

**UNIVERSIDAD DE EL SALVADOR**  
**FACULTAD DE INGENIERIA Y ARQUITECTURA**  
**ESCUELA DE INGENIERIA DE SISTEMAS INFORMATICOS**  
**ESPECIALIZACION: INGENIERO DE DATOS**  
**MODULO 3: PROCESAMIENTO DE DATOS**  
**PROFESOR: ING. RENE FABRICIO QUINTANILLA**



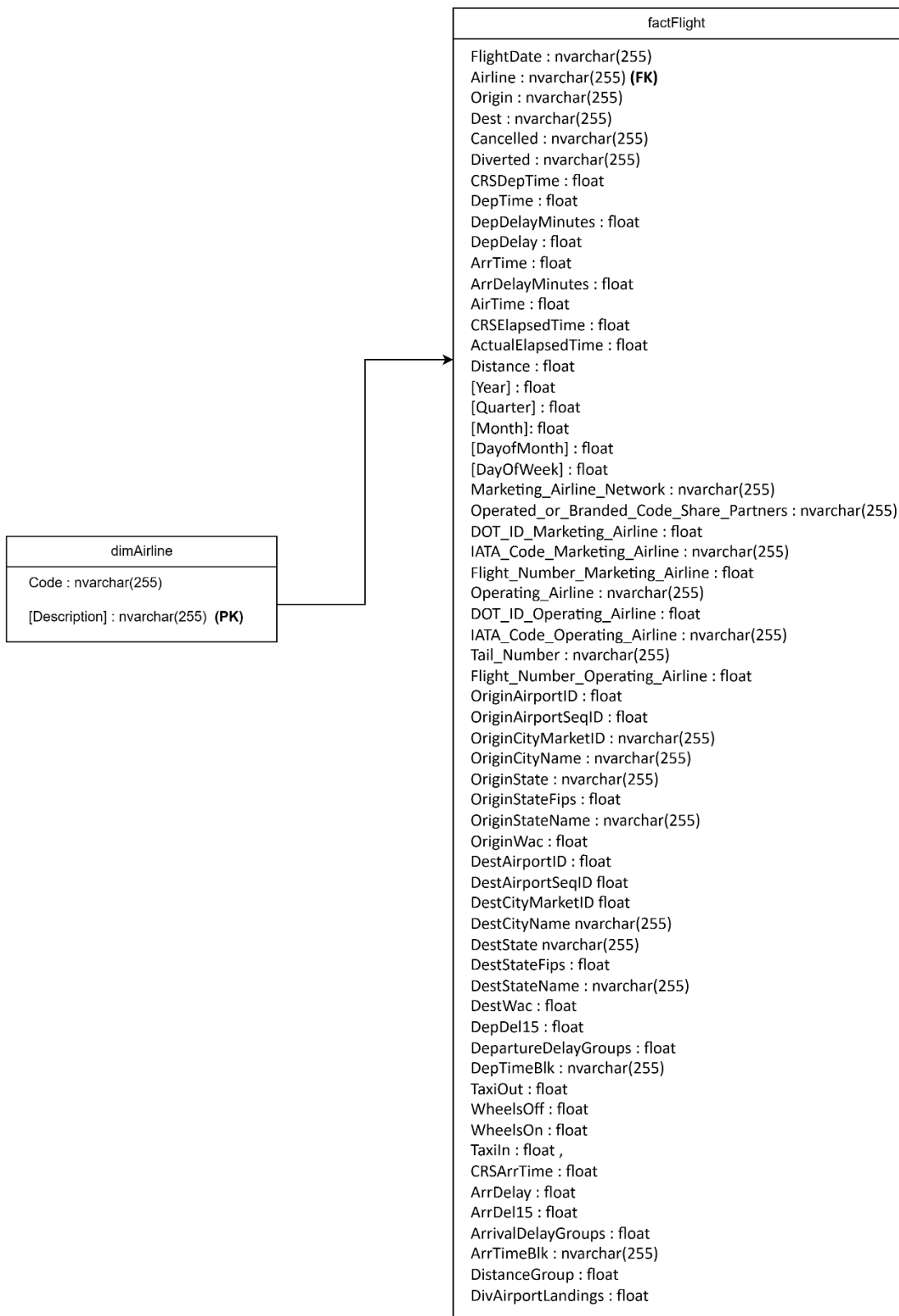
**EVALUACION ORDINARIA**

**ALUMNO: SAMANTHA ALEXANDRA ROMERO MARTÍNEZ**

**CARNET: RM16068**

**FECHA: 25 DE OCTUBRE DE 2024**

1. (25%) Diseñar una base de datos para almacenar el datawarehouse con el que se analizará esta información



## 2. (20%) Cree una estructura de carpetas para procesar la información

Nombre	Fecha de modificación	Tipo	Tamaño
Airlines	23/10/2024 21:42	Carpeta de archivos	
Exámen Módulo 3 - Ingenieria de Datos UES	24/10/2024 15:41	Carpeta de archivos	
Flights2018	23/10/2024 19:49	Carpeta de archivos	
Flights2019	23/10/2024 19:50	Carpeta de archivos	
Flights2020	23/10/2024 19:52	Carpeta de archivos	
Flights2021	23/10/2024 19:53	Carpeta de archivos	
Flights2022	23/10/2024 19:55	Carpeta de archivos	

En la carpeta Airlines está el archivo Excel de **Airlines.csv** y **Airlines2.xlsx**

En la carpeta Exámen Módulo 3 – Ingenieria de Datos UES se encuentra el proyecto de Integration Services : **Exámen Módulo 3 - Ingenieria de Datos UES.sln**

En las carpetas Flight2018, Flight2019, Flight2020, Flight2021 y Flight2022, se encuentran los archivos **Combined\_Flights\_2018.csv**, **Combined\_Flights\_2019.csv**, **Combined\_Flights\_2020.csv**, **Combined\_Flights\_2021.csv** y **Combined\_Flights\_2022.csv**, respectivamente

## 3. (30%) Cree una base de datos en SQL Server siguiendo el diseño de la base de datos elaborada en el punto 1.

Se utilizaron los siguientes scripts para la creación de las tablas en SQL Management Studio

**Para la creación de las tablas `dbo.st_Flights{año}`**

```
CREATE TABLE [dbo].[st_Flights2018](
    [FlightDate] [nvarchar](255) NULL,
    [Airline] [nvarchar](255) NULL,
    [Origin] [nvarchar](255) NULL,
    [Dest] [nvarchar](255) NULL,
    [Cancelled] [nvarchar](255) NULL,
    [Diverted] [nvarchar](255) NULL,
    [CRSDepTime] [nvarchar](255) NULL,
    [DepTime] [nvarchar](255) NULL,
    [DepDelayMinutes] [nvarchar](255) NULL,
    [DepDelay] [nvarchar](255) NULL,
    [ArrTime] [nvarchar](255) NULL,
    [ArrDelayMinutes] [nvarchar](255) NULL,
    [AirTime] [nvarchar](255) NULL,
    [CRSElapsedTime] [nvarchar](255) NULL,
```

```

[ActualElapsedTime] [nvarchar](255) NULL,
[Distance] [nvarchar](255) NULL,
[Year] [nvarchar](255) NULL,
[Quarter] [nvarchar](255) NULL,
[Month] [nvarchar](255) NULL,
[DayOfMonth] [nvarchar](255) NULL,
[DayOfWeek] [nvarchar](255) NULL,
[Marketing_Airline_Network] [nvarchar](255) NULL,
[Operated_or_Branded_Code_Share_Partners] [nvarchar](255) NULL,
[DOT_ID_Marketing_Airline] [nvarchar](255) NULL,
[IATA_Code_Marketing_Airline] [nvarchar](255) NULL,
[Flight_Number_Marketing_Airline] [nvarchar](255) NULL,
[Operating_Airline] [nvarchar](255) NULL,
[DOT_ID_Operating_Airline] [nvarchar](255) NULL,
[IATA_Code_Operating_Airline] [nvarchar](255) NULL,
[Tail_Number] [nvarchar](255) NULL,
[Flight_Number_Operating_Airline] [nvarchar](255) NULL,
[OriginAirportID] [nvarchar](255) NULL,
[OriginAirportSeqID] [nvarchar](255) NULL,
[OriginCityMarketID] [nvarchar](255) NULL,
[OriginCityName] [nvarchar](255) NULL,
[OriginState] [nvarchar](255) NULL,
[OriginStateFips] [nvarchar](255) NULL,
[OriginStateName] [nvarchar](255) NULL,
[OriginWac] [nvarchar](255) NULL,
[DestAirportID] [nvarchar](255) NULL,
[DestAirportSeqID] [nvarchar](255) NULL,
[DestCityMarketID] [nvarchar](255) NULL,
[DestCityName] [nvarchar](255) NULL,
[DestState] [nvarchar](255) NULL,
[DestStateFips] [nvarchar](255) NULL,
[DestStateName] [nvarchar](255) NULL,
[DestWac] [nvarchar](255) NULL,
[DepDel15] [nvarchar](255) NULL,
[DepartureDelayGroups] [nvarchar](255) NULL,
[DepTimeBlk] [nvarchar](255) NULL,
[TaxiOut] [nvarchar](255) NULL,
[WheelsOff] [nvarchar](255) NULL,
[WheelsOn] [nvarchar](255) NULL,
[TaxiIn] [nvarchar](255) NULL,
[CRSArrTime] [nvarchar](255) NULL,
[ArrDelay] [nvarchar](255) NULL,
[ArrDel15] [nvarchar](255) NULL,
[ArrivalDelayGroups] [nvarchar](255) NULL,
[ArrTimeBlk] [nvarchar](255) NULL,
[DistanceGroup] [nvarchar](255) NULL,
[DivAirportLandings] [nvarchar](255) NULL
);

```

#### Para la creación de las tablas **dbo.Flights{año}**

```

CREATE TABLE [dbo].[Flights2018](
    [FlightDate] [nvarchar](255) NULL,
    [Airline] [nvarchar](255) NULL,
    [Origin] [nvarchar](255) NULL,
    [Dest] [nvarchar](255) NULL,

```

```

[Cancelled] [nvarchar](255) NULL,
[Diverted] [nvarchar](255) NULL,
[CRSDepTime] [float] NULL,
[DepTime] [float] NULL,
[DepDelayMinutes] [float] NULL,
[DepDelay] [float] NULL,
[ArrTime] [float] NULL,
[ArrDelayMinutes] [float] NULL,
[AirTime] [float] NULL,
[CRSElapsedTime] [float] NULL,
[ActualElapsedTime] [float] NULL,
[Distance] [float] NULL,
[Year] [float] NULL,
[Quarter] [float] NULL,
[Month] [float] NULL,
[DayOfMonth] [float] NULL,
[DayOfWeek] [float] NULL,
[Marketing_Airline_Network] [nvarchar](255) NULL,
[Operated_or_Branded_Code_Share_Partners] [nvarchar](255) NULL,
[DOT_ID_Marketing_Airline] [float] NULL,
[IATA_Code_Marketing_Airline] [nvarchar](255) NULL,
[Flight_Number_Marketing_Airline] [float] NULL,
[Operating_Airline] [nvarchar](255) NULL,
[DOT_ID_Operating_Airline] [float] NULL,
[IATA_Code_Operating_Airline] [nvarchar](255) NULL,
[Tail_Number] [nvarchar](255) NULL,
[Flight_Number_Operating_Airline] [float] NULL,
[OriginAirportID] [float] NULL,
[OriginAirportSeqID] [float] NULL,
[OriginCityMarketID] [nvarchar](255) NULL,
[OriginCityName] [nvarchar](255) NULL,
[OriginState] [nvarchar](255) NULL,
[OriginStateFips] [nvarchar](255) NULL,
[OriginStateName] [nvarchar](255) NULL,
[OriginWac] [nvarchar](255) NULL,
[DestAirportID] [float] NULL,
[DestAirportSeqID] [float] NULL,
[DestCityMarketID] [float] NULL,
[DestCityName] [nvarchar](255) NULL,
[DestState] [nvarchar](255) NULL,
[DestStateFips] [nvarchar](255) NULL,
[DestStateName] [nvarchar](255) NULL,
[DestWac] [float] NULL,
[DepDel15] [nvarchar](255) NULL,
[DepartureDelayGroups] [float] NULL,
[DepTimeBlk] [nvarchar](255) NULL,
[TaxiOut] [float] NULL,
[WheelsOff] [nvarchar](255) NULL,
[WheelsOn] [float] NULL,
[TaxiIn] [float] NULL,
[CRSArrTime] [float] NULL,
[ArrDelay] [float] NULL,
[ArrDel15] [float] NULL,
[ArrivalDelayGroups] [float] NULL,
[ArrTimeBlk] [nvarchar](255) NULL,
[DistanceGroup] [float] NULL,
[DivAirportLandings] [nvarchar](255) NULL

```

);

Para la creación de las vistas dbo.uvw\_Flights{año}

```
CREATE view [dbo].[uvw_Flights2018] as
select
    FlightDate,
        convert(nvarchar, try_parse(FlightDate as date using 'es'), 112) as
        Airline,
        Origin,
        Dest,
        Cancelled,
        Diverted,
        try_convert(float, CRSDepTime) as CRSDepTime,
        try_convert(float, DepTime) as DepTime,
        try_convert(float, DepDelayMinutes) as DepDelayMinutes,
        try_convert(float, DepDelay) as DepDelay,
        try_convert(float, ArrTime) as ArrTime,
        try_convert(float, ArrDelayMinutes) as ArrDelayMinutes,
        try_convert(float, AirTime) as AirTime,
        try_convert(float, CRSElapsedTime) as CRSElapsedTime,
        try_convert(float, ActualElapsedTime) as ActualElapsedTime,
        try_convert(float, Distance) as Distance,
        try_convert(float, [Year]) as [Year],
        try_convert(float, [Quarter]) as [Quarter],
        try_convert(float, [Month]) as [Month],
        try_convert(float, [DayofMonth]) as [DayofMonth],
        try_convert(float, [DayOfWeek]) as [DayOfWeek],
        Marketing_Airline_Network,
        Operated_or_Branded_Code_Share_Partners,
        try_convert(float, DOT_ID_Marketing_Airline) as
DOT_ID_Marketing_Airline,
        IATA_Code_Marketing_Airline,
        try_convert(float, Flight_Number_Marketing_Airline) as
Flight_Number_Marketing_Airline,
        Operating_Airline,
        try_convert(float, DOT_ID_Operating_Airline) as
DOT_ID_Operating_Airline,
        IATA_Code_Operating_Airline,
        Tail_Number,
        try_convert(float, Flight_Number_Operating_Airline) as
Flight_Number_Operating_Airline,
        try_convert(float, OriginAirportID) as OriginAirportID,
        try_convert(float, OriginAirportSeqID) as OriginAirportSeqID,
        OriginCityMarketID,
        OriginCityName,
        OriginState,
        OriginStateFips,
        OriginStateName,
        OriginWac,
        try_convert(float, DestAirportID) as DestAirportID,
        try_convert(float, DestAirportSeqID) as DestAirportSeqID,
        try_convert(float, DestCityMarketID) as DestCityMarketID,
        DestCityName,
        DestState,
        DestStateFips,
        DestStateName,
        try_convert(float, DestWac) as DestWac,
```

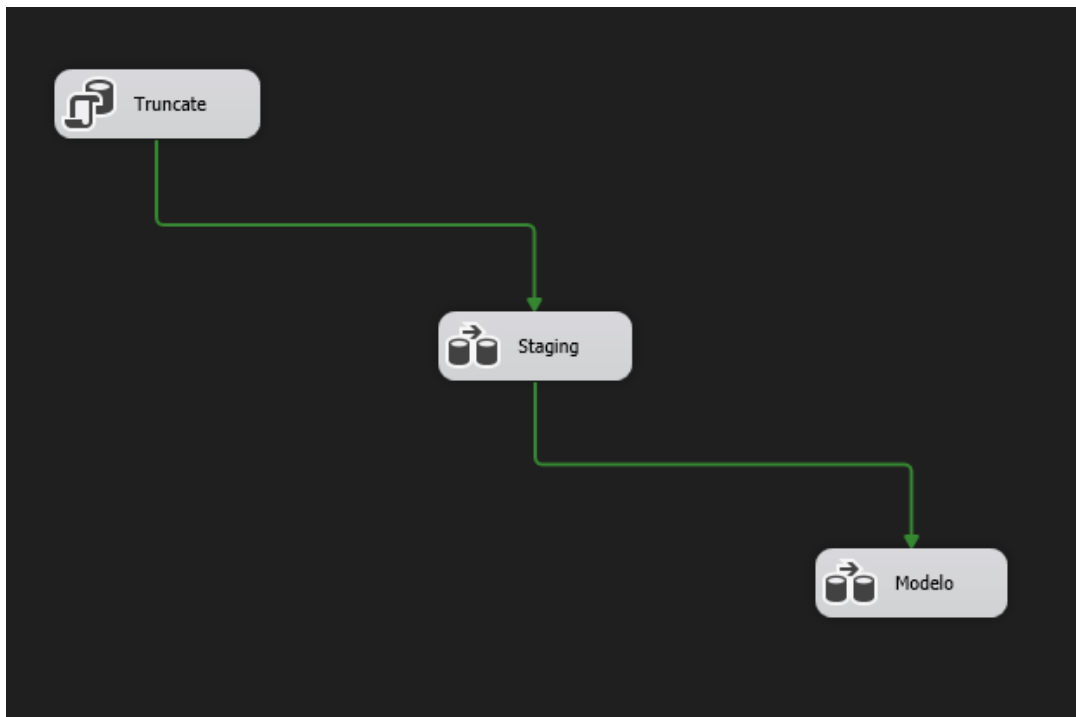
```

DepDel15,
try_convert(float, DepartureDelayGroups) as DepartureDelayGroups,
DepTimeBlk,
try_convert(float, TaxiOut) as TaxiOut,
WheelsOff,
try_convert(float, WheelsOn) as WheelsOn,
try_convert(float, TaxiIn) as TaxiIn,
try_convert(float, CRSArrTime) as CRSArrTime,
try_convert(float, ArrDelay) as ArrDelay,
try_convert(float, ArrDel15) as ArrDel15,
try_convert(float, ArrivalDelayGroups) as ArrivalDelayGroups,
ArrTimeBlk,
try_convert(float, DistanceGroup) as DistanceGroup,
DivAirportLandings
from dbo.st_Flights2018

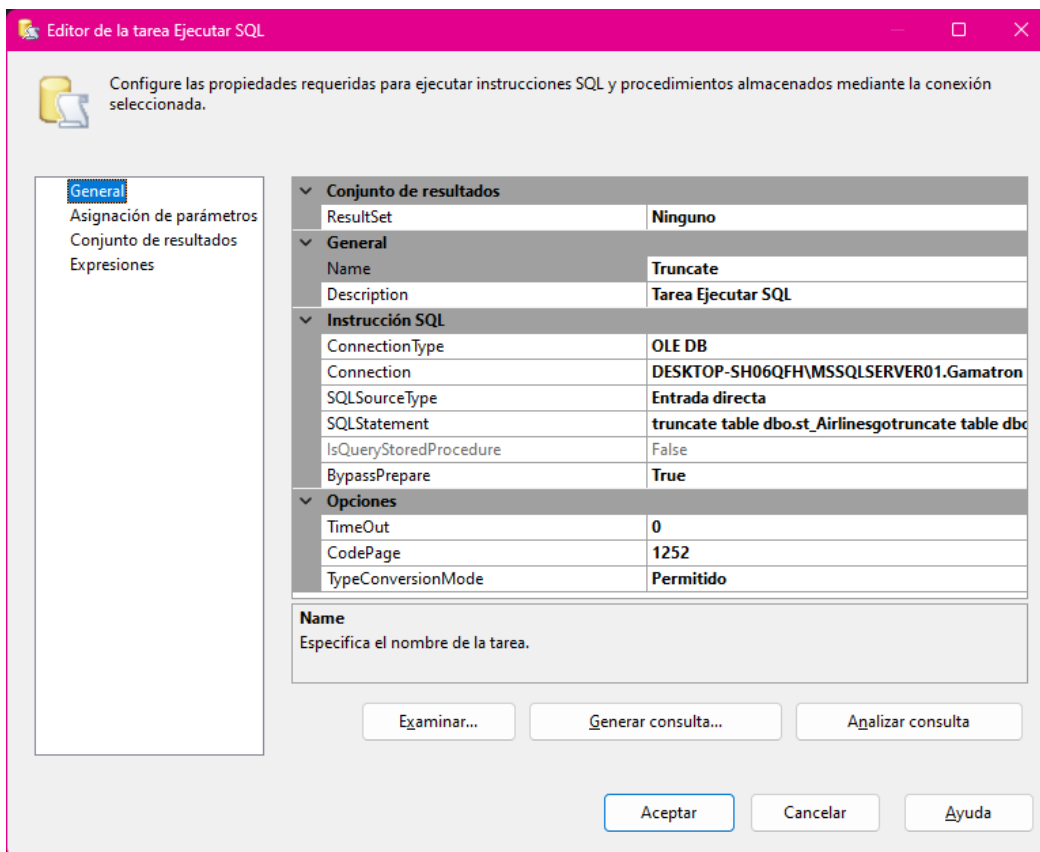
```

**4. (25%) Construya procesos ETL para la carga de información en la carpeta designada y su procesamiento para prepararlo para su carga en SQL Server.**

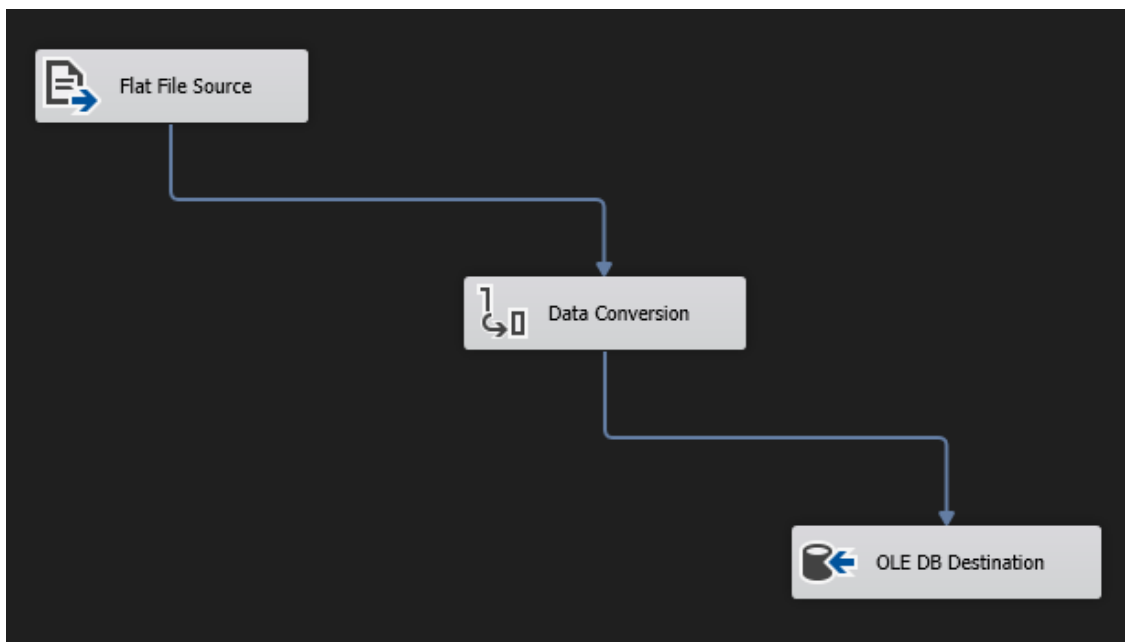
Primero se construyó en Flujo de control con los siguientes elementos



Dentro de la Tarea Ejecutar SQL (Truncate) se hizo la conexión a la base de SQL, y se creó un SQLStatement.



Dentro de Data Flow Task (Staging) se creó el siguiente flujo de datos





Dentro de Flat File Source se indicó la ruta del archivo que usará para llenar las tablas de SQL

Editor del administrador de conexiones de archivos planos

Nombre del administrador de conexiones:

Descripción:

**General**  
Columnas  
Opciones av  
Vista previa

Seleccione un archivo y especifique sus propiedades y formato.

Nombre de archivo:

Configuración regional:  ☐ Unicode

Página de códigos:

Formato:

Calificador de texto:

Delimitador de filas de encabezados:

Filas de encabezados que se omitirán:

☒ Nombres de columna de la primera fila de datos

En Data Conversion los atributos y se convierten a Unicode string [DT\_WSTR] 255

**Data Conversion Transformation Editor**

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

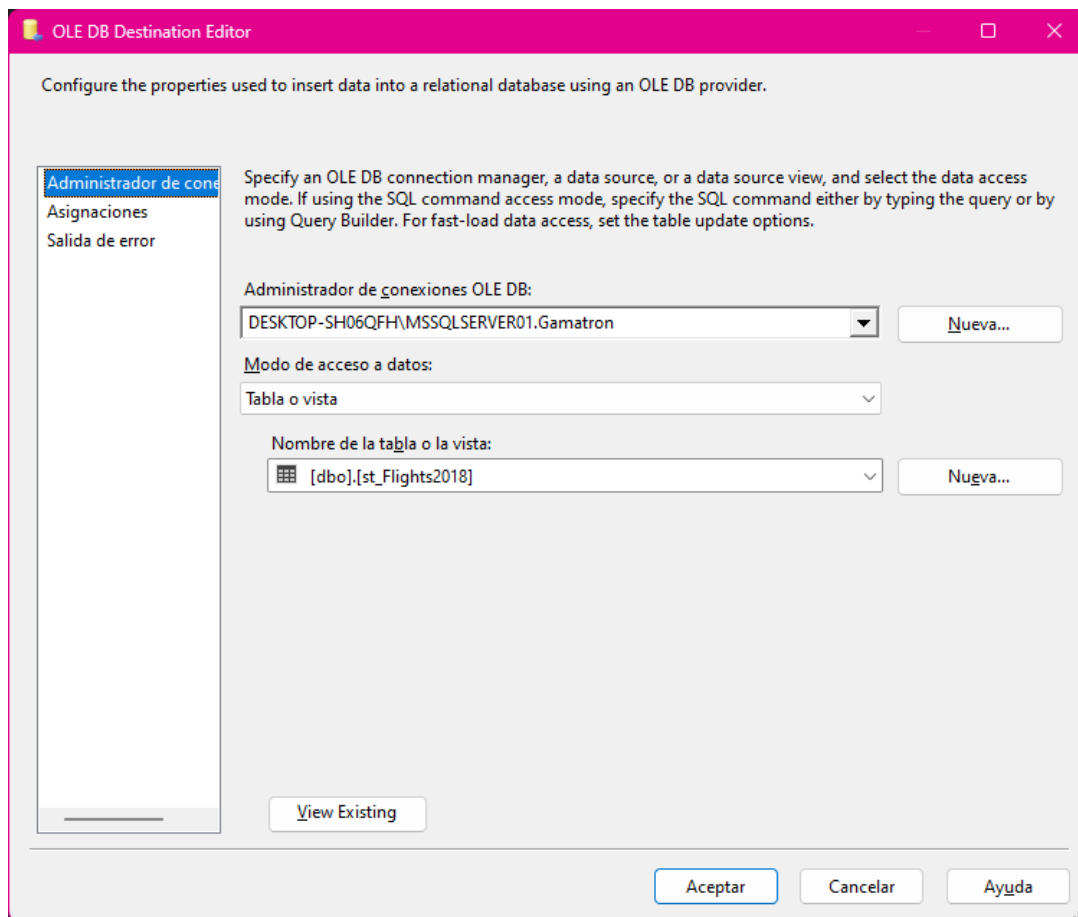
Columnas de entrada...

- ☒ Nombre
- ☒ FlightDate
- ☒ Airline
- ☒ Origin
- ☒ Dest
- ☒ Cancelled

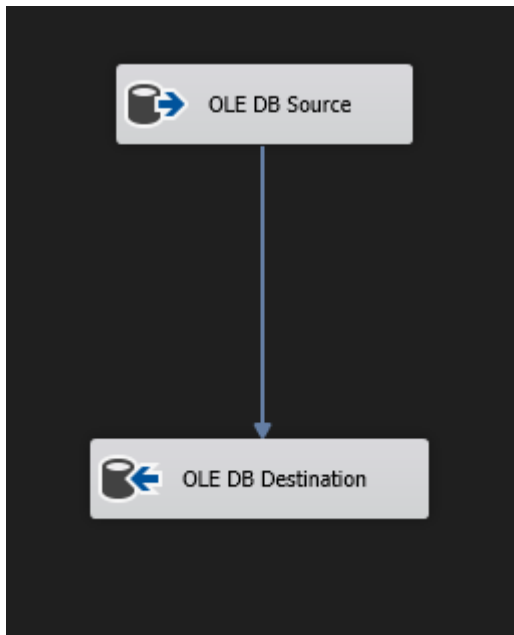
Input Column	Output Alias	Data Type	Length	Precision	Scale	Code P
FlightDate	Copy of FlightDate	Unicode string [DT_WSTR]	255			
Airline	Copy of Airline	Unicode string [DT_WSTR]	255			
Origin	Copy of Origin	Unicode string [DT_WSTR]	255			
Dest	Copy of Dest	Unicode string [DT_WSTR]	255			
Cancelled	Copy of Cancelled	Unicode string [DT_WSTR]	255			
Diverted	Copy of Diverted	Unicode string [DT_WSTR]	255			
CRSDepTime	Copy of CRSDepTime	Unicode string [DT_WSTR]	255			
DepTime	Copy of DepTime	Unicode string [DT_WSTR]	255			
DepDelayMinutes	Copy of DepDelayMin...	Unicode string [DT_WSTR]	255			
DepDelay	Copy of DepDelay	Unicode string [DT_WSTR]	255			
ArrTime	Copy of ArrTime	Unicode string [DT_WSTR]	255			
ArrDelayMinutes	Copy of ArrDelayMinu...	Unicode string [DT_WSTR]	255			
AirTime	Copy of AirTime	Unicode string [DT_WSTR]	255			

Configure Error Output... OK Cancel Ayuda

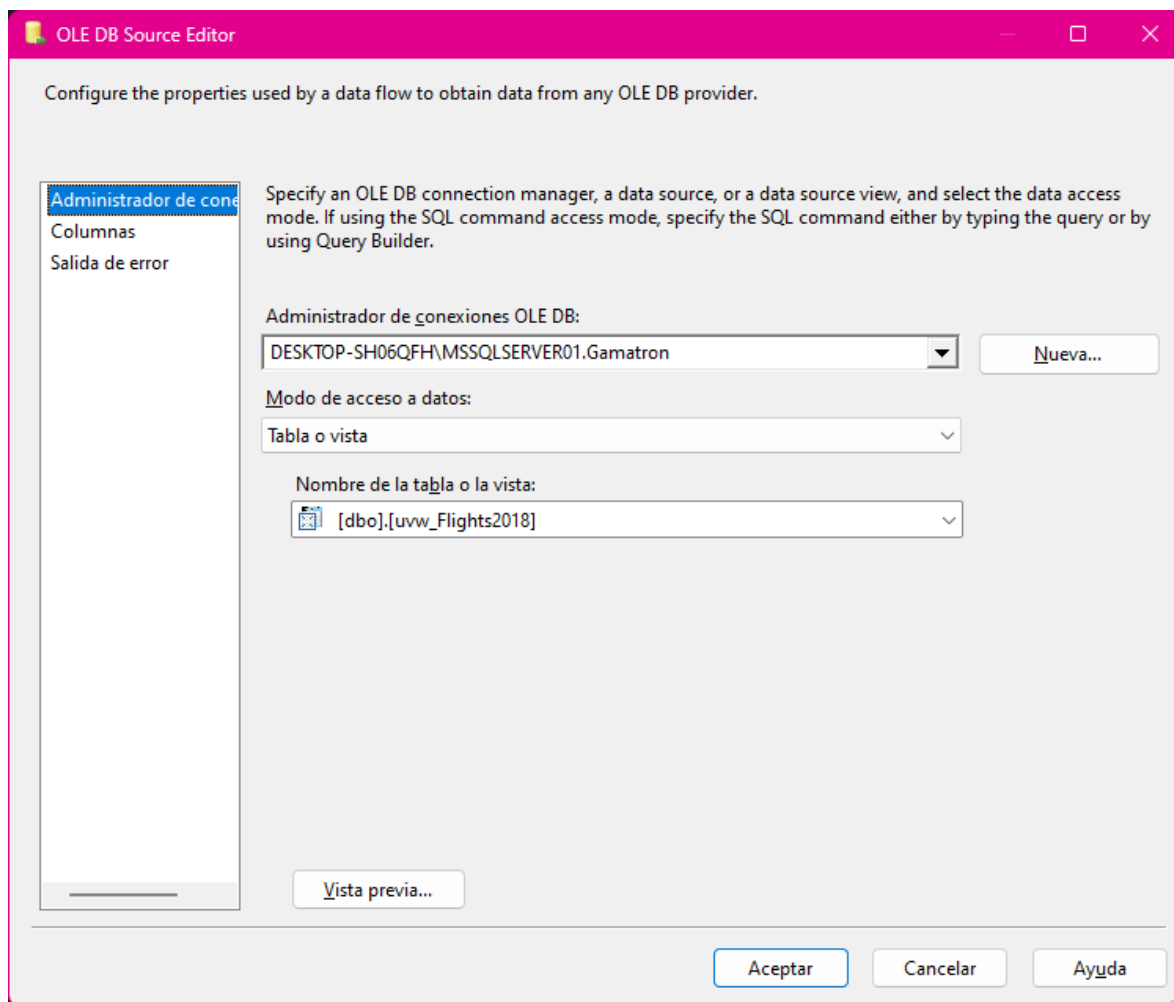
En OLE DB Destination se selecciona la tabla a la que queremos mandar los datos raw



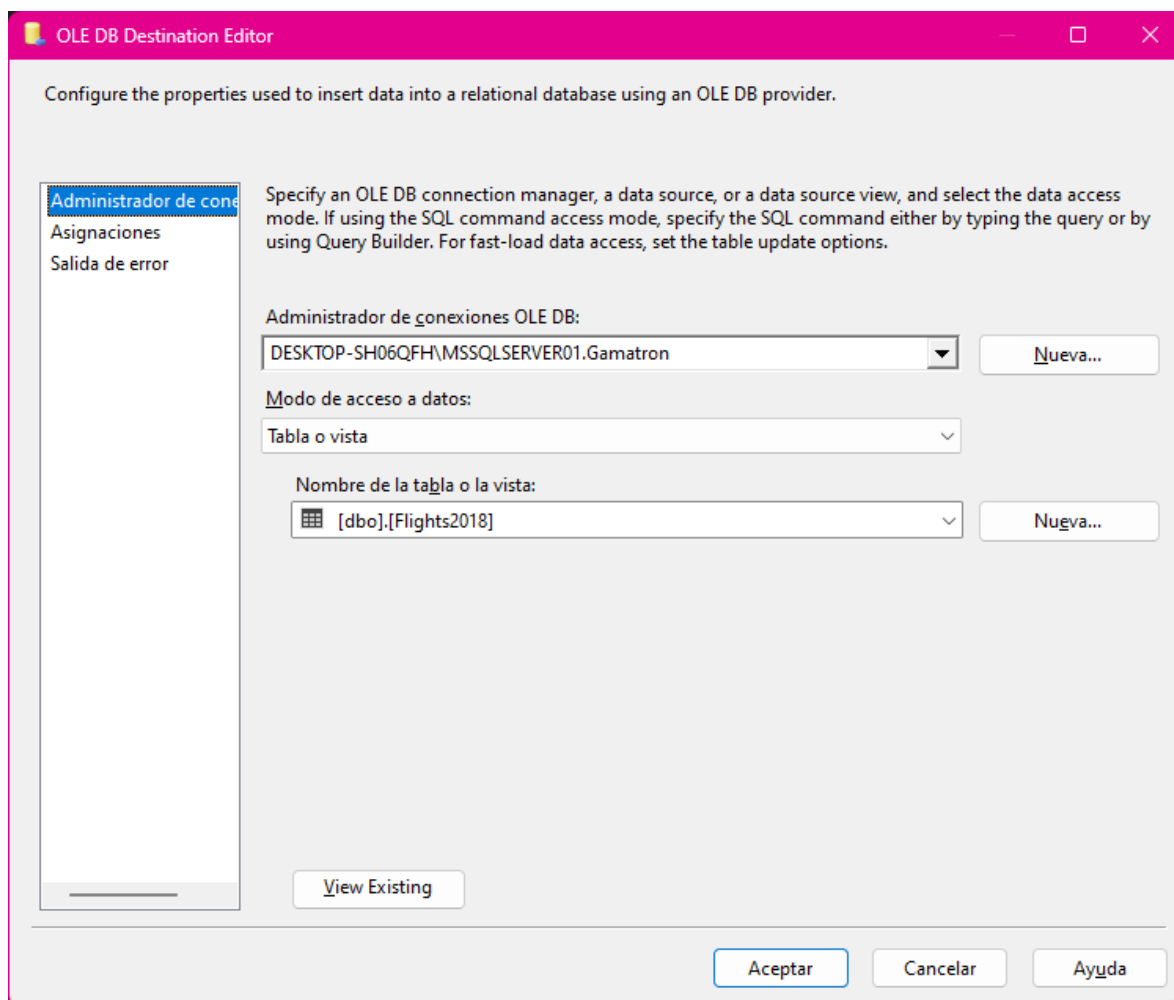
Luego regresando al flujo de control, dentro del segundo Data Flow Task (Modelo) se creó el siguiente Flujo de datos



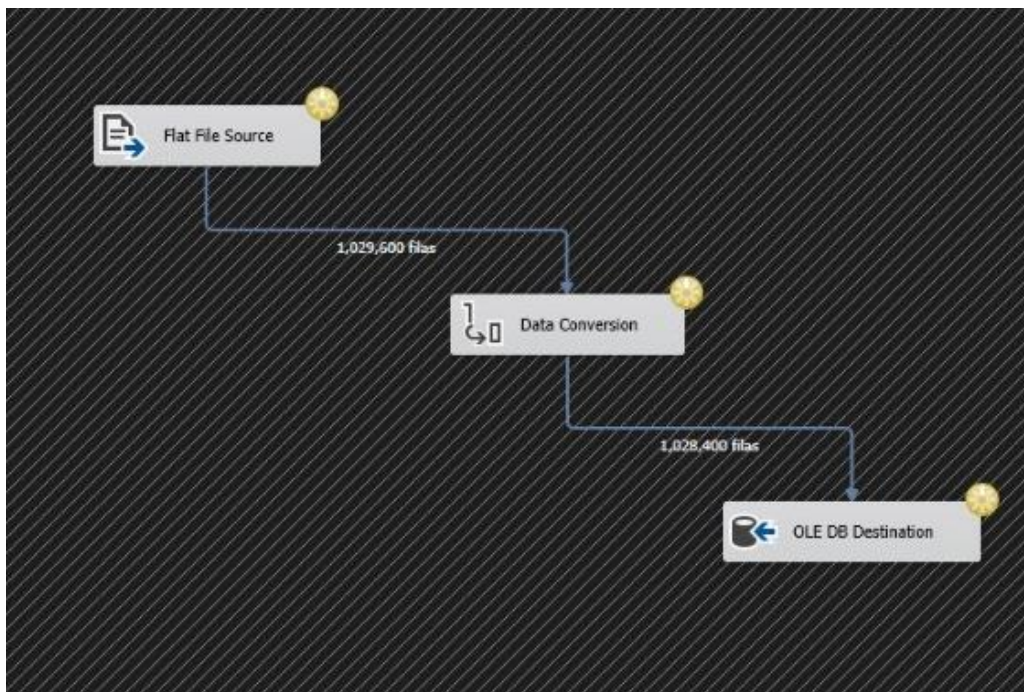
Dentro de OLE DB Source se selecciona la vista la cual ocuparemos para setear el tipo de atributo



Y en OLE DB Destination se selecciona la tabla a la cual mandaremos los datos ya con su respectivo tipo y que se pueda utilizar en herramientas como Power BI



## Ejecución de del Data Flow Task (Staging)



## Ejecución del Data Flow Task (Modelo)

