

## Actividad Migración y testing 2



Nicolas Ballesteros Olarte

C.C 1028862982

[nicolasballesteros900@gmail.com](mailto:nicolasballesteros900@gmail.com)

Fecha 16/08/2025

Servicio Nacional De Aprendizaje

Tecnológico En Análisis y de Desarrollo de Software

Instructor: Paola Balle

## Introducción

El presente informe describe el proceso de migración de datos realizado en el marco de la formación en el área de Análisis y Desarrollo de Software del SENA. Se abordan los objetivos, las acciones realizadas, los requisitos cumplidos y los desafíos superados durante la migración. Asimismo, se presentan evidencias gráficas y resultados de las pruebas realizadas, finalizando con conclusiones individuales de los integrantes del equipo.

## Objetivos

### Objetivo general

Realizar el proceso de migración de datos hacia la base de datos ICFES de forma correcta, garantizando la integridad, consistencia y disponibilidad de la información..

### Objetivos específicos

- Documentar las acciones realizadas en cada fase del proceso de migración.
- Verificar el cumplimiento de los requisitos de migración mediante sentencias SQL y evidencias gráficas.
- Identificar y superar los desafíos técnicos y personales surgidos durante el proceso.

## Acciones

1. Creación de la base de datos en Workbench.
2. Definición de las tablas según el modelo.
3. Cambio del formato del archivo donde se encontraban los datos a CSV.
4. Comprobación de que los datos sean acordes al diccionario de datos.
5. Carga de datos mediante sentencias **LOAD DATA INFILE**.
6. Verificación de la integridad de los datos.
7. Normalización y ajustes en las tablas.
8. Generación de consultas SQL para cumplir con los requisitos de migración.

## Elementos

icfes\_saber\_11

ESTU\_CONSECUTIVO VARCHAR(255)

PERIODO VARCHAR(255)

ESTU\_TIPO\_DOCUMENTO VARCHAR(255)

ESTU\_PAIS\_RESIDE VARCHAR(255)

ESTU\_GENERO VARCHAR(255)

ESTU\_NACIMIENTO\_DIA INT(11)

ESTU\_NACIMIENTO\_MES INT(11)

ESTU\_NACIMIENTO\_ANNO INT(11)

ESTU\_EDAD INT(11)

ESTU\_LIMITA\_BAJAVISION VARCHAR(255)

ESTU\_LIMITA\_SORDOCEGUERA VARCHAR(255)

ESTU\_LIMITA\_COGNITIVA VARCHAR(255)

ESTU\_LIMITA\_INVIDENTE VARCHAR(255)

ESTU\_LIMITA\_MOTRIZ VARCHAR(255)

ESTU\_LIMITA\_SORDOINTERPRETE VARCHAR(255)

ESTU\_LIMITA\_SORDONOIDERPRETE VARCHAR(255)

ESTU\_ETNIA INT(11)

ESTU\_COD\_RESIDE\_MCPPIO INT(11)

ESTU\_RESIDE\_MCPPIO VARCHAR(255)

ESTU\_RESIDE\_DEPTO VARCHAR(255)

ESTU\_ZONA\_RESIDE VARCHAR(255)

ESTU\_AREA\_RESIDE VARCHAR(255)

IND\_COD\_ICFES\_TERMINO INT(11)

COLE\_COD\_ICFES INT(11)

COLE\_COD\_DANE\_INSTITUCION VARCHAR(255)

COLE\_NOMBRE\_SEDE VARCHAR(255)

COLE\_CALEDARIO VARCHAR(255)

COLE\_GENERO VARCHAR(255)

COLE\_NATURALEZA VARCHAR(255)

COLE\_BILINGUE VARCHAR(255)

84 more...

## Requisitos

### 1. Mostrar los mejor puntaje por cada materia

```
SELECT 'Lenguaje' AS Materia, t.ESTU_CONSECUTIVO,
t.PUNT LENGUAJE AS Puntaje
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT LENGUAJE
    FROM icfes_saber_11
    ORDER BY PUNT LENGUAJE DESC
    LIMIT 1
) t
UNION ALL
SELECT 'Matemáticas', t.ESTU_CONSECUTIVO, t.PUNT_MATEMATICAS
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT_MATEMATICAS
    FROM icfes_saber_11
    ORDER BY PUNT_MATEMATICAS DESC
    LIMIT 1
) t
UNION ALL
SELECT 'C. Sociales', t.ESTU_CONSECUTIVO, t.PUNT_C_SOCIALES
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT_C_SOCIALES
    FROM icfes_saber_11
    ORDER BY PUNT_C_SOCIALES DESC
    LIMIT 1
) t
UNION ALL
SELECT 'Filosofía', t.ESTU_CONSECUTIVO, t.PUNT_FILOSOFIA
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT_FILOSOFIA
    FROM icfes_saber_11
    ORDER BY PUNT_FILOSOFIA DESC
    LIMIT 1
) t
UNION ALL
SELECT 'Biología', t.ESTU_CONSECUTIVO, t.PUNT_BIOLOGIA
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT_BIOLOGIA
    FROM icfes_saber_11
```

```

ORDER BY PUNT_BIOLOGIA DESC
LIMIT 1
) t
UNION ALL
SELECT 'Química', t.ESTU_CONSECUTIVO, t.PUNT_QUIMICA
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT_QUIMICA
    FROM icfes_saber_11
    ORDER BY PUNT_QUIMICA DESC
    LIMIT 1
) t
UNION ALL
SELECT 'Física', t.ESTU_CONSECUTIVO, t.PUNT_FISICA
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT_FISICA
    FROM icfes_saber_11
    ORDER BY PUNT_FISICA DESC
    LIMIT 1
) t
UNION ALL
SELECT 'Inglés', t.ESTU_CONSECUTIVO, t.PUNT_INGLES
FROM (
    SELECT ESTU_CONSECUTIVO, PUNT_INGLES
    FROM icfes_saber_11
    ORDER BY PUNT_INGLES desc
    LIMIT 1
) t;

```

Materia	ESTU_CONSECUTIVO	Puntaje
Lenguaje	SABER1120211029172	91
Matemáticas	SABER1120211000595	117
C. Sociales	SABER1120211080758	90
Filosofía	SABER1120211008794	103
Biología	SABER1120211079417	100
Química	SABER1120211008782	92
Física	SABER1120211012075	124
Inglés	SABER1120211000342	100

## 2. Colegio con el mejor puntaje promedio

```
SELECT
```

```

COLE_NOMBRE_SEDE,
ROUND(
    AVG(
        (IFNULL(PUNT_LENGUAJE,0) +
        IFNULL(PUNT_MATEMATICAS,0) +
        IFNULL(PUNT_C_SOCIALES,0) +
        IFNULL(PUNT_FILOSOFIA,0) +
        IFNULL(PUNT_BIOLOGIA,0) +
        IFNULL(PUNT_QUIMICA,0) +
        IFNULL(PUNT_FISICA,0) +
        IFNULL(PUNT_INGLES,0)) / 8
    ) AS Promedio_Colegio
FROM icfes_saber_11
GROUP BY COLE_NOMBRE_SEDE
ORDER BY Promedio_Colegio DESC
LIMIT 1;

```

	COLE_NOMBRE_SEDE	Promedio_Colegio
►	COLEGIO FRANCISCO DE ALDANA	76

### 3. Cantidad de personas con nivel de inglés superior a B1

```

SELECT COUNT(*) AS Cantidad_Superior_B1
FROM icfes_saber_11
WHERE DESEMP_INGLES = 'B+';

```

	Cantidad_Superior_B1
►	5045

4. Municipio o ciudad con mayor cantidad de personas con nivel de inglés superior a B1

```
SELECT
    ESTU_MCPIO_PRESENTACION AS Municipio,
    COUNT(*) AS Cantidad_Superior_B1
FROM icfes_saber_11
WHERE DESEMP_INGLES = 'B+'
GROUP BY ESTU_MCPIO_PRESENTACION
ORDER BY Cantidad_Superior_B1 DESC
LIMIT 1;
```

	Municipio	Cantidad_Superior_B1
►	BOGOTÁ D.C.	2313

5. Los cinco colegios no bilingües con mejor puntaje promedio en inglés

```
SELECT
    COLE_NOMBRE_SEDE,
    round(AVG (PUNT_INGLES)) AS Promedio_Ingles
FROM icfes_saber_11
WHERE COLE_BILINGUE = '0'
GROUP BY COLE_NOMBRE_SEDE
ORDER BY Promedio_Ingles DESC
LIMIT 5;
```

	COLE_NOMBRE_SEDE	Promedio_Ingles
►	LICEO DE LONDRES	100
	COLEGIO AGUSTINIANO DE SAN NICOLAS	100
	COL MARYMOUNT	100
	GIMNASIO SANTA ANA	92
	INSTITUTO JUVENIL COLOMBIANO	92

6. Persona con mejor puntaje en matemáticas por género

```
SELECT
    ESTU_CONSECUTIVO,
    ESTU_GENERO,
    MAX(PUNT_MATEMATICAS) AS PUNT_MATEMATICAS_MAX
FROM icfes_saber_11
GROUP BY ESTU_GENERO;
```

ESTU_CONSECUTIVO	ESTU_GENERO	PUNT_MATEMATICAS_MAX
SABER1120211000001	M	117
SABER1120211000003	F	117

## Desafíos

El primer desafío fue convertir los datos a un formato que permitiera registrarlos de manera rápida en la base de datos, ya que se encontraban en un archivo TXT. Además, en dicho archivo los atributos o los títulos de las columnas no coincidían con el diccionario de datos, o había más de los necesarios, por lo que fue necesario comparar y verificar que los datos correspondieran.

## Pruebas

Se realizaron pruebas en cada fase del proceso:

- Qué se probó: la correcta carga de datos desde CSV a MySQL.
- Que cada consulta correspondiera con los datos de la base de datos.
- Comparar el diccionario de datos con los datos ingresados a la base de datos.
- Resultados: los datos se migraron exitosamente, confirmando integridad y consistencia.

## Conclusiones

- Como conclusión, considero que lo más difícil fue saber por dónde empezar y cómo convertir los datos a un formato adecuado para que la base de datos los aceptara. Después de lograr insertar los datos, el resto fue fácil de completar para cumplir con los requisitos.