

MANUAL DE USUARIO

Proyecto ETL: Análisis de Datos y Cálculo de Facturación

1. INTRODUCCIÓN

Este proyecto implementa un proceso ETL (Extract, Transform, Load) que permite:

- Cargar datos desde una base de datos SQLite.
- Realizar un análisis exploratorio de datos, generando estadísticas y visualizaciones.
- Calcular la facturación de diversas empresas aplicando reglas contractuales específicas.
- Exportar los resultados a archivos Excel.
- Enviar los reportes generados por correo electrónico mediante el servidor SMTP de Gmail.

Este manual está diseñado para guiarte en la instalación, configuración y uso del sistema.

2. REQUISITOS DEL SISTEMA

- Antes de comenzar, asegúrate de tener lo siguiente:
- Python 3.x instalado.
- Librerías de Python:
 - Pandas

- Numpy
- matplotlib

(Además, se utilizan las librerías estándar sqlite3, logging, smtplib, email y datetime)

- Archivo de base de datos SQLite llamado "database.sqlite" ubicado en la carpeta "Datos/" con las tablas "apicall" y "commerce".
- Cuenta de Gmail para enviar reportes. Si usas autenticación de dos factores, necesitarás configurar una contraseña de aplicación.

3. INSTALACIÓN Y CONFIGURACIÓN

a) Clonar el repositorio

- Abre una terminal y ejecuta:

```
git clone git@github.com:Asterios1/Prueba_tecnica_analista_junior_bancolombia.git
```

```
cd Prueba_tecnica_analista_junior_bancolombia
```

b) Instalar las dependencias

- Si no tienes instaladas las librerías necesarias, ejecuta:

```
pip install pandas numpy matplotlib
```

c) Estructura del proyecto

Asegúrate de contar con la siguiente estructura de directorios:

- /Datos/

database.sqlite (Archivo de la base de datos SQLite)

- /reportes/

(Directorio donde se guardarán los reportes Excel)

- /analisis_resultados/

(Directorio para guardar las visualizaciones generadas)

- /logs/

(Directorio para los archivos de log)

- main.py (Script principal)

- README y/o manual (Documentación del proyecto)

4. EJECUCIÓN DEL PROGRAMA

a) Ejecutar el script principal

- Navega a la carpeta del proyecto y ejecuta:

```
python main.py
```

b) Interacción en consola

Durante la ejecución, se te solicitará lo siguiente:

- Selección de meses: Ingresa los números de los meses que deseas analizar (por ejemplo, "7,8" para julio y agosto). Se mostrará el nombre de cada mes para confirmar la selección.
- Ingreso de correos: Se te pedirá la cantidad de destinatarios y luego deberás ingresar cada dirección de correo, validando el formato básico.
- Configuración del correo: Si no se han definido las credenciales en el script, se solicitará ingresar tu dirección de Gmail y contraseña (o contraseña de aplicación).

c) Generación y exportación de reportes

- El programa creará dos archivos Excel:
 - Reporte de Análisis de Datos: Contiene estadísticas, detección de valores faltantes, distribuciones y visualizaciones.
 - Reporte de Facturación: Resume el cálculo de facturación por empresa, aplicando tarifas, descuentos e IVA.

• Los reportes se guardarán en la carpeta "reportes/" y se nombrarán según los meses analizados y la fecha de ejecución.

d) Envío de reportes por correo

- Una vez generados, los reportes se enviarán a las direcciones de correo ingresadas utilizando el servidor SMTP de Gmail.

5. DESCRIPCIÓN DEL FLUJO DE TRABAJO

El proceso ETL se divide en las siguientes etapas:

- Carga de datos: Conexión a la base de datos SQLite y extracción de información de las tablas "apicall" y "commerce".
- Análisis exploratorio: Cálculo de estadísticas descriptivas, detección de valores faltantes y análisis de distribuciones (por ejemplo, estado de las llamadas y estado de los comercios). Se generan visualizaciones como gráficos de barras, pastel y líneas.
- Cálculo de facturación: Agrupación de datos por comercio, conteo de llamadas exitosas y fallidas, y aplicación de reglas de facturación específicas (tarifas, descuentos, IVA).
- Exportación de resultados: Los resultados se guardan en archivos Excel con varias hojas para facilitar la interpretación.
- Envío de reportes: Los archivos Excel se envían por correo utilizando Gmail.

6. CONFIGURACIÓN DEL CORREO ELECTRÓNICO

Para el envío de correos:

- Asegúrate de que la cuenta de Gmail permita el acceso a aplicaciones externas o utiliza una contraseña de aplicación (si tienes autenticación en dos pasos).

- Durante la ejecución, ingresa tus credenciales cuando se te solicite, si no están predefinidas en el script.

7. SOLUCIÓN DE PROBLEMAS

a) Error de conexión a la base de datos:

- Verifica que el archivo "database.sqlite" esté en la carpeta "Datos/" y que la ruta en el script sea correcta.

b) Errores en el análisis o facturación:

- Consulta el archivo de log en la carpeta "logs/" para obtener detalles del error y revisa la integridad de los datos en la base de datos.

c) Problemas con el envío de correo:

- Asegúrate de ingresar correctamente las credenciales de Gmail.
- Revisa la configuración de seguridad de tu cuenta para permitir el acceso desde aplicaciones externas.