



Taxi\_Bot

# ENTREGA **FINAL**

Realizado por Julián Quintero

# TABLA DE CONTENIDO

• Presentación	01
• Contexto Empresa	02
• Análisis de Mejora	03
• Información Manejada	04
• Kpi's	05
• Procesamiento de Datos	06
• Valor de la Información	07
• Conclusiones	08

# CONTEXTO

# EMPRESA TAXIMED

Taximed ha estado operando en un entorno tradicional usando herramientas como Excel y llamadas telefónicas para la gestión de sus operaciones diarias. Sin embargo, con el crecimiento de la demanda y la necesidad de gestionar grandes cantidades de datos de forma más eficiente, la empresa ha decidido optar por modernizarse y comenzar a usar la aplicación de viajes y una base de datos SQL para mejorar la gestión y el análisis de la información.



# DATA QUE UTILIZAN

## 01

### Usuarios

---

- > Cedula
- > Nombre
- > Dirección
- > Telefono, email
- > Viajes
- > Calificaciones a Conductores

## 02

### Conductores

---

- > Nombre
- > Cedula
- > Vehiculo
- > Historial de Viajes
- > Calificacion Promedio

## 03

### Viajes

---

- > Fecha
- > Origen, Destino y Distancia
- > Duración Viaje.
- > Costo Viaje
- > Método de Pago



# KPI'S

1

Satisfacción del usuario (promedio de calificación del usuario para cada conductor y viaje).

2

Distancia promedio por viaje (utilizado para optimizar las rutas).

3

Rentabilidad por conductor (ingresos generados por conductor en un período determinado).



# KPI'S

4

Crecimiento en la cantidad de viajes (número total de viajes realizados por mes/trimestre).

5

Promedio de ingresos por viaje (medición del ingreso promedio por cada servicio prestado).





# KPI'S

usuario_id	nombre	direccion	telefono	metodo_pago	promedio_califica
1558944834	Aaron Allen III	31 Trujillo RestGinatn, MH 940	307725556	Tarjeta	4,8
1989947539	Aaron Brown	n Dam Suite 973Valdezchester, A	309554180	Tarjeta	3,2
1495277925	Aaron Butler	ackson LakesWallacechester, IL C	308018477	Tarjeta	3,5
528806159	Aaron Campbell	son Mountains Apt. 005Cainshi	318610945	Efectivo	4
95091041	Aaron Chung	urgess OrchardPort Lauren, AR	308727719	Efectivo	4,8
833610132	Aaron Coleman	ez Square Apt. 131Lindaview, VA	316481408	Efectivo	4,5
1103150258	Aaron Delgado	n Gardens Apt. 581Smithfort, W	303804820	Efectivo	4,9
1032857060	Aaron Evans	on Spurs Suite 221East Jacob, k	315629968	Efectivo	3,5
1952662970	Aaron Farrell	3 Laura SpringLake Ryan, GA 005	313767941	Tarjeta	4
1072295890	Aaron Fernandez	79 Eddie CampDavidville, CT 183	310933412	Efectivo	4,6
530747768	Aaron Frost	on Mission Apt. 234Bridgesview	310110982	Tarjeta	4
718265755	Aaron Graham	ea Village Apt. 080North Regina	316906985	Tarjeta	3
600490607	Aaron Greene	ohnson RidgesWarrenstad, OH 6	321903261	Efectivo	4,8
1343776245	Aaron Higgins	Sydney StreamClarkberg, PA 6	318894956	Tarjeta	3,9
469341120	Aaron Hoffman	as GardensNorth Anthony, MT	317699270	Efectivo	3,4
1480212165	Aaron Jensen	os EstatesSouth Emilyburgh, NC	309368864	Efectivo	4,4
1150709418	Aaron Johnson	Wells Suite 758North Tracymou	314683900	Tarjeta	4,4
1951996763	Aaron Joseph	ndsay HarborSouth Carlberg, MI	311170101	Tarjeta	3,9
1496560709	Aaron Lucas	ond Cove Apt. 907Stevenmout	310161860	Efectivo	3,2
1477871468	Aaron Mathis	3 Carr TrailGregoryfurt, MS 998	316948608	Efectivo	3,3
361917881	Aaron Mckay	Mountain Apt. 947New Jonathan	321439415	Tarjeta	3,3
1718352229	Aaron Mckay	s Parkways Suite 973Karenland	313528432	Tarjeta	4,9
1517867730	Aaron Miller	on Spurs Apt. 075Port Erica, MT	307028858	Efectivo	3,6
1596894770	Aaron Ramsey	Meadow Suite 277Mckeestad, I	313353784	Efectivo	4,8
982694823	Aaron Reed	Bush MallNorth Sheilafurt, MD	318299635	Efectivo	3,9
1039246634	Aaron Rogers	h FlatSouth Michaelchester, ND	318338873	Tarjeta	4
1281279292	Aaron Rogers	da Ridge Suite 731Parkview, PA	306905742	Efectivo	4,8
813250864	Aaron Scott	h Flat Suite 411New Tracynport	316549646	Tarjeta	4,7

## ANTES

# COMO SE PROCESAN LOS DATOS?

- Ingreso manual de datos: Los datos de los viajes, conductores, y usuarios se introducen manualmente en hojas de cálculo de Excel.
- Análisis de datos: Usan funciones y gráficos en Excel para realizar análisis básicos, como el cálculo de promedios, totales y gráficos de rendimiento (aunque esto puede ser ineficiente con grandes volúmenes de datos).
- Informes periódicos: La empresa genera informes sobre la cantidad de viajes, ingreso calificaciones, etc., de forma manual, lo que puede ser propenso a errores y lleva mucho tiempo.



**AHORA**

# COMO SE PROCESAN LOS DATOS?

- Automatización del ingreso de datos: Al adoptar una aplicación la información se ingresará de forma automática, Con la base de datos SQL
- Consultas SQL: Los analistas y gerentes podrán ejecutar consultas SQL para obtener insights rápidos y detallados de la información almacenada.
- Dashboards en tiempo real: Usando herramientas de Business Intelligence (como Power BI o Tableau), la empresa puede visualizar los KPIs en tiempo real, ayudando a los gerentes a tomar decisiones basadas en datos.
- Optimización de rutas y asignación de conductores: La aplicación puede incluir algoritmos que optimicen las rutas para los conductores en función del tráfico, la ubicación de los vehículos, y la demanda de los usuarios.

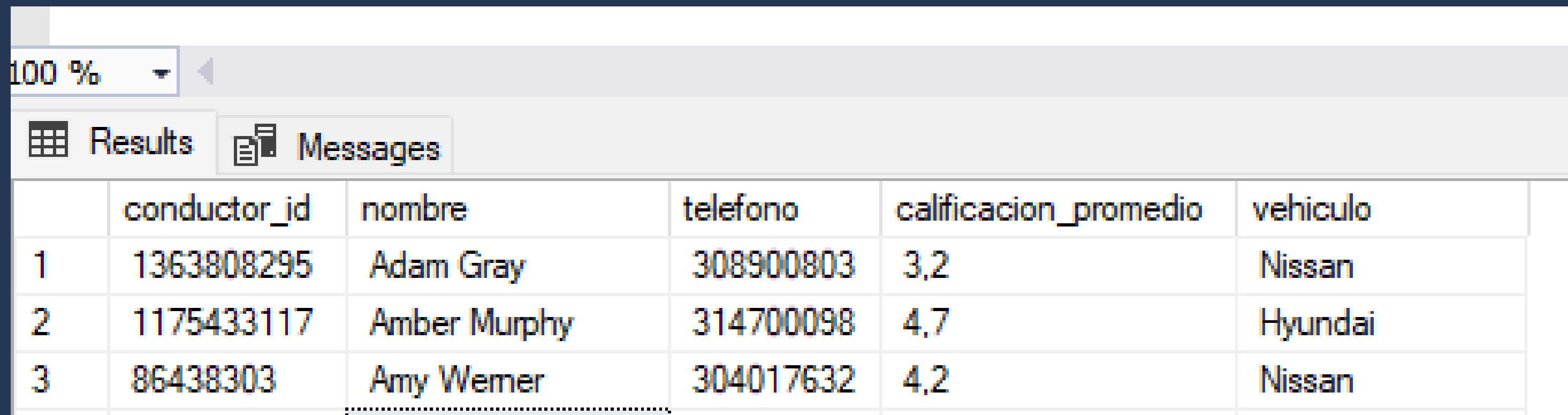
ANTES

# COMO SE PROCESAN LOS DATOS?

usuario_id	nombre	direccion	telefono	metodo_pago	promedio_calificacion
1558944834	Aaron Allen III	31 Trujillo RestGinaton, MH 940	307725556	Tarjeta	4,8
1989947539	Aaron Brown	n Dam Suite 973Valdezchester, A	309554180	Tarjeta	3,2
1495277925	Aaron Butler	ackson LakesWallacechester, IL C	308018477	Tarjeta	3,5
528806159	Aaron Campbell	son Mountains Apt. 005Cainshi	318610945	Efectivo	4
95091041	Aaron Chung	urgess OrchardPort Lauren, AR	308727719	Efectivo	4,8
833610132	Aaron Coleman	ez Square Apt. 131Lindaview, VA	316481408	Efectivo	4,5
1103150258	Aaron Delgado	n Gardens Apt. 581Smithfort, W	303804820	Efectivo	4,9
1032857060	Aaron Evans	on Spurs Suite 221East Jacob, K	315629968	Efectivo	3,5
1952662970	Aaron Farrell	3 Laura SpringLake Ryan, GA 005	313767941	Tarjeta	4
1072295890	Aaron Fernandez	79 Eddie CampDavidville, CT 185	310933412	Efectivo	4,6
530747768	Aaron Frost	on Mission Apt. 234Bridgesview	310110982	Tarjeta	4
718265755	Aaron Graham	ea Village Apt. 080North Regina	316906985	Tarjeta	3
600490607	Aaron Greene	ohnson RidgesWarrenstad, OH 6	321903261	Efectivo	4,8
1343776245	Aaron Higgins	Sydney StreamClarkberg, PA 6	318894956	Tarjeta	3,9

AHORA

# COMO SE PROCESAN LOS DATOS?



A screenshot of a data processing interface. At the top, there is a zoom control set to '100 %'. Below it are two tabs: 'Results' (active) and 'Messages'. The 'Results' tab displays a table with six columns: an index column, 'conductor\_id', 'nombre', 'telefono', 'calificacion\_promedio', and 'vehiculo'. The table contains three rows of data.

	conductor_id	nombre	telefono	calificacion_promedio	vehiculo
1	1363808295	Adam Gray	308900803	3,2	Nissan
2	1175433117	Amber Murphy	314700098	4,7	Hyundai
3	86438303	Amy Werner	304017632	4,2	Nissan



# COMO SE HIZO



Investigación



Análisis

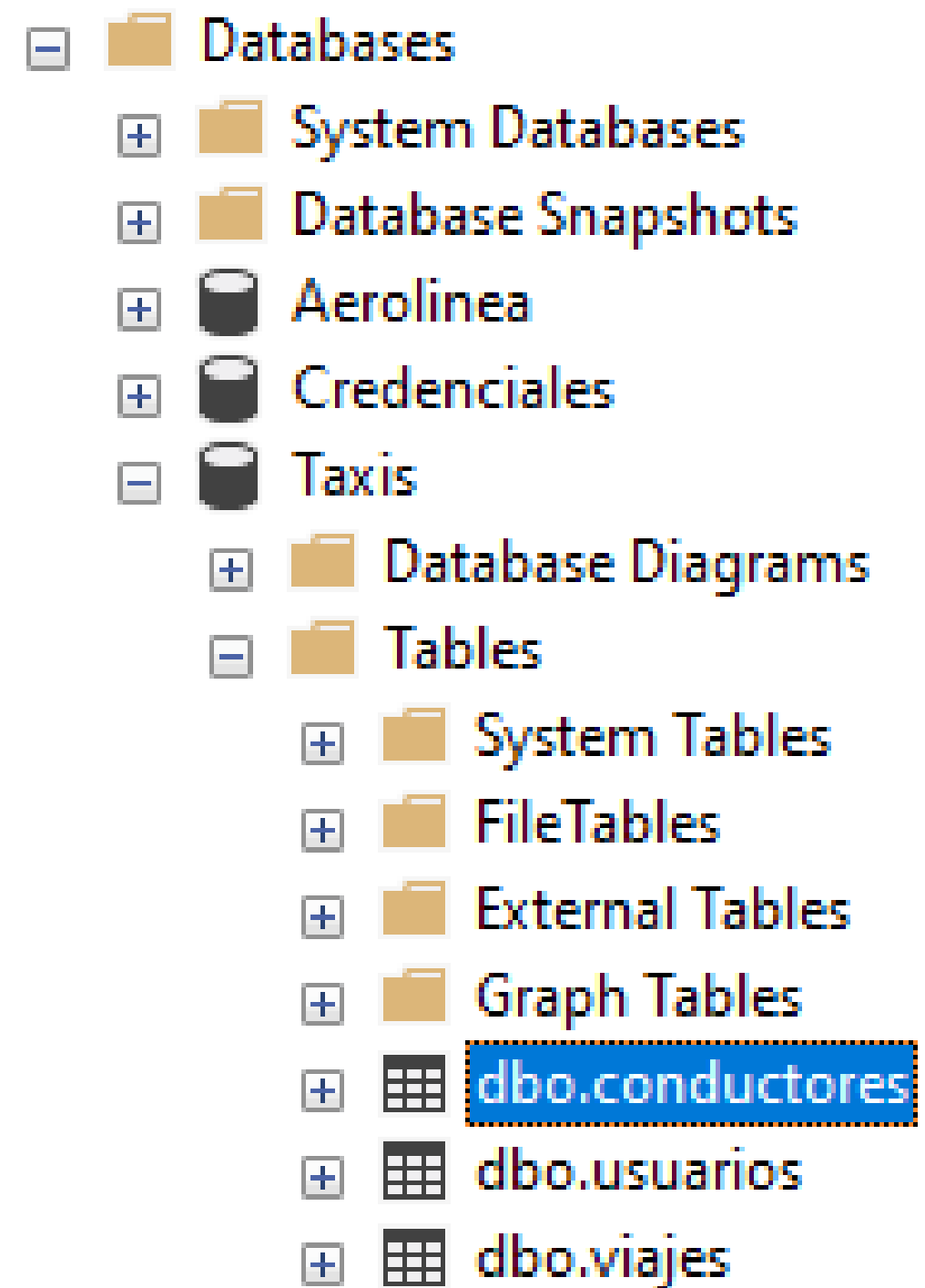


Diseño



Implementación

# COMO SE HIZO



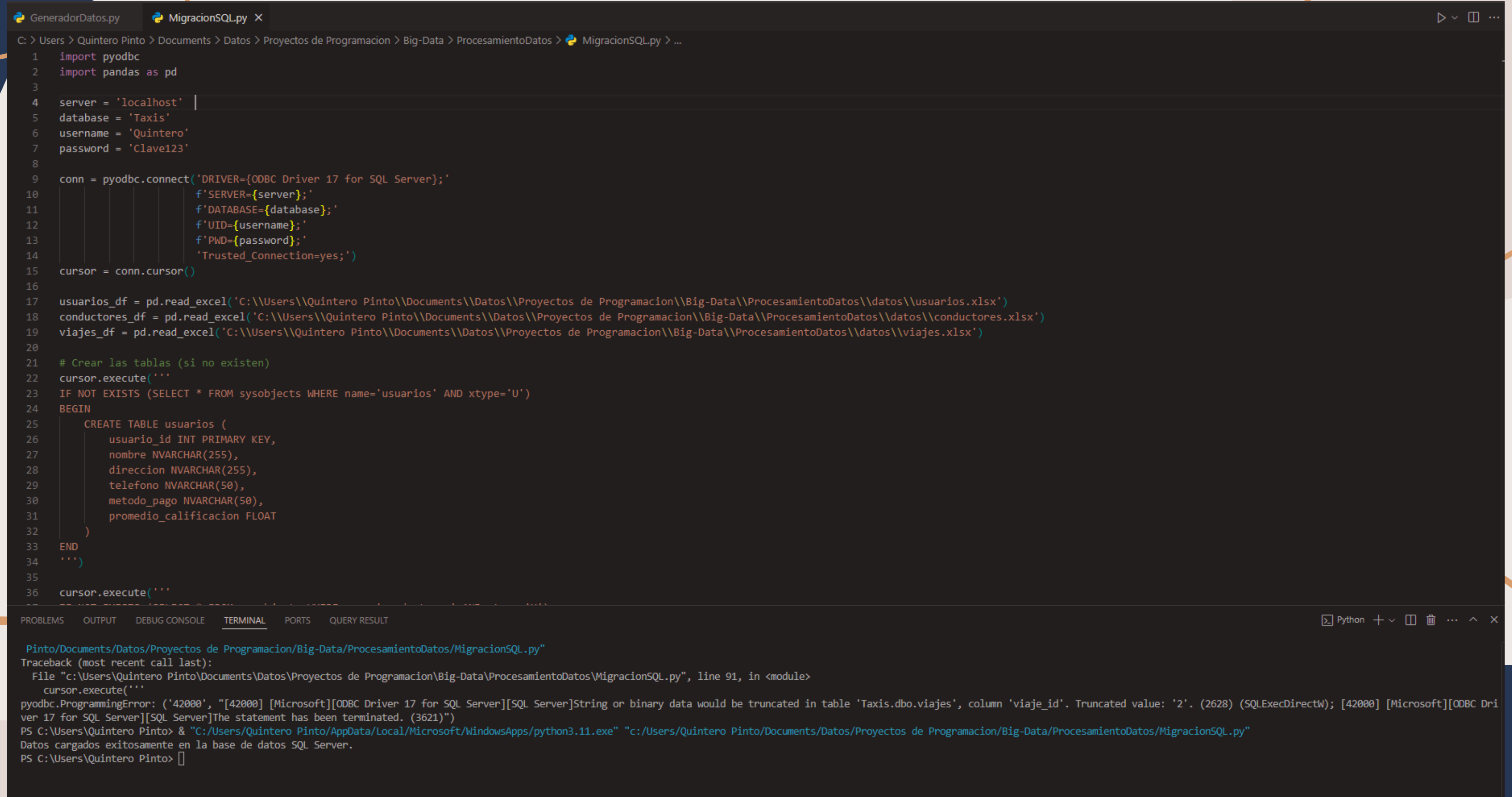
# COMO SE HIZO

```
Analysis.py X
C: > Users > Quintero > Desktop > Proyectos Programación > Analisis de datos python > Analysis.py > ...
14 df_final = pd.concat(dataframes, ignore_index=True)
15
16 df_final = df_final.drop_duplicates()
17
18 df_final = df_final.groupby('cliente_id', as_index=False).first()
19
20 df_final.to_csv('Archivo_Unificado.csv', index=False)
21
22 #Depuracion de la columna edad
23
24 df_final['edad'] = pd.to_numeric(df_final['edad'], errors='coerce') # Convertir a numérico, NaN para errores
25
26 df_final['edad'] = df_final['edad'].round().astype('Int64') # Redondear y convertir a tipo entero (Int64 para manejar NaN)
27
28 edades_validas = df_final['edad'][(df_final['edad'] >= 0) & (df_final['edad'] <= 120)]
29
30 promedio_edad = edades_validas.mean()
31
32 df_final['edad'] = df_final['edad'].fillna(promedio_edad) # Reemplazar NaN con el promedio
33
34 df_final['edad'] = df_final['edad'].apply(
35     lambda x: promedio_edad if x < 0 or x > 120 else x
36 )
37
38 #Depuracion de la columna Salario
39
40 df_final['salario'] = df_final['salario'].replace('???', pd.NA)
41
42 df_final['salario'] = pd.to_numeric(df_final['salario'], errors='coerce')
43
```

Ln 1, Col 1 Spaces: 4 UTF-8 CRLF {} Python 3.13.0 64-bit



# COMO SE HIZO



The image shows a Visual Studio Code editor with two tabs: 'GeneradorDatos.py' and 'MigracionSQL.py'. The active tab is 'MigracionSQL.py', which contains a Python script for connecting to a SQL Server database and loading data from Excel files. The script defines connection parameters (server, database, username, password) and uses the pyodbc library to establish a connection. It then reads three Excel files: 'usuarios.xlsx', 'conductores.xlsx', and 'viajes.xlsx'. The script includes SQL commands to create tables if they don't exist and to insert data into the 'usuarios' table. The terminal at the bottom shows the execution of the script, displaying a traceback error and a message indicating that the data was successfully loaded into the SQL Server database.

```
C: > Users > Quintero Pinto > Documents > Datos > Proyectos de Programacion > Big-Data > ProcesamientoDatos > MigracionSQL.py > ...

1  import pyodbc
2  import pandas as pd
3
4  server = 'localhost'
5  database = 'Taxis'
6  username = 'Quintero'
7  password = 'Clave123'
8
9  conn = pyodbc.connect('DRIVER={ODBC Driver 17 for SQL Server};'
10                        f'SERVER={server};'
11                        f'DATABASE={database};'
12                        f'UID={username};'
13                        f'PWD={password};'
14                        'Trusted_Connection=yes;')
15  cursor = conn.cursor()
16
17  usuarios_df = pd.read_excel('C:\\Users\\Quintero Pinto\\Documents\\Datos\\Proyectos de Programacion\\Big-Data\\ProcesamientoDatos\\datos\\usuarios.xlsx')
18  conductores_df = pd.read_excel('C:\\Users\\Quintero Pinto\\Documents\\Datos\\Proyectos de Programacion\\Big-Data\\ProcesamientoDatos\\datos\\conductores.xlsx')
19  viajes_df = pd.read_excel('C:\\Users\\Quintero Pinto\\Documents\\Datos\\Proyectos de Programacion\\Big-Data\\ProcesamientoDatos\\datos\\viajes.xlsx')
20
21  # Crear las tablas (si no existen)
22  cursor.execute('''
23  IF NOT EXISTS (SELECT * FROM sysobjects WHERE name='usuarios' AND xtype='U')
24  BEGIN
25      CREATE TABLE usuarios (
26          usuario_id INT PRIMARY KEY,
27          nombre NVARCHAR(255),
28          direccion NVARCHAR(255),
29          telefono NVARCHAR(50),
30          metodo_pago NVARCHAR(50),
31          promedio_calificacion FLOAT
32      )
33  END
34  ''')
35
36  cursor.execute('''
37  INSERT INTO usuarios (usuario_id, nombre, direccion, telefono, metodo_pago, promedio_calificacion)
38  SELECT usuario_id, nombre, direccion, telefono, metodo_pago, promedio_calificacion
39  FROM usuarios_df
40  ''')
```

PROBLEMS OUTPUT DEBUG CONSOLE **TERMINAL** PORTS QUERY RESULT

Python + - [ ] [ ] ... ^ x

Pinto/Documents/Datos/Proyectos de Programacion/Big-Data/ProcesamientoDatos/MigracionSQL.py"

Traceback (most recent call last):

File "c:\Users\Quintero Pinto\Documents\Datos\Proyectos de Programacion\Big-Data\ProcesamientoDatos\MigracionSQL.py", line 91, in <module>

cursor.execute(''

pyodbc.ProgrammingError: ('42000', "[42000] [Microsoft][ODBC Driver 17 for SQL Server][SQL Server]String or binary data would be truncated in table 'Taxis.dbo.viajes', column 'viaje\_id'. Truncated value: '2'. (2628) (SQLExecDirectW); [42000] [Microsoft][ODBC Driver 17 for SQL Server][SQL Server]The statement has been terminated. (3621)")

PS C:\Users\Quintero Pinto> & "C:/Users/Quintero Pinto/AppData/Local/Microsoft/WindowsApps/python3.11.exe" "c:/Users/Quintero Pinto/Documents/Datos/Proyectos de Programacion/Big-Data/ProcesamientoDatos/MigracionSQL.py"

Datos cargados exitosamente en la base de datos SQL Server.

PS C:\Users\Quintero Pinto> [ ]

# VALOR DE LOS DATOS

- Optimización de rutas: Usando algoritmos de análisis de datos en tiempo real, la empresa puede optimizar las rutas de los taxis, reduciendo el tiempo de viaje y mejorando la eficiencia operativa. Esto no solo mejora la experiencia del usuario, sino que también reduce los costos operativos.
- Personalización de la experiencia del usuario: Los datos de los usuarios y su historial de viajes permiten crear recomendaciones personalizadas, ofertas y descuentos, mejorando la retención de clientes y la lealtad a la marca.
- Monitoreo en tiempo real: Con la nueva infraestructura, los gerentes pueden monitorear la disponibilidad de los conductores y el estado de los vehículos en tiempo real, lo que permite una mejor asignación de recursos y reducción de tiempos de espera.

# VALOR DE LOS DATOS

- Análisis predictivo: Al contar con datos históricos de viajes, la empresa puede realizar análisis predictivo para identificar patrones de demanda, prever picos de viajes (por ejemplo, en días festivos o durante ciertas horas del día) y asignar recursos adecuadamente.
- Mejora de la rentabilidad: Al analizar los costos y las ganancias por viaje, la empresa puede identificar qué conductores y rutas son más rentables, lo que les permite optimizar la asignación de vehículos y mejorar la rentabilidad general.
- Toma de decisiones informadas: Gracias al acceso a datos actualizados y detallados en tiempo real, los gerentes pueden tomar decisiones informadas sobre la estrategia de la empresa, como el ajuste de tarifas, la expansión de la flota de taxis, o la promoción de ciertos servicios.



# VALOR DE LOS DATOS


fecha	hora	distancia_km	tiempo_minutos
8/05/2024	17:47:17	30,32	115
8/05/2024	23:12:52	15,73	52
8/05/2024	16:35:21	29,71	47
8/05/2024	20:31:36	28,89	70
8/05/2024	19:46:38	32,15	71
8/05/2024	00:05:18	18,7	102
8/05/2024	03:18:38	43,56	84
8/05/2024	10:40:50	32,82	22
8/05/2024	19:43:48	18,48	9
8/05/2024	22:21:16	39,55	49
8/05/2024	07:55:20	8,5	115
8/05/2024	19:11:25	45,13	54
8/05/2024	23:52:01	26,27	113
8/05/2024	19:34:56	23,07	69
8/05/2024	12:53:29	46,39	80
8/05/2024	07:55:14	40,2	84
8/05/2024	19:23:28	47,98	38
8/05/2024	01:32:09	14,76	27



# CONCLUSIONES

La migración de Excel a una solución basada en SQL y una aplicación de viajes ofrece grandes ventajas en términos de eficiencia operativa, toma de decisiones basada en datos y optimización de la experiencia del cliente. Los KPI's se vuelven más accesibles y la información en tiempo real permite que la empresa de taxis ajuste su estrategia de manera ágil y eficaz.

Este cambio también permite que la empresa aproveche el análisis predictivo, la automatización de procesos, y el análisis de datos en tiempo real, generando valor tanto para la empresa como para sus clientes.





# MUCHAS GRACIAS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vulputate nulla at ante rhoncus, vel efficitur felis condimentum. Proin odio odio.



911-234-5678



[www.unsitiogenial.es](http://www.unsitiogenial.es)



# Página de Recursos

