

Título del Proyecto: Proceso ETL para Análisis de Vuelos

1. Resumen Ejecutivo

Este proyecto se centra en la implementación de un proceso ETL (Extracción, Transformación y Carga) para analizar datos relacionados con vuelos y pasajeros. Los estudiantes desarrollarán una aplicación de consola en Python que permitirá gestionar el proceso ETL completo, desde la extracción de datos hasta la generación de reportes específicos, utilizando SQL Server como base de datos principal para el almacenamiento y análisis de la información.

3. Objetivos del Aprendizaje

3.1 Objetivo General

El estudiante será capaz de implementar un proceso ETL completo, utilizando Python y SQL Server, para extraer, transformar y analizar datos de vuelos, demostrando la capacidad de procesar información y generar resultados analíticos significativos.

3.2 Objetivos Específicos

1. Utilizar el lenguaje Python para desarrollar una aplicación de consola que gestione el proceso ETL completo.
2. Aplicar técnicas de limpieza y transformación de datos para garantizar la calidad de la información.
3. Implementar un modelo de datos en SQL Server que permita el almacenamiento y análisis eficiente de la información de vuelos.

4. Enunciado del Proyecto

4.1 Descripción del problema a resolver

Se requiere procesar y analizar datos relacionados con vuelos y pasajeros mediante un proceso ETL. Los datos necesitan ser extraídos de archivos fuente, transformados según reglas específicas y cargados en un modelo de datos optimizado para el análisis.

4.2 Alcance del proyecto

La aplicación debe incluir las siguientes funcionalidades:

- Borrar modelo existente
- Crear nuevo modelo de datos
- Extraer información de archivos fuente

- Cargar información transformada al modelo
- Realizar consultas analíticas específicas

4.3 Requerimientos técnicos

- Desarrollo en Python para la aplicación de consola
- Utilización de SQL Server para el almacenamiento de datos
- Capacidad de procesar archivos de entrada específicos
- Implementación de consultas analíticas predefinidas

4.4 Entregables

1. Aplicación de consola en Python
2. Scripts SQL para la creación y gestión del modelo de datos
3. Documentación técnica que incluya:
 - Diagrama del modelo realizado
 - Manual técnico con descripción del modelo (tablas de hechos y dimensiones)
4. Resultados de las consultas solicitadas:
 - Conteo de registros por tabla
 - Análisis de pasajeros por género
 - Estadísticas de nacionalidades por fecha
 - Conteo de vuelos por país
 - Top 5 aeropuertos por número de pasajeros
 - Análisis de estados de vuelo
 - Top 5 países más visitados
 - Top 5 continentes más visitados
 - Análisis de edades por género
 - Conteo de vuelos por mes-año

5. Metodología

1. Desarrollo de la aplicación:
 - Implementación de la interfaz de consola
 - Desarrollo de funcionalidades ETL
 - Creación de scripts SQL
2. Modelado de datos:
 - Diseño del modelo dimensional
 - Implementación en SQL Server
 - Validación del modelo
3. Implementación de consultas:
 - Desarrollo de consultas analíticas
 - Validación de resultados
 - Generación de reportes

7. Cronograma

Tipo	Fecha Inicio	Fecha Fin
Asignación de Proyecto	-	-
Elaboración	-	Domingo 04/08/2024 23:59
Calificación	-	Por definir

8. Evaluación

La evaluación considerará:

- Funcionalidad completa de la aplicación de consola
- Correcta implementación del modelo de datos
- Precisión en las consultas y análisis
- Calidad de la documentación técnica
- Cumplimiento de las restricciones especificadas

Notas importantes:

- No se aceptarán entregas después de la fecha límite
- Las copias recibirán una nota de 0 y serán reportadas a la Escuela
- El entregable debe seguir el formato: [SS2]Practica1_carnet.rar o repositorio correspondiente