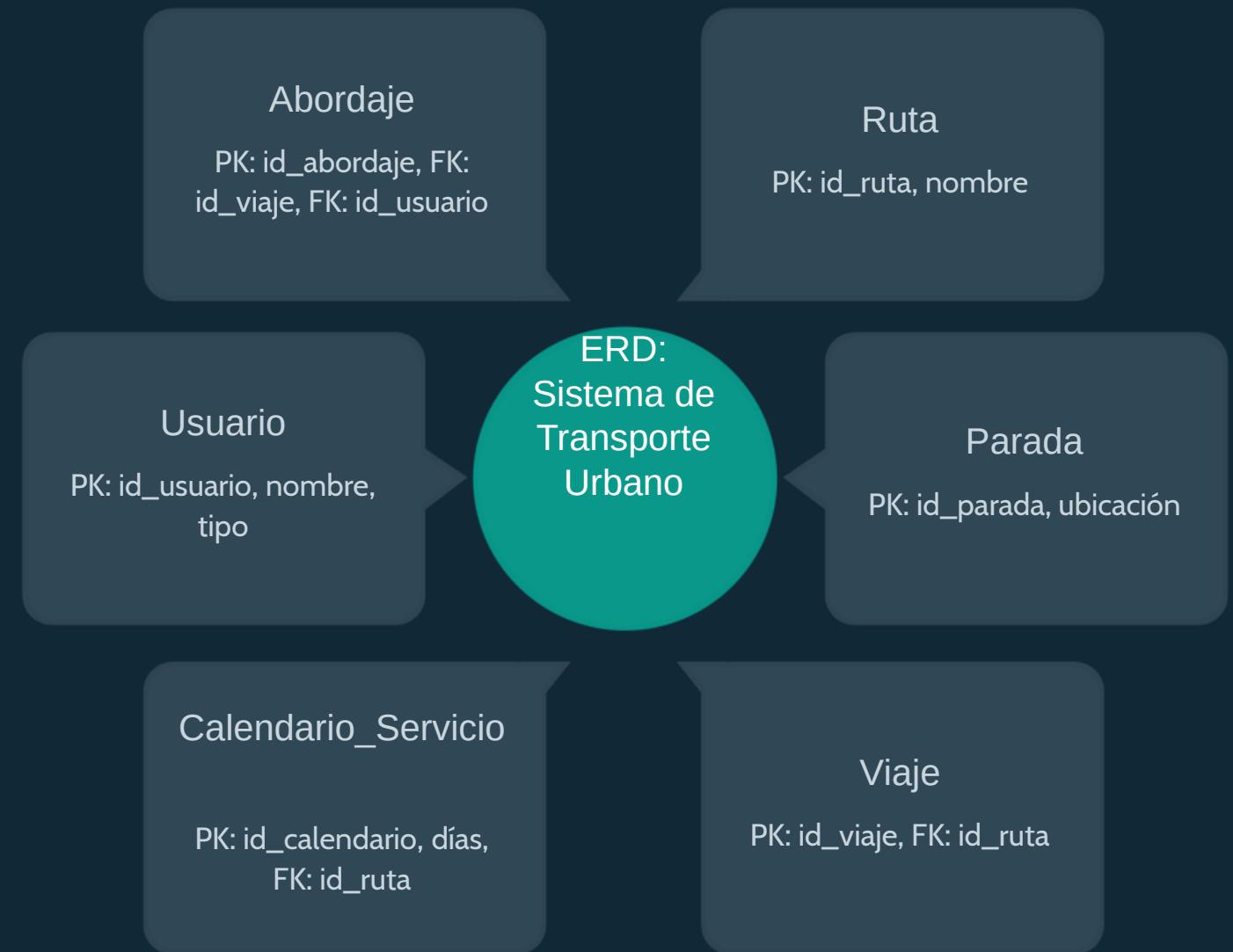


Lógica de Conteo de Demanda

- Demanda contabilizada por eventos de dirección = 'sube'.
- Flujo a bordo acumulado por tramo mediante funciones de ventana.
- Costeo por ruta-fecha basado en pasajeros que suben diariamente.

Este enfoque nos permite tener una visión clara y estructurada de los datos de transporte, asegurando que las métricas calculadas sean relevantes y accionables para la planificación urbana.

Diseño de Tablas y Relaciones para Consistencia



Integridad y Reglas Clave

- **CHECK**: Asegurar que el orden de las paradas sea positivo y las tarifas no sean negativas.
- **UNIQUE**: Definir claves únicas para evitar duplicidades en pares como (id_ruta, orden) y (id_costo, tipo_costo).
- **Claves Foráneas (FKs)**: Establecer relaciones con CASCADE/RESTRICT para mantener la integridad referencial.

Tipos de Consultas Desarrolladas para el Análisis

Agregación y Ranking



Ocupación por ruta(horas pico).

Series Temporales Discretas



Análisis de demanda por franjas horarias de 30 minutos.

Funciones de Ventana



Determinación de pasajeros a bordo por cada tramo.

Costeo y Rentabilidad



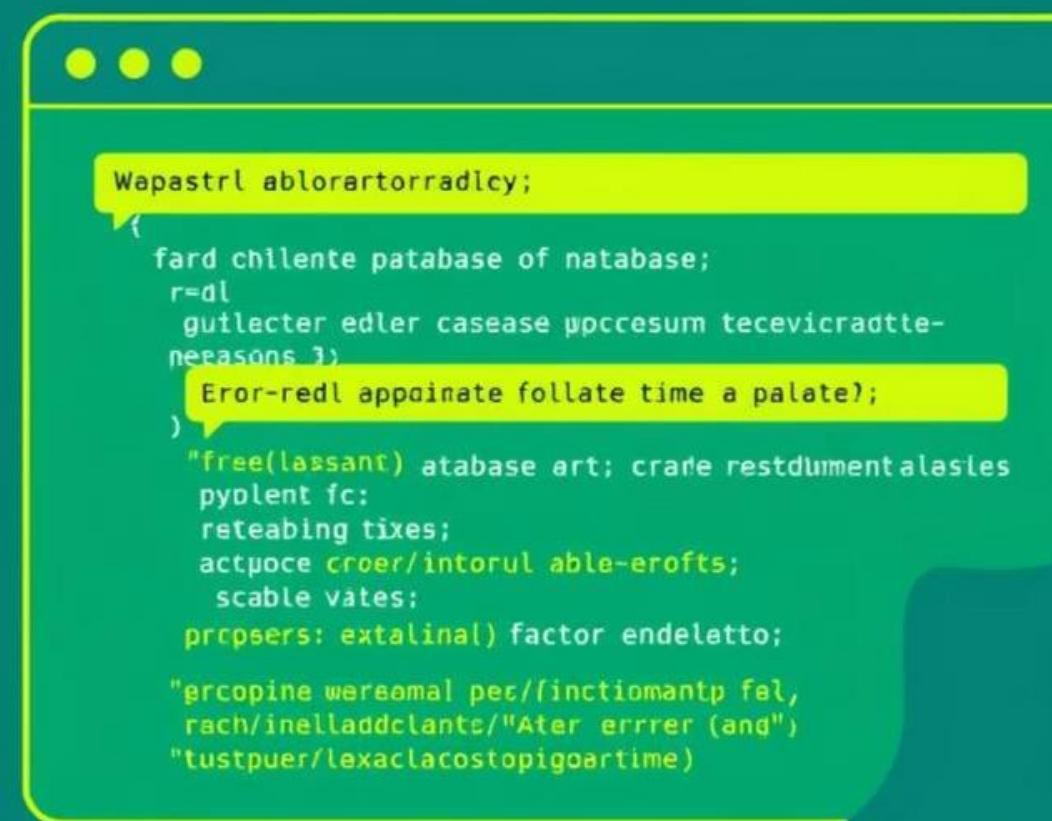
Costos,ingresos y márgenes por ruta-fecha.

Duración de Viaje



Tiempo entre la primera salida y la última llegada.

Obstáculos Superados en el Proceso de Desarrollo



Se reescribieron consultas con CTEs para evitar incompatibilidades.

Se implementó una restricción UNIQUE en (id_viaje, secuencia) para evitar paradas repetidas en un mismo viaje.

Se aplicaron claves foráneas compuestas para mantener la integridad entre las rutas y paradas asignadas en cada viaje.

Mejora del rendimiento de las consultas a través de la creación de índices estratégicos en claves foráneas y columnas de filtrado.

Respuestas a Preguntas Clave y Metodología

0 1 Rutas con Mayor Ocupación en Horas Pico

Contamos los eventos de 'sube' entre 06:00-09:00 y 17:00-20:00 por cada ruta. El resultado es un ranking de rutas/viajes con mayor demanda en estos períodos críticos.

0 2 Horarios con Menor Uso

Agregamos la demanda por slots de 30 minutos (ej. HOUR(ts_evento)). Identificamos así los períodos de menor uso, lo que permite ajustar frecuencias para optimizar recursos.

0 3 Métricas de Optimización

- **Duración promedio programada:** por viaje y ruta.
- **Headway (frecuencia):** promedio de frecuencia_min.
- **Costo por pasajero:** ($\text{SUM}(\text{costos}) / \#\text{subidas}$) por ruta-fecha.
- **Pasajeros a bordo por tramo:** detección de cuellos de botella.

Resultados — Preguntas de análisis (dataset de ejemplo)

¿Qué rutas tienen mayor ocupación en horas pico? → R01 (Terminal Norte–Centro, ida) es la única con abordajes y encabeza con 2 subidas en hora pico (06:00–06:30).

¿Qué horarios presentan menor uso del transporte? → Con el dataset de ejemplo, solo hay subidas entre 06:00–06:30 (2). Todos los demás *slots*/horas del día registran 0 subidas (menor uso).

Principios de Diseño de la Base de Datos

Modelo Normalizado

Separación clara entre datos de planificación (horarios), demanda (eventos de pasajeros) y costos, para evitar redundancias y facilitar la gestión.

Reglas y FK Compuestas

Garantía de la consistencia de los datos mediante el uso de claves foráneas compuestas y reglas de validación, especialmente entre rutas y sus paradas.

Índices Optimizados

Reducción significativa de la latencia en consultas analíticas recurrentes, al aplicar índices eficientes en claves foráneas y campos de filtrado.

Ventanas y CTEs

Simplificación de cálculos complejos, como acumulados y rankings, mediante el uso de Common Table Expressions (CTEs) y funciones de ventana.

PRÓXIMOS PASOS

Resultados Clave y Futuras Mejoras

Hemos logrado generar rankings de rutas por demanda en horas pico, identificar horarios de bajo uso, calcular duraciones programadas y estimar costos por pasajero. Estos resultados proporcionan una base sólida para la toma de decisiones.



Registrar Tiempos Reales

Integrar datos de tiempos de llegada y salida reales para evaluar la puntualidad y el rendimiento operativo.



Capacidad de Vehículos

Incluir la capacidad de los vehículos para calcular el factor de ocupación real y optimizar la asignación de recursos.



Dashboard de BI

Desarrollar un dashboard interactivo con KPIs y alertas operativas para una monitorización continua y proactiva.

Reflexión final

Una buena base de datos no solo organiza información, sino que también permite tomar decisiones más eficientes y prevenir errores en sistemas complejos.

Aprendizajes :

- Aplicamos SQL para mejorar un sistema de transporte urbano.
- Diseñamos bases de datos relacionales, usando claves, restricciones y normalización.
- Practicamos subconsultas, JOINs e índices, asegurando integridad de los datos.