

Análisis de las transacciones de recarga del
sistema de transporte publico de Bogotá para la
mejora de los tiempos de espera de los usuarios
en la fila.

Cendy Lozano, César Pedreros, Fabian Vargas

Facultad de Ingeniería y Ciencias Básicas,

Universidad Central

Maestría en Analítica de Datos

Curso de Bases de Datos

Bogotá, Colombia

¹*clozanog1@ucentral.edu.co*, ²*cpedrerosc@ucentral.edu.co*, ³*hvargasg1@ucentral.edu.co*

October 8, 2022

Contents

1	Introducción (Max 250 Palabras) - (<i>Primera entrega</i>)	4
2	Características del proyecto de investigación (Max 500 Palabras) - (<i>Primera entrega</i>)	4
2.1	Titulo del proyecto de investigación (Max 100 Palabras) - (<i>Primera entrega</i>)	4
2.2	Objetivo general (Max 100 Palabras) - (<i>Primera entrega</i>)	5
2.2.1	Objetivos especificos (Max 100 Palabras) - (<i>Primera entrega</i>)	5
2.3	Alcance (Max 200 Palabras) - (<i>Primera entrega</i>)	5
2.4	Pregunta de investigación (Max 100 Palabras) - (<i>Primera entrega</i>)	5
2.5	Hipotesis (Max 100 Palabras) - (<i>Primera entrega</i>)	5
3	Reflexiones sobre el origen de datos e información (Max 400 Palabras) - (<i>Primera entrega</i>)	6
3.1	¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (<i>Primera entrega</i>)	6
3.2	¿Cuales son las consideraciones legales o eticas del uso de la información? (Max 100 Palabras) - (<i>Primera entrega</i>)	6
3.3	¿Cuales son los retos de la información y los datos que utilizara en la base de datos en terminos de la calidad y la consolidación? (Max 100 Palabras) - (<i>Primera entrega</i>)	6
3.4	¿Que espera de la utilización de un sistema de Bases de Datos para su proyecto? (Max 100 Palabras) - (<i>Primera entrega</i>)	7
4	Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos)(<i>Primera entrega</i>)	8
4.1	Características del SMBD (Sistema Manejador de Bases de Datos) para el proyecto (<i>Primera entrega</i>)	8
4.2	Diagrama modelo de datos (<i>Primera entrega</i>)	9
4.3	Imágenes de la Base de Datos (<i>Primera entrega</i>)	10
4.4	Código SQL - lenguaje de definición de datos (DDL) (<i>Primera entrega</i>)	12
4.5	Código SQL - Manipulación de datos (DML) (<i>Primera entrega</i>)	13
4.6	Código SQL + Resultados: Vistas (<i>Primera entrega</i>)	14
4.7	Código SQL + Resultados: Triggers (<i>Primera entrega</i>)	14
4.8	Código SQL + Resultados: Funciones (<i>Primera entrega</i>)	14
4.9	Código SQL + Resultados: procedimientos almacenados (<i>Primera entrega</i>)	14
5	Bases de Datos No-SQL (<i>Segunda entrega</i>)	15
5.1	Diagrama Bases de Datos No-SQL (<i>Segunda entrega</i>)	15
5.2	SMBD utilizado para la Base de Datos No-SQL (<i>Segunda entrega</i>)	15

6	Aplicación de ETL (Extract, Transform, Load) y Bodega de Datos (<i>Tercera entrega</i>)	16
6.1	Ejemplo de aplicación de ETL y Bodega de Datos (<i>Tercera entrega</i>)	16
7	Lecciones aprendidas (<i>Tercera entrega</i>)	17
8	Bibliografía	18

1 Introducción (Max 250 Palabras) - (*Primera entrega*)

La movilidad de las personas en las grandes ciudades del mundo siempre ha sido un gran reto, en Bogotá de los años 60 a los 90 se contaba con un sistema de transporte privado que no cubría las necesidades de sus habitantes. Diferentes administraciones distritales propusieron planes para mejorar los aspectos de la movilidad en la ciudad, sin embargo, fue hasta el año 2000 cuando se logró implementar la primera troncal de sistema Transmilenio S.A.[Andrade01] que se extendía por la toda la avenida Caracas. El sistema Transmilenio ha tenido muchos contradictores sin embargo se mantiene en la actualidad con 11 troncales, 138 estaciones y 9 portales distribuidos por toda la ciudad y con una troncal adicional en pleno proceso de construcción como lo es la troncal de la calle 68.

Sin embargo, es evidente que en diferentes sectores de la ciudad el sistema se encuentra saturado, los usuarios deben realizar largas filas para poder ingresar a él, lo cual ha generado malestar, inconformismo y descontento con el servicio prestado. La ciudadana requiere con urgencia soluciones que mitiguen este tipo de situaciones en las que los ciudadanos invierten su tiempo en realizar largas filas para poder realizar la recarga de los pasajes en sus tarjetas.

La situación mencionada se ha visto reflejada en una reducción considerable en los ingresos del sistema por el aumento del fenómeno de evasión de pago de pasajes (colados) sumado al periodo de pandemia que incrementó el desplome de recursos con los que el sistema masivo de transporte puede operar.

Este proyecto tiene como objetivo realizar un análisis descriptivo de la cantidad y valor de las transacciones de recarga del sistema masivo de transporte Transmilenio buscando de determinar cuáles son los puntos de menos frecuencia para darle la mejor opción al usuario para que realice su recarga.

2 Características del proyecto de investigación (Max 500 Palabras) - (*Primera entrega*)

El presente proyecto ha sido una investigación estructurada siguiendo de forma paulatina lo aprendido en el curso de Bases de Datos y colaborativa (sin ánimo de lucro) entre un grupo de profesionales de diversas áreas y que a la vez son estudiantes de la maestría de Analítica de Datos y una compañía privada que de forma muy atenta nos permitió usar data clasificada para analizar un proceso que es de los más críticos dentro del sistema de transporte masivo Transmilenio, como lo es del de recaudo y recargas de pasajes. La investigación usó únicamente data asociada a las troncales más importantes del sistema, taquillas, dispositivos de recarga y cantidad de recargas sin incluir información sensible de usuarios del sistema.

2.1 Título del proyecto de investigación (Max 100 Palabras) - (*Primera entrega*)

Análisis de las transacciones de recarga del sistema de transporte público de Bogotá para la mejora de los tiempos de espera de los usuarios en la fila.

2.2 Objetivo general (Max 100 Palabras) - (*Primera entrega*)

Realizar un análisis descriptivo de la cantidad y valor de las transacciones y recargas del sistema masivo de transporte Transmilenio en sus zonas y taquillas con el fin de determinar cuáles son los puntos de menos frecuencia para darle la mejor opción al usuario para que realice su recarga.

2.2.1 Objetivos especificos (Max 100 Palabras) - (*Primera entrega*)

- Recopilar la información de las transacciones de recarga realizadas en las diferentes taquillas del sistema Transmilenio de los ultimo 5 años.
- Realizar un análisis descriptivo de la información de recargas realizado en las diferentes estaciones del sistema Transmilenio.
- Identificar el comportamiento de recarga de las diferentes estaciones del sistema por troncal.
- Realizar un pronóstico de recargas para el año 2023.

2.3 Alcance (Max 200 Palabras) - (*Primera entrega*)

El presente proyecto tiene como alcance en primera medida, realizar el análisis de la información disponible de las recargas que se realizaron en el sistema de transporte masivo Transmilenio entre los años 2017 y 2022 de forma que permita tener un diagnóstico de la situación y comportamiento que se ha venido presentando del proceso de recargas del sistema. Una vez terminada esta fase diagnóstica se busca realizar un modelo que permita generar un pronóstico de las recargas para el año 2023.

2.4 Pregunta de investigación (Max 100 Palabras) - (*Primera entrega*)

¿Es posible mejorar los tiempos de recarga presencial en las diferentes estaciones del sistema Transmilenio por medio de un análisis que le permita a los usuarios determinar a través de un modelo cuales son las estaciones menos congestionadas para realizar la recarga de los pasajes?

2.5 Hipotesis (Max 100 Palabras) - (*Primera entrega*)

El modelo tipo pronóstico que se desarrolle permitirá predecir con un 95 por ciento de certidumbre el comportamiento que tendrán las recargas en las diferentes estaciones del sistema de forma que permita al usuario mejorar su experiencia respecto al tiempo de recarga.

3 Reflexiones sobre el origen de datos e información (Max 400 Palabras) - (*Primera entrega*)

Como es sabido, gran parte de la información que se cuenta sobre el sistema de transporte masivo Transmilenio está disponible en la red y en particular en la página de corporativa de <https://www.transmilenio.gov.co/>. Sin embargo, la información puntual del proceso de recargas no es información pública y el acceso a ella la tienen privados que son los operadores del recaudo como lo es Recaudo Bogotá. La oportunidad de acceso a esta información fue posible ya que uno de los integrantes del grupo de trabajo es funcionario de esta última y a partir de un proceso interno le dieron aprobación para poder tener disponible la información estadística de las recargas entre los años 2017-2022. Tomando en cuenta lo anterior, se hizo un acuerdo de manejo de confidencialidad con la persona integrante del equipo de forma que la data base para el proyecto fuera filtrada inicialmente por ella y que nos fuera entregado para el objetivo del proyecto la información menos sensible y que nos sirva para la generación del modelo de pronóstico.

3.1 ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (*Primera entrega*)

El origen de la data utilizada en este proyecto son los sistemas de gestión y monitoreo del sistema de recaudo de la compañía Recaudo Bogotá.

3.2 ¿Cuales son las consideraciones legales o eticas del uso de la información? (Max 100 Palabras) - (*Primera entrega*)

Como se mencionó en uno de los apartados anteriores, fue posible el acceso a la información base para el correcto desarrollo del proyecto a partir de la aprobación de los líderes del servicio de Recaudo Bogotá, aprovechando que uno de los integrantes del grupo es funcionaria de la mencionada compañía. Sin embargo, es importante acotar que hubo un acuerdo de confidencialidad entre todos los estudiantes del equipo, de forma que la información tomada de la data original no tuviera datos sensible o clasificados y que se usaran únicamente campos que tuvieran información numérica que permitiera filtrar cálculos estadísticos para generar el futuro modelo sin afectar aspectos legales como el Habeas Data, infracciones a derechos de autor, copias ilegales o éticos como hacer un uso indebido de la información para autobeneficio económico.

3.3 ¿Cuales son los retos de la información y los datos que utilizara en la base de datos en terminos de la calidad y la consolidación? (Max 100 Palabras) - (*Primera entrega*)

A partir del desarrollo ejecutado sobre la data original es bueno tomar en cuenta que la misma correspondía a más de 12 millones de registros los cuales nos fueron entregados en archivos .csv y que tenían un peso aproximado a 1 Tb. Por tanto uno de los retos iniciales por parte de los dueños de la información

fue la generación de la misma, ya que de los 4 archivos .csv facilitados por los líderes de Recaudo Bogotá les tomó aproximadamente 8 horas cada archivo.

A continuación el paso a seguir (segundo reto) fue la revisión y selección de los campos que nos permitirá hacer el análisis y generar el futuro pronóstico de la amplia cantidad de información que se tenía originalmente.

El siguiente reto fue la carga de los archivos en el SGDB para continuar con la creación de tablas de forma tal que se pudiesen normalizar y definir relaciones optimas para la extracción de la data.

3.4 ¿Que espera de la utilización de un sistema de Bases de Datos para su proyecto? (Max 100 Palabras) - (*Primera entrega*)

Nuestro principal objetivo del uso de BD en el proyecto es la optimización en el manejo y acceso ordenado a la información clave para nuestro futuro modelo de pronóstico, asegurando que la información no va a estar duplicada de forma que se afecten nuestros resultados.

4 Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos) *(Primera entrega)*

4.1 Características del SMBD (Sistema Manejador de Bases de Datos) para el proyecto *(Primera entrega)*

El SMBD que usamos para el desarrollo de nuestro proyecto fue MySQL 8.0 cuyas características las podemos resumir de la siguiente manera:

- Rendimiento: MySQL 8.0 ofrece un mejor rendimiento en las siguientes áreas: cargas de trabajo de lectura / escritura, cargas de trabajo intensivas en E / S y cargas de trabajo de alta competencia.
- NoSQL: MySQL desde la versión 5.7 proporcionar la función de almacenamiento NoSQL. Esta función también se ha mejorado considerablemente en la versión 8.0, permitiendo eliminar la necesidad de una base de datos de documentos NoSQL separada.
- Funciones de Ventana: A partir de MySQL 8.0, se ha agregado un nuevo concepto llamado función de ventana, que se puede usar para implementar varios métodos de consulta nuevos. Las funciones de ventana son similares a las funciones agregadas como SUM () y COUNT (), pero no combina varias filas de resultados de consulta en una fila, sino que vuelve a colocar los resultados en varias filas. Es decir, la función de ventana no requiere GROUP BY.
- Fiabilidad: InnoDB ahora admite la atomicidad de la tabla DDL, es decir, el DDL en la tabla InnoDB también puede lograr la integridad de la transacción, ya sea fallando al retroceder o confirmando con éxito, de modo que no se produzca el éxito parcial de DDL. Además, también es compatible con la función a prueba de choques. Los metadatos se almacenan en un único diccionario de datos de transacciones.
- Seguridad: Mejoras en OpenSSL, nueva autenticación predeterminada, roles SQL, seguridad de contraseña, autorización.
- Requerimientos de instalación:
 - 512 Mb de memoria Ram.
 - 1024 Mb maquina virtual.
 - 1 GB de espacio de disco duro.
 - Sistema operativo: Windows, Linux y Unix.
 - Arquitectura del sistema 32/64 bit.

4.2 Diagrama modelo de datos (Primera entrega)

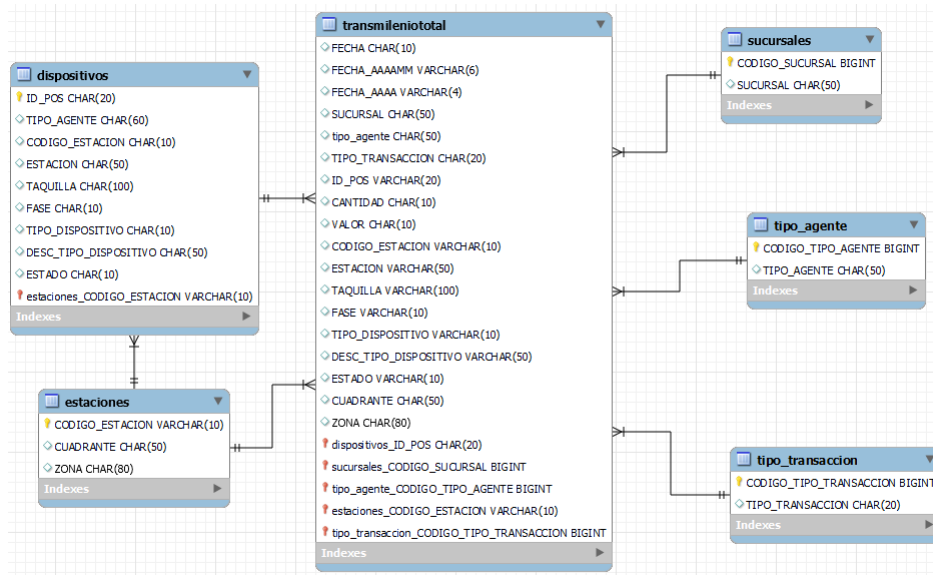


Figure 1: Modelo Base de Datos

4.3 Imágenes de la Base de Datos (*Primera entrega*)

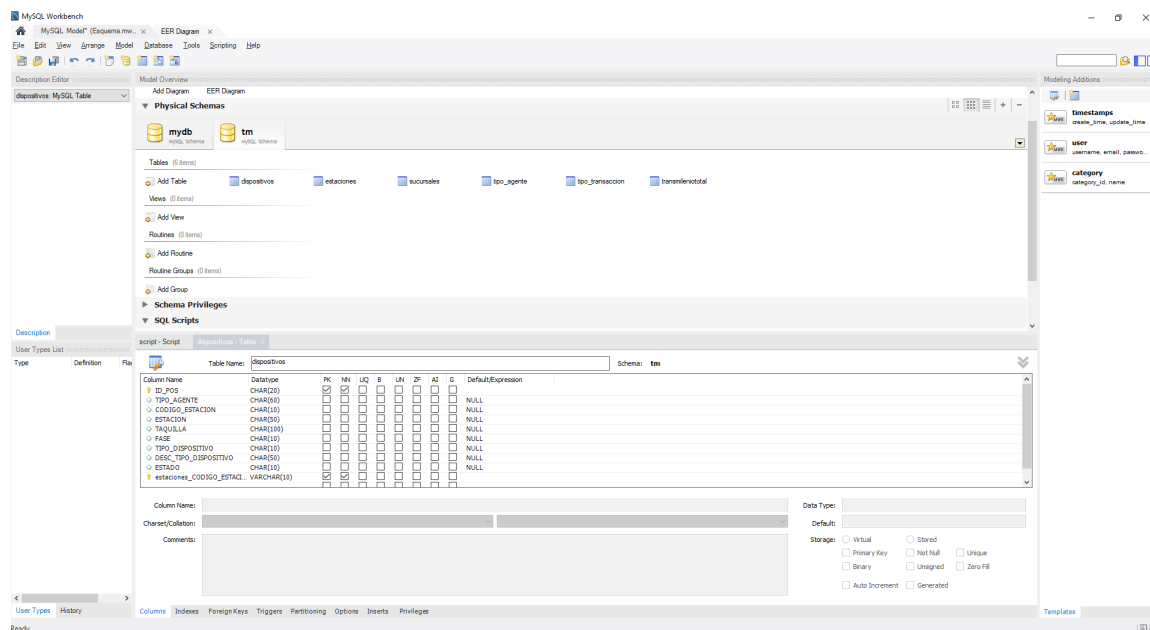


Figure 2: Tabla Dispositivos

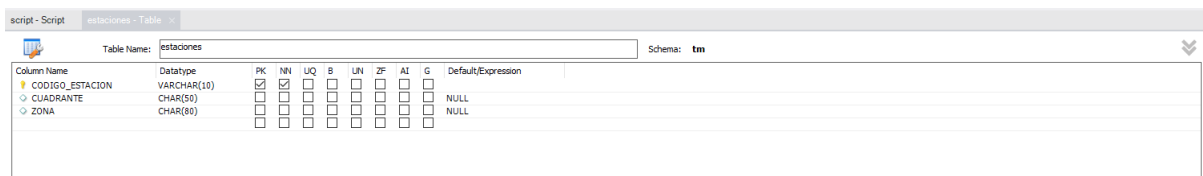


Figure 3: Tabla Estaciones

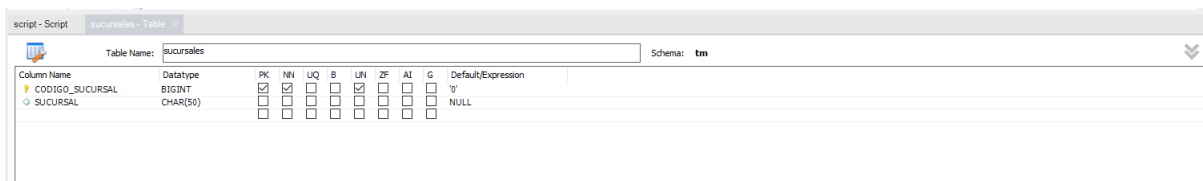


Figure 4: Tabla Sucursales

script - Script tipo_agente - Table

Table Name: Schema: **tm**

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
↳ CODIGO_TIPO_AGENTE	BIGINT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0'
↳ TIPO_AGENTE	CHAR(50)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL

Figure 5: Tabla Tipo Agente

script - Script tipo_transaccion - Table

Table Name: Schema: **tm**

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
↳ CODIGO_TIPO_TRANSACCION	BIGINT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0'
↳ TIPO_TRANSACCION	CHAR(20)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL

Figure 6: Tabla Tipo Transacción

Physical Schemas

mydb tm

Tables (5 items)

dispositivos estaciones sucursales tipo_agente tipo_transaccion transmileniototal

Views (0 items)

script - Script transmileniototal - Table

Table Name: Schema: **tm**

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
↳ fecha	CHAR(10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ FECHA_AAAA	VARCHAR(4)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ SUCURSAL	CHAR(50)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ tipo_agente	CHAR(50)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ TIPO_TRANSACCION	CHAR(20)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ ID_POS	VARCCHAR(20)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ CANTIDAD	CHAR(10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ VALOR	CHAR(10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ CODIGO_ESTACION	VARCCHAR(10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ ESTACION	VARCCHAR(30)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ TAQUILLA	VARCCHAR(100)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ FASE	VARCCHAR(10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ TIPO_DISPOSITIVO	VARCCHAR(10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ DESC_TIPO_DISPOSITIVO	VARCCHAR(50)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ ESTADO	VARCCHAR(10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ CUADRANTE	CHAR(50)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ ZONA	CHAR(80)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
↳ dispositivos_ID_POS	CHAR(20)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
↳ sucursales_CODIGO_SUCUR...	BIGINT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
↳ tipo_agente_CODIGO_TIPO...	BIGINT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
↳ estaciones_CODIGO_ESTACI...	VARCCHAR(10)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
↳ tipo_transaccion_CODIGO_T...	BIGINT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Figure 7: Tabla Transmilenio Total

4.4 Código SQL - lenguaje de definición de datos (DDL) (Primera entrega)

```
LOAD DATA INFILE
'C:\\ProgramData\\MySQL\\MySQL_Server_8.0\\Uploads\\
Estaciones.txt'
INTO TABLE ESTACIONES
FIELDS TERMINATED BY ',';
```

```
CREATE TABLE TRANSMILENIOTEMP (
FECHA                                CHAR(10),
SUCURSAL                            CHAR(50),
TIPO_AGENTE                          CHAR(50),
NOMBREAGENTE                        CHAR(80),
ESTACION                            CHAR(50),
ACCESO_ESTACION                     CHAR(100),
ID_POS                              CHAR(20),
TIPO_TRANSACCION                    CHAR(20),
CANTIDAD                            CHAR(10),
VALOR                               CHAR(20)
);
COMMIT;
```

```
CREATE TABLE DISPOSITIVOS (
ID_POS                              CHAR(20),
TIPO_AGENTE                          CHAR(60),
CODIGO_ESTACION                     CHAR(10),
ESTACION                            CHAR(50),
TAQUILLA                            CHAR(100),
FASE                                CHAR(10),
TIPO_DISPOSITIVO                     CHAR(10),
DESC_TIPO_DISPOSITIVO                CHAR(50),
ESTADO                              CHAR(10)
);
COMMIT;
```

```
CREATE TABLE TIPO_AGENTETEMP AS (
SELECT DISTINCT TIPO_AGENTE FROM TRANSMILENIODEF
);
COMMIT;
```

4.5 Código SQL - Manipulación de datos (DML) (Primera entrega)

```
SELECT * FROM TRANSMILENIOTOTAL;##10260365
SELECT COUNT(*) FROM TRANSMILENIOTOTAL;##10260365
```

```
SELECT SUM(CANTIDAD) CANTIDAD,SUM(VALOR) VALOR,
FECHA_AAAAMM, ESTACION, TAQUILLA,
DESC.TIPO_DISPOSITIVO ,CUADRANTE,ZONA
FROM TRANSMILENIOTOTAL
WHERE CUADRANTE IS NOT NULL
GROUP BY FECHA_AAAAMM, ESTACION, TAQUILLA,
DESC.TIPO_DISPOSITIVO ,CUADRANTE,ZONA;
```

```
SELECT
TRANSMILENIODEFNESTNDISP.FECHA, TRANSMILENIODEFNESTNDISP.FECHA_AAAAMM,
TRANSMILENIODEFNESTNDISP.FECHA_AAAA,
SUCURSALES.SUCURSAL, TIPO_AGENTE.tipo_agente ,
tipo_transaccion.TIPO_TRANSACCION,
TRIM(LEADING '0' FROM TRANSMILENIODEFNESTNDISP.ID_POS)
ID_POS, TRANSMILENIODEFNESTNDISP.CANTIDAD,
TRANSMILENIODEFNESTNDISP.VALOR,
TRIM(LEADING '0' FROM DISPOSITIVOS.CODIGO_ESTACION)
CODIGO_ESTACION, TRIM(DISPOSITIVOS.ESTACION)
ESTACION, TRIM(DISPOSITIVOS.TAQUILLA) TAQUILLA,
TRIM(DISPOSITIVOS.FASE)
FASE, TRIM(DISPOSITIVOS.TIPO_DISPOSITIVO)
TIPO_DISPOSITIVO, TRIM(DISPOSITIVOS.DESC_TIPO_DISPOSITIVO)
DESC_TIPO_DISPOSITIVO,
TRIM(DISPOSITIVOS.ESTADO) ESTADO
FROM
TRANSMILENIODEFNESTNDISP INNER JOIN SUCURSALES ON
TRIM(TRANSMILENIODEFNESTNDISP.CODIGO_SUCURSAL)=
TRIM(SUCURSALES.CODIGO_SUCURSAL)
INNER JOIN TIPO_AGENTE ON
TRIM(TRANSMILENIODEFNESTNDISP.CODIGO_TIPO_AGENTE)=
TRIM(TIPO_AGENTE.CODIGO_TIPO_AGENTE)
INNER JOIN TIPO_TRANSACCION ON
TRIM(TRANSMILENIODEFNESTNDISP.CODIGO_TIPO_TRANSACCION)=
TRIM(TIPO_TRANSACCION.CODIGO_TIPO_TRANSACCION)
LEFT JOIN DISPOSITIVOS ON TRIM(LEADING '0' FROM
TRANSMILENIODEFNESTNDISP.ID_POS)=TRIM(LEADING '0' FROM
DISPOSITIVOS.ID_POS)
);
COMMIT;
```

```
INSERT INTO TRANSMILENIODEF (  
SELECT * FROM TRANSMILENIOTEMP  
);  
COMMIT;
```

```
UPDATE TRANSMILENIOTEMP SET  
FECHA=REPLACE(FECHA, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
SUCURSAL=REPLACE(SUCURSAL, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
TIPO_AGENTE=REPLACE(TIPO_AGENTE, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
NOMBREAGENTE=REPLACE(NOMBREAGENTE, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
ESTACION=REPLACE(ESTACION, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
ACCESO_ESTACION=REPLACE(ACCESO_ESTACION, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
ID_POS=REPLACE(ID_POS, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
TIPO_TRANSACCION=REPLACE(TIPO_TRANSACCION, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
CANTIDAD=REPLACE(CANTIDAD, ' ' , ' ');  
UPDATE TRANSMILENIOTEMP SET  
VALOR=REPLACE(VALOR, ' ' , ' ');  
COMMIT;
```

- 4.6 Código SQL + Resultados: Vistas (*Primera entrega*)
- 4.7 Código SQL + Resultados: Triggers (*Primera entrega*)
- 4.8 Código SQL + Resultados: Funciones (*Primera entrega*)
- 4.9 Código SQL + Resultados: procedimientos almacenados (*Primera entrega*)

5 Bases de Datos No-SQL (*Segunda entrega*)

5.1 Diagrama Bases de Datos No-SQL (*Segunda entrega*)

5.2 SMBD utilizado para la Base de Datos No-SQL (*Segunda entrega*)

6 Aplicación de ETL (Extract, Transform, Load) y Bodega de Datos *(Tercera entrega)*

6.1 Ejemplo de aplicación de ETL y Bodega de Datos *(Tercera entrega)*

7 Lecciones aprendidas (*Tercera entrega*)

8 Bibliografía

M. ANDRADE *El sistema integrado de transporte masivo Transmilenio: Recopilación, Estudio y Análisis de los componentes del sistema; Proyecto para la elaboración de unas memorias de la primera fase de operación del sistema.*, 2003

M. CORRAL *Plan de inversión para la disminución de evasores del pago en las estaciones del sistema de transporte masivo Transmilenio*, 2021

L. RÍOS *Análisis de datos de sentimientos enfocados al servicio de transporte masivo Transmilenio S.A aplicando tecnologías Big Data*, 2019