



# Azure + Power BI

## Pet-project

July 2022  
Yerzhan Sapenov

# Contents

General information

User story

Power BI developer story

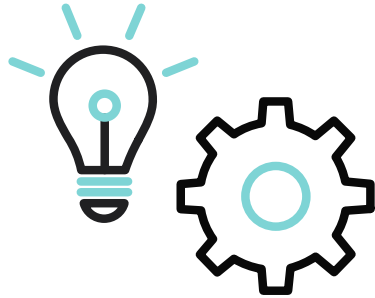
Project story

Appendix

# General information

# Description

## Pet-project



### **Description:**

The client company provides services in the real estate sector, specifically – the sale of houses.

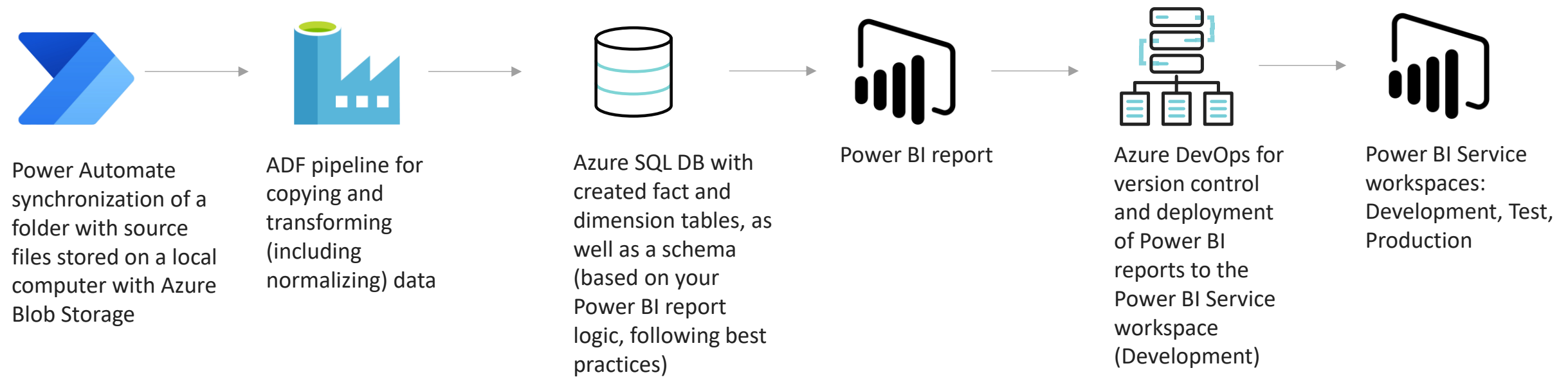
### **Previous Situation:**

The previous method of data collection involved Excel files saved in \*.csv format, containing descriptions and prices of specific properties. There was also a reference table with codes and descriptions of various property features. Reports were then generated from these files in Excel format, which was quite inconvenient and time-consuming.

### **Achieved Result:**

The user can now upload files to a designated folder on their computer, run Power Automate, and the file will be copied to a specific location in cloud storage as \*.csv files containing property prices and descriptions. From there, an automated process records the transformed (including normalized) data into an Azure database, which is connected to a Power BI report that visualizes this data.

# High-level process



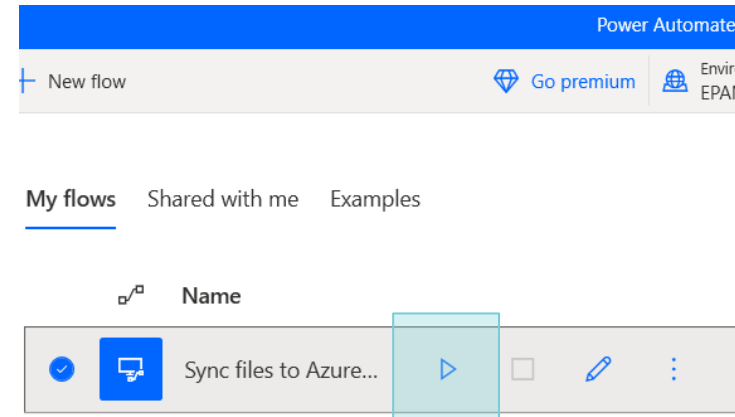
# User story

## The process for user contains two steps

1. Create or update local folder with required files

Name	Status
 train.csv	
 train2.csv	
 train3.csv	

2. Click Power Automate button to synchronize local folder with Azure Blob Storage (Settings for Power Automate are in the Appendix)

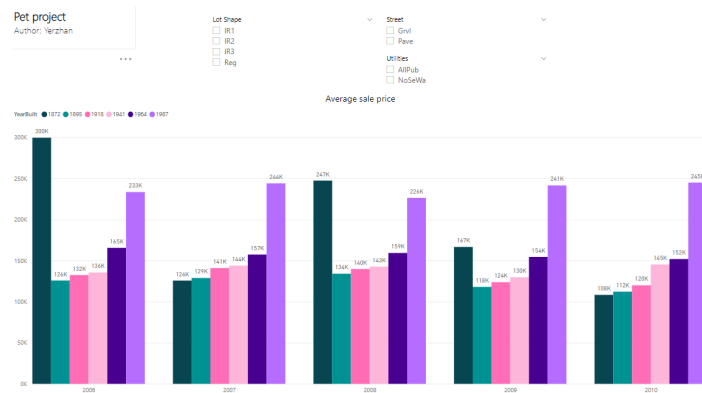


# Power BI developer story

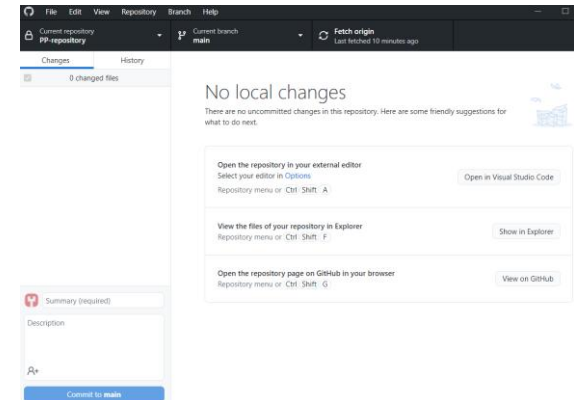


# The process for developer contains four steps

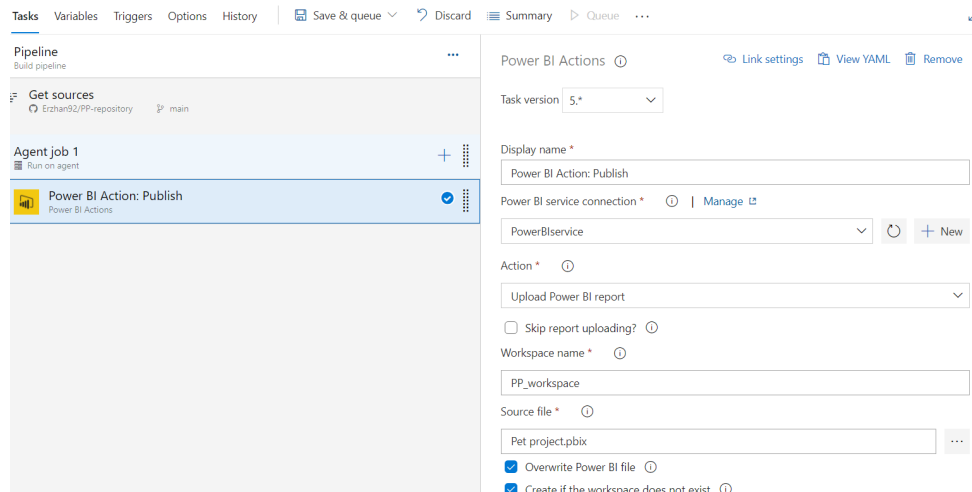
## 1. Create or update local folder with Power BI file



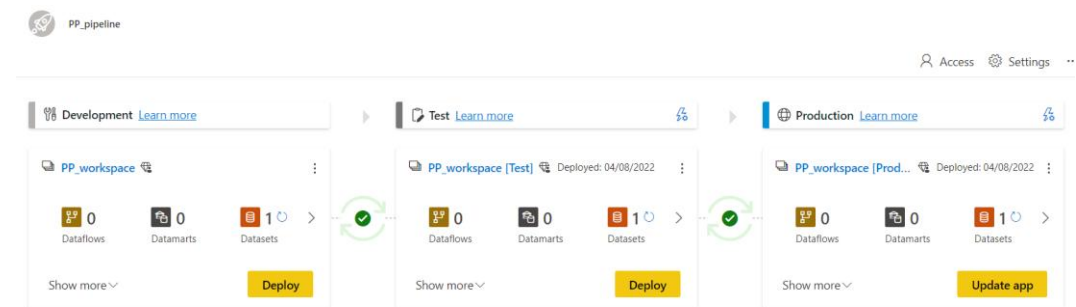
## 2. Update version with GitHub desktop



## 3. Activate or wait trigger for DevOps pipeline



## 4. Deploy to production workspace and update App



# Project story

# The project has 13 user stories and 35 tasks

1

3

Preparation work for Pet project

State 

● Done

✓ 0/3

+ Add Task

☐

✓

Go through the presentation

☐

✓

Download and install all software

☐

✓

Create accounts and receive necessary subscriptions

2

8

Get acquaint with all software and online portals

State 

● Done

✓ 0/4

+ Add Task

☐

✓

Azure portal

☐

✓

Azure DevOps

☐

✓

GitHub

☐

✓

Storage Explorer

3

13

Put initial file "train" in Blob

State 

● Done

✓ 0/3

+ Add Task

☐

✓

Create Blob

☐

✓

Create Storage account

☐

✓

Upload File

4

17

Create Azure SQL database

State 

● Done

✓ 0/3

+ Add Task

☐

✓

Create SQL Server

☐

✓

Create SQL database

☐

✓

Set Firewall

5

21

Create simple ADF copy pipeline

State 

● Done

✓ 0/4

+ Add Task

☐

✓

Create table in SQL database

☐

✓

Change the column names of original train file started with numbers

☐

✓

Create copy dataflow and insert in pipeline

☐

✓

Run and check pipeline

6

26

Create simple PowerBI report

State 

● Done

✓ 0/5

+ Add Task

☐

✓

Connect to database

☐

✓

publish report

☐

✓

create deployment pipeline in PowerBI service

☐

✓

create App in production environment

☐

✓

Make data updating settings

## The project has 13 user stories and 35 tasks

- 7

32

Create DevOps project

State ● Done

✓ 0/2

▼

+ Add Task

☐ ✓ Connect ADF to DevOps

☐ ✓ Extension installation
- 8

34

Create GitHub project

State ● Done

✓ 0/2

+ Add Task

☐ ✓ Connect with DevOps

☐ ✓ Sync repository with Azure DevOps
- 9

38

Create Azure DevOps release pipeline to PowerBI

State ● Done

✓ 0/3

+ Add Task

☐ ✓ Provide access to PowerBI Service

☐ ✓ Create App in Azure Active Directory

☐ ✓ Sync GitHub and Azure DevOps
- 
- 10

41

Normalize data in ADF pipeline

State ● Done

✓ 0/2

+ Add Task

☐ ✓ Create DIM tables in SQL DB

☐ ✓ Create dataflows
- 11

44

Create incremental data update in ADF pipeline

State ● Done

✓ 0/2

+ Add Task

☐ ✓ use parameters and variables

☐ ✓ test by adding new files
- 12

46

Create Power Automate sync process for users with using AzCopy utility

State ● Done
- 13

48

Update PowerBI report

State ● Done

# Appendix

Power Automate

Azure Blob storage

Azure SQL database

ADF pipeline and dataflows

Power BI report

GitHub project

DevOps pipeline

# A1. Power Automate

Synchronization process consists of 4 steps and uses AzCopy utility

The screenshot displays a Power Automate flow editor. On the left, a vertical list of steps is numbered 1 to 4. Step 1 is 'Set variable' for 'UploadFolder'. Step 2 is 'Set variable' for 'AzCopy'. Step 3 is 'Set variable' for 'AzureBlob'. Step 4 is 'Run PowerShell script'. On the right, a 'Variables' pane is open, showing 'Input / output variables' (0) and 'Flow variables' (5). The flow variables listed are AzCopy, AzureBlob, PowershellOut..., ScriptError, and UploadFolder.

**Step 1: Set variable**  
Assign to variable **UploadFolder** the value 'C:\Users\Yerzhan\_Sapenov\OneDrive - EPAM\pet project\initial\files'

**Step 2: Set variable**  
Assign to variable **AzCopy** the value 'C:\Users\Yerzhan\_Sapenov\Documents\Pet\_project\azcopy\_windows\_amd64\_10.16...

**Step 3: Set variable**  
Assign to variable **AzureBlob** the value 'https://pp1storageaccountname.blob.core.windows.net/pp-blob?...'

**Step 4: Run PowerShell script**  
Run PowerShell script and store its output into **PowershellOutput** and its error into **ScriptError**

**Variables Pane:**  
Input / output variables: 0  
Flow variables: 5  
Listed variables: AzCopy, AzureBlob, PowershellOut..., ScriptError, UploadFolder

The corresponding script in Power Automate

```
SET UploadFolder TO '$"C:\Users\Yerzhan_Sapenov\OneDrive - EPAM\pet project\initial\files"'
```

```
SET AzCopy TO '$"C:\Users\Yerzhan_Sapenov\Documents\Pet_project\azcopy_windows_amd64_10.16.0\azcopy.exe"'
```

```
SET AzureBlob TO '$"https://pp1storageaccountname.blob.core.windows.net/pp-blob?sp=racwdli&st=2022-08-02T05:08:33Z&se=2022-09-02T13:08:33Z&sv=2021-06-08&sr=c&sig=AcRORUtAm8gqX4kHeva0ecC2lhDI%%2Fv6Wi2knLJJvMAY%%3D"'
```

```
Scripting.RunPowershellScript Script: '$"%AzCopy% sync \"%UploadFolder%\" \"%AzureBlob%\" --recursive\" ScriptOutput=> PowershellOutput ScriptError=> ScriptError'
```

## A2. Azure Blob storage

Microsoft Azure

Upgrade

Search resources, services, and docs (G+)

Home > Storage accounts > pp1storageaccountname | Containers >

pp-blob  
Container

Search (Ctrl+)

UploadChange access levelRefreshDeleteChange tierAcquire leaseBreak leaseView snapshotsCreate snapshot

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Access policy

Properties

Metadata

Authentication method: Access key (Switch to Azure AD User Account)

Location: pp-blob

Search blobs by prefix (case-sensitive)

Show deleted blobs

Add filter

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
<input type="checkbox"/> train.csv	8/3/2022, 3:14:50 PM	Hot (Inferred)		Block blob	451.31 KiB	Available
<input type="checkbox"/> train2.csv	8/3/2022, 5:18:36 PM	Hot (Inferred)		Block blob	3.66 KiB	Available
<input type="checkbox"/> train3.csv	8/3/2022, 5:19:09 PM	Hot (Inferred)		Block blob	3.66 KiB	Available

## A3. Azure SQL database

The screenshot displays the Microsoft Azure portal interface. At the top, the navigation bar includes the Microsoft Azure logo, an Upgrade button, a search bar, and user information (ysapenov@gmail.com, DEFAULT DIRECTORY). The breadcrumb trail shows the path: Home > SQL databases > pp-database (pp-sqlserver/pp-database).

The main content area is titled "SQL databases" and shows a list of databases. The selected database, "pp-database (pp-sqlserver/pp-database)", is highlighted. The left sidebar contains navigation options: Overview, Activity log, Tags, Diagnose and solve problems, Getting started, and Query editor (preview). Below these are sections for Power Platform (Power BI, Power Apps, Power Automate) and Settings (Compute + storage, Connection strings, Properties, Locks).

The "Query editor (preview)" window is open, showing the "pp-database (adminlogin)" connection. A message states: "Showing limited object explorer here. For full capability please open SSDT." The object explorer on the right lists the following items:

- Tables
  - dbo.dim\_lotshape
  - dbo.dim\_street
  - dbo.dim\_Uilities
  - dbo.fact\_table
  - dbo.pptable\_test
- Views
- Stored Procedures

The query editor area on the right shows a single query labeled "Query 1" with a line number "1". Below the query editor, there are tabs for "Results" and "Messages", and a search bar labeled "Search to filter items...".



## A4. ADF pipeline and dataflows

The screenshot displays the Azure Data Factory (ADF) interface. On the left, the 'Factory Resources' pane shows a tree view with 'Pipelines' (1) and 'Datasets' (4). The 'UseNewFile' pipeline is selected. The 'Activities' pane lists various activity categories, including 'Move & transform', 'Azure Data Explorer', 'Azure Function', 'Batch Service', 'Databricks', 'Data Lake Analytics', 'General', 'HDInsight', 'Iteration & conditionals', 'Machine Learning', and 'Power Query'. The main canvas shows a pipeline named 'UseNewFile' with the following sequence of activities: 'Get Metadata' (containing 'GetFolderMetadata'), 'ForEach' (containing 'ForEachFile' with 3 activities), and four 'Data flow' activities: 'StreetDim', 'UtilitiesDim', 'LotShapeDim', and 'FactReplace'. The 'Parameters' tab is active at the bottom, showing a '+ New' button.

## A4. ADF pipeline and dataflows

Factory Resources

Filter resources by name

Pipelines1

UseNewFile

Datasets6

Data flows4

FactReplace

LotShapeDim

StreetDim

UtilitiesDim

Power Query1

Templates1

UseNewFile

Activities

Search activities

Move & transform

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

SavedSave as templateValidateDebugAdd triggerData flow debug

Get Metadata

ForEach

Data flow

Data flow

Data flow

Data flow

FactReplace

Parameters

Variables

Settings

Output

Pipeline run ID: 8d295864-6d51-4b0f-9b89-f0e7f55eae93

View debug run consumption

Name	Type	Run start	Duration	Status	Integrat
FactReplace	Data flow	2022-08-03T11:27:06.60356	00:00:38	Succeeded	AutoRes
LotShapeDim	Data flow	2022-08-03T11:26:25.89872	00:00:37	Succeeded	AutoRes
UtilitiesDim	Data flow	2022-08-03T11:25:46.06568	00:00:36	Succeeded	AutoRes
StreetDim	Data flow	2022-08-03T11:24:12.89835	00:01:31	Succeeded	AutoRes
Set variable2	Set variable	2022-08-03T11:24:06.75757	00:00:01	Succeeded	

## A4. ADF pipeline and dataflows

**Factory Resources**

Filter resources by name

Pipelines1

UseNewFile

Datasets6

DimLotShape

DimStreet

DimUtilities

FactTable

TrainDataKaggle

TrainDataKaggle\_dynamic

Data flows4

FactReplace

LotShapeDim

StreetDim

UtilitiesDim

Power Query1

StreetDimUseNewFileFactReplace

SavedValidateData flow debug

SourceFactTable

StreetJoin

selectAllExcept...

LotShapeJoin

selectAllExceptL...

UtilitiesJoin

selectAllExcept...

AlterRow1

FactReplaceTable

DimTableStreet

DimTableLotSh...

DimTableUtilities

Add Source

ParametersSettings

+ New

Delete

<input type="checkbox"/>	Name	Type	Default value	
<input type="checkbox"/>	parameterFRFileNa	abc string	Enter expression... ANY	<div></div>

## A4. ADF pipeline and dataflows

### Factory Resources

Filter resources by name

Pipelines1

UseNewFile

Datasets6

DimLotShape

DimStreet

DimUtilities

FactTable

TrainDataKaggle

TrainDataKaggle\_dynamic

Data flows4

FactReplace

LotShapeDim

StreetDim

UtilitiesDim

StreetDim

UseNewFile

LotShapeDim

Saved Validate Data flow debug

SourceFactTable

selectLotShape

DistinctLotShape

LotShapePrevious

unionOldNew

DistinctWithNew

LotShapeKeyGener...

LotShapeDimTa...

DistinctLotShape

Parameters

Settings

New

Delete

<input type="checkbox"/>	Name	Type	Default value	
<input type="checkbox"/>	parameterLFileNam	abc string	Enter expression... ANY	

## A4. ADF pipeline and dataflows

### Factory Resources

Filter resources by name

Pipelines1

UseNewFile

Datasets6

DimLotShape

DimStreet

DimUtilities

FactTable

TrainDataKaggle

TrainDataKaggle\_dynamic

Data flows4

FactReplace

LotShapeDim

StreetDim

UtilitiesDim

StreetDim

UseNewFile

LotShapeDim

Saved

Validate

Data flow debug

SourceFactTable

Import data from TrainDataKaggle\_dynamic

selectStreet

Renaming SourceFactTable to selectStreet with columns 'StreetName'

DistinctStreet

Aggregating data by 'StreetName' producing columns 'StreetID'

StreetPrevious

Import data from DimStreet

unionOldNew

Combining rows from transformation 'StreetPrevious, DistinctStreet'

DistinctWithNew

Aggregating data by 'StreetName' producing columns 'StreetID'

StreetKeyGeneration

Adding new key StreetID starting from 1 with step 1

StreetDimTable

Export data to DimStreet

DistinctStreet

Parameters

Settings

New

Delete

<input type="checkbox"/>	Name	Type	Default value
<input type="checkbox"/>	parameterFileName	abc string	<div>Enter expression...</div> ANY

## A4. ADF pipeline and dataflows

**Factory Resources**

Filter resources by name

Pipelines1

UseNewFile

Datasets6

DimLotShape

DimStreet

DimUtilities

FactTable

TrainDataKaggle

TrainDataKaggle\_dynamic

Data flows4

FactReplace

LotShapeDim

StreetDim

UtilitiesDim

StreetDimUseNewFileLotShapeDimUtilitiesDim

SavedValidateData flow debug

SourceFactTable

Import data from TrainDataKaggle\_dynamic

selectUtilities

Renaming SourceFactTable to selectUtilities with columns 'UtilitiesName'

DistinctUtilities

Aggregating data by 'UtilitiesName' producing columns 'UtilitiesID'

UtilitiesPrevious

Import data from DimUtilities

unionOldNew

Combining rows from transformation 'UtilitiesPrevious, DistinctUtilities'

DistinctWithNew

Aggregating data by 'UtilitiesName' producing columns 'UtilitiesID'

UtilitiesKeyGenera...

Adding new key UtilitiesID starting from 101 with step 1

UtilitiesDimTable

Export data to DimUtilities

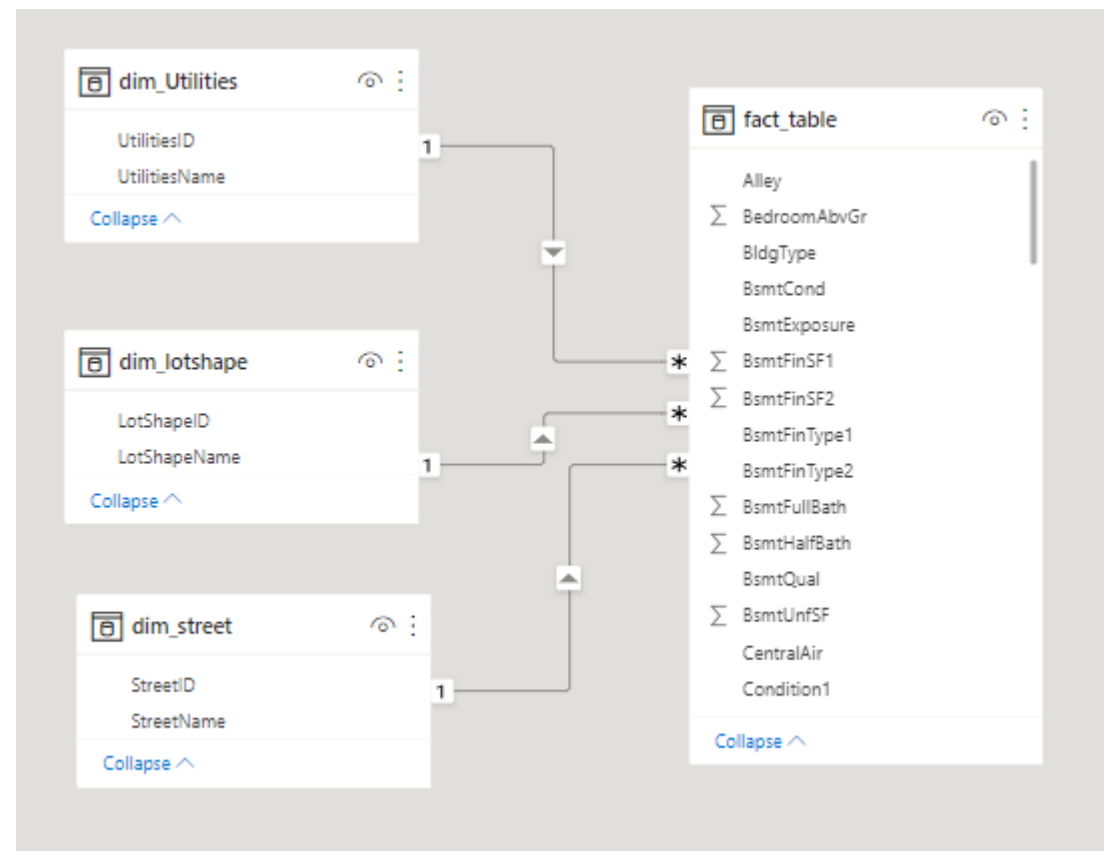
DistinctUtilities

ParametersSettings


NewDelete

<input type="checkbox"/>	Name	Type	Default value
<input type="checkbox"/>	parameterUFileNan	abc string	Enter expression... ANY

## A5. Power BI report



# A6. GitHub project




Pull requests


Issues


Marketplace

Explore




+




 Erzhan92 / PP-repository


Private

 Unwatch

1


 Fork


0


 Star


0


<> Code


 Issues


 Pull requests

 Actions

 Projects

 Security

 Insights

 Settings

 main

1 branch

 0 tags

Go to file

Add file

Code

 Erzhan92 Update Pet project.pbix

✖ 1200bba

yesterday

 6 commits

Add a README with an overview of your project.

Add a README

About



Pet project

 0 stars

 1 watching

 0 forks

Releases

No releases published

Create a new release



## A7. DevOps pipeline

Tasks

Variables

Triggers

Options

History

Save & queue

Discard

Summary

Queue

Pipeline

Build pipeline

Get sources

Erzhan92/PP-repository

main

Agent job 1

Run on agent

Power BI Action: Publish

Power BI Actions

Power BI Actions

Link settings

View YAML

Remove

Task version

5.\*

Display name \*

Power BI Action: Publish

Power BI service connection \*

PowerBIService

Manage

Action \*

Upload Power BI report

Skip report uploading?

Workspace name \*

PP\_workspace

Source file \*

Pet project.pbix

Overwrite Power BI file

Create if the workspace does not exist

<epam>

25

# Thank you!