

Original project materials on my device

Archive	08.12.2023 11:31	Folder plików	
Complaints	08.12.2023 10:52	Folder plików	
Queries	08.12.2023 11:05	Folder plików	
SSIS files	08.12.2023 10:59	Folder plików	
ZIP codes backup	08.12.2023 11:26	Folder plików	
P2 - Logical model	30.08.2023 16:53	Chrome HTML Do...	20 KB
P2 - Physical model	04.09.2023 14:08	Chrome HTML Do...	22 KB
P2_ComplaintsPBI_v01	08.12.2023 12:09	Microsoft.Microso...	47 180 KB
zip_code_database	28.08.2023 18:37	Plik w formacie w...	4 544 KB

Archive – Folder to which data source .csv files go after extraction to staging layer

Complaints – Folder in which data source .csv files are stored before extraction to staging layer

Queries – Folder in which sql queries are stored

SSIS files – Folder in which SSIS project files are stored (originally it is stored in other place on my computer)

ZIP codes backup – Official US ZIP codes list in .csv format, based on which Address dimension is build and related to fact table ZIP codes attribute from data source

P2 – Logical data model – PDF file with first version logical data model (may not include changes done in final data warehouse)

P2 – Physical data model – PDF file with first version physical data model (may not include changes done in final data warehouse)

P2_ComplaintsPBI_v01 – Power BI report

zip_code_database – Official US ZIP codes list in .csv format

Requirements to run project:

Specific file paths are stored in SSIS Connection manager. They need to be adjusted if SSIS packages have to be run on other devices. Moreover databases named 'db_complaints_staging' and 'db_complaints_staging' are needed to be created beforehand.

Specific file paths include:

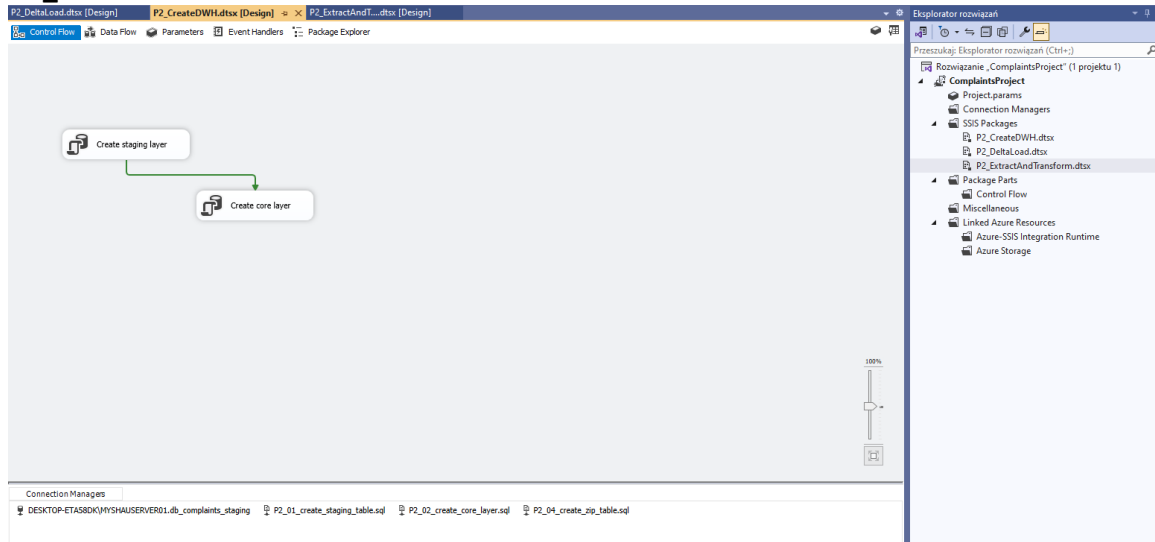
1. zip_code_database and complaints data source .csv files as Connection managers, used variables and parameters (used for including possibilities of storing more than one .csv file) based on them
2. Queries – In Control Flow SQL Tasks SQLTypeConnection are File connections
3. Archive folder

SSIS components

SSIS project is made out of 3 packages:

- P2_CreateDWH.dtsx – Which runs queries: 'P2_01_create_staging_table.sql' and 'P2_02_create_core_layer.sql' to create all necessary tables
- P2_ExtractAndTransform.dtsx – Which is responsible for initial load
- P2_DeltaLoad.dtsx – Which is responsible for delta loads

P2_CreateDWH.dtsx



Create staging layer

Execute SQL Task Editor

Configure the properties required to run SQL statements and stored procedures using the selected connection.

General

Property	Value
Name	Create staging layer
Description	Execute SQL Task
Options	
TimeOut	0
CodePage	1250
TypeConversionMode	Allowed
Result Set	
ResultSet	None
SQL Statement	
ConnectionType	OLE DB
Connection	DESKTOP-ETASBOK\MYSHAUSER\01_db_complaints_staging
SQLSourceType	File connection
FileConnection	P2_01_create_staging_table.sql
IsQueryStoredProcedure	False
BypassPrepare	True

Name
Specifies the name of the task.

Browse... Build Query... Parse Query

OK Cancel Help

Create core layer

Execute SQL Task Editor

Configure the properties required to run SQL statements and stored procedures using the selected connection.

General

Property	Value
Name	Create core layer
Description	Execute SQL Task
Options	
TimeOut	0
CodePage	1250
TypeConversionMode	Allowed
Result Set	
ResultSet	None
SQL Statement	
ConnectionType	OLE DB
Connection	DESKTOP-ETASBOK\MYSHAUSER\01_db_complaints_staging
SQLSourceType	File connection
FileConnection	P2_02_create_core_layer.sql
IsQueryStoredProcedure	False
BypassPrepare	True

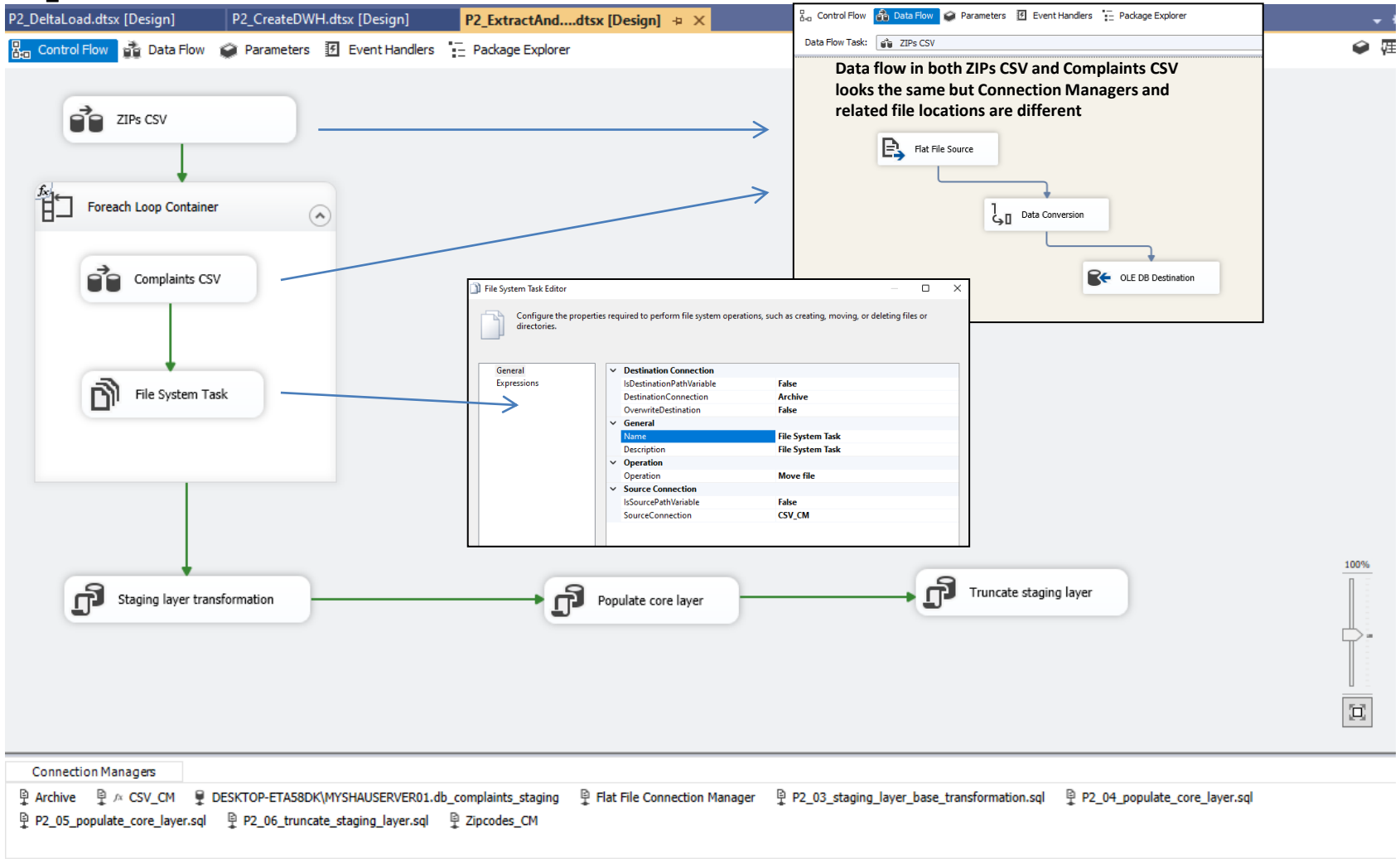
Name
Specifies the name of the task.

Browse... Build Query... Parse Query

OK Cancel Help

SSIS components

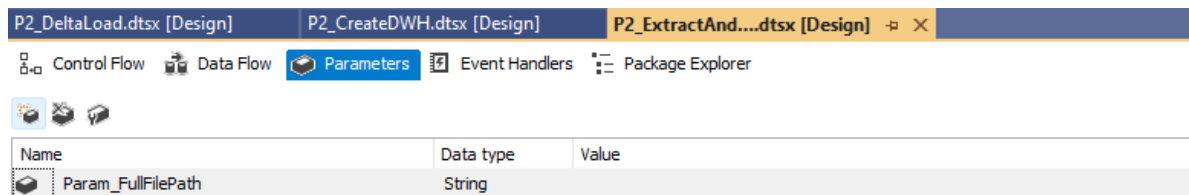
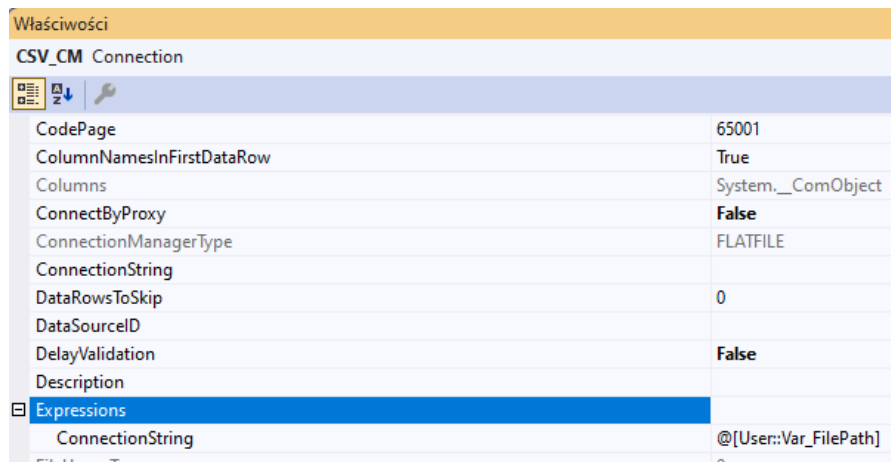
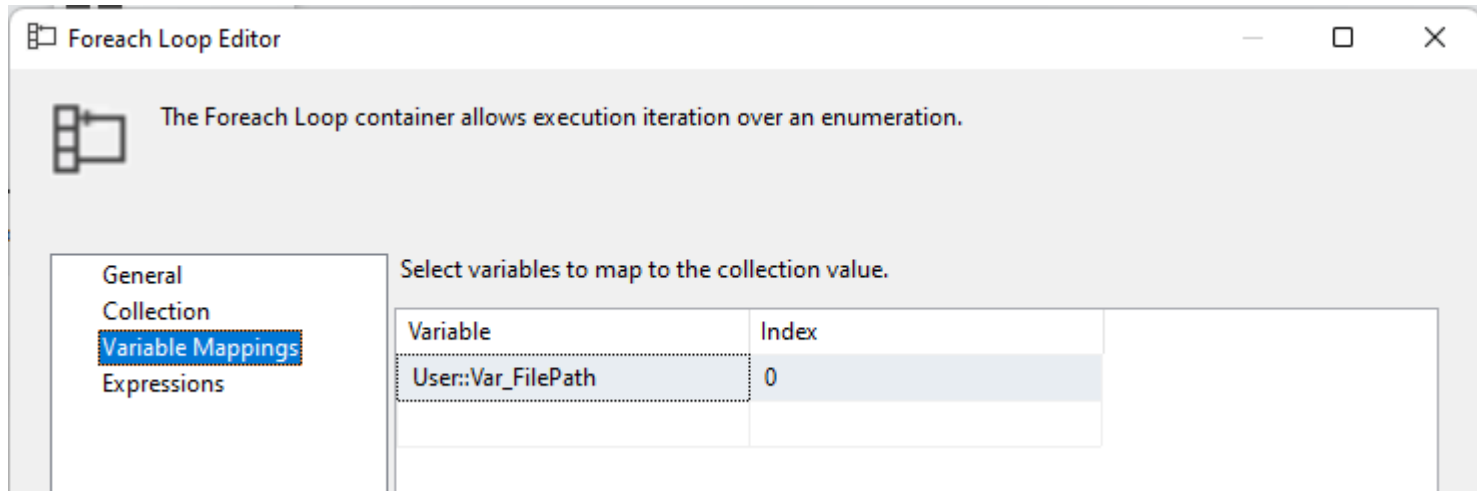
P2_ExtractAndTransform.dtsx



Package extracts data from zipcodes .csv file and complaints .csv file or files to tbl_Staging and tbl_Zipcodes in Staging layer, then queries are run to transform data and populate fact table and dimensions tables with proper data.

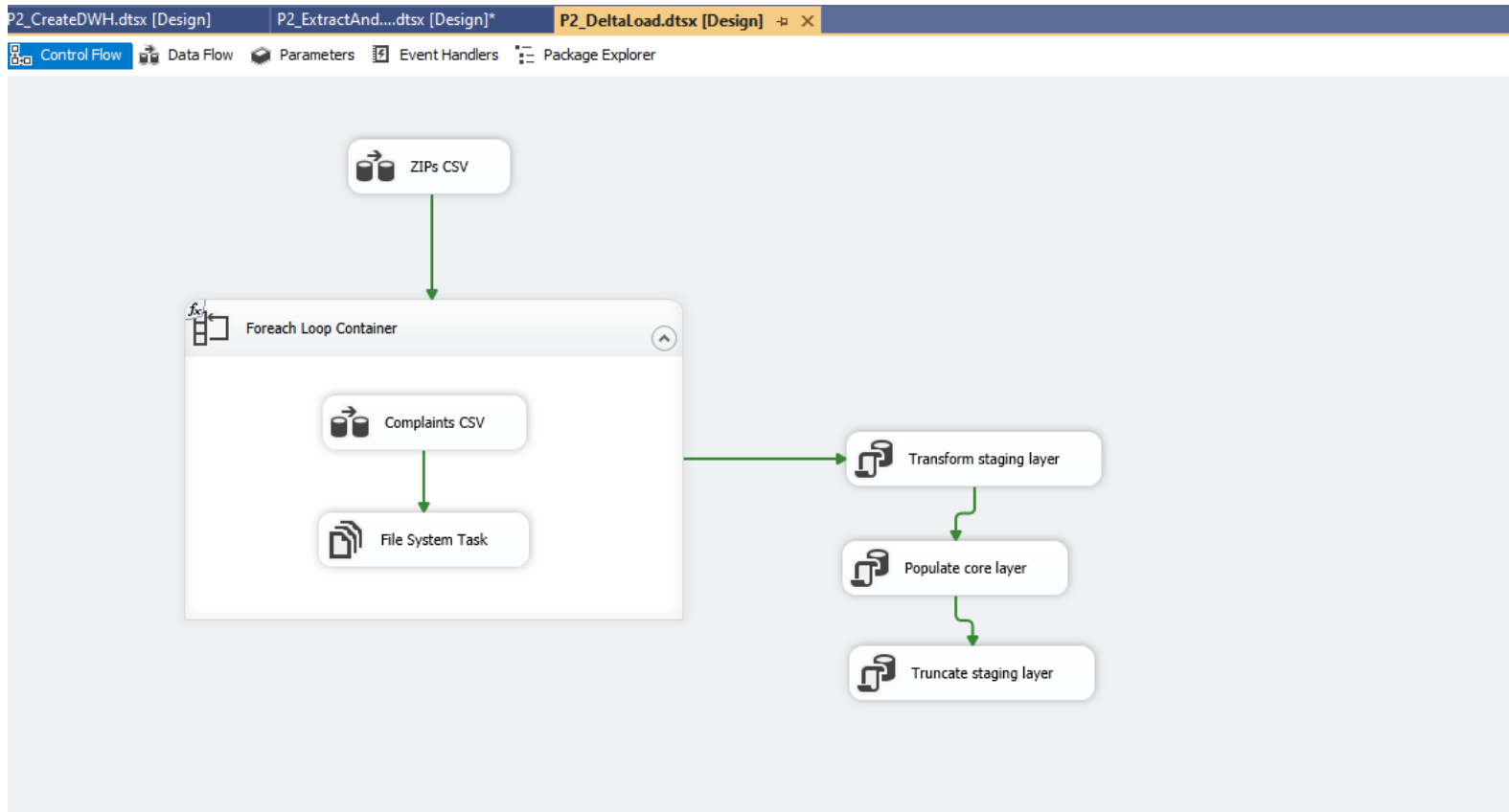
SSIS components

P2_ExtractAndTransform.dtsx – properties and parameters related to project



SSIS components

P2_DeltaLoad.dtsx



Package does the same what **P2_ExtractAndTransform.dtsx** only query in Populate core layer is different to make any new changes include the fact that there already is some data in core layer.