

**WYDZIAŁ
ELEKTROTECHNIKI
I INFORMATYKI
POLITECHNIKI RZESZOWSKIEJ**

Wielowymiarowa Analiza Danych

Analiza Lotów

Dawid Stachiewicz

Jakub Soboń

Katarzyna Strzępek

Rzeszów, 2025

1. Spis treści

1. Spis treści	2
1. Wstęp.....	4
2. Etap Pierwszy	5
2.1. Import Danych.....	12
2.2. Poprawki Danych w SMS (to do zrobienia w SSMS)	18
3. Etap 3 - Tabele wymiarów (dim_)	19
3.1. Tworzenie tabeli wymiaru dim_date (wymiar czasu)	20
3.1.1. Wypełnienie dim_date datami	20
3.2. Tworzenie wymiaru dim_airline	21
3.3. Tworzenie wymiaru dim_aircraft.....	22
3.4. Wymiar dim_airport	22
3.5. Tworzenie wymiaru dim_weather.....	23
4. Etap Trzeci - Budowa widoku tabeli faktów	23
5. Etap Czwarty - Stworzenie i załadowanie głównej tabeli fact_flights	24
6. Etap Piąty - Fundamenty pod kostkę OLAP	27
6.1. Tworzenie Data Source	29
6.2. Tworzenie wymiarów (dimension)	30
6.3. Widok źródłowy (Data Source View) i relacje	32
6.4. Kostka OLAP (Cube)	33
6.5. Tworzenie Hierarchi.....	35
6.6. Dodatkowe hierarchie: Weather, Aircraft, Airport, Airline	37
6.7. Sprawdzenie poprawności kostki OLAP w zakładce Browser	38
6.8. Stworzenie 4 scenariuszy/KPI.....	40
6.8.1. Obliczona miara: Średnie opóźnienie odlotów (Avg Departure Delay)	40
6.8.2. Obliczona miara: Średni czas lotu względem wieku samolotu (Avg Flight Duration per Aircraft Age)	41
6.8.3. Obliczona miara: Indeks opóźnień systemowych względem przewoźników (NAS vs Carrier Delay Index).....	44
6.8.4. Obliczona miara: Udział przewoźnika w opóźnieniach operacyjnych (Carrier Share of Delay).....	45
7. Data Mining	46
7.1. Model 1 – Drzewo Decyzyjne.....	47
7.2. Model 2 – Klasteryzacja.....	53
7.3. Model 3 – Sieć Neuronowa	62
7.4. Model 4 – Asociation Rules (zależności).....	65

7.5.	Model 5 - szeregi czasowe	68
8.	Power BI.....	70
8.1.	Radar typów opóźnień.....	72
8.2.	Analiza opóźnień względem dni tygodnia	73
8.3.	Mapa opóźnień lotniczych w USA.....	74
8.4.	Absorpcja opóźnień podczas lotu – analiza różnicy między opóźnieniem w odlocie a przylocie	
	75	
8.5.	Zależność między opóźnieniem odlotu i przylotu w zależności od miasta i czasu trwania lotu	
	76	
8.6.	Analiza wpływu średniej temperatury (Tavg) na opóźnienia przylotów i odlotów w różnych liniach	
lotniczych	78	
8.7.	Udział opóźnień lotów według typu i dnia tygodnia.....	79

1. Wstęp

Projekt dotyczy analizy danych dotyczących cywilnego ruchu lotniczego w Stanach Zjednoczonych w roku 2023. Głównym celem jest zbadanie zjawiska opóźnień lotów, ich możliwych przyczyn oraz identyfikacja wzorców poprzez integrację danych pogodowych, lokalizacyjnych oraz operacyjnych. Analiza prowadzona jest z wykorzystaniem narzędzi SQL Server Management Studio, Visual Studio (SSIS) oraz Tableau. Projekt wpisuje się w ramy wielowymiarowej analizy danych - obejmuje przetwarzanie, integrację i wizualizację danych w celu generowania mierzalnych wskaźników (KPI).

Celem projektu jest zidentyfikowanie głównych czynników wpływających na opóźnienia lotów oraz przedstawienie ich zależności z warunkami pogodowymi, lokalizacją i specyfiką lotnisk. Planujemy pokazać, które czynniki mają największy wpływ na punktualność lotów, jakie są wzorce opóźnień w zależności od pory roku, lotniska czy przewoźnika, a także, gdzie występują największe problemy operacyjne. Wyniki zaprezentujemy w formie przejrzystych raportów oraz interaktywnych wizualizacji.

Opis danych

W projekcie wykorzystano pięć kluczowych zbiorów danych, z których każdy wnosi istotne informacje analityczne:

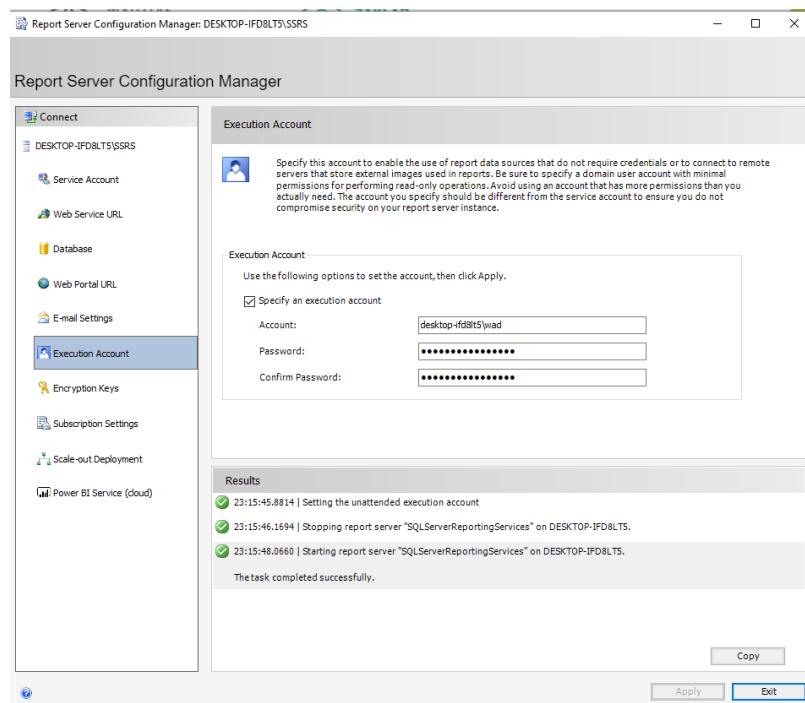
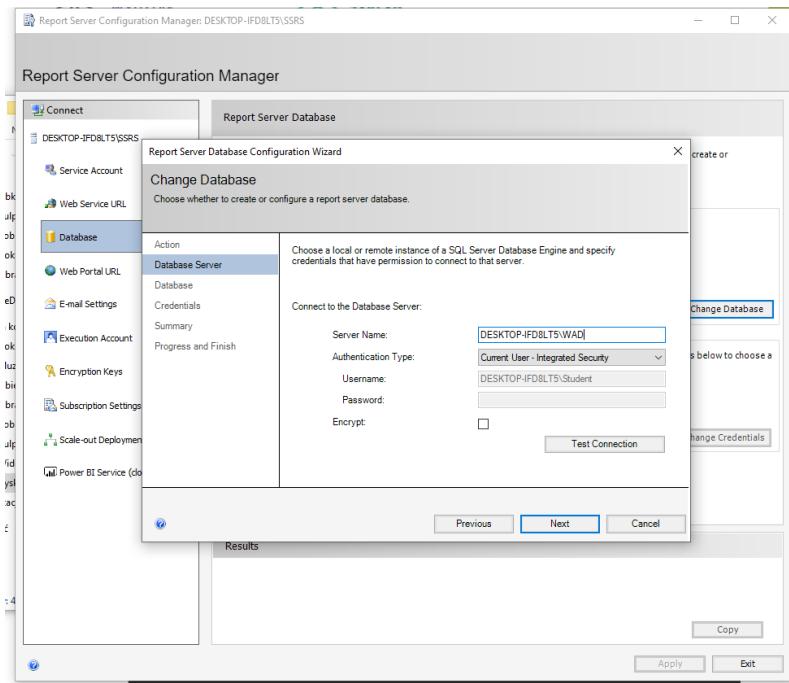
- **US_flights_2023.csv** - główny zbiór zawierający szczegółowe dane o ponad **6,5 miliona lotów** krajowych w USA w 2023 roku, w tym godziny odlotów i przylotów, opóźnienia, linie lotnicze, lotniska początkowe i docelowe oraz identyfikatory samolotów. Zbiór ten stanowi podstawę całej analizy.
- **Cancelled_Diverted_2023.csv** - wyodrębniona tabela zawierająca dane dotyczące anulowanych i przekierowanych lotów. Oddzielenie tych rekordów ma na celu uniknięcie zakłóceń w statystycznym ujęciu opóźnień oraz umożliwienie osobnej analizy przyczyn takich przypadków.
- **weather_meteo_by_airport.csv** - zbiór pogodowy z danymi meteorologicznymi dla każdego lotniska i dnia, w tym: opady, śnieg, ciśnienie atmosferyczne, temperatury (min./max./średnia), prędkość i kierunek wiatru. Dane te zostały pobrane z biblioteki Meteostat i są kluczowe dla analizy wpływu warunków atmosferycznych na punktualność lotów.
- **airports_geolocation.csv** - tabela z informacjami o lokalizacji lotnisk: identyfikatory, współrzędne geograficzne (szerokość i długość geograficzna) oraz nazwy. Pozwala na geograficzne mapowanie lotów i integrację z danymi pogodowymi.
- **maj us flight - january 2024.csv** - dodatkowy zbiór z danymi o lotach z początku roku 2024, uwzględniony w celu porównania trendów oraz rozszerzenia analizy o najnowszy okres.

Datasets zostały pobrane ze strony Kaggle :

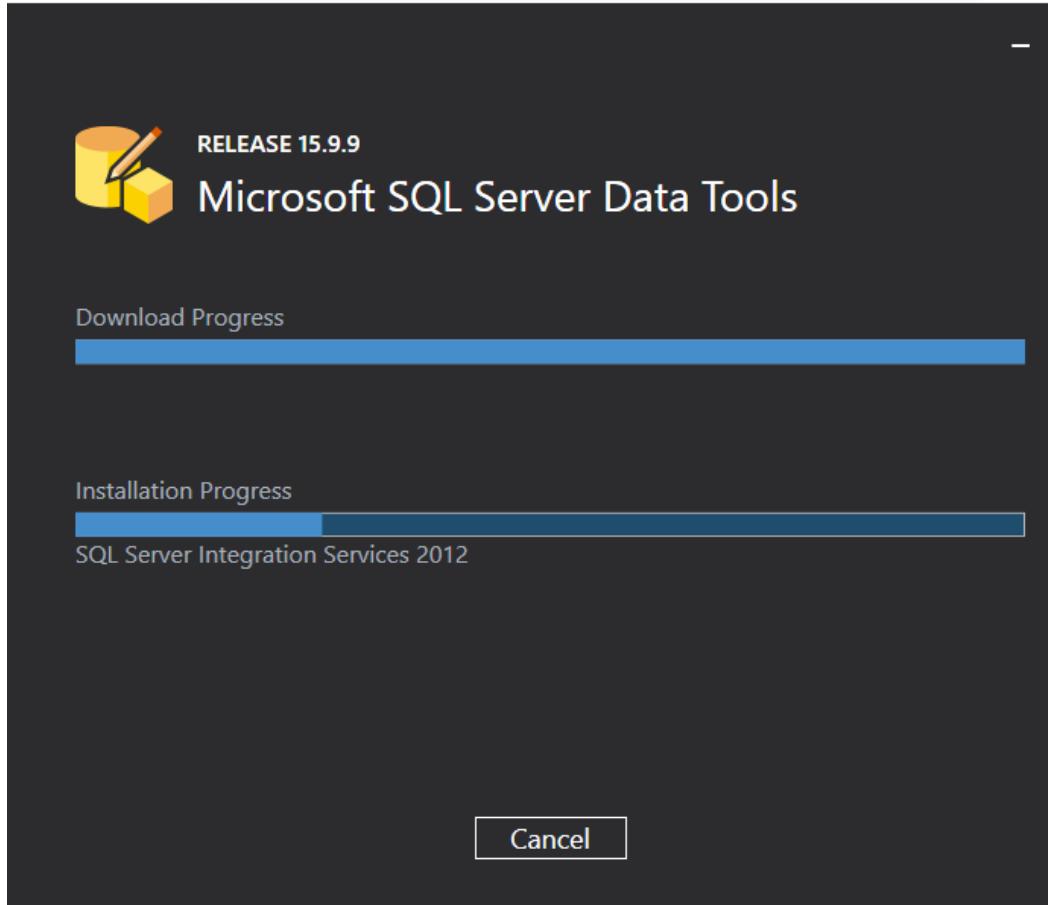
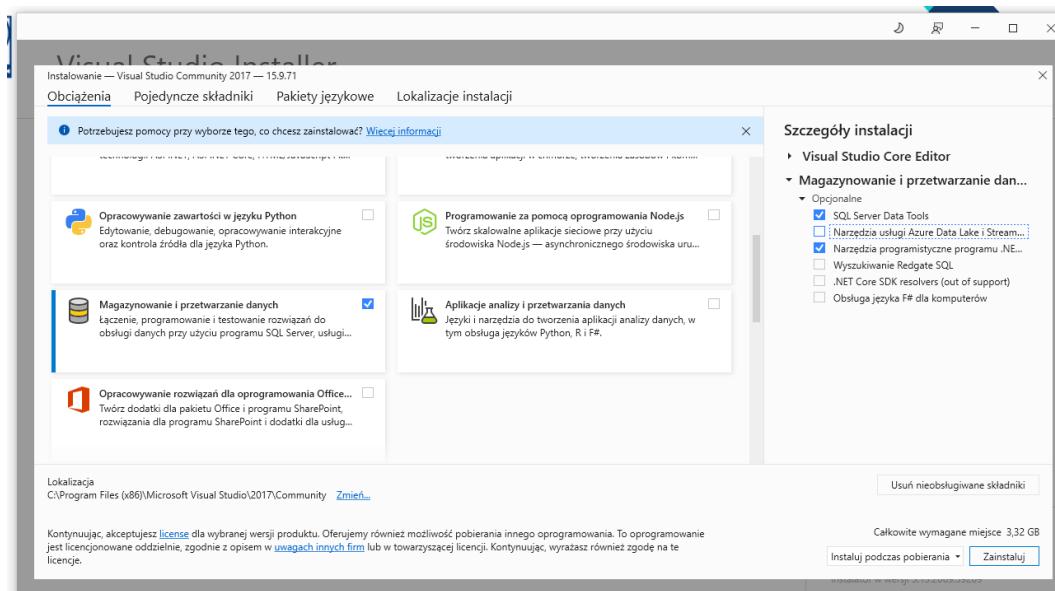
<https://www.kaggle.com/datasets/bordanova/2023-us-civil-flights-delay-meteo-and-aircraft>.

2. Etap Pierwszy

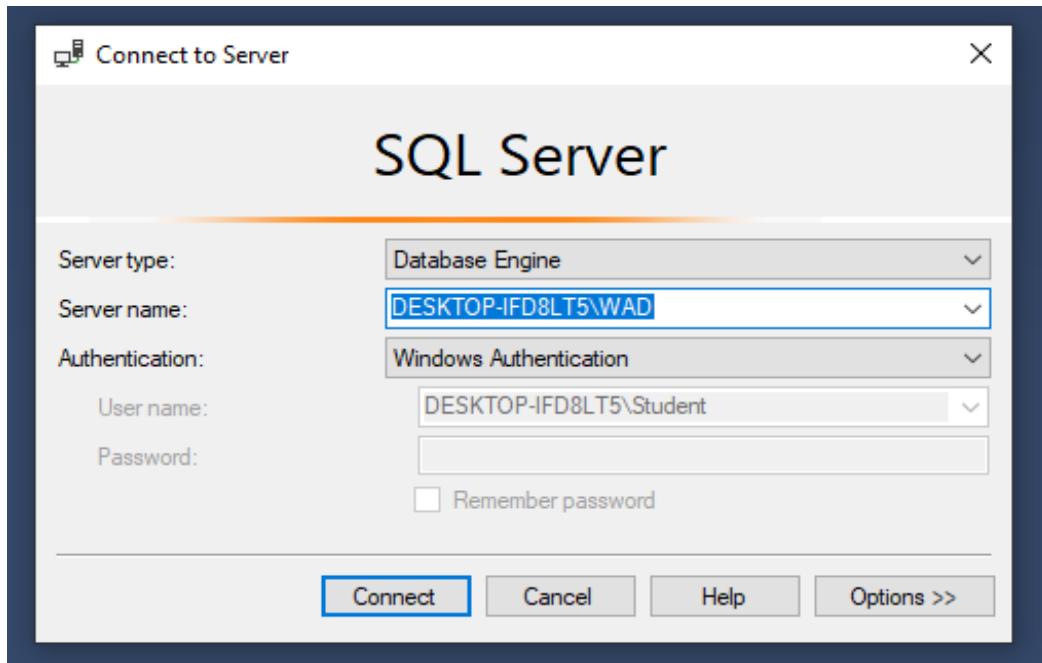
Pierwszym krokiem jest przygotowanie środowiska pracy. Pobrano SQL Server Management Studio oraz Visual Studio. Nazwa Serwera to: **DESKTOP-IFD8LT5\WAD**



Podczas instalacji VS wybrano dodatek **Magazynowanie i przetwarzanie danych** oraz **Microsoft SQL Server Data Tools**.



Pierwszą fundamentalną czynnością wykonaną podczas projektu było nawiązanie połączenia z SMS.



Kolejnym było utworzenie bazy danych o nazwie Lot na której odbywają się wszystkie operacje. Dodano jej również uproszczony model zapisywania danych, aby przyśpieszyć operacje na danych.

The screenshot shows the SQL Server Management Studio interface. In the Object Explorer, the connection path is visible: DESKTOP-IFD8LT5\WAD (SQL Server 15.0.2000.5 - DESKTOP-IFD8LT5) \ Databases \ lot. The SQL Query window contains the command: `Create database lot;` and `use lot;`. The Database Properties dialog for 'lot' is open, showing the following configuration:

Collation:	Polish_CI_AS
Recovery model:	Simple
Compatibility level:	SQL Server 2019 (150)
Containment type:	None
Other options:	Auto Close: False Auto Create Incremental Statistics: False Auto Create Statistics: True Auto Shrink: False Auto Update Statistics: True Auto Update Statistics Asynchronously: False
Containment:	Default Fulltext Language LCID: 1033 Default Language: English Nested Triggers Enabled: True Transform Noise Words: False Two Digit Year Cutoff: 2049
Cursor:	Close Cursor on Commit Enabled: False Default Cursor: GLOBAL
Database Scoped Configurations:	Legacy Cardinality Estimation: OFF Legacy Cardinality Estimation For Secondary: PRIMARY Max DOP: 0
Auto Close:	(empty)

Następnie został utworzony nowy użytkownik LotUser z odpowiednimi prawami, aby móc dokonywać operacji w bazie danych.

Login - New

Select a page: General, Server Roles, User Mapping, Securables, Status

Script, Help

Users mapped to this login:

Map	Database	User	Default Schema
<input checked="" type="checkbox"/>	lot	LotUser	
<input type="checkbox"/>	master		
<input type="checkbox"/>	model		
<input type="checkbox"/>	msdb		
<input type="checkbox"/>	ReportServer		
<input type="checkbox"/>	ReportServerTempDB		
<input type="checkbox"/>	tempdb		

Connection: Server: DESKTOP-IFD8LT5\WAD, Connection: DESKTOP-IFD8LT5\Student, View connection properties

Progress: Ready

Guest account enabled for: lot

Database role membership for: lot

<input type="checkbox"/> db_accessadmin
<input type="checkbox"/> db_backupoperator
<input type="checkbox"/> db_datareader
<input type="checkbox"/> db_datawriter
<input type="checkbox"/> db_ddladmin
<input type="checkbox"/> db_denydatareader
<input type="checkbox"/> db_denydatawriter
<input checked="" type="checkbox"/> db_owner
<input type="checkbox"/> db_securityadmin
<input checked="" type="checkbox"/> public

OK, Cancel

Tak jak w przerabianym sprawozdaniu, tabele zostały utworzone ręcznie w SSMS.

```
-- Tabela: lotniska_lokalizacja
CREATE TABLE lotniska_lokalizacja (
    IATA_CODE VARCHAR(50) PRIMARY KEY,
    AIRPORT VARCHAR(150),
    CITY VARCHAR(150),
    STATE VARCHAR(50),
    COUNTRY VARCHAR(150),
    LATITUDE FLOAT,
    LONGITUDE FLOAT
);

-- Tabela: pogoda_lotniska
CREATE TABLE pogoda_lotniska (
    time DATE,
    airport_id VARCHAR(50),
    tavg FLOAT,
    tmin FLOAT,
    tmax FLOAT,
    prcp FLOAT,
    snow FLOAT,
);

Commands completed successfully.

Completion time: 2025-04-12T22:13:00.5523449+02:00
```

```

-- Tabela: lotniska_lokalizacja
CREATE TABLE lotniska_lokalizacja (
    IATA_CODE VARCHAR(50) PRIMARY KEY,
    AIRPORT VARCHAR(150),
    CITY VARCHAR(150),
    STATE VARCHAR(50),
    COUNTRY VARCHAR(150),
    LATITUDE FLOAT,
    LONGITUDE FLOAT
);

-- Tabela: pogoda_lotniska
CREATE TABLE pogoda_lotniska (
    time DATE,
    airport_id VARCHAR(50),
    tavg FLOAT,
    tmin FLOAT,
    tmax FLOAT,
    prcp FLOAT,
    snow FLOAT,
    wdir FLOAT,
    wspd FLOAT,
    pres FLOAT,
    PRIMARY KEY (time, airport_id),
    FOREIGN KEY (airport_id) REFERENCES lotniska_lokalizacja(IATA_CODE)
);

```

```

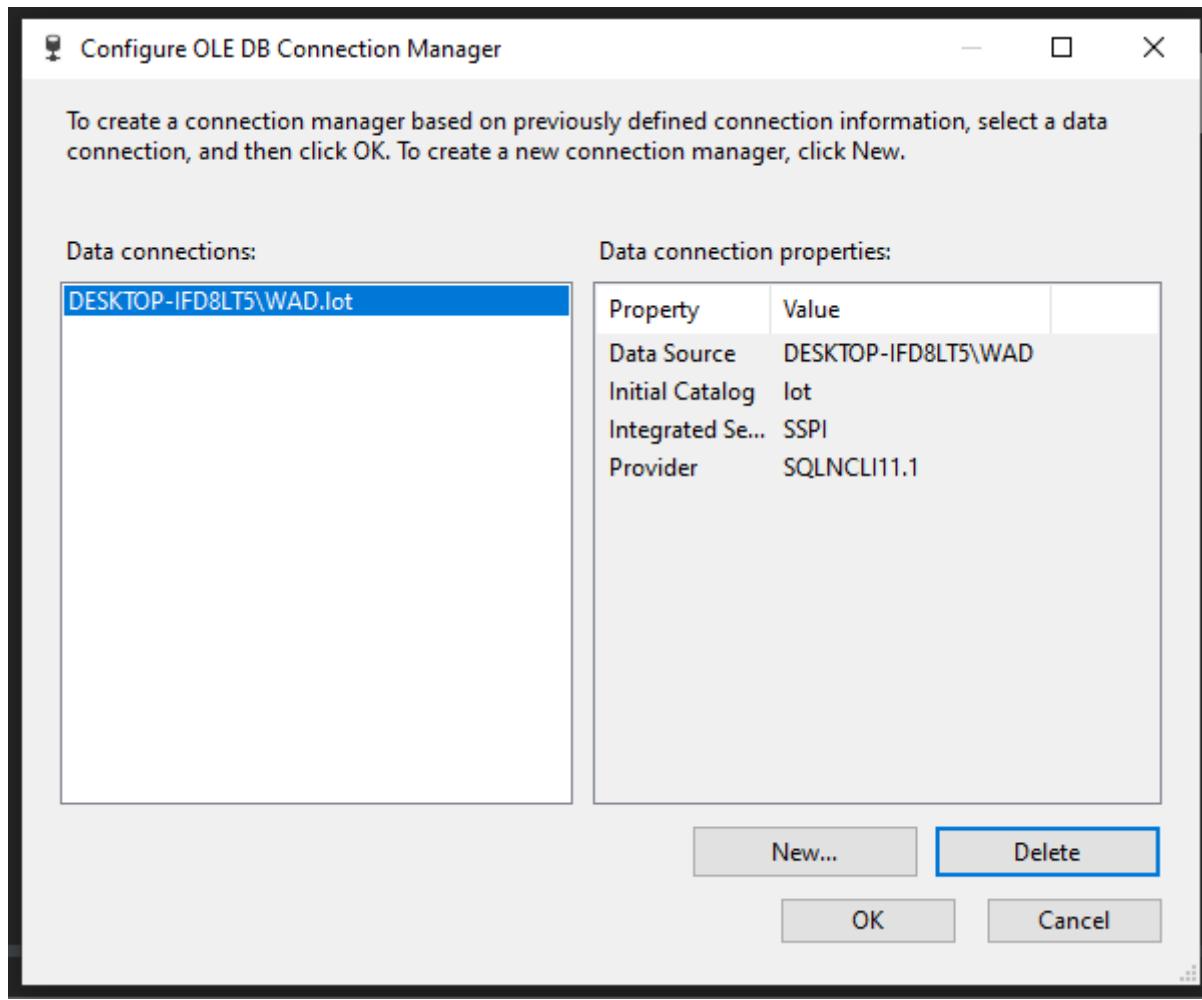
CREATE TABLE loty_usa_2023 (
    id BIGINT IDENTITY(1,1) PRIMARY KEY,
    FlightDate DATE,
    Day_Of_Week INT,
    Airline VARCHAR(100),
    Tail_Number VARCHAR(20),
    Dep_Airport VARCHAR(5) COLLATE Polish_CI_AS,
    Dep_CityName VARCHAR(100),
    DepTime_label VARCHAR(50),
    Dep_Delay FLOAT,
    Dep_Delay_Tag INT,
    Dep_Delay_Type VARCHAR(50),
    Arr_Airport VARCHAR(5) COLLATE Polish_CI_AS,
    Arr_CityName VARCHAR(100),
    Arr_Delay FLOAT,
    Arr_Delay_Type VARCHAR(50),
    Flight_Duration FLOAT,
    Distance_type VARCHAR(50),
    Delay_Carrier FLOAT,
    Delay_Weather FLOAT,
    Delay_NAS FLOAT,
    Delay_Security FLOAT,
    Delay_LastAircraft FLOAT,
    Manufacturer VARCHAR(100),
    Model VARCHAR(50),
    Aircraft_age INT,
);

```

```
CREATE TABLE loty_styczen_2024 (
    id BIGINT IDENTITY(1,1) PRIMARY KEY,
    FlightDate DATE,
    Day_Of_Week INT,
    Airline VARCHAR(100),
    Tail_Number VARCHAR(20),
    Dep_Airport VARCHAR(5) COLLATE Polish_CI_AS,
    Dep_CityName VARCHAR(100),
    Deptime_label VARCHAR(50),
    Dep_Delay FLOAT,
    Dep_Delay_Tag INT,
    Dep_Delay_Type VARCHAR(50),
    Arr_Airport VARCHAR(5) COLLATE Polish_CI_AS,
    Arr_CityName VARCHAR(100),
    Arr_Delay FLOAT,
    Arr_Delay_Type VARCHAR(50),
    Flight_Duration FLOAT,
    Distance_type VARCHAR(50),
    Delay_Carrier FLOAT,
    Delay_Weather FLOAT,
    Delay_NAS FLOAT,
    Delay_Security FLOAT,
    Delay_LastAircraft FLOAT,
    Manufacturer VARCHAR(100),
    Model VARCHAR(50),
    Aircraft_age INT
```

```
CREATE TABLE loty_odwolane_2023 (
    id BIGINT IDENTITY(1,1) PRIMARY KEY,
    FlightDate DATE,
    Day_Of_Week INT,
    Airline VARCHAR(100),
    Tail_Number VARCHAR(20),
    Cancelled BIT,
    Diverted BIT,
    Dep_Airport VARCHAR(5) COLLATE Polish_CI_AS,
    Dep_CityName VARCHAR(100),
    Deptime_label VARCHAR(50),
    Dep_Delay FLOAT,
    Dep_Delay_Tag INT,
    Dep_Delay_Type VARCHAR(50),
    Arr_Airport VARCHAR(5) COLLATE Polish_CI_AS,
    Arr_CityName VARCHAR(100),
    Arr_Delay FLOAT,
    Arr_Delay_Type VARCHAR(50),
    Flight_Duration FLOAT,
    Distance_type VARCHAR(50),
    Delay_Carrier FLOAT,
    Delay_Weather FLOAT,
    Delay_NAS FLOAT,
    Delay_Security FLOAT,
    Delay_LastAircraft FLOAT
);
```

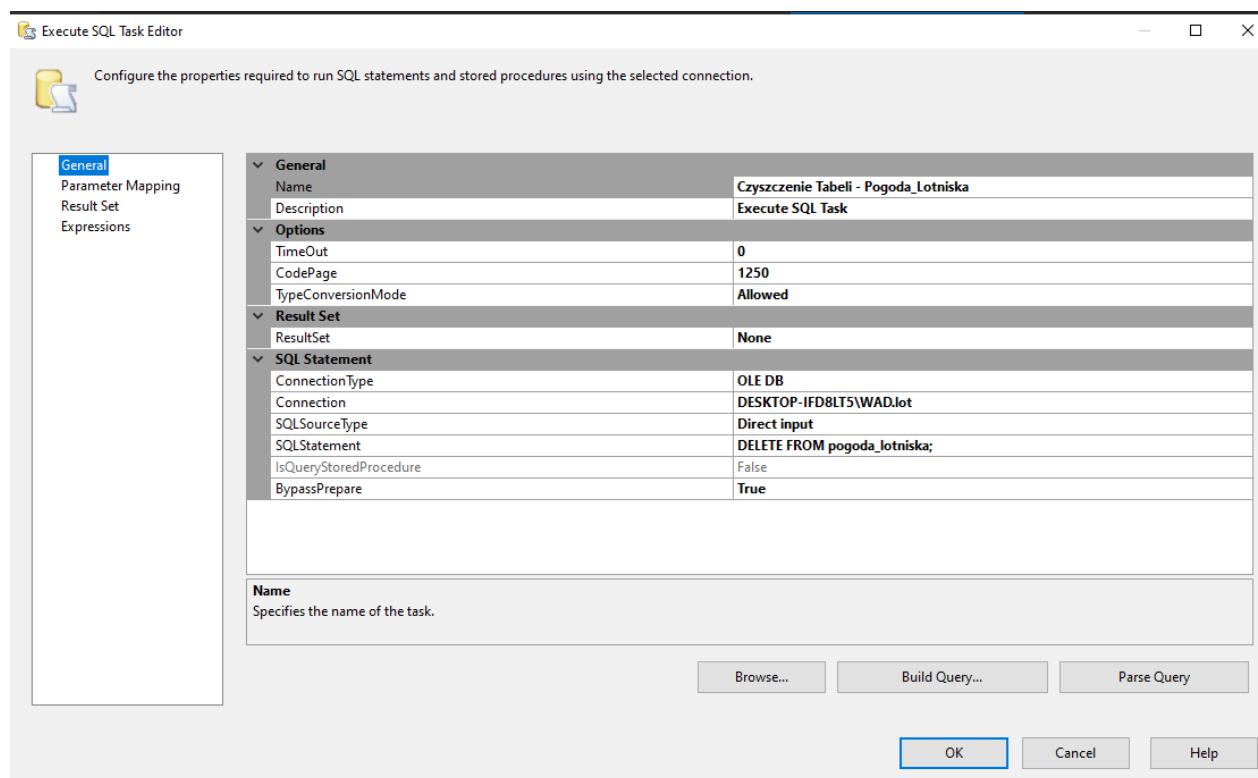
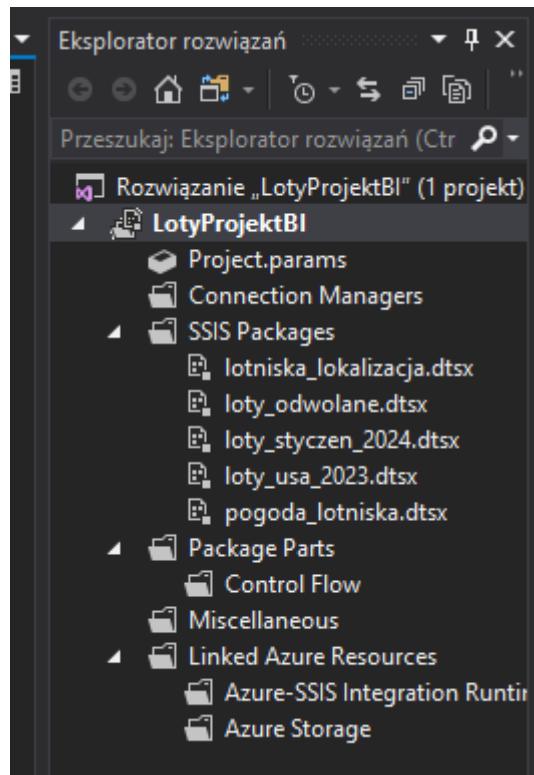
W Visual Studio po utworzeniu nowego projektu, od razu nawiązano kontakt z SSMS.

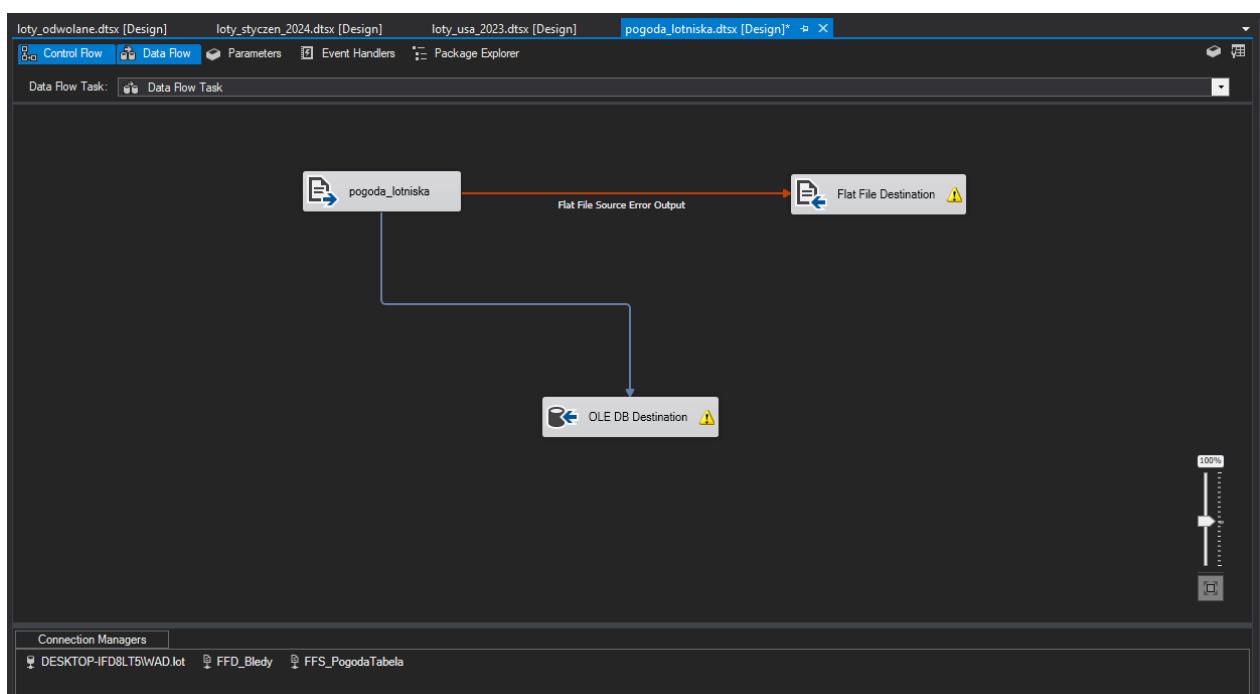
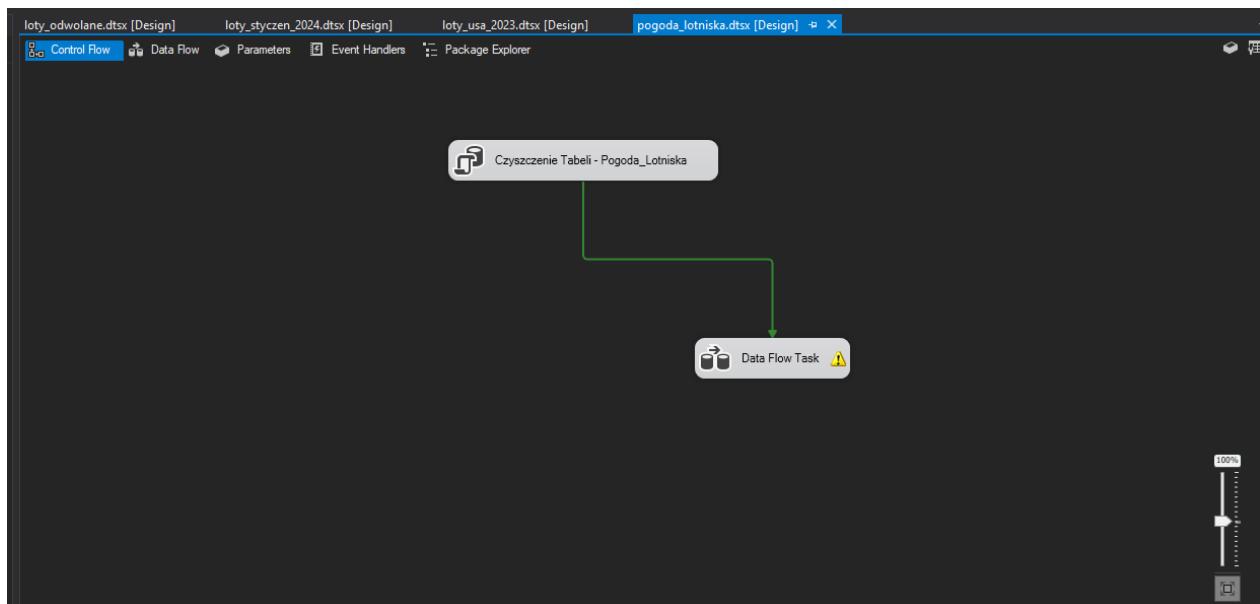


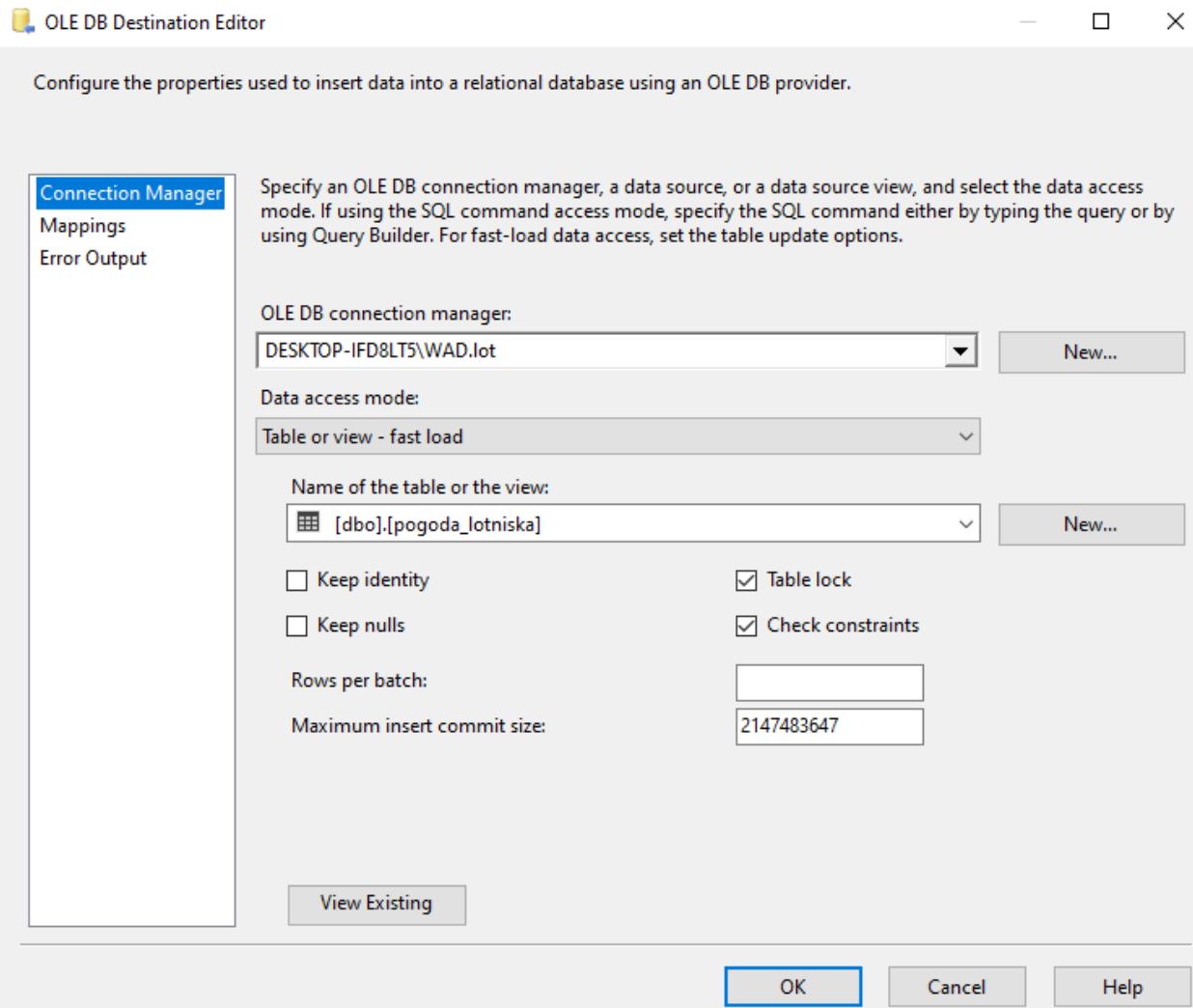
2.1. Import Danych

Poniższe zrzuty ekranu przedstawiają system importu danych do bazy. Najpierw Moduł Execute SQL czyści tabelę, a data flow task w każdym przypadku pobiera dane z pliku i umieszcza je w bazie danych.

Poniżej umieszczamy fragmenty z procesu ładowania danych do utworzonych tabel.







Wynik tych operacji przedstawiają poniższe zrzuty ekranu.

Object Explorer

Connect ▾ DESKTOP-ID8LT5\WAD (SQL Server 15.0.2000.5 - DESKTOP-ID8LT5)

- Databases
 - System Databases
 - Database Snapshots
 - lot
 - Database Diagrams
 - Tables
 - System Tables
 - FileTables
 - External Tables
 - Graph Tables
 - dbo.lotniska_kalizacja
 - dbo.loty_odwolane_2023
 - dbo.loty_styczen_2024
 - dbo.loty_usa_2023
 - dbo.pogoda_lotnika
 - Views
 - External Resources
 - Synonyms
 - Programmability
 - Service Broker
 - Storage
 - Security
 - ReportServer
 - ReportServerTempDB
 - Security
 - Server Objects
 - Replication
 - PolyBase
 - Always On High Availability
 - Management
 - Integration Services Catalogs
 - SQL Server Agent (Agent XPs disabled)
 - XEvent Profiler

SQLQuery1.sql - DE...DBLT5\Student (60)* + x

```
select * from lotniska_kalizacja
```

Results

IATA_CODE	AIRPORT	CITY	STATE	COUNTRY	LATITUDE	LONGITUDE
1 ABE	Lehigh Valley International Airport	Allentown	PA	USA	40.65236	-75.4404
2 ABI	Abilene Regional Airport	Abilene	TX	USA	32.41132	-99.6819
3 ABQ	Albuquerque International Sunport	Albuquerque	NM	USA	35.04022	-106.60919
4 ABR	Aberdeen Regional Airport	Aberdeen	SD	USA	45.44906	-98.42183
5 ABY	Southwest Georgia Regional Airport	Albany	GA	USA	31.53552	-84.19447
6 ACK	Nantucket Memorial Airport	Nantucket	MA	USA	41.25305	-70.06018
7 ACT	Waco Regional Airport	Waco	TX	USA	31.61129	-97.23052
8 ACV	Arcata Airport	Arcata/Eureka	CA	USA	40.97812	-124.10862
9 ACY	Atlantic City International Airport	Atlantic City	NJ	USA	39.3642	-74.4231
10 ADK	Adak Airport	Adak	AK	USA	51.87796	-176.64603
11 ADQ	Kodiak Airport	Kodiak	AK	USA	57.74997	-152.49386
12 AEX	Alexandria International Airport	Alexandria	LA	USA	31.32737	-92.54856
13 AGS	Augusta Regional Airport (Bush Field)	Augusta	GA	USA	33.36996	-81.9645
14 AKN	King Salmon Airport	King Salmon	AK	USA	58.6768	-156.64922
15 ALB	Albany International Airport	Albany	NY	USA	42.74812	-73.80298
16 ALO	Waterloo Regional Airport	Waterloo	IA	USA	42.55708	-92.40034
17 ALW	Walla Walla Regional Airport	Walla Walla	WA	USA	46.0646	-118.343
18 AMA	Rick Husband Amarillo International Airport	Amarillo	TX	USA	35.21937	-101.70593
19 ANC	Ted Stevens Anchorage International Airport	Anchorage	AK	USA	61.17432	-149.99619
20 APN	Alpena County Regional Airport	Alpena	MI	USA	45.07807	-83.56029
21 ASE	Aspen-Pitkin County Airport	Aspen	CO	USA	39.22316	-106.86885
22 ATL	Hartsfield-Jackson Atlanta International Airport	Atlanta	GA	USA	33.64044	-84.42694
23 ATW	Appleton International Airport	Appleton	WI	USA	44.25741	-88.51948
24 AUS	Austin-Bergstrom International Airport	Austin	TX	USA	30.19453	-97.66987
25 AVL	Asheville Regional Airport	Asheville	NC	USA	35.4333	-82.0333
26 AVP	Wilkes-Barre/Scranton International Airport	Wilkes-Barre/Scranton	PA	USA	41.3333	-75.0667
27 AZA	Phoenix Mesa Gateway Airport	Phoenix	AZ	USA	33.4484	-112.074
28 AZO	Kalamazoo/Battle Creek International Airport	Kalamazoo	MI	USA	42.23488	-85.55206
29 BDL	Bradley International Airport	Windsor Locks	CT	USA	41.93887	-72.68323
30 BET	Bethel Airport	Bethel	AK	USA	60.77978	-161.838
31 BFF	Western Nebraska Regional Airport	Scottsbluff	NE	USA	41.8666	-103.672
32 BFL	Meadows Field	Bakersfield	CA	USA	35.4336	-119.05677
33 BGM	Greater Binghamton Airport	Binghamton	NY	USA	42	-75.15
34 BGR	Bangor International Airport	Bangor	ME	USA	44.80744	-68.82814
35 BHM	Birmingham-Shuttlesworth International Airport	Birmingham	AL	USA	33.56294	-86.75355
36 BIH	Bishop Airport	Bishop	CA	USA	37.3667	-118.3667
37 BIL	Billings Logan International Airport	Billings	MT	USA	45.80766	-108.54286
38 BIS	Bismarck Municipal Airport	Bismarck	ND	USA	46.77411	-100.74672
39 BJI	Bemidji Regional Airport	Bemidji	MN	USA	47.50942	-94.93372
40 BLI	Bellingham International Airport	Bellingham	WA	USA	48.79275	-122.53753
41 BLV	Saint Louis Mid America Airport	Bellefonte	IL	USA	38.6245	-90.1509
42 BMI	Central Illinois Regional Airport at Bloomington...	Bloomington	IL	USA	40.47799	-88.91595
43 BNA	Nashville International Airport	Nashville	TN	USA	36.12448	-86.67818
44 BOI	Boise Airport (Boise Air Terminal)	Boise	ID	USA	43.56444	-116.22278
45 BOS	Gen. Edward Lawrence Logan International ...	Boston	MA	USA	42.36435	-71.00518
46 BPT	Jack Brooks Regional Airport (Southeast T...	Beaumont/Port Arthur	TX	USA	29.95083	-94.02069
47 BQK	Brunswick Golden Isles Airport	Brunswick	GA	USA	31.25903	-81.46631
48 BQN	Rafael Hernández Airport	Aguadilla	PR	USA	18.49486	-67.12944
49 BRD	Brainerd Lakes Regional Airport	Brainerd	MN	USA	46.39785	-94.13723
50 BRO	Brownsville/South Padre Island International...	Brownsville	TX	USA	25.90683	-97.42586
51 BRW	Wiley Post-Will Rogers Memorial Airport	Barrow	AK	USA	71.28545	-156.766
52 BTM	Bert Mooney Airport	Butte	MT	USA	45.9548	-112.49746
53 BTR	Baton Rouge Metropolitan Airport	Baton Rouge	LA	USA	30.53316	-91.14963
54 RTV	Rutland International Airport	Rutland	VT	USA	44.473	-73.15031

SQLQuery1.sql - DE...DBLT5\Student (60)* + x

```
SELECT TOP 100 *
FROM loty_usa_2023;
```

Results

id	RightDate	Day_of_Week	Airline	Tail_Number	Dep_Airport	Dep_CityName	Dep_Time_Job	Dep_Delay	Dep_Delay_Tag	Dep_Delay_Type	Ar_Airport	Ar_CityName	Ar_Delay	Ar_Delay_Type	Right_Duration	Distance_type	Delay_Carrier	Delay_Weather	Delay_NAS	Delay_Security	Delay_LastAircraft	Manufacturer	Model	Aircraft_Age
1	674305	2023-01-02	1	Endeavor Air	N605LR	BDL	Portland, OR	-3	0	Low <5min	LGA	New York, NY	-12	Low <5min	56	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16	
2	674306	2023-01-03	2	Endeavor Air	N605LR	BDL	Portland, OR	-5	0	Low <5min	LGA	New York, NY	-8	Low <5min	62	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	10	
3	674307	2023-01-04	3	Endeavor Air	N331PK	BDL	Portland, OR	-5	0	Low <5min	LGA	New York, NY	-21	Low <5min	49	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	17	
4	674308	2023-01-05	4	Endeavor Air	N505KJ	BDL	Portland, OR	-6	0	Low <5min	LGA	New York, NY	-17	Low <5min	54	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	10	
5	674309	2023-01-06	5	Endeavor Air	N337PK	BDL	Portland, OR	-1	0	Low <5min	LGA	New York, NY	-16	Low <5min	50	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	10	
6	674310	2023-01-07	6	Endeavor Air	N336PK	BDL	Portland, OR	-10	0	Low <5min	LGA	New York, NY	-13	Low <5min	62	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	10	
7	674311	2023-01-08	7	Endeavor Air	N311PK	LGA	New York, NY	-8	0	Low <5min	CVG	Cincinnati, OH	-31	Low <5min	117	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	10	
8	674312	2023-01-09	8	Endeavor Air	N304KJ	LGA	New York, NY	-10	0	Low <5min	CVG	Cincinnati, OH	-25	Low <5min	125	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16	
9	674313	2023-01-10	9	Endeavor Air	N305PK	LGA	New York, NY	-10	0	Low <5min	CVG	Cincinnati, OH	-15	Low <5min	130	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16	
10	674314	2023-01-11	10	Endeavor Air	N491PK	LGA	New York, NY	-7	0	Low <5min	BGM	Binghamton, NY	-3	Low <5min	63	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	4	
11	674315	2023-01-12	10	Endeavor Air	N479PK	LGA	New York, NY	-15	0	Low <5min	BGM	Binghamton, NY	6	Low <5min	60	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	5	
12	674316	2023-01-13	11	Endeavor Air	N135EV	LGA	New York, NY	-15	1	Medium >15min	BGM	Binghamton, NY	61	High >30min	85	Short Haul >15000	0	0	26	0	CANADAIR REGIONAL JET	CRJ	15	
13	674317	2023-01-14	12	Endeavor Air	N137PK	LGA	New York, NY	-12	1	High >30min	BGM	Binghamton, NY	122	High >30min	53	Short Haul >15000	50	0	0	82	CANADAIR REGIONAL JET	CRJ	16	
14	674318	2023-01-15	13	Endeavor Air	N515KU	LGA	New York, NY	-5	0	Low <5min	BGM	Binghamton, NY	34	Medium >15min	98	Short Haul >15000	0	0	34	0	CANADAIR REGIONAL JET	CRJ	17	
15	674319	2023-01-16	14	Endeavor Air	N520KU	LGA	New York, NY	-5	1	Low <5min	BGM	Binghamton, NY	4	Low <5min	98	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16	
16	674320	2023-01-17	15	Endeavor Air	N515KU	LGA	New York, NY	-3	0	Low <5min	BGM	Binghamton, NY	14	Low <5min	76	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16	
17	674321	2023-01-18	2	Endeavor Air	N302KU	LGA	New York, NY	-15	1	Low <5min	BGM	Binghamton, NY	40	Medium >15min	84	Short Haul >15000	0	0	40	0	CANADAIR REGIONAL JET	CRJ	17	
18	674322	2023-01-19	3	Endeavor Air	N491PK	LGA	New York, NY	-7	0	Low <5min	BGM	Binghamton, NY	-2	Low <5min	64	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	4	
19	674323	2023-01-20	4	Endeavor Air	N691CA	LGA	New York, NY	-15	1	High >30min	BGM	Binghamton, NY	673	High >30min	56	Short Haul >15000	0	667	0	6	CANADAIR REGIONAL JET	CRJ	17	
20	674324	2023-01-21	5	Endeavor Air	N904KU	LGA	New York, NY	-10	0	Low <5min	BGM	Binghamton, NY	36	Medium >15min	105	Short Haul >15000	0	0	36	0	CANADAIR REGIONAL JET	CRJ	17	
21	674325	2023-01-22	7	Endeavor Air	N310PK	LGA	New York, NY	-29	1	Medium >15min	BGM	Binghamton, NY	47	Medium >15min	77	Short Haul >15000	29	0	18	0	CANADAIR REGIONAL JET	CRJ	10	
22	674326	2023-01-23	1	Endeavor Air	N197KU	LGA	New York, NY	-27	1	Medium >15min	BGM	Binghamton, NY	22	Medium >15min	54	Short Haul >15000	0	0	0	22	CANADAIR REGIONAL JET	CRJ	16	
23	674327	2023-01-24	2	Endeavor Air	N182GJ	LGA	New York, NY	-3	0	Low <5min	BGM	Binghamton, NY	-18	Low <5min	44	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16	
24	674328	2023-01-25	3	Endeavor Air	N490PK	LGA	New York, NY	-50	1	Medium >15min	BGM	Binghamton, NY	61	High >30min	70	Short Haul >15000	25	0	11	0	CANADAIR REGIONAL JET	CRJ	4	
25	674329	2023-01-26	4	Endeavor Air	N134EV	LGA	New York, NY	-13	0	Low <5min	BGM	Binghamton, NY	-3	Low <5min	69	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	15	
26	674330	2023-01-27	5	Endeavor Air	N515KU	LGA	New York, NY	-11	0	Low <5min	BGM	Binghamton, NY	4	Low <5min	66	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16	
27	674331	2023-01-28	7	Endeavor Air	N489PK	LGA	New York, NY	-19	0	Low <5min	BGM	Binghamton, NY	-17	Low <5min	61	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	4	
28	674332	2023-01-29	10	Endeavor Air	N491PK	LGA	New York, NY	-10	0	Low <5min	BGM	Binghamton, NY	6	Low <5min	63	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	4	
29	674333	2023-01-31	2	Endeavor Air	N489PK	LGA	New York, NY	-16	1	Medium >15min	BGM	Binghamton, NY	10	Low <5min	53	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	4	
30	674334	2023-02-01	7	Endeavor Air	N904DU	DUL	Duluth, MN	Night	6	1	Low <5min	MSP	Minneapolis, MN	24	Medium >15min	94	Short Haul >15000	6	0	18	0	CANADAIR REGIONAL JET	CRJ	17
31	674335	2023-02-02	1	Endeavor Air	N197PK	DUL	Duluth, MN	Night	-8	0	Low <5min	MSP	Minneapolis, MN	-4	Low <5min	80	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16
32	674336	2023-02-03	1	Endeavor Air	N181GJ	ATL	Atlanta, GA	Afternoon	-3	0	Low <5min	FAT	Fayetteville, NC	-12	Low <5min	67	Short Haul >15000	0	0	0	0	CANADAIR REGIONAL JET	CRJ	16

Object Explorer

Connect ▾ DESKTOP-IDF8LT5\WAD (SQL Server 15.0.2000.5 - DESKTOP-IDF8LT5)

- Databases
 - System Databases
 - Database Snapshots
- Iot
 - Database Diagrams
 - Tables
 - System Tables
 - FileTables
 - External Tables
 - Graph Tables
 - dbo.lotniska_lokalizacja
 - dbo.loty_odwoalone_2023
 - dbo.loty_styczen_2024
 - dbo.loty_usa_2023
 - dbo.pogoda_lotniska
 - Views
 - External Resources
 - Synonyms
 - Programmability
 - Service Broker
 - Storage
 - Security
- ReportServer
- ReportServerTempDB
- Security
- Server Objects
- Replication
- PolyBase
- Always On High Availability
- Management
- Integration Services Catalogs
 - SQL Server Agent (Agent XPs disabled)
- XEvent Profiler

SQLQuery1.sql - DES...t (60) Executing...*

```
select * from pogoda_lotniska;
```

Results Messages

time	airport_id	tavg	tmin	tmax	prcp	snow	wdir	wspd	pres
1 2023-01-01	ABE	8.1	2.2	11.7	0	0	278	9.7	1013.8
2 2023-01-01	ABI	16.4	8.9	26.1	0	0	184	15.8	1009
3 2023-01-01	ABQ	6.2	2.2	10	4.3	0	111	11.5	1009.1
4 2023-01-01	ABR	-11.2	-16.6	-3.2	0	180	267	6.8	1016.1
5 2023-01-01	ABY	17.5	10.6	23.9	0	0	333	3.2	1020.1
6 2023-01-01	ACK	9.2	7	12.2	6.3	0	266	22.4	1009.5
7 2023-01-01	ACT	17.8	10	26.1	0	0	172	16.6	1011.6
8 2023-01-01	ACV	6.3	2.2	11.1	0	0	55	8.4	1016
9 2023-01-01	ACY	12.7	1.7	15	0	0	281	12.6	1014.1
10 2023-01-01	ADK	1.8	0	3.3	4.8	0	344	27.7	989
11 2023-01-01	ADQ	4.2	3.3	5	4.3	0	77	15.5	964.7
12 2023-01-01	AEX	17.3	10.6	21.7	0	0	172	6.5	1016
13 2023-01-01	AGS	16.4	7.8	20	0	0	238	9.4	1019.4
14 2023-01-01	AKN	2.2	-3.9	3.9	0	0	13	16.9	964.3
15 2023-01-01	ALB	4.2	0.6	7.6	0.2	0	274	13.9	1012.5
16 2023-01-01	ALO	-0.2	-6.6	4.4	0	50	350	1.8	1014.2
17 2023-01-01	ALW	-1	-2	1.1	0	0	254	7.4	1019.4
18 2023-01-01	AMA	13.6	8.3	21.7	0	0	233	19.4	1005.5
19 2023-01-01	ANC	-2.3	-2.7	1.7	3.6	610	351	10.4	971.7
20 2023-01-01	APN	1.1	-0.5	2.8	1.8	0	201	8.3	1013.1
21 2023-01-01	ASE	2.4	0.6	4.4	3.3	0	161	15.1	1003
22 2023-01-01	ATL	14.1	8.9	21.1	0	0	336	3.7	1019
23 2023-01-01	ATW	1.3	0	4	0	0	233	11.5	1011.4
24 2023-01-01	AUS	14.8	5	25.6	0	0	156	12.2	1012.3
25 2023-01-01	AVL	9.7	5	18.3	0	0	342	3.6	1018.8
26 2023-01-01	AVP	10.4	5	14.4	0	0	250	8.3	1013.3
27 2023-01-01	AZA	14.2	8.9	16.7	10.9	0	129	15.5	1006.4
28 2023-01-01	AZO	4.4	2.2	7.2	4	0	213	10.5	1014.3

SQLQuery1.sql - DES...t (60) Executing...*

```
select * from loty_odwoalone_2023;
```

Results Messages

id	FlightDate	Day_Of_Week	Airline	Tai_Number	Canceled	Diverted	Dep_Airport	Dep_CityName	DepTime_Label	Dep_Delay	Dep_Delay_Tag	Dep_Delay_Type	Arr_Airport	Arr_CityName	Arr_Delay	Arr_Delay_Type	Flight_Duration	Distance_Type	Delay_Caller	Delay_Weather	Delay_NAS	Delay_Security	Delay_LastAircraft	
1	2023-01-25	3	Endeavor Air	N691CA	1	0	JFK	New York, NY	Evening	0	0	No Departure Delay	ITH	Ithaca-Confard, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
2	2023-01-03	2	Endeavor Air	N937XJ	1	0	MSP	Minneapolis, MN	Morning	0	0	No Departure Delay	PIT	Pittsburgh, PA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
3	2023-01-03	2	Endeavor Air	N937XJ	1	0	PIT	Pittsburgh, PA	Afternoon	0	0	No Departure Delay	MSP	Minneapolis, MN	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
4	2023-01-11	3	Endeavor Air	N526XJ	1	0	BWI	Baltimore, MD	Afternoon	0	0	No Departure Delay	JFK	New York, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
5	2023-01-25	3	Endeavor Air	N335PQ	0	1	BWI	Baltimore, MD	Afternoon	2	0	No Departure Delay	JFK	New York, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
6	2023-01-11	3	Endeavor Air	N526XJ	1	0	JFK	New York, NY	Afternoon	0	0	No Departure Delay	BWI	Baltimore, MD	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
7	2023-01-18	3	Endeavor Air	N678CA	1	0	MSP	Minneapolis, MN	Evening	0	0	No Departure Delay	CID	Cedar Rapids/Iowa City, IA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
8	2023-01-23	1	Endeavor Air	N183GJ	1	0	LGA	New York, NY	Afternoon	0	0	No Departure Delay	SDF	Louisville, KY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0	
9	9	2023-01-03	2	Endeavor Air	N330PQ	1	0	CID	Cedar Rapids/Iowa City, IA	Morning	0	0	No Departure Delay	MSP	Minneapolis, MN	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
10	10	2023-01-16	1	Endeavor Air	N301PQ	1	0	BGR	Bangor, ME	Night	0	0	No Departure Delay	LGA	New York, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
11	11	2023-01-25	3	Endeavor Air	N303PQ	1	0	ROU	Raleigh/Durham, NC	Evening	0	0	No Departure Delay	JFK	New York, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
12	12	2023-01-11	3	Endeavor Air	N367CA	0	1	EVV	Evergreen, CO	Morning	-4	0	No Departure Delay	ATL	Atlanta, GA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
13	13	2023-01-03	2	Endeavor Air	N937XJ	1	0	MSP	Minneapolis, MN	Morning	0	0	No Departure Delay	CVG	Cincinnati, OH	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
14	14	2023-01-09	1	Endeavor Air	N482PK	1	0	LGA	New York, NY	Morning	0	0	No Departure Delay	CVG	Cincinnati, OH	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
15	15	2023-01-11	3	Endeavor Air	N478PK	1	0	LGA	New York, NY	Evening	0	0	No Departure Delay	RDU	Raleigh/Durham, NC	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
16	16	2023-01-03	2	Endeavor Air	N932XJ	1	0	JFK	New York, NY	Afternoon	0	0	No Departure Delay	MSP	Minneapolis, MN	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
17	17	2023-01-23	1	Endeavor Air	N504XJ	1	0	RIC	Richmond, VA	Afternoon	0	0	No Departure Delay	LGA	New York, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
18	18	2023-01-25	3	Endeavor Air	N593BA	1	0	DTW	Detroit, MI	Afternoon	-4	0	No Departure Delay	MBS	Saginaw/Bay City/Midland, MI	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
19	19	2023-01-25	3	Endeavor Air	N692BA	1	0	MBS	Saginaw/Bay City/Midland, MI	Afternoon	0	0	No Departure Delay	DTW	Detroit, MI	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
20	20	2023-01-22	7	Endeavor Air	N8877A	1	0	ATL	Atlanta, GA	Morning	0	0	No Departure Delay	MGM	Montgomery, AL	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
21	21	2023-01-30	1	Endeavor Air	N8968B	1	0	ATL	Atlanta, GA	Morning	0	0	No Departure Delay	MGM	Montgomery, AL	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
22	22	2023-01-22	7	Endeavor Air	N8877A	1	0	MGM	Montgomery, AL	Morning	0	0	No Departure Delay	ATL	Atlanta, GA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
23	23	2023-01-30	1	Endeavor Air	N8968B	1	0	MGM	Montgomery, AL	Morning	0	0	No Departure Delay	ATL	Atlanta, GA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
24	24	2023-01-30	1	Endeavor Air	N478PK	0	1	LGA	New York, NY	Morning	-6	0	No Departure Delay	PIT	Pittsburgh, PA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
25	25	2023-01-03	2	Endeavor Air	N330PQ	1	0	MSP	Minneapolis, MN	Evening	0	0	No Departure Delay	CID	Cedar Rapids/Iowa City, IA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
26	26	2023-01-23	1	Endeavor Air	N134EV	1	0	PIT	Pittsburgh, PA	Morning	0	0	No Departure Delay	LGA	New York, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
27	27	2023-01-25	3	Endeavor Air	N526XJ	1	0	JFK	New York, NY	Evening	0	0	No Departure Delay	SAV	Savannah, GA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
28	28	2023-01-03	2	Endeavor Air	N906XJ	1	0	MOT	Mnnt, ND	Afternoon	0	0	No Departure Delay	MSP	Minneapolis, MN	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
29	29	2023-01-03	2	Endeavor Air	N906XJ	1	0	MSP	Minneapolis, MN	Morning	0	0	No Departure Delay	MOT	Mnnt, ND	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
30	30	2023-01-25	3	Endeavor Air	N307PQ	1	0	BNA	Nashville, TN	Evening	0	0	No Departure Delay	JFK	New York, NY	0	No Arrival Delay	0	Short Haul	0	0	0	0	0
31	31	2023-01-03	2	Endeavor Air	N933XJ	0	1	DTW	Detroit, MI	Morning	-2	0	No Departure Delay	PHL	Philadelphia, PA	0	No Arrival Delay	0	Short Haul	0	0	0	0	0

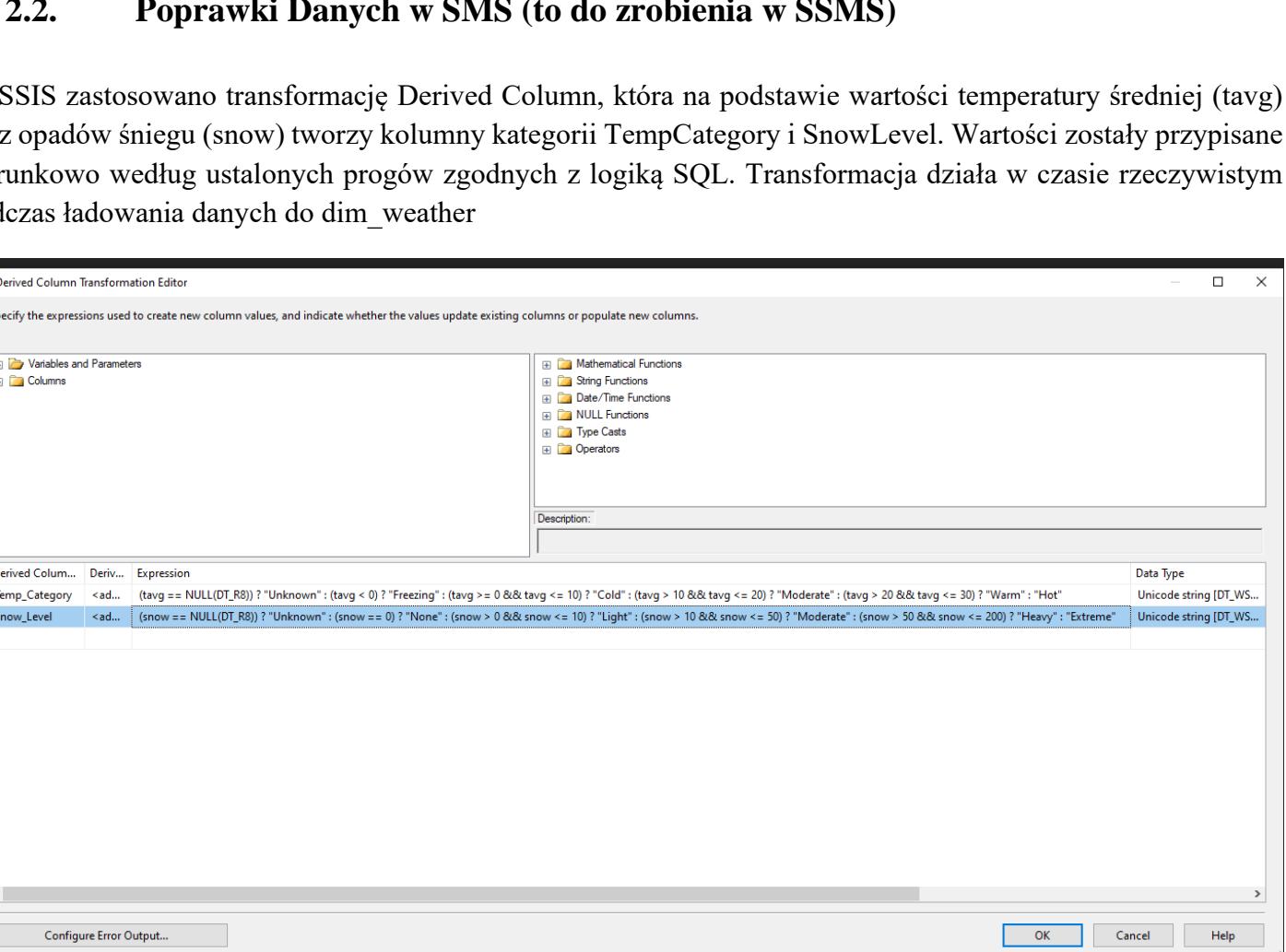
SQLQuery1.sql - DES...t (80) Executing... 4 X

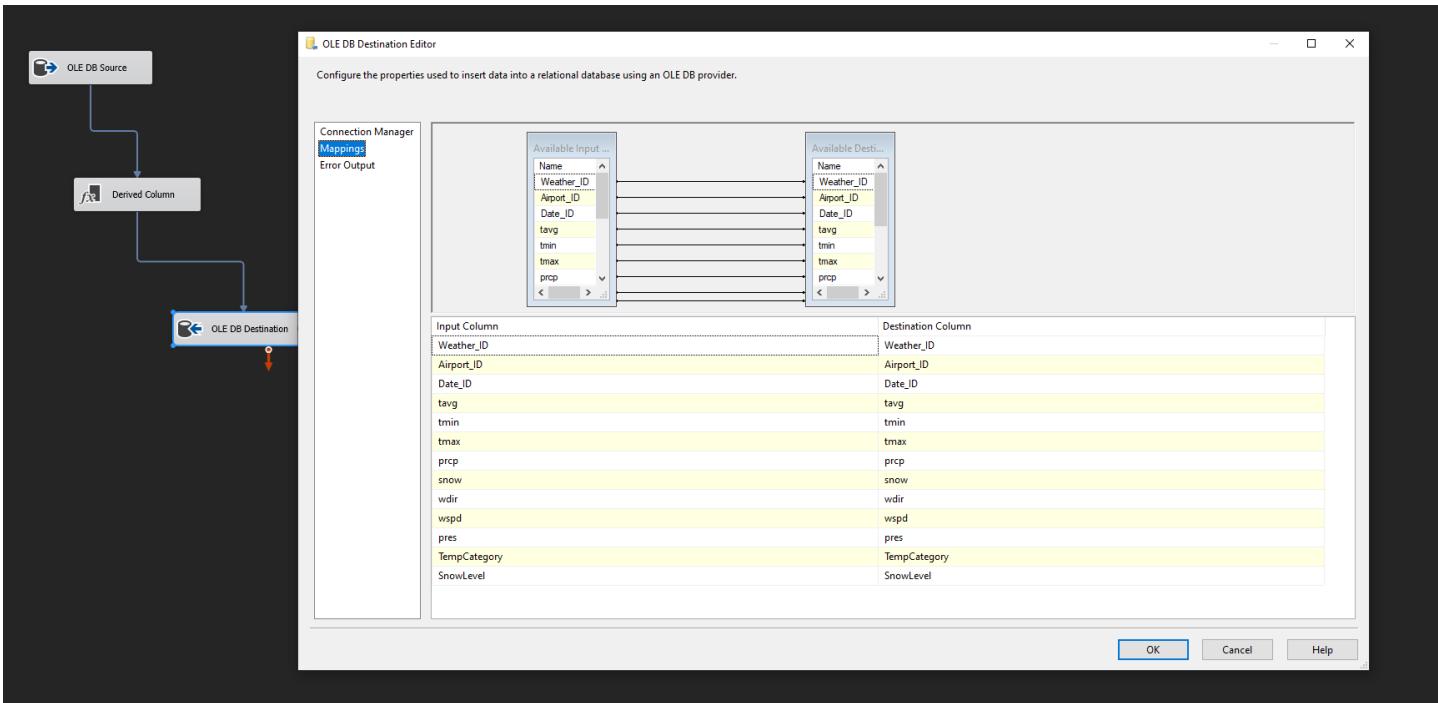
```
select * from loty_styczen_2024
```

109 %

Results [Messages]

	id	RunDate	Day_of_Week	Airline	Tail_Number	Dep_Airport	Dep_CityName	DepTime_Label	Dep_Delay	Dep_Delay_Tag	Dep_Delay_Type	Arr_Report	Arr_CityName	Arr_Delay	Arr_Delay_Type	Flight_Duration	Distance_Type	Delay_Caller	Delay_Weather	Delay_NAS	Delay_Security	Delay_LastAircraft	Manufacturer	Model	Aircraft_Age
1	1	2023-01-02	1	Endeavor Air	N155LR	BDL	Hartford, CT	Morning	-3	0	Low <5min	LGA	New York, NY	-12	Low <5min	55	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
2	2	2023-01-03	2	Endeavor Air	N155PQ	BDL	Hartford, CT	Morning	-5	0	Low <5min	LGA	New York, NY	-8	Low <5min	62	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
3	3	2023-01-04	3	Endeavor Air	N155X	BDL	Hartford, CT	Morning	-5	0	Low <5min	LGA	New York, NY	-21	Low <5min	49	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 10	10	
4	4	2023-01-05	4	Endeavor Air	N155XU	BDL	Hartford, CT	Morning	-4	0	Low <5min	LGA	New York, NY	-17	Low <5min	54	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 17	17	
5	5	2023-01-06	5	Endeavor Air	N155PQ	BDL	Hartford, CT	Morning	-1	0	Low <5min	LGA	New York, NY	-16	Low <5min	50	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 10	10	
6	6	2023-01-07	6	Endeavor Air	N155PQ	BDL	Hartford, CT	Morning	-10	0	Low <5min	LGA	New York, NY	-13	Low <5min	62	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 10	10	
7	7	2023-01-14	7	Endeavor Air	N151PQ	LGA	New York, NY	Afternoon	-8	0	Low <5min	CVG	Cincinnati, OH	-31	Low <5min	117	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 10	10	
8	8	2023-01-21	8	Endeavor Air	N151XU	LGA	New York, NY	Afternoon	-10	0	Low <5min	CVG	Cincinnati, OH	-25	Low <5min	125	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
9	9	2023-01-28	9	Endeavor Air	N155PQ	LGA	New York, NY	Afternoon	-5	0	Low <5min	CVG	Cincinnati, OH	-15	Low <5min	130	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 10	10	
10	10	2023-01-09	1	Endeavor Air	N1451PX	LGA	New York, NY	Evening	-7	0	Low <5min	BGM	Brighton, NY	-3	Low <5min	63	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 4	4	
11	11	2023-01-10	2	Endeavor Air	N1451PX	LGA	New York, NY	Evening	-15	0	Low <5min	BGM	Brighton, NY	6	Low <5min	80	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 5	5	
12	12	2023-01-11	3	Endeavor Air	N135EV	LGA	New York, NY	Evening	35	1	Medium >15min	BGM	Brighton, NY	51	High >5min	85	Short Haul >1500M	0	0	26	0	CANADAIR REGIONAL JET	CRJ 15	15	
13	13	2023-01-12	4	Endeavor Air	N151PQ	LGA	New York, NY	Evening	132	1	High >5min	BGM	Brighton, NY	132	High >5min	59	Short Haul >1500M	50	0	0	0	82	CANADAIR REGIONAL JET	CRJ 16	16
14	14	2023-01-13	5	Endeavor Air	N151XU	LGA	New York, NY	Evening	-5	0	Low <5min	BGM	Brighton, NY	34	Medium >15min	98	Short Haul >1500M	0	0	34	0	CANADAIR REGIONAL JET	CRJ 17	17	
15	15	2023-01-15	7	Endeavor Air	N152XU	LGA	New York, NY	Evening	5	1	Low <5min	BGM	Brighton, NY	4	Low <5min	58	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
16	16	2023-01-16	1	Endeavor Air	N151XU	LGA	New York, NY	Evening	-3	0	Low <5min	BGM	Brighton, NY	14	Low <5min	76	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
17	17	2023-01-17	2	Endeavor Air	N152XU	LGA	New York, NY	Evening	15	1	Low <5min	BGM	Brighton, NY	40	Medium >15min	84	Short Haul >1500M	0	0	40	0	CANADAIR REGIONAL JET	CRJ 17	17	
18	18	2023-01-18	3	Endeavor Air	N1451PX	LGA	New York, NY	Evening	-7	0	Low <5min	BGM	Brighton, NY	-2	Low <5min	64	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 4	4	
19	19	2023-01-19	4	Endeavor Air	N151CA	LGA	New York, NY	Evening	576	1	High >5min	BGM	Brighton, NY	573	High >5min	56	Short Haul >1500M	0	0	667	0	CANADAIR REGIONAL JET	CRJ 17	17	
20	20	2023-01-20	5	Endeavor Air	N155XU	LGA	New York, NY	Evening	-10	0	Low <5min	BGM	Brighton, NY	36	Medium >15min	105	Short Haul >1500M	0	0	36	0	CANADAIR REGIONAL JET	CRJ 17	17	
21	21	2023-01-22	7	Endeavor Air	N151PQ	LGA	New York, NY	Evening	29	1	Medium >15min	BGM	Brighton, NY	132	High >5min	77	Short Haul >1500M	29	0	18	0	0	CANADAIR REGIONAL JET	CRJ 10	10
22	22	2023-01-23	1	Endeavor Air	N151XU	LGA	New York, NY	Evening	27	1	Medium >15min	BGM	Brighton, NY	22	Medium >15min	54	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
23	23	2023-01-24	2	Endeavor Air	N152XU	LGA	New York, NY	Evening	-3	0	Low <5min	BGM	Brighton, NY	-18	Low <5min	44	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
24	24	2023-01-25	3	Endeavor Air	N155PQ	LGA	New York, NY	Evening	50	1	Medium >15min	BGM	Brighton, NY	-61	High >5min	70	Short Haul >1500M	25	0	11	0	CANADAIR REGIONAL JET	CRJ 4	4	
25	25	2023-01-26	4	Endeavor Air	N151EV	LGA	New York, NY	Evening	-13	0	Low <5min	BGM	Brighton, NY	-3	Low <5min	69	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 15	15	
26	26	2023-01-27	5	Endeavor Air	N151CA	LGA	New York, NY	Evening	-11	0	Low <5min	BGM	Brighton, NY	-4	Low <5min	66	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
27	27	2023-01-29	7	Endeavor Air	N1451PX	LGA	New York, NY	Evening	-19	0	Low <5min	BGM	Brighton, NY	-17	Low <5min	61	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 4	4	
28	28	2023-01-30	1	Endeavor Air	N1451PX	LGA	New York, NY	Evening	-10	0	Low <5min	BGM	Brighton, NY	-6	Low <5min	63	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 4	4	
29	29	2023-01-31	2	Endeavor Air	N1451PX	LGA	New York, NY	Evening	16	1	Medium >15min	BGM	Brighton, NY	-10	Low <5min	53	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 4	4	
30	30	2023-01-31	7	Endeavor Air	N151PQ	DLU	Duluth, MN	Night	6	1	Low <5min	MSP	Minneapolis, MN	24	Medium >15min	34	Short Haul >1500M	6	0	18	0	CANADAIR REGIONAL JET	CRJ 17	17	
31	31	2023-01-02	1	Endeavor Air	N151XU	DLU	Duluth, MN	Night	-8	0	Low <5min	MSP	Minneapolis, MN	-6	Low <5min	60	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
32	32	2023-01-09	1	Endeavor Air	N151XU	ATL	Atlanta, GA	Afternoon	-3	0	Low <5min	FAY	Fayetteville, NC	-12	Low <5min	67	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
33	33	2023-01-10	2	Endeavor Air	N152XU	ATL	Atlanta, GA	Afternoon	5	0	Low <5min	FAY	Fayetteville, NC	-12	Low <5min	69	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
34	34	2023-01-11	3	Endeavor Air	N151CA	ATL	Atlanta, GA	Afternoon	2	0	Low <5min	FAY	Fayetteville, NC	-7	Low <5min	71	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
35	35	2023-01-12	4	Endeavor Air	N151CA	ATL	Atlanta, GA	Afternoon	16	1	Medium >15min	FAY	Fayetteville, NC	-11	Low <5min	71	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 21	21	
36	36	2023-01-13	5	Endeavor Air	N152PQ	ATL	Atlanta, GA	Afternoon	-3	0	Low <5min	FAY	Fayetteville, NC	-8	Low <5min	71	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 10	10	
37	37	2023-01-14	6	Endeavor Air	N151PQ	ATL	Atlanta, GA	Afternoon	-1	0	Low <5min	FAY	Fayetteville, NC	-4	Low <5min	73	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 10	10	
38	38	2023-01-15	7	Endeavor Air	N151PQ	ATL	Atlanta, GA	Afternoon	-3	0	Low <5min	FAY	Fayetteville, NC	-7	Low <5min	72	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
39	39	2023-01-16	1	Endeavor Air	N151XU	ATL	Atlanta, GA	Afternoon	-3	0	Low <5min	FAY	Fayetteville, NC	-9	Low <5min	70	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 17	17	
40	40	2023-01-17	2	Endeavor Air	N151PQ	ATL	Atlanta, GA	Afternoon	-5	0	Low <5min	FAY	Fayetteville, NC	-7	Low <5min	74	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
41	41	2023-01-18	3	Endeavor Air	N151XU	ATL	Atlanta, GA	Afternoon	5	0	Low <5min	FAY	Fayetteville, NC	-17	Low <5min	64	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 17	17	
42	42	2023-01-19	4	Endeavor Air	N151XU	ATL	Atlanta, GA	Afternoon	-5	0	Low <5min	FAY	Fayetteville, NC	-11	Low <5min	70	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 15	15	
43	43	2023-01-20	5	Endeavor Air	N152PQ	ATL	Atlanta, GA	Afternoon	0	0	Low <5min	FAY	Fayetteville, NC	-7	Low <5min	69	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 11	11	
44	44	2023-01-21	6	Endeavor Air	N152PQ	ATL	Atlanta, GA	Afternoon	-6	0	Low <5min	FAY	Fayetteville, NC	-23	Low <5min	59	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 11	11	
45	45	2023-01-22	7	Endeavor Air	N151XU	ATL	Atlanta, GA	Afternoon	-2	0	Low <5min	FAY	Fayetteville, NC	-2	Low <5min	76	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 17	17	
46	46	2023-01-23	1	Endeavor Air	N155LR	ATL	Atlanta, GA	Afternoon	-1	0	Low <5min	FAY	Fayetteville, NC	-10	Low <5min	67	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 16	16	
47	47	2023-01-24	2	Endeavor Air	N155PQ	ATL	Atlanta, GA	Afternoon	-5	0	Low <5min	FAY	Fayetteville, NC	-15	Low <5min	66	Short Haul >1500M	0	0	0	0	CANADAIR REGIONAL JET	CRJ 11	11	





3. Etap 3 - Tabele wymiarów (dim_)

W tym etapie dokonano logicznego rozdzielenia danych na tabele faktów i tabele wymiarów, zgodnie z klasycznym modelem *schema gwiazdy* (star schema). Dane opisowe zostały przeniesione do osobnych tabel wymiarów (dim_*), a tabela faktów (fact_flights_with_weather) zawiera wyłącznie klucze techniczne i miary liczbowe.

Celem tej operacji było uporządkowanie struktury danych w sposób przyjazny dla systemów analitycznych typu OLAP i Power BI.

Dlaczego to zrobiliśmy:

- Zgodność z Analysis Services (OLAP):** SSAS wymaga jasnego podziału między miary (np. opóźnienia, liczba lotów), a wymiary (np. linia lotnicza, data, typ samolotu).
- Poprawa wydajności:** Dzięki usunięciu powtarzalnych danych tekstowych (np. nazwy lotnisk), model danych jest lżejszy i bardziej zoptymalizowany.
- Logika analityczna:** Wymiarowe tabele pozwalają na tworzenie hierarchii (np. Rok → Miesiąc → Dzień) oraz łatwe filtrowanie i grupowanie danych.
- Skalowalność i elastyczność:** W przyszłości możliwe będzie dodanie nowych atrybutów (np. „czy lotnisko jest międzynarodowe”) bez ingerencji w dane faktów.

3.1. Tworzenie tabeli wymiaru dim_date (wymiar czasu)

- grupowanie lotów po miesiącach, dniach tygodnia, kwartałach,
- łączenie dat z loty_usa_2023.FlightDate,
- analizowanie sezonowości, weekendy itd.

```
SQLQuery1.sql - DE...D8LT5\Student (58)* ✎ X
CREATE TABLE dim_date (
    Date_ID INT PRIMARY KEY,           -- np. 20230101
    Full_Date DATE,
    Year INT,
    Month INT,
    Month_Name VARCHAR(20),
    Day INT,
    Day_Of_Week INT,                  -- 1 = Poniedziałek
    Weekday_Name VARCHAR(15),
    Is_Weekend BIT
);
```

3.1.1. Wypełnienie dim_date datami

Wprowadzić do tabeli dim_date dane od 2023-01-01 do 2024-12-31, z automatycznie wyliczonymi właściwościami (rok, miesiąc, nazwa dnia tygodnia, itd.).

```
SQLQuery1.sql - DE...D8LT5\Student (58)* ✎ X
WITH Dates AS (
    SELECT CAST('2023-01-01' AS DATE) AS d
    UNION ALL
    SELECT DATEADD(DAY, 1, d) FROM Dates WHERE d < '2024-12-31'
)
INSERT INTO dim_date (
    Date_ID, Full_Date, Year, Month, Month_Name, Day, Day_Of_Week, Weekday_Name, Is_Weekend
)
SELECT
    CONVERT(INT, FORMAT(d, 'yyyyMMdd')) AS Date_ID,
    d AS Full_Date,
    DATEPART(YEAR, d),
    DATEPART(MONTH, d),
    DATENAME(MONTH, d),
    DATEPART(DAY, d),
    DATEPART(WEEKDAY, d),
    DATENAME(WEEKDAY, d),
    CASE WHEN DATENAME(WEEKDAY, d) IN ('Saturday', 'Sunday') THEN 1 ELSE 0 END
FROM Dates
OPTION (MAXRECURSION 800);
```

SQLQuery1.sql - DE...D8LT5\Student (58)* ↗ X

```
SELECT TOP 5 * FROM dim_date ORDER BY Full_Date;
```

109 %

	Date_ID	Full_Date	Year	Month	Month_Name	Day	Day_Of_Week	Weekday_Name	Is_Weekend
1	20230101	2023-01-01	2023	1	January	1	1	Sunday	1
2	20230102	2023-01-02	2023	1	January	2	2	Monday	0
3	20230103	2023-01-03	2023	1	January	3	3	Tuesday	0
4	20230104	2023-01-04	2023	1	January	4	4	Wednesday	0
5	20230105	2023-01-05	2023	1	January	5	5	Thursday	0

3.2. Tworzenie wymiaru dim_airline

- zawiera **unikalne nazwy przewoźników** (Airline_Name),
- przypisuje im **numer ID** (Airline_ID),
- umożliwia późniejsze **łączenie** z tabelą faktów i filtrowanie raportów po nazwie przewoźnika.

SQLQuery1.sql - DE...D8LT5\Student (58)* ↗ X

```
CREATE TABLE dim_airline (
    Airline_ID INT IDENTITY(1,1) PRIMARY KEY,
    Airline_Name VARCHAR(100) UNIQUE
);
```

SQLQuery1.sql - DE...D8LT5\Student (58)* ↗ X

```
SELECT * FROM dim_airline ORDER BY Airline_Name;
```

109 %

	Airline_ID	Airline_Name
1	15	Alaska Airlines Inc.
2	14	Allegiant Air
3	13	American Airlines Inc.
4	7	American Eagle Airlines Inc.
5	3	Delta Air Lines Inc
6	10	Endeavor Air
7	6	Frontier Airlines Inc.
8	4	Hawaiian Airlines Inc.
9	1	JetBlue Airways
10	12	PSA Airlines
11	11	Republic Airways
12	2	Skywest Airlines Inc.
13	8	Southwest Airlines Co.
14	5	Spirit Air Lines
15	9	United Air Lines Inc.

3.3. Tworzenie wymiaru dim_aircraft

- wpływ wieku maszyny na opóźnienia,
- który model najczęściej się spóźnia,
- czy opóźnienia zależą od producenta.

```
SQLQuery1.sql - DE...D8LT5\Student (58)* ✎ X
CREATE TABLE dim_aircraft (
    Aircraft_ID INT IDENTITY(1,1) PRIMARY KEY,
    Tail_Number VARCHAR(20),
    Manufacturer VARCHAR(100),
    Model VARCHAR(50),
    Aircraft_Age INT,
    UNIQUE (Tail_Number, Manufacturer, Model)
);
```

```
SQLQuery1.sql - DE...D8LT5\Student (58)* ✎ X
SELECT TOP 5 * FROM dim_aircraft ORDER BY Aircraft_ID;
```

109 %

	Aircraft_ID	Tail_Number	Manufacturer	Model	Aircraft_Age
1	1	N236NN	EMBRAER	170/175	8
2	2	N439UA	AIRBUS	A320	27
3	3	N68807	BOEING	737 NG	11
4	4	N489UA	AIRBUS	A320	22
5	5	N549CA	CANADAIR REGIONAL JET	CRJ	16

3.4. Wymiar dim_airport

Zmieniamy tylko nazwę obecnej tabeli

```
SQLQuery1.sql - DE...D8LT5\Student (58)* ✎ X
EXEC sp_rename 'lotniska_lokalizacja', 'dim_airport';
```

3.5. Tworzenie wymiaru dim_weather

The screenshot shows a SQL query window titled "SQLQuery1.sql - DE...D8LT5\Student (58)*". The code creates a table named "dim_weather" with the following columns and constraints:

```
CREATE TABLE dim_weather (
    Weather_ID INT IDENTITY(1,1) PRIMARY KEY,
    Airport_ID VARCHAR(5),
    Date_ID INT,
    tavg FLOAT,
    tmin FLOAT,
    tmax FLOAT,
    prcp FLOAT,
    snow FLOAT,
    wdir FLOAT,
    wspd FLOAT,
    pres FLOAT,
    UNIQUE (Airport_ID, Date_ID)
);
```

4. Etap Trzeci - Budowa widoku tabeli faktów

W tym etapie został utworzony widok SQL (vw_flight_fact_input), który pełni rolę warstwy pośredniej między tabelą źródłową *loty_usa_2023*, a tabelą faktów *fact_flights_with_weather*. Jego zadaniem jest przygotowanie danych w formacie zgodnym z modelem schematu gwiazdy.

- **Zgodność z tabelą faktów:** Tabela *fact_flights* powinna zawierać wyłącznie klucze obce oraz miary liczbowe. Widok umożliwia to rozdzielenie w sposób zautomatyzowany.
- **Centralizacja logiki mapowania:** Zamiast pisać wiele zapytań z *JOIN-ami* do każdej analizy, tworzymy jeden widok, który łączy dane z wymiarów i źródła w spójną całość.
- **Eliminacja duplikacji i błędów:** Widok gwarantuje jednolite przypisanie ID do każdej encji (np. tej samej linii lotniczej), co zwiększa spójność modelu.

```

SQLQuery1.sql - DE...D8LT5\Student (58)* ✎ X
CREATE VIEW vw_flight_fact_input AS
SELECT
    1.id AS Flight_ID,
    -- klucze do wymiarów
    CONVERT(INT, FORMAT(1.FlightDate, 'yyyyMMdd')) AS Dep_Date_ID,
    CONVERT(INT, FORMAT(1.FlightDate, 'yyyyMMdd')) AS Arr_Date_ID,
    1.Dep_Airport AS Dep_Airport_ID,
    1.Arr_Airport AS Arr_Airport_ID,
    a.Airline_ID,
    ac.Aircraft_ID,
    -- miary
    1.Dep_Delay,
    1.Arr_Delay,
    1.Flight_Duration,
    1.Delay_Carrier,
    1.Delay_Weather,
    1.Delay_NAS,
    1.Delay_Security,
    1.Delay_LastAircraft,
    CAST(0 AS BIT) AS Cancelled, -- tylko dla loty_usa_2023
    CAST(0 AS BIT) AS Diverted
FROM loty_usa_2023 1
JOIN dim_airline a ON 1.Airline = a.Airline_Name
JOIN dim_aircraft ac
    ON 1.Tail_Number = ac.Tail_Number
    AND 1.Manufacturer = ac.Manufacturer
    AND 1.Model = ac.Model;

```

SQLQuery1.sql - DE...D8LT5\Student (58)* ✎ X

```

SELECT TOP 10 * FROM vw_flight_fact_input;

```

90 %

Results Messages

	Flight_ID	Dep_Date_ID	Arr_Date_ID	Dep_Airport_ID	Arr_Airport_ID	Airline_ID	Aircraft_ID	Dep_Delay	Arr_Delay
1	6743405	20230102	20230102	BDL	LGA	10	2507	-3	-12
2	6743406	20230103	20230103	BDL	LGA	10	2507	-5	-8
3	6743407	20230104	20230104	BDL	LGA	10	5053	-5	-21
4	6743408	20230105	20230105	BDL	LGA	10	1327	-6	-17
5	6743409	20230106	20230106	BDL	LGA	10	2350	-1	-16
6	6743410	20230107	20230107	BDL	LGA	10	5101	-10	-13
7	6743411	20230114	20230114	LGA	CVG	10	3134	-8	-31
8	6743412	20230121	20230121	LGA	CVG	10	985	-10	-25
9	6743413	20230128	20230128	LGA	CVG	10	5101	-5	-15
10	6743414	20230109	20230109	LGA	BGM	10	15	-7	-3

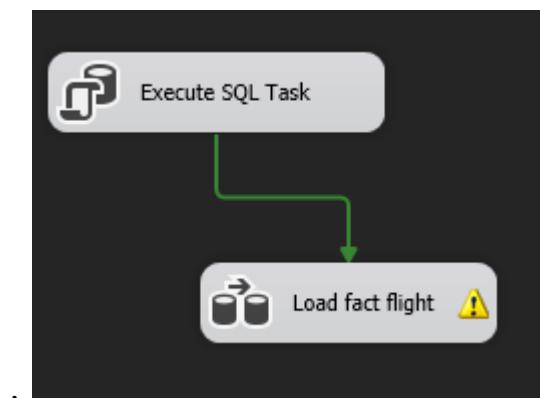
5. Etap Czwarty - Stworzenie i załadowanie głównej tabeli fact_flights

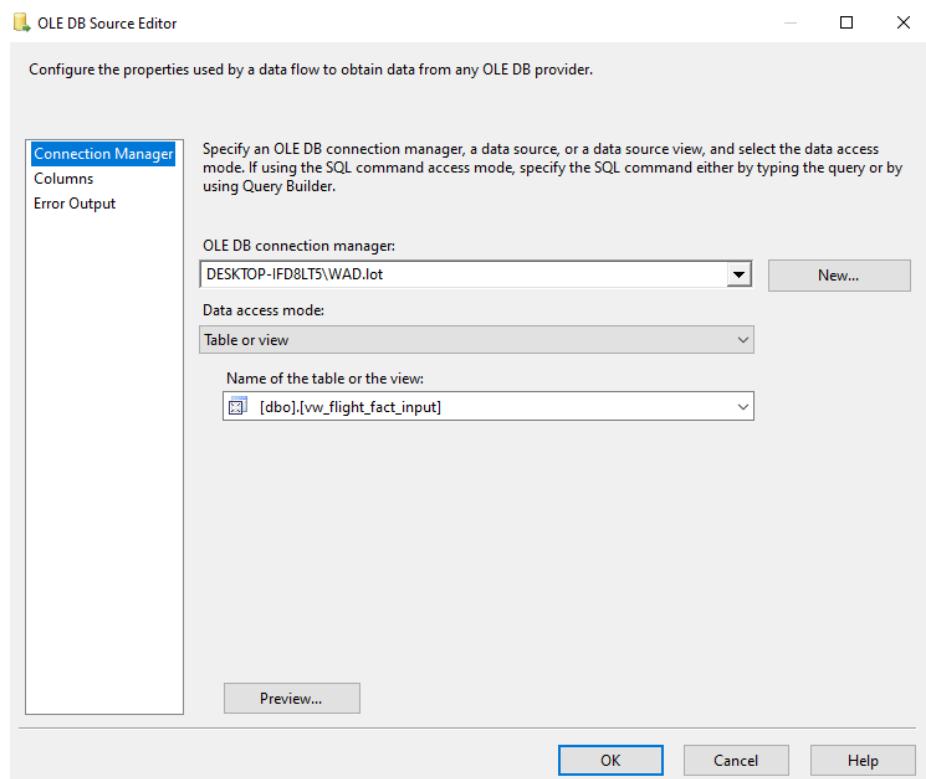
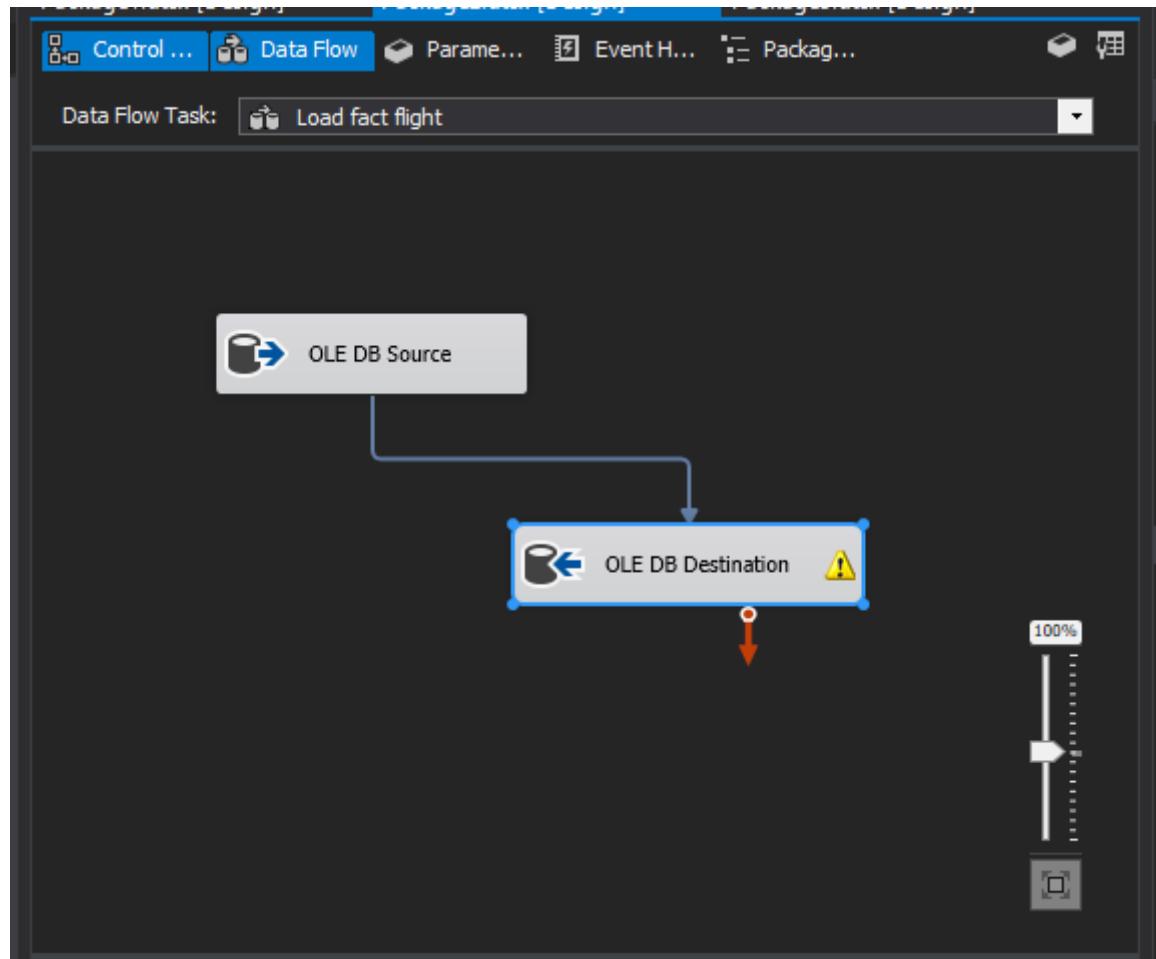
W tym kroku została utworzona centralna tabela modelu gwiazdy - tzw. **tabela faktów**. To właśnie ona stanowi bazę analityczną, na której opiera się każda analiza w SSAS i Power BI.

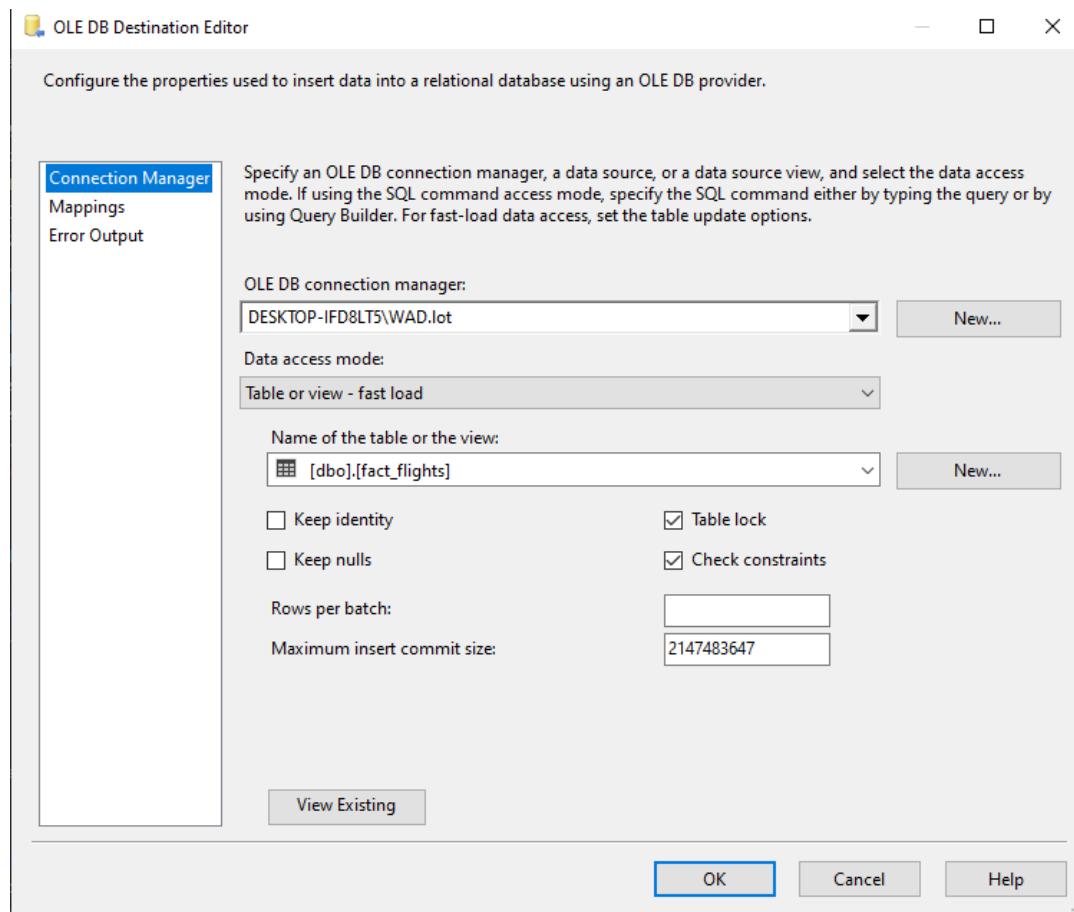
- Średnie opóźnienie przy startie wg dnia tygodnia lub miesiąca,
- Liczba lotów zrealizowanych przez konkretną linię lotniczą,
- Porównanie długości lotów wg typu samolotu,
- Wpływ warunków pogodowych (np. śnieg, niskie ciśnienie) na opóźnienia,
- Analizy anulowanych lotów w zależności od pory roku czy przewoźnika.

Bez tej tabeli **nie można wykonać żadnej analizy** - nie da się nic zsumować, porównać ani pogrupować. To fundament całej hurtowni danych - każda kostka OLAP i każdy raport w Power BI odwołuje się właśnie do **tabeli faktów**.

```
SQLQuery1.sql - DE...D8LT5\Student (58)* ↗ X
CREATE TABLE fact_flights (
    Flight_ID BIGINT PRIMARY KEY,
    -- Klucze do wymiarów
    Dep_Date_ID INT,
    Arr_Date_ID INT,
    Dep_Airport_ID VARCHAR(5),
    Arr_Airport_ID VARCHAR(5),
    Airline_ID INT,
    Aircraft_ID INT,
    -- Miary
    Dep_Delay FLOAT,
    Arr_Delay FLOAT,
    Flight_Duration FLOAT,
    Delay_Carrier FLOAT,
    Delay_Weather FLOAT,
    Delay_NAS FLOAT,
    Delay_Security FLOAT,
    Delay_LastAircraft FLOAT,
    Cancelled BIT,
    Diverted BIT
);
```







```
SQLQuery1.sql - DE...D8LT5\Student (58)*  X
SELECT COUNT(*) AS [Załadowanych wierszy] FROM fact_flights;
```

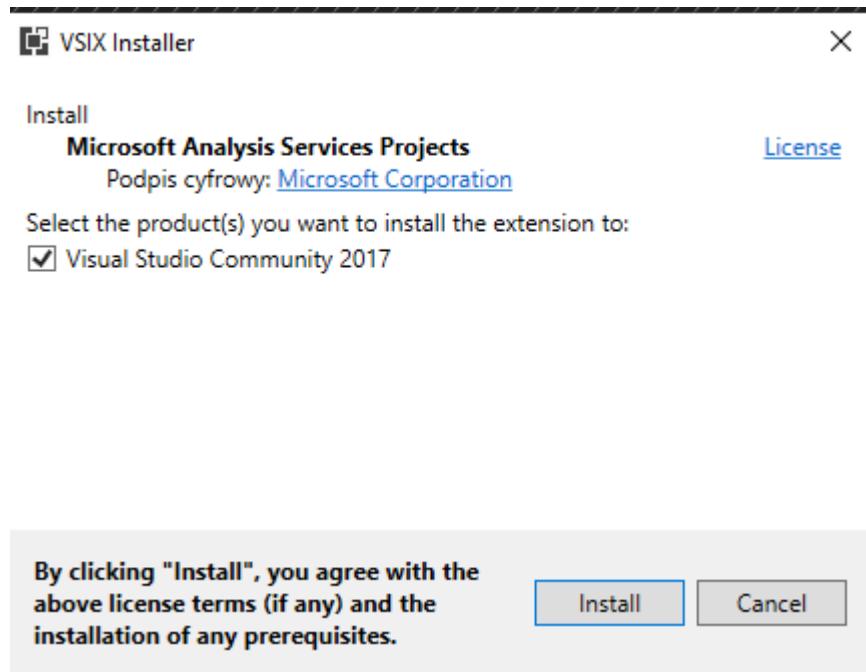
90 %

Results Messages

Załadowanych wierszy
1 6743404

6. Etap Piąty - Fundamenty pod kostkę OLAP

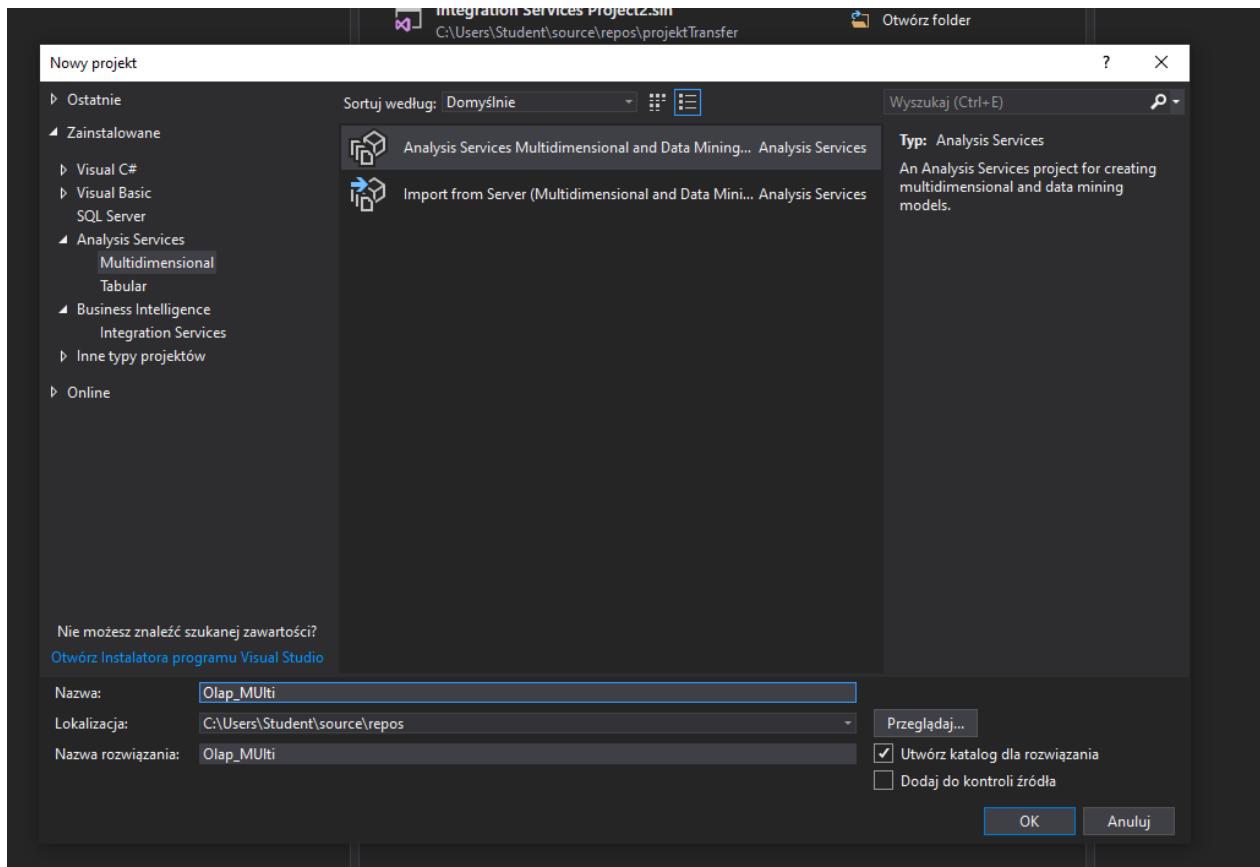
Instalujemy środowisko pod Kostkę OLAP.



The screenshot shows the SQL Server Configuration Manager interface. The left pane shows a tree view with "SQL Server Configuration Manager (Local)" expanded, showing "SQL Server Services", "SQL Server Network Configuration (32bit)", "SQL Native Client 11.0 Configuration (32bit)", "SQL Server Network Configuration", and "SQL Native Client 11.0 Configuration". The right pane displays a table of services:

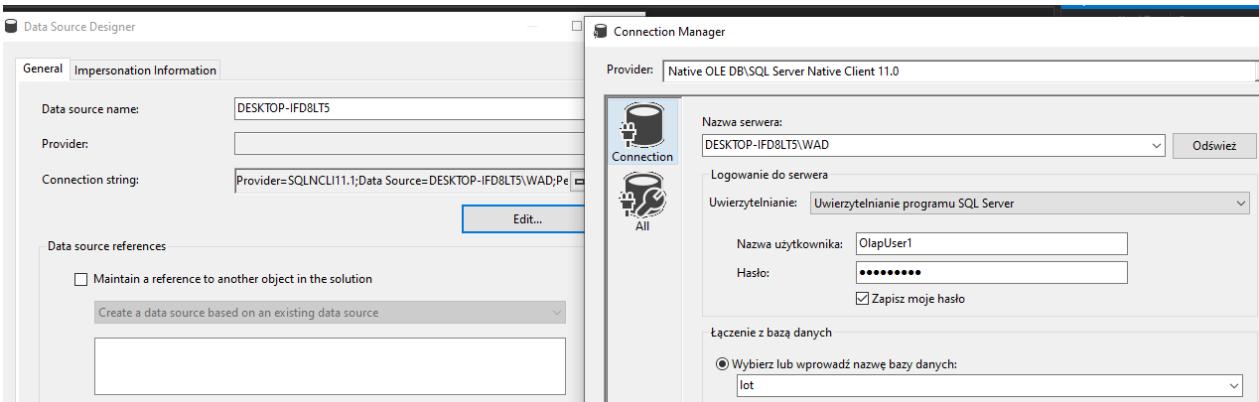
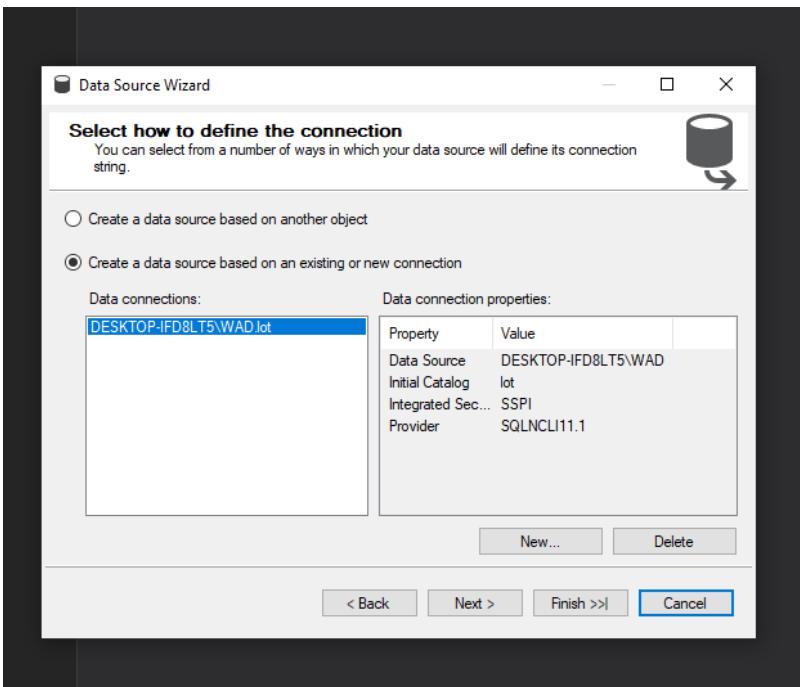
Name	State	Start Mode	Log On As	Process ID	Service Type
SQL Server Analysis Services (WAD)	Running	Automatic	NT Service\MSOLA...	3164	Analysis Server
SQL Server (WAD)	Running	Automatic	NT Service\MSSQLS...	3256	SQL Server
SQL Server Agent (WAD)	Stopped	Manual	NT Service\SQLAge...	0	SQL Agent
SQL Server Browser	Running	Automatic	NT AUTHORITY\LO...	3044	

Po zainstalowaniu wszystkich komponentów tworzymy nowy projekt pod kostkę OLAP.



6.1. Tworzenie Data Source

Pierwszym krokiem było utworzenie połączenia z bazą danych zawierającą dane o lotach (`loty_usa_2023`). Dzięki temu SSAS (SQL Server Analysis Services) może sięgnąć do rzeczywistych danych, na których oparty będzie cały model OLAP. Data Source to "brama wejściowa" do danych - bez niej nie możemy nic załadować ani przetwarzać.



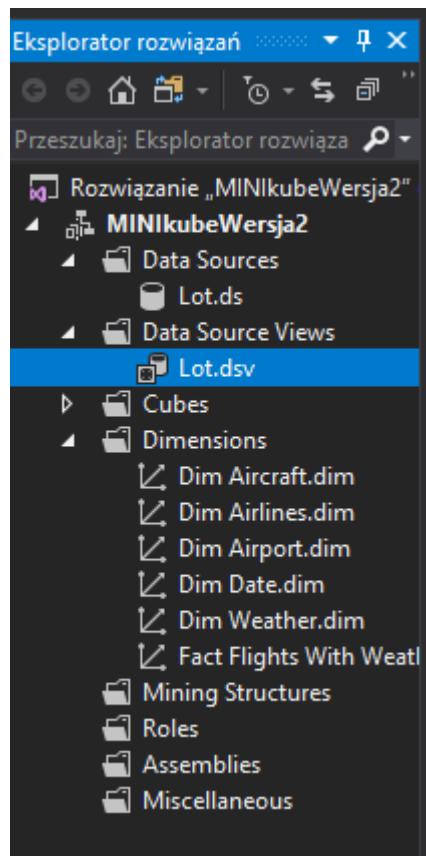
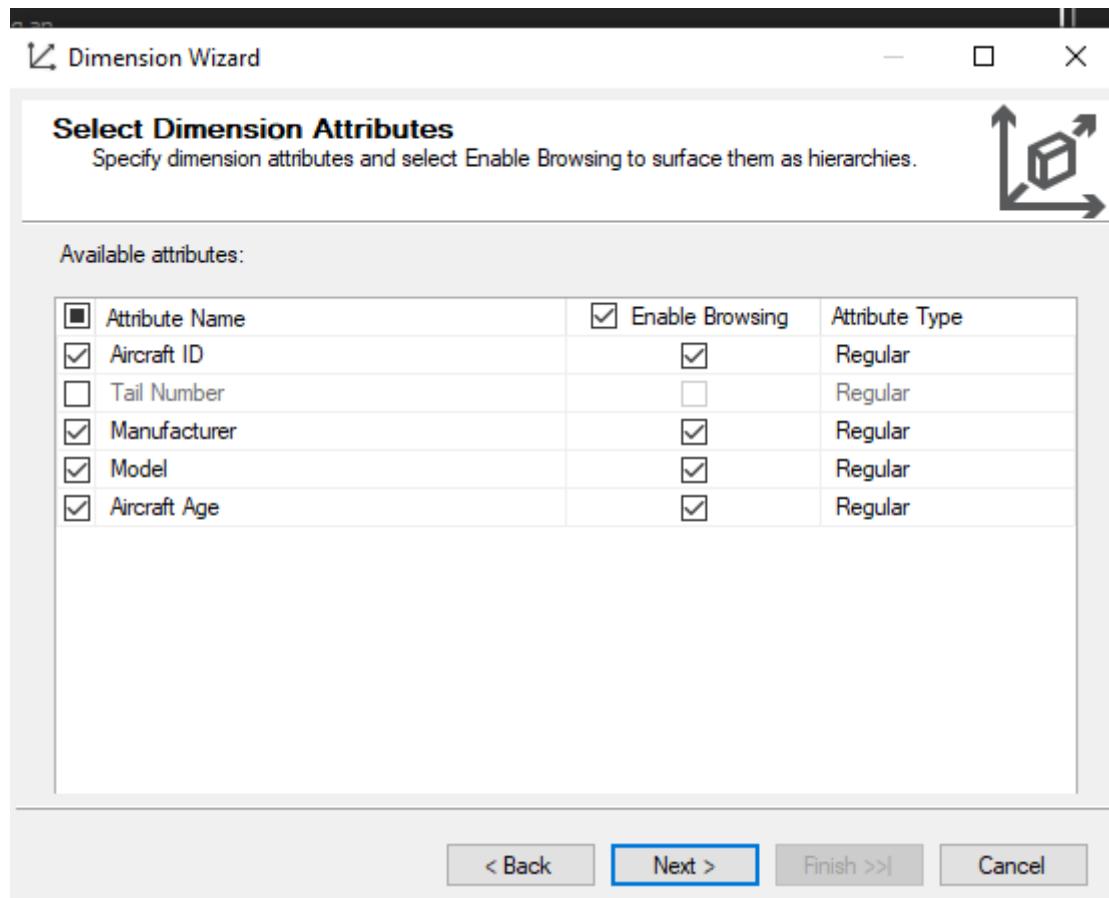
6.2. Tworzenie wymiarów (dimension)

W tym etapie utworzone zostały wszystkie **wymiary analityczne**, czyli:

- dim_date
- dim_airport
- dim_airline
- dim_aircraft
- dim_weather

Po co tworzymy dimension?

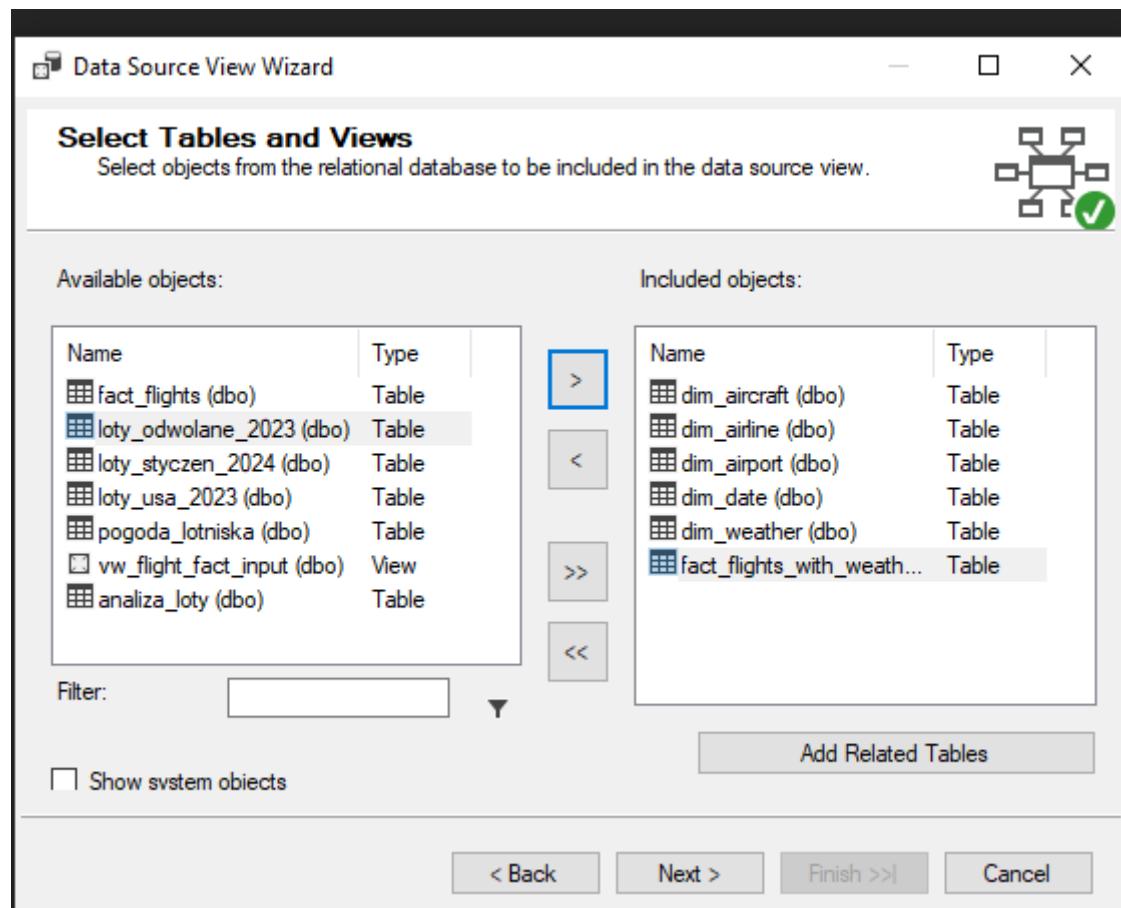
- Pozwalają analizować dane z różnych perspektyw (np. wg czasu, przewoźnika, lotniska).
- Służą do budowania relacji z faktami.
- Umożliwiają tworzenie **hierarchii** (np. Rok → Miesiąc → Dzień), filtrów, a także segmentacji danych.
- Dzięki nim dane stają się **czytelne i opisowe** - np. zamiast ID linii lotniczej pokazujemy jej nazwę.

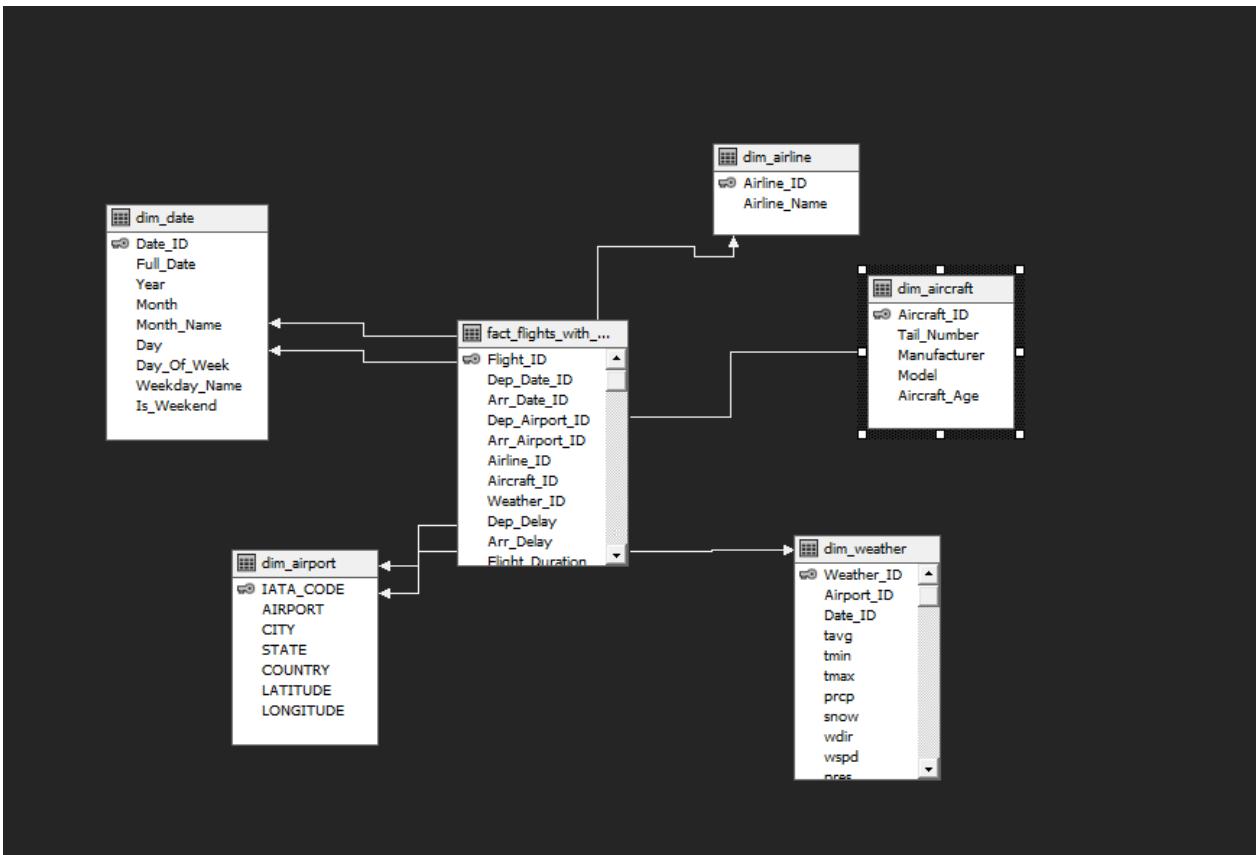


6.3. Widok źródłowy (Data Source View) i relacje

Po utworzeniu połączenia i wymiarów, wszystkie potrzebne tabele zostały przeciągnięte do **Data Source View** (DSV). Jest to graficzny widok struktury danych, który umożliwia:

- wizualne tworzenie relacji między faktami a wymiarami (klucz główny → klucz obcy),
- sprawdzenie spójności modelu,
- przygotowanie danych do dalszego przetwarzania w kostce OLAP.





6.4. Kostka OLAP (Cube)

Na końcu stworzona została kostka OLAP o nazwie Lot, która:

- łączy tabele faktów z wymiarami,
- definiuje **miary** (np. opóźnienia, liczba lotów, czas trwania lotu),
- udostępnia dane do analizy w narzędziach typu Excel / Power BI.

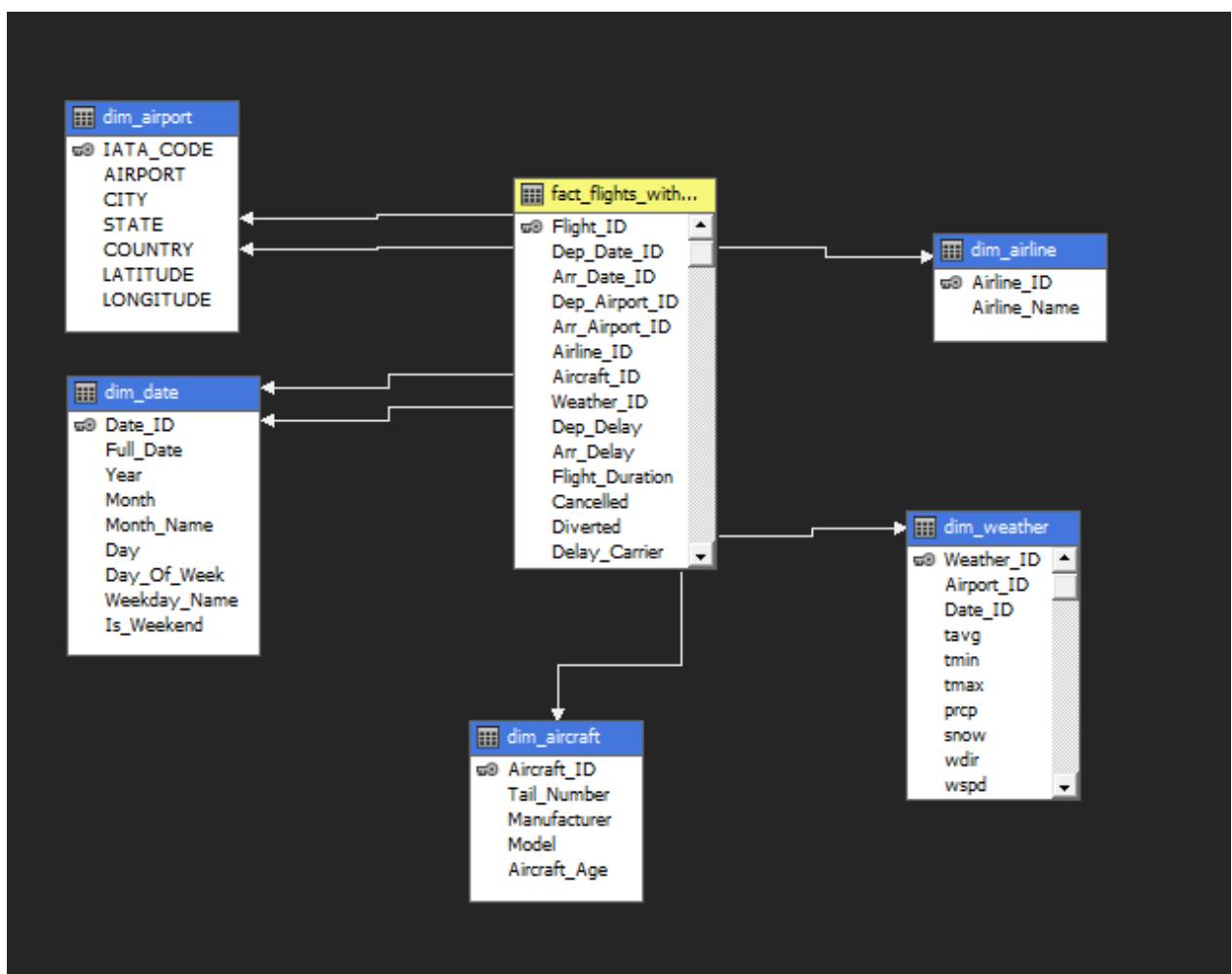
Po co nam kostka OLAP?

- Umożliwia dynamiczną analizę danych (agregacje, porównania, KPI),
- Ułatwia szybkie zapytania analityczne (działa szybciej niż zapytania SQL),
- Pozwala eksplorować dane interaktywnie (np. średnie opóźnienia wg miesiąca, przewoźnika czy typu samolotu),
- To ostateczny model, na którym bazuje Power BI.

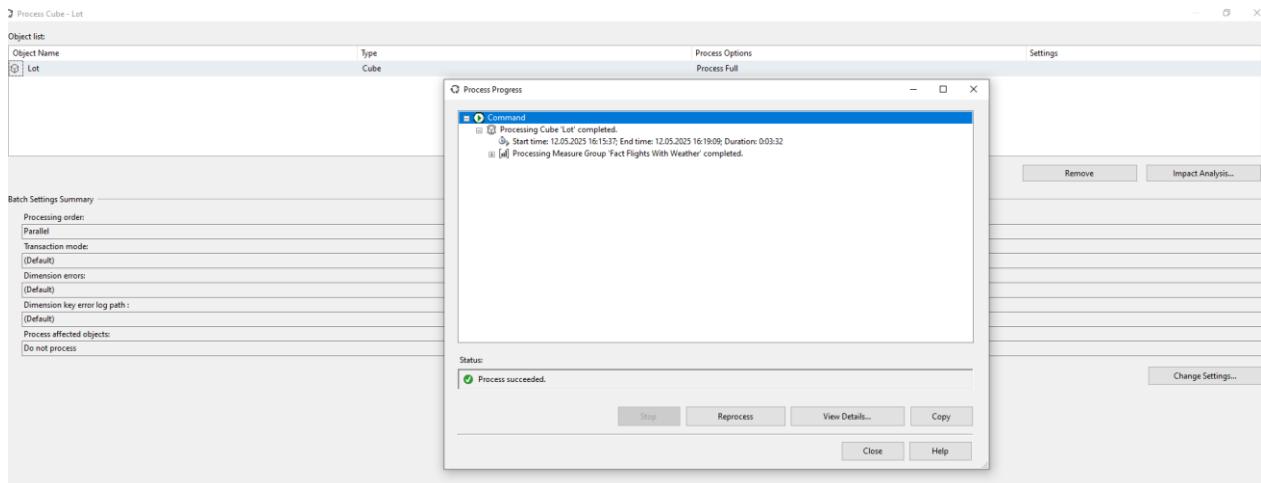
Measure Groups

Dimensions

	Fact Flights With Weather
Fact Flights With Weather...	Flight ID
Dim Aircraft	Aircraft ID
Dim Date (Dep Date)	Date ID
Dim Date (Arr Date)	Date ID
Dim Airline	Airline ID
Dim Airport (Dep Airport)	IATA CODE
Dim Airport (Arr Airport)	IATA CODE
Dim Weather	Weather ID



Uruchamiamy kostkę.



6.5. Tworzenie Hierarchii

W ramach budowy wymiarów w modelu OLAP utworzono hierarchie atrybutów, które pozwalają na wielopoziomową analizę danych. Najbardziej rozbudowaną i krytyczną strukturą była hierarchia **wymiaru daty (dim_date)**, ponieważ to ona stanowi podstawę dla większości analiz czasowych.

Jakie hierarchie utworzono:

Hierarchia czasu - Date Hierarchy

Zawierała następujące poziomy:

- Year
- Month
- Month Name
- Day

Dzięki niej możliwa była analiza danych w układzie:

- rocznym (np. porównanie lat),
- miesięcznym (np. suma opóźnień w styczniu),
- dziennym (np. analiza weekendów, konkretnych dat).

Hierarchia tygodniowa - Weekday Hierarchy

- Weekday Name (np. Monday, Tuesday...)
- Is Weekend (wartość logiczna: Tak/Nie)

Pozwala na porównanie np. opóźnień w tygodniu vs. weekendy.

Zmiany, jakie musielismy nałożyć:

1. Ustawienie NameColumn

Aby zamiast liczb 1-12 wyświetlały się rzeczywiste nazwy miesięcy (January, February...), przypisano w atrybucie Month:

- KeyColumn = Month
- NameColumn = Month_Name

Source	
CustomRollupColumn	(none)
CustomRollupPropertiesColumn	(none)
KeyColumns	dim_date.Month_Key (WChar)
NameColumn	dim_date.Month_Name (WChar)
ValueColumn	(none)

Analogicznie, dla Day i Weekday Name zadbano o to, by nazwy były czytelne i jednoznaczne w analizie.

2. Dodanie relacji między atrybutami

W zakładce **Attribute Relationships** zdefiniowano zależności:

Date ID → Year → Month → Month Name → Day

oraz osobne:

Date ID → Weekday Name

Date ID → Is Weekend

Pozwoliło to na lepsze działanie hierarchii, szybsze agregacje i brak ostrzeżeń przy procesowaniu.

- **Problemy napotkane podczas pracy:**

Duplikaty Month Name

Podczas przypisania KeyColumns = Year + Month do Month Name, SSAS zwracał błędy duplikatów:

„A duplicate attribute key has been found... for Month Name = January.”

Było to spowodowane tym, że Year + Month nie było wystarczająco unikalne — ponieważ w danych występowały dziesiątki dni przypisanych do jednego miesiąca.

Rozwiązanie: zamiast tego ustawiono **KeyColumn = Date_ID**, a jako NameColumn pozostawiono Month_Name. W ten sposób każda data była unikalna, ale browser wyświetlał poprawne nazwy.

Niepoprawne grupowanie Month Name w Browser

Mimo widocznych nazw miesięcy, w zakładce Browser wyświetlały się dziesiątki wierszy z nazwą „January”. Problem rozwiązało przez **zastosowanie Month Name tylko wewnątrz zdefiniowanej hierarchii**, a nie jako samodzielnego atrybutu.

Efekt końcowy:

Po poprawkach hierarchie działają poprawnie - pozwalają na:

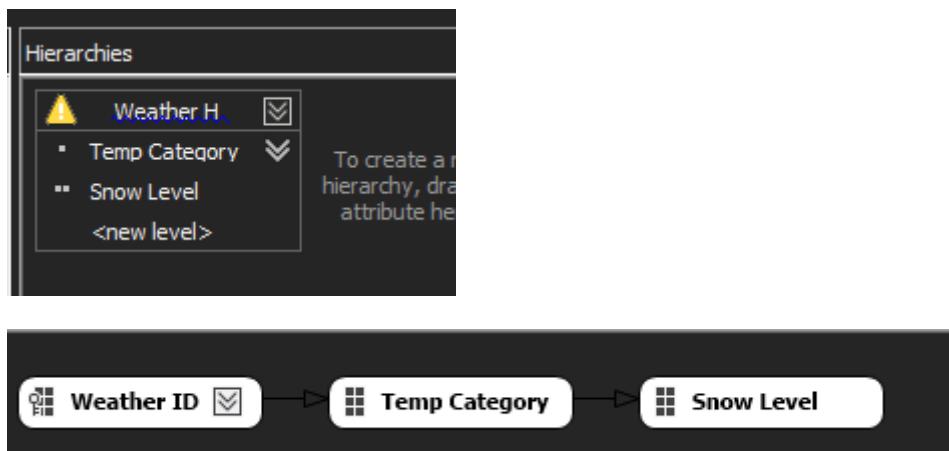
- **drill-down** z poziomu roku do dnia,
- **dynamiczne filtrowanie** po poziomach (np. analizuj tylko Q1, tylko weekendy, tylko June),

6.6. Dodatkowe hierarchie: Weather, Aircraft, Airport, Airline

Poza hierarchią daty, w projekcie utworzono także uproszczone hierarchie w innych wymiarach. Ich celem było ułatwienie filtrowania, grupowania danych oraz przejrzystość analizy w SSAS i Power BI.

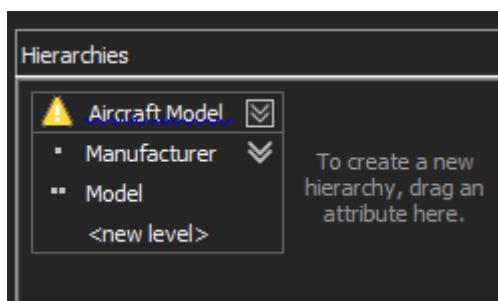
Hierarchia Weather

- Date_ID → Airport_ID → TAVG / PRCP / SNOW
- Pozwala analizować wpływ warunków pogodowych (temperatura, opady, śnieg) na opóźnienia i anulacje, w zależności od lotniska i dnia.



Hierarchia Aircraft

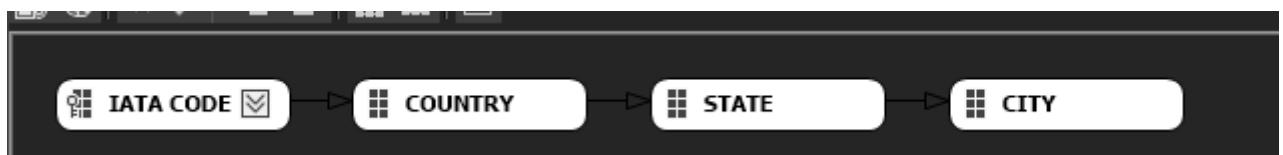
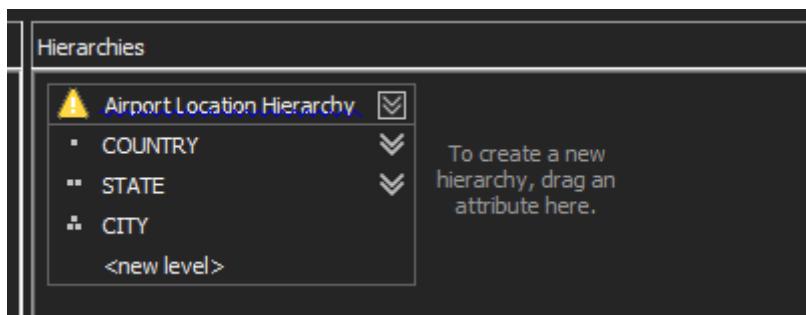
- Manufacturer → Model → Aircraft_Age
- Umożliwia analizę opóźnień z uwzględnieniem typu samolotu i jego wieku - np. czy starsze maszyny są bardziej awaryjne lub opóźnione.





Hierarchia Airport

- State → City → IATA_CODE
- Dzięki niej można prowadzić analizy lokalizacyjne: opóźnienia wg miasta, stanu, czy konkretnego lotniska.



Wszystkie te hierarchie zostały wdrożone w celu poprawy nawigacji w kostce OLAP oraz umożliwienia zaawansowanej analizy danych w sposób intuicyjny i hierarchiczny.

6.7. Sprawdzenie poprawności kostki OLAP w zakładce Browser

Po zakończeniu budowy kostki OLAP i przetworzeniu danych (Deploy + Process Full), przeprowadzono testy poprawności działania w zakładce **Browser** w Visual Studio.

Celem tego etapu było sprawdzenie, czy:

- kostka poprawnie agreguje dane,
- relacje między faktami a wymiarami działają zgodnie z oczekiwaniemi,
- miary obliczają się prawidłowo,
- dane można filtrować i analizować względem różnych wymiarów (czas, przewoźnik, pogoda, itd.).

Lot.cube [Design] < Dim Date.dim [Design]

File Cube Structure Dimension Usage Calculations KPIs Actions Partitions Aggregations Perspectives Translations Browsers

Language: Default

Dimension Hierarchy Operator Filter Expression Parameters

<Select dimension>

Drag levels or measures here to add to the query.

Lot

- Measures
 - Fact Flights With Weather
 - Arr Delay
 - Cancelled Flights
 - Delay Carrier
 - Delay Last Aircraft
 - Delay NAS
 - Delay Security
 - Delay Weather
 - Diverted Flights
 - Flight Duration
 - KPIs
 - Arr Airport
 - Arr Date
 - Dep Airport
 - Dep Date

Dane wyjściowe

Poprawnie agreguje.

Month Display	Delay Carrier
January	2860559
February	2392270
March	3051310
April	3018861
May	2612269
June	4010319
July	4415505
August	3452931
September	2677700
October	2350720
November	1924411
December	2095219

6.8. Stworzenie 4 scenariuszy/KPI

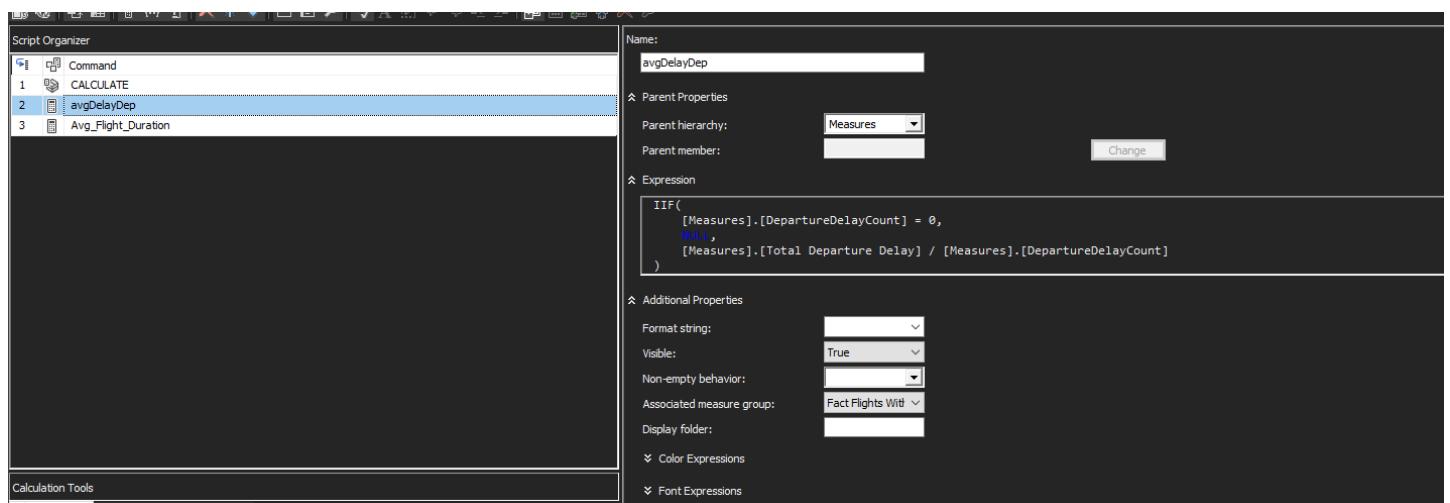
6.8.1. Obliczona miara: Średnie opóźnienie odlotów (Avg Departure Delay)

W zakładce *Calculations* utworzono miernik Avg Departure Delay, który oblicza średnie opóźnienie przy odlocie jako stosunek sumy opóźnień (Total Departure Delay) do liczby opóźnionych lotów (DepartureDelayCount). Miara została przypisana do grupy Fact Flights With Weather.

W zakładce *Browser* przetestowano działanie miernika na wymiarze Dep Airport.AIRPORT. Wartości liczbowe pojawiły się prawidłowo, co potwierdza, że miara działa poprawnie i może być wykorzystywana w dalszej analizie KPI.

Integracja z modelami data mining:

- Wykorzystano w **drzewie decyzyjnym** jako zmienna wejściowa do klasyfikacji lotów wysokiego ryzyka (np. reguła: "*DepDelay > 45 minut → 89% szans na opóźnienie przylotu*").
- W **klasteryzacji** służyło do wyodrębnienia grupy lotów z chronicznymi opóźnieniami (klaster nr 6 w wynikach).



Dimension	Hierarchy	Operator	Filter Expression
<Select dimension>			
AIRPORT	avgDelayDep	Total Departure Delay	DepartureDelayCount
Aberdeen Regional Airport	6,50773558368495	4627	711
Abilene Regional Airport	6,34228187919463	8505	1341
Abraham Lincoln Capital Airport	22,227621483376	8691	391
Adak Airport	-2,46875	-237	96
Akron-Canton Regional Airport	8,24234068627451	26903	3264
Albany International Airport	9,82518442271402	114542	11658
Albert J. Ellis Airport	12,3117338003503	14060	1142
Albuquerque International Sunport	9,98186739999141	232308	23273
Alexandria International Airport	11,9090398652442	21210	1781
Alpena County Regional Airport	19,9403669724771	8694	436
Appleton International Airport	11,8820259419395	57711	4857
Arcata Airport	14,5181538461538	23592	1625
Arnold Palmer Regional Airport	13,9647495361781	7527	539
Asheville Regional Airport	13,867221482344	153316	11056
Aspen-Pitkin County Airport	22,8030453492221	137776	6042
Atlantic City International Airport	9,86522911051213	29280	2968
Augusta Regional Airport (Bush Field)	12,9256700746062	46778	3619
Austin-Bergstrom International Airport	12,8570756831826	1175741	91447
Baltimore-Washington International Airport	15,6467400454584	1486941	95032
Bangor International Airport	9,49823243482103	42989	4526
Barnstable Municipal Airport	24,4114583333333	4687	192
Baton Rouge Metropolitan Airport	9,11810102899906	38989	4276
Bellingham International Airport	6,7977011942529	17742	2610
Bemidji Regional Airport	4,03741007194245	2806	695
Benedum Airport	25,2333333333333	5299	210
Bert Mooney Airport	-5,7027027027027	-3798	666
Bethel Airport	15,0754985754986	10583	702
Bill and Hillary Clinton National Airport (Ad...	9,94756944444444	114596	11520
Billings Logan International Airport	6,872668997669	23587	3432
Birmingham-Shuttlesworth International Air...	10,1912008149752	160063	15706
Bishop Airport	29,344398340249	7072	241
Bishop International Airport	7,86848514369216	16154	2053
Bismarck Municipal Airport	11,7917393833624	40540	3438
Blue Grass Airport	9,43930488116535	73872	7826
Bob Hope Airport (Hollywood Burbank Air...	8,05538218853949	234903	29161
Boise Airport (Boise Air Terminal)	6,80700419327453	146686	70074

6.8.2. Obliczona miara: Średni czas lotu względem wieku samolotu (Avg Flight Duration per Aircraft Age)

Drugi KPI został zaprojektowany w celu zbadania wpływu wieku samolotu na czas trwania lotu. W tym celu utworzono miernik, który dzieli całkowity czas lotów (Flight Duration) przez liczbę lotów (Fact Flights With Weather Count) dla każdej grupy wiekowej samolotów (Aircraft Age).

Wyniki wyświetlono w zakładce *Browser* w przekroju wymiaru Aircraft Age, dzięki czemu możliwa była analiza zależności między wiekiem floty a efektywnością operacyjną. Dla najstarszych samolotów zaobserwowano wyraźnie wyższy średni czas trwania lotu, co może wskazywać na ograniczenia techniczne starszych jednostek lub częstsze konieczności dostosowania tras i procedur. Z kolei młodsze maszyny osiągają krótsze czasy, co może sugerować większą wydajność i nowocześniejsze parametry operacyjne.

Integracja z modelami data mining:

- W sieci neuronowej zmienna ta pomogła zidentyfikować nieliniowe zależności (np. samoloty 10-15 lat mają wyższą wariancję czasów lotów).
- W regułach asocjacyjnych wykryto wzorzec: "*AircraftAge > 20 lat AND TempCategory = 'Cold' → FlightDuration zwiększone o 18%*".

The screenshot shows the Microsoft Analysis Services Script Organizer interface. The top menu bar includes options like Cube Structure, Dimension Usage, Calculations, KPIs, Actions, Partitions, Aggregations, Perspectives, Translations, and Browser. The left sidebar is titled "Script Organizer" and lists four items under "CALCULATE": 1. AvgFlightDurationperAircraftAge, 2. avgDelayDep, and 3. Avg_Flight_Duration. The main workspace displays the properties for a new calculated member:

- Name:** AvgFlightDurationperAircraftAge
- Parent Properties:**
 - Parent hierarchy: Measures
 - Parent member: (Change)
- Expression:**

```
IIF(
    [Measures].[Fact Flights With Weather Count] = 0,
    NULL,
    [Measures].[Flight Duration] / [Measures].[Fact Flights With Weather Count]
)
```
- Additional Properties:**
 - Format string: (dropdown)
 - Visible: True
 - Non-empty behavior: (dropdown)
 - Associated measure group: Fact Flights Witl (dropdown)
 - Display folder: (dropdown)
- Color Expressions:** (collapsible section)
- Font Expressions:** (collapsible section)

At the bottom left, there's a "Calculation Tools" section with tabs for Metadata, Elements, and Translations.

Cube Structure Dimension Usage Calculations KPIs Actions Partitions Aggregations Perspectives Translations

Language: Default

Edit as Text Import... MDX

Lot

Metadata Search Model <All>

Lot

- Measures
 - Fact Flights With Weather
 - Arr Delay
 - avgDelayDep
 - AvgFlightDurationperAircraftAge
 - Cancelled Flights
 - Delay Carrier
 - Delay Last Aircraft
 - Delay NAS
 - Delay Security
 - Delay Weather
 - Dep Delay
 - DepartureDelayCount
 - Diverted Flights
 - Fact Flights With Weather Cou
 - Flight Duration
 - Total Departure Delay
 - Avg_Flight_Duration
 - KPIs
 - Arr Airport
 - Arr Date
 - Dep Airport
 - Dep Date
 - Dim Aircraft
 - Aircraft Age
 - Aircraft ID
 - Manufacturer
 - Model
 - Aircraft Model
 - Dim Airline
 - Dim Weather
 - Fact Flights With Weather

Dimension Hierarchy

<Select dimension>

Aircraft Age	AvgFlightDurationperAircraftAge
1	170,641259075739
2	169,511583005685
3	174,393928863812
4	159,965387935304
5	152,554663381731
6	150,66424083738
7	152,935200459415
8	144,582509626454
9	141,345247598617
10	142,15984733561
11	145,226015213104
12	148,665032571225
13	132,784998091479
14	138,392055974455
15	131,82383187013
16	121,267905700977
17	119,857995420447
18	123,340344703017
19	113,85377056829
20	114,223795473582
21	114,081498136264
22	124,88317622492
23	126,39134331742
24	129,040323814868
25	136,994579330044
26	142,481819355047
27	205,04821615746
28	169,196111740279
29	187,967936152338
30	191,197117670727
31	143,350246496427
32	134,775900671687
33	143,356463019089
34	220,86696111171
38	177,261306532663
39	185,384496124031
48	170,978389830508
56	125,480225988701
57	129,383777239709

6.8.3. Obliczona miara: Indeks opóźnień systemowych względem przewoźników (NAS vs Carrier Delay Index)

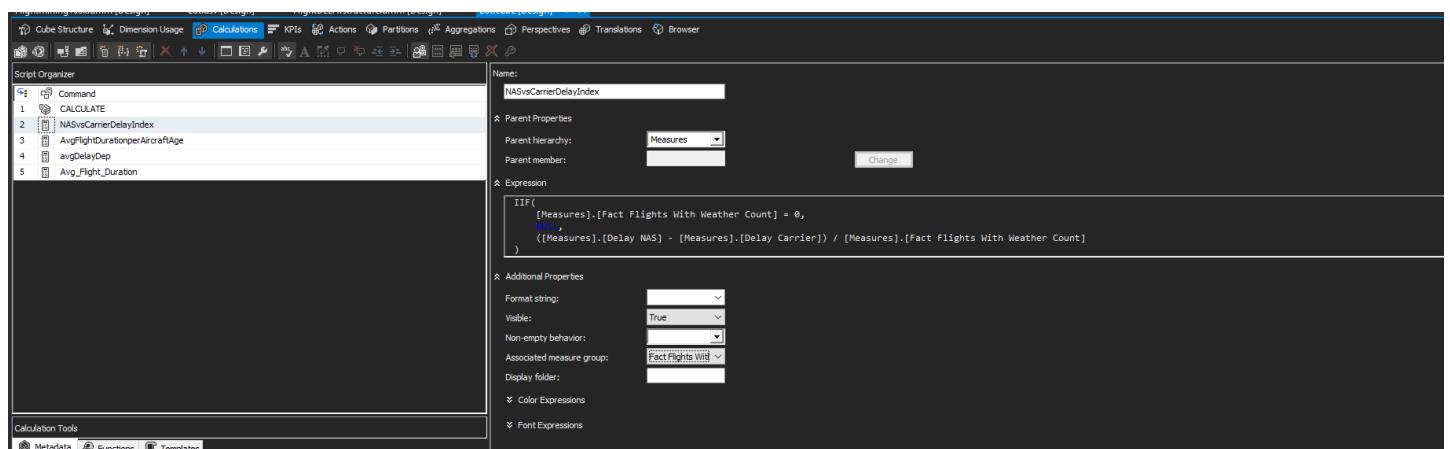
Trzeci KPI został opracowany w celu porównania wpływu dwóch kluczowych źródeł opóźnień: systemowych (NAS) i operacyjnych (Carrier). Miara została zdefiniowana jako różnica sumarycznych opóźnień NAS i Carrier, podzielona przez liczbę lotów, dzięki czemu otrzymujemy uśredniony wskaźnik dominującego źródła problemu.

Wyniki pozytywne wskazują na przeważający wpływ systemu (np. infrastruktura, zarządzanie ruchem lotniczym), natomiast wartości ujemne sugerują, że to linie lotnicze częściej odpowiadają za opóźnienia (np. problemy operacyjne, rotacje, załogi).

Miara została uruchomiona w zakładce *Browser* w przekroju linii lotniczych, lotnisk i miesięcy. Wyniki pozwalają łatwo zidentyfikować, kiedy i gdzie występują najczęstsze zakłócenia oraz kto za nie odpowiada. Dodatkowo stanowią solidną podstawę do dalszej eksploracji danych i raportów kierunkowych.

Integracja z modelami data mining:

- W **szeregach czasowych** indeks posłużył do prognozowania tygodniowych szczytów opóźnień systemowych (np. wzrost o 22% w piątki).
- W **klasteryzacji** wyodrębniono grupę lotów z dominującym wpływem NAS (klaster 9: "DelayNAS > 60 minut w 78% przypadków").



Dimension		Hierarchy	
<Select dimension>			
Airline Name	AIRPORT	Month Display	NASvsCarrierDelayIndex
Alaska Airlines Inc.	Adak Airport	January	0
Alaska Airlines Inc.	Adak Airport	February	0
Alaska Airlines Inc.	Adak Airport	March	0
Alaska Airlines Inc.	Adak Airport	April	4,625
Alaska Airlines Inc.	Adak Airport	May	0
Alaska Airlines Inc.	Adak Airport	June	0
Alaska Airlines Inc.	Adak Airport	July	-16,111111111111
Alaska Airlines Inc.	Adak Airport	August	-2,875
Alaska Airlines Inc.	Adak Airport	September	-0,6666666666666667
Alaska Airlines Inc.	Adak Airport	October	-37,1428571428571
Alaska Airlines Inc.	Adak Airport	November	-2,57142857142857
Alaska Airlines Inc.	Adak Airport	December	-6,375
Alaska Airlines Inc.	Albuquerque Inter...	January	-1,09836065573771
Alaska Airlines Inc.	Albuquerque Inter...	February	-7,29090909090909
Alaska Airlines Inc.	Albuquerque Inter...	March	-4,704545454545
Alaska Airlines Inc.	Albuquerque Inter...	April	0,641025641025641
Alaska Airlines Inc.	Albuquerque Inter...	May	-6,46511627906977
Alaska Airlines Inc.	Albuquerque Inter...	June	-1,14583333333333
Alaska Airlines Inc.	Albuquerque Inter...	July	-0,469387755102041
Alaska Airlines Inc.	Albuquerque Inter...	August	-4,17721518987342
Alaska Airlines Inc.	Albuquerque Inter...	September	-3,49206349206349
Alaska Airlines Inc.	Albuquerque Inter...	October	-0,396551724137931
Alaska Airlines Inc.	Albuquerque Inter...	November	-3,06122448979592
Alaska Airlines Inc.	Albuquerque Inter...	December	-3,87096774193548
Alaska Airlines Inc.	Austin-Bergstrom ...	January	0,421276595744681
Alaska Airlines Inc.	Austin-Bergstrom ...	February	-0,673170731707317
Alaska Airlines Inc.	Austin-Bergstrom ...	March	0,699588477366255
Alaska Airlines Inc.	Austin-Bergstrom ...	April	1,94893617021277
Alaska Airlines Inc.	Austin-Bergstrom ...	May	2,65182186234818
Alaska Airlines Inc.	Austin-Bergstrom ...	June	2,37154150197628
Alaska Airlines Inc.	Austin-Bergstrom ...	July	-1,71480144404332
Alaska Airlines Inc.	Austin-Bergstrom ...	August	-0,0863309352517986
Alaska Airlines Inc.	Austin-Bergstrom ...	September	1,26521739130435
Alaska Airlines Inc.	Austin-Bergstrom ...	October	-0,6666666666666667
Alaska Airlines Inc.	Austin-Bergstrom ...	November	-0,0561224489795918
Alaska Airlines Inc.	Austin-Bergstrom ...	December	1,49438202247191

6.8.4. Obliczona miara: Udział przewoźnika w opóźnieniach operacyjnych (Carrier Share of Delay)

Czwarty KPI mierzy, jaka część wszystkich opóźnień przypisywanych przewoźnikom (Delay Carrier) jest generowana przez daną linię lotniczą. Miara została zdefiniowana jako udział opóźnień danego przewoźnika w całkowitej liczbie opóźnień typu Carrier dla wszystkich przewoźników.

Dzięki temu możliwa jest identyfikacja najbardziej problematycznych linii lotniczych z punktu widzenia opóźnień operacyjnych. Wskaźnik ten szczególnie dobrze nadaje się do analizy rankingowej i może stanowić podstawę do działań naprawczych lub pogłębionych analiz przyczyn.

Miara została wdrożona i przetestowana w przekroju Airline Name w zakładce *Browser*, gdzie uzyskano oczekiwane, zróżnicowane wartości procentowe.

Integracja z modelami data mining:

- W **drzewie decyzyjnym** zmienna posłużyła do klasyfikacji przewoźników na kategorie: "Niski" (<5%), "Średni" (5-15%), "Wysoki" (>15%) udział.
- W **regulach asocjacyjnych** wykryto zależność: "*CarrierShare > 15% AND WeatherCategory = 'Rainy'* → 92% szans na opóźnienie > 1 godziny".

Airline Name	AvgprzewoznikDelay
Alaska Airlines Inc.	0,0212495389689093
Allegiant Air	0,0210928930963775
American Airlines Inc.	0,185528491506271
American Eagle Airlines Inc.	0,0171024535143836
Delta Air Lines Inc	0,161938816376788
Endeavor Air	0,0193827538774658
Frontier Airlines Inc.	0,0373586207177462
Hawaiian Airlines Inc.	0,0138110830698139
JetBlue Airways	0,0830655112486997
PSA Airlines	0,0208156003569954
Republic Airways	0,0178143732928798
Skywest Airlines Inc.	0,128705452234425
Southwest Airlines Co.	0,123578361975825
Spirit Air Lines	0,0441605396167767
United Air Lines Inc.	0,104395510146642

The screenshot shows the Script Organizer interface. On the left, a list of calculated members is displayed, including 'Command', 'CALCULATE', 'NASvsCarrierDelayIndex', 'AvgFlightDurationperAircraftAge', 'avgDelayDep', and 'Avg_Flight_Duration'. The 'avgDelayDep' member is currently selected. On the right, a detailed properties pane is open for the 'avgDelayDep' member. It shows the following details:

- Name:** DelaySNOW
- Parent Properties:** Parent hierarchy: Measures
- Expression:**

```
IIF(
    [Dim Weather].[SnowLevel1].CURRENTMEMBER IS [Dim Weather].[SnowLevel1].[High],
    [Measures].[avgDelayDep],
    NULL
)
```
- Additional Properties:**
 - Format string: (dropdown menu)
 - Visible: True
 - Non-empty behavior: (dropdown menu)
 - Associated measure group: Fact Flights With
 - Display folder: (dropdown menu)
 - Color Expressions: (dropdown menu)
 - Font Expressions: (dropdown menu)

7. Data Mining

W ramach naszego projektu data mining (eksploracja danych) został wykorzystany do identyfikacji wzorców i zależności wpływających na opóźnienia lotów w USA.

Zastosowaliśmy pięć różnych modeli analitycznych:

- **Drzewo decyzyjne** – model klasyfikacyjny pozwalający przewidzieć, czy dany lot będzie opóźniony, na podstawie takich zmiennych jak opóźnienia przewoźnika, systemowe, czas trwania lotu i warunki pogodowe.

- **Klasteryzacja** – technika grupowania lotów o podobnych cechach (np. długość trasy, poziom opóźnień) w segmenty, co pozwoliło lepiej zrozumieć, które grupy lotów są bardziej narażone na opóźnienia.
- **Sieć neuronowa** – model predykcyjny, który umożliwia bardziej precyzyjne przewidywanie opóźnień dzięki analizie nieliniowych zależności między zmiennymi.
- **Reguły asocjacyjne** – metoda wykrywająca powtarzalne wzorce, np. że długie opóźnienia przy startie prawie zawsze skutkują spóźnionym przylotem.
- **Szeregi czasowe** – analiza trendów sezonowych w opóźnieniach, dzięki czemu mogliśmy przewidzieć ich wzrost w okresach świątecznych i wakacyjnych.

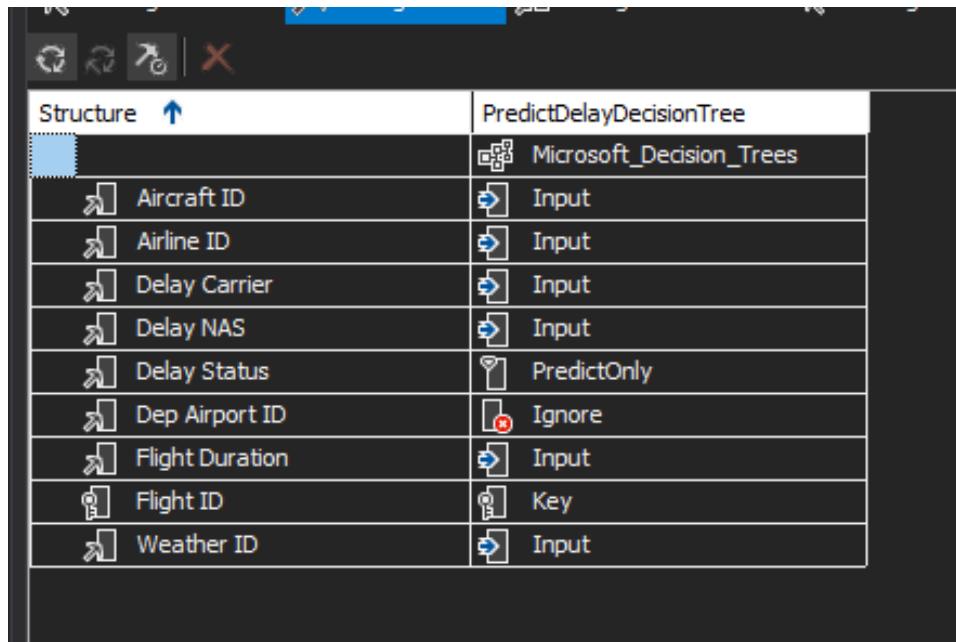
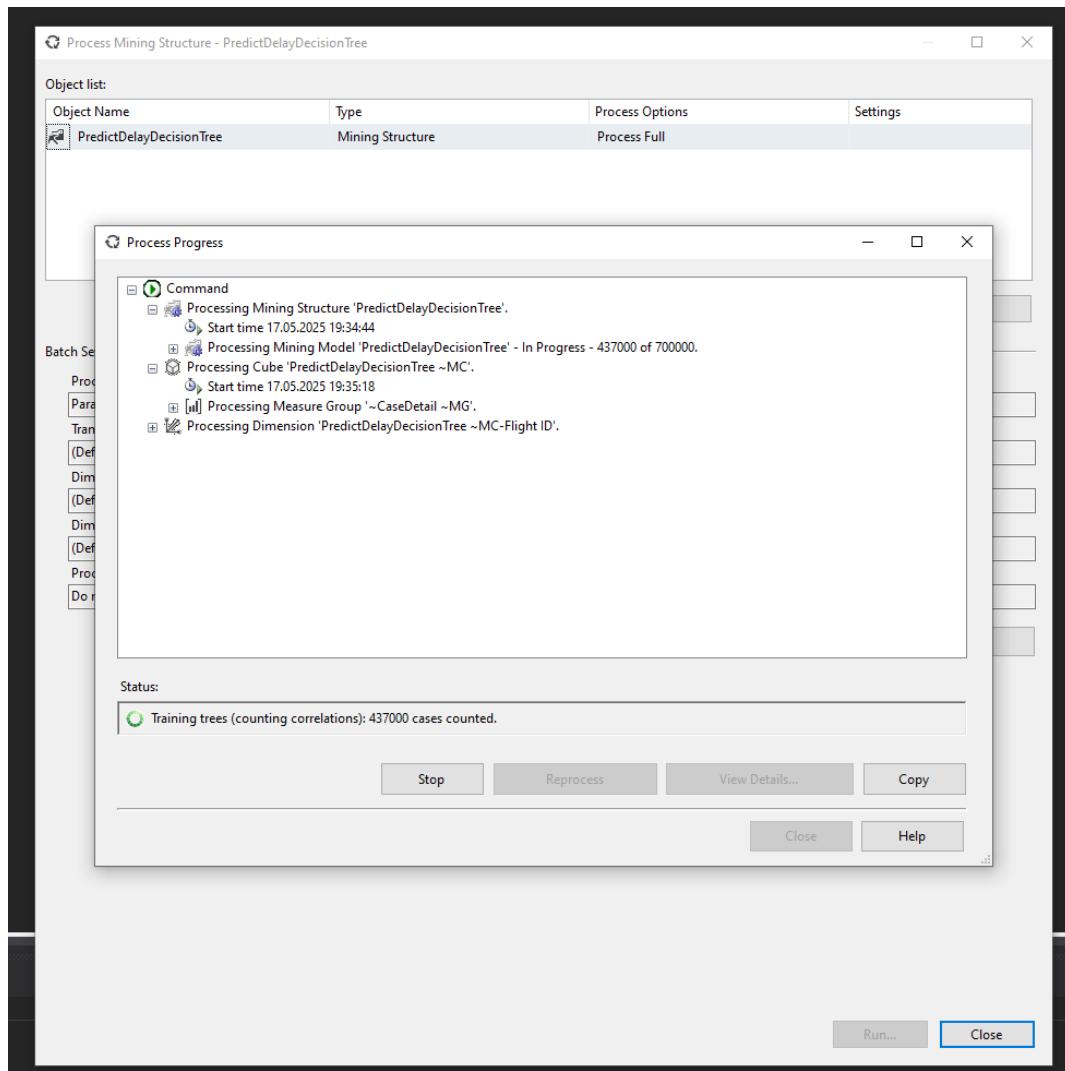
Eksploracja danych była kluczowym elementem naszego projektu, pozwalającym nie tylko opisać, ale i prognozować oraz segmentować zjawisko opóźnień. Dzięki temu możliwe stało się wyciągnięcie praktycznych wniosków, które mogą zostać wykorzystane w zarządzaniu ruchem lotniczym, planowaniu operacyjnym czy poprawie komunikacji z pasażerami.

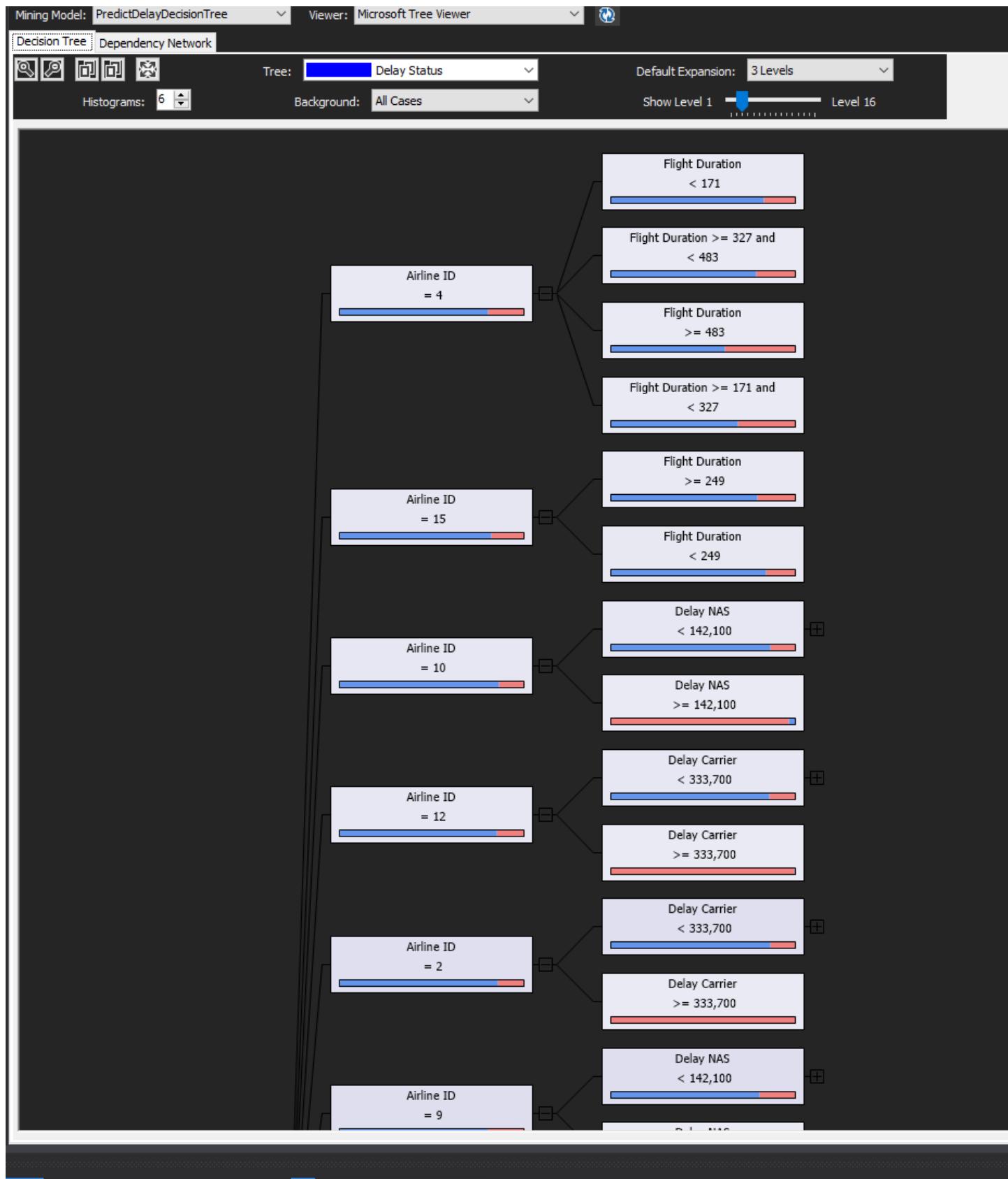
7.1. Model 1 – Drzewo Decyzyjne

```

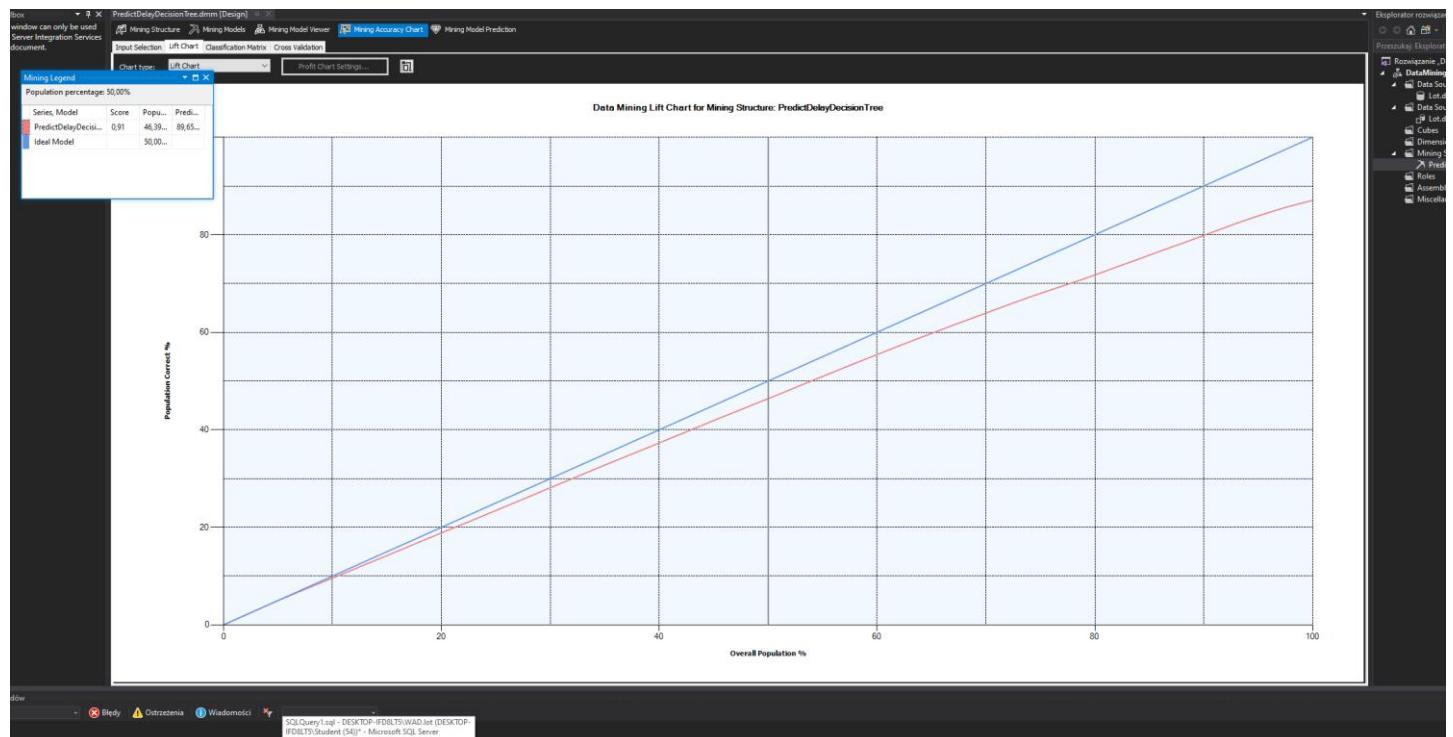
SQLQuery2.sql - DE...D8LT5\Student (66))          SQLQuery1.sql - DE...D8LT5\Student (54))*
[ ] SELECT TOP 1000000
    Flight_ID,
    CASE
        WHEN Dep_Delay >= 15 THEN 'Delayed'
        ELSE 'OnTime'
    END AS DelayStatus,
    Airline_ID,
    Dep_Airport_ID,
    Aircraft_ID,
    Weather_ID,
    Dep_Delay,
    Arr_Delay,
    Flight_Duration,
    Delay_Carrier,
    Delay_NAS
    INTO dm_sample_delay_prediction
    FROM fact_flights_with_weather
    WHERE Dep_Delay IS NOT NULL
    ORDER BY NEWID();

```





Input Selection	Lift Chart	Classification Matrix	Cross Validation
Columns of the classification matrices correspond to actual values; rows correspond to predicted values			
Counts for PredictDelayDecisionTree on Delay Status:			
Predicted	Delayed (Actual)	OnTime (Actual)	
Delayed	26389	4149	
OnTime	34693	234769	



Screenshot of the Microsoft SQL Server Data Mining interface showing the Mining Model configuration screen.

Left pane: Mining Model list, currently selected: PredictDelayDecisionTree.

Right pane: Select Input Table(s) dialog box showing the dm_sample_delay_prediction table with its columns listed.

Bottom pane: Source selection grid showing the input fields and their mapping to the PredictDelayDecisionTree model.

Source	Field	Alias	Show	Group	And/Or	Criteria/Argument
dm_sample_delay_prediction	Airline_ID		<input checked="" type="checkbox"/>			
dm_sample_delay_prediction	Delay_NAS		<input checked="" type="checkbox"/>			
dm_sample_delay_prediction	Delay_Carrier		<input checked="" type="checkbox"/>			
dm_sample_delay_prediction	Flight_ID		<input checked="" type="checkbox"/>			
dm_sample_delay_prediction	Flight_Duration		<input checked="" type="checkbox"/>			
dm_sample_delay_prediction	Weather_ID		<input checked="" type="checkbox"/>			
Prediction Function	PredictProbability	PredictedStatus	<input checked="" type="checkbox"/>			PredictDelayDecisionTree.[Delay Status]
Prediction Function	PredictProbability	PredictionConfidence	<input checked="" type="checkbox"/>			PredictDelayDecisionTree.[Delay Status]

F:\Data\DelayDecisionTree\airline [Design] - Microsoft Data Mining Base

Airline_ID	Delay_NAS	Delay_Carrier	Flight_ID	Flight_Duration	Weather_ID	PredictedStatus	PredictionConfidence
3	0	0	10802740	206	82501	OnTime	0,9215350628...
3	0	0	101989111	223	75080	OnTime	0,9215350628...
13	6	0	11260667	121	91073	Delayed	0,8163265306...
3	0	0	10841113	65	85781	OnTime	0,9215350628...
2	0	0	9356331	192	54921	OnTime	0,9219975711...
3	0	0	7950592	111	31690	OnTime	0,9215350628...
5	0	0	9748345	213	57341	OnTime	0,8905666094...
5	0	0	7003744	129	8030	OnTime	0,8905666094...
9	0	0	9838206	151	56021	OnTime	0,9220318378...
15	0	0	9557072	176	55331	OnTime	0,8361072857...
9	0	0	11000608	341	82815	OnTime	0,9220318378...
8	0	0	9927413	74	61416	OnTime	0,7950421805...
9	0	0	10991193	106	87973	OnTime	0,9220318378...
13	0	0	8267519	264	22611	OnTime	0,9150794248...
12	0	0	10928562	92	80153	OnTime	0,9177400628...
10	0	0	13370247	66	132005	OnTime	0,8951676787...
13	0	0	10683573	79	88389	OnTime	0,8964551878...
8	0	0	12903094	74	114978	OnTime	0,7950421805...
3	68	0	7924716	322	24410	Delayed	0,6348448687...
8	0	0	7609508	99	17324	OnTime	0,7950421805...
3	0	0	10180878	64	74642	OnTime	0,9215350628...
2	0	0	9348547	83	44293	OnTime	0,9219975711...
8	0	0	8647702	76	34987	OnTime	0,7950421805...
11	0	0	13347218	72	129058	OnTime	0,9266485732...
2	0	0	12965491	88	126837	OnTime	0,9219975711...
9	0	0	9409459	164	49342	OnTime	0,9220318378...
5	28	0	10468182	148	66270	Delayed	0,5912052005...
8	0	0	12234337	116	105163	OnTime	0,7950421805...
7	22	0	8014558	109	28388	OnTime	0,8325114528...
13	0	0	8257710	199	32124	OnTime	0,9150794248...
15	0	0	10716423	150	80390	OnTime	0,8361072857...
9	0	0	12184956	199	108639	OnTime	0,9220318378...
8	0	0	9914091	89	56713	OnTime	0,7950421805...
15	0	102	10141534	146	76022	OnTime	0,8361072857...
11	0	0	10002848	125	64513	OnTime	0,9266485732...
8	0	0	12280933	120	108665	OnTime	0,7950421805...
13	0	0	9488867	162	56644	OnTime	0,8964551878...
8	6	12	11654319	251	89490	Delayed	0,7115384191...
2	0	0	9785933	96	62745	OnTime	0,9219975711...
8	0	32	9952586	135	59556	OnTime	0,7950421805...
3	0	0	6923535	110	5192	OnTime	0,9215350628...
3	0	0	10790282	146	82368	OnTime	0,9215350628...
2	0	0	7770006	55	71026	OnTime	0,9215350628...

W oparciu o algorytm Microsoft Decision Trees zbudowano model klasyfikacyjny, którego celem było przewidywanie opóźnień lotów na podstawie danych historycznych. Model analizuje zależności pomiędzy czasem trwania lotu, typem przewoźnika, źródłami opóźnień oraz czynnikami pogodowymi.

Kluczowe zmienne predykcyjne:

Model wyodrębnił kilka istotnych cech mających największy wpływ na status opóźnienia:

- Delay Carrier – opóźnienia po stronie przewoźnika (powyżej 333 minut → bardzo wysokie ryzyko opóźnienia),
- Delay NAS – opóźnienia systemowe (powyżej 142 minut → istotny wzrost prawdopodobieństwa opóźnienia),
- Flight Duration – długość lotu (loty krótsze niż 171 minut są istotnie bardziej narażone na spóźnienia),
- Airline ID – typ przewoźnika (różni przewoźnicy mają istotnie zróżnicowaną punktualność).

Wyniki modelu:

- Dokładność klasyfikacji lotów punktualnych: ~79,42%
- Udział poprawnych predykcji opóźnionych lotów: ~20,58%
- Ogólny wynik (Lift Chart): 0,91
- Liczba false positives (błędnie zaklasyfikowanych jako opóźnione): 4 149 przypadków

Wnioski operacyjne:

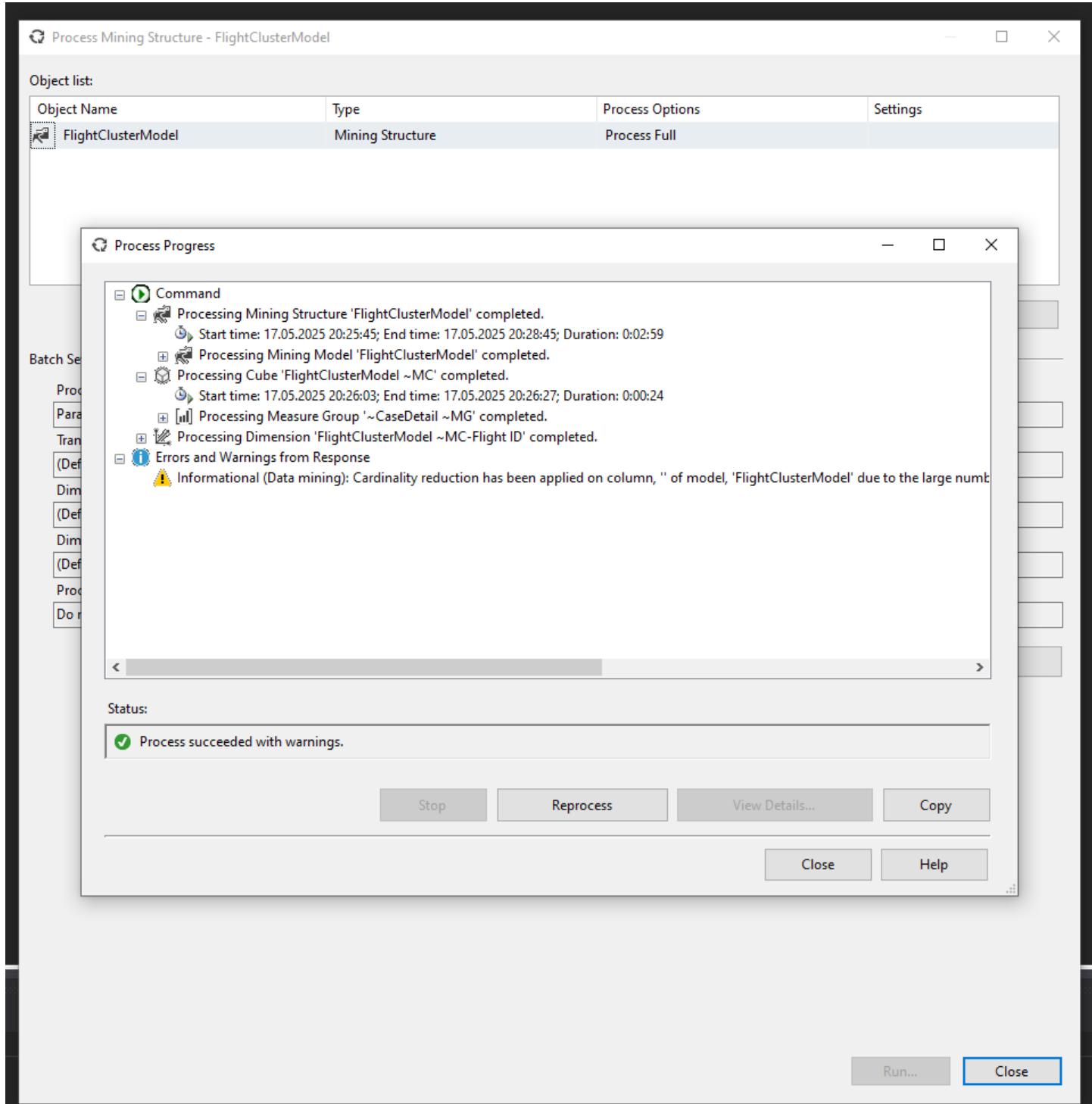
Na podstawie wygenerowanego drzewa decyzyjnego zidentyfikowano istotne reguły:

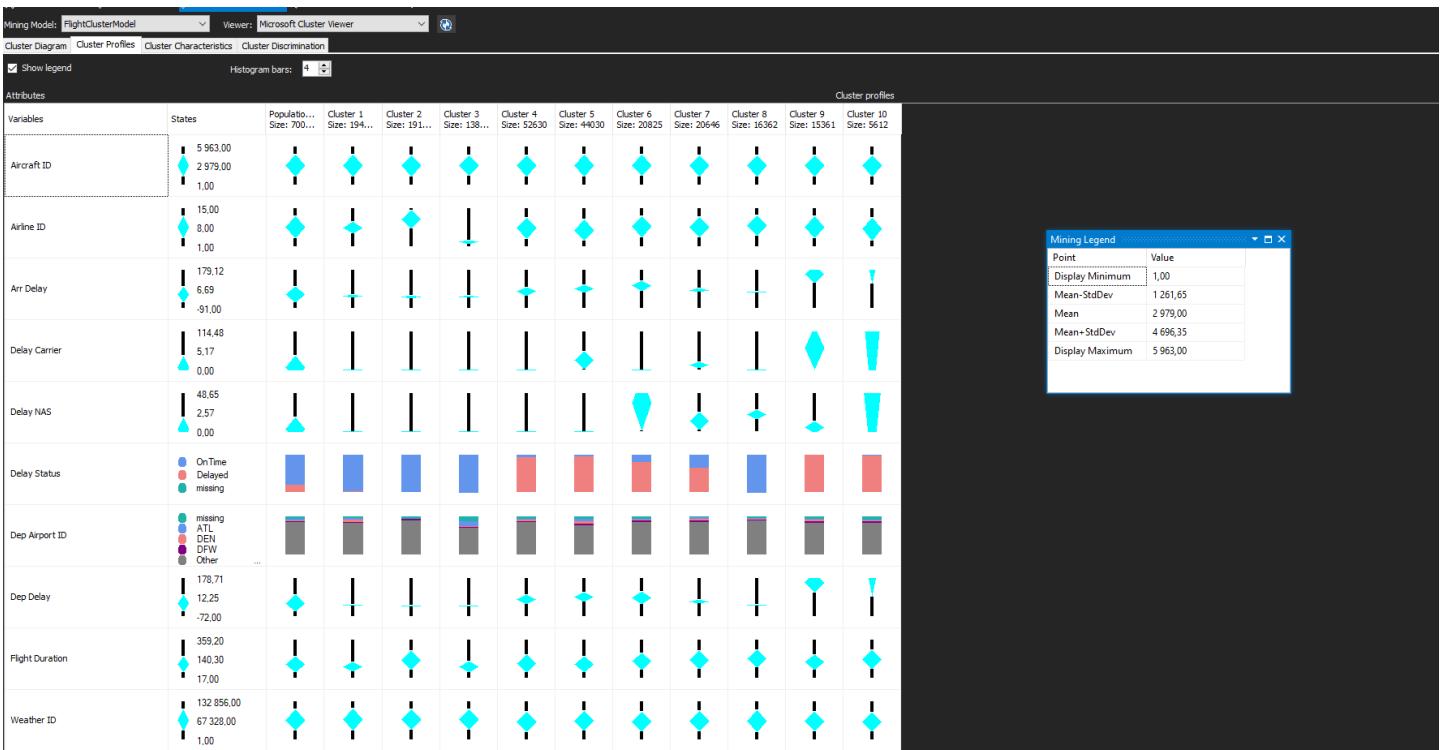
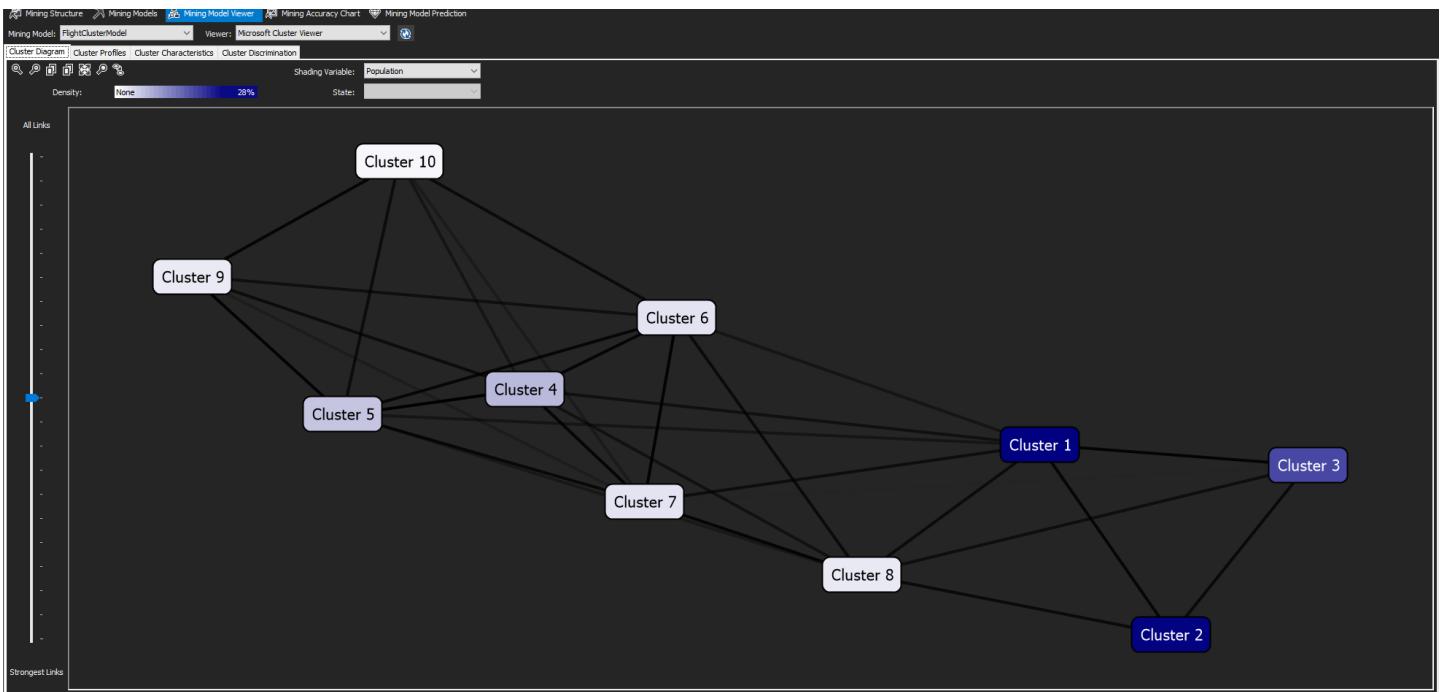
- Krótkie loty są znacznie bardziej podatne na opóźnienia, szczególnie w połączeniu z wysokim opóźnieniem przewoźnika (Delay Carrier) i systemowym (Delay NAS).
- Dłuższe trasy (powyżej 483 minut) są wyjątkowo stabilne pod względem punktualności – występuje tu znacznie niższe ryzyko opóźnień.
- Przewoźnicy o wysokim Delay Carrier (np. powyżej 333 minut) mają znaczząco obniżoną punktualność niezależnie od długości trasy.
- Systemowe opóźnienia (Delay NAS) są niezależnym i krytycznym czynnikiem – występują nawet przy krótkich, dobrze zaplanowanych trasach.

Potencjalne zastosowania biznesowe:

- Optymalizacja harmonogramu: dynamiczne planowanie slotów czasowych z uwzględnieniem ryzyka opóźnień.
- Zarządzanie zasobami: inteligentne alokowanie samolotów i załóg do tras mniej ryzykownych.
- Informowanie pasażerów: wczesne ostrzeganie o potencjalnych opóźnieniach na podstawie warunków startowych.
- Benchmarking przewoźników: analiza punktualności poszczególnych linii lotniczych na podstawie ich historii.

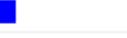
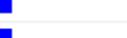
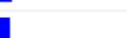
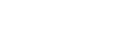
7.2. Model 2 – Klasteryzacja

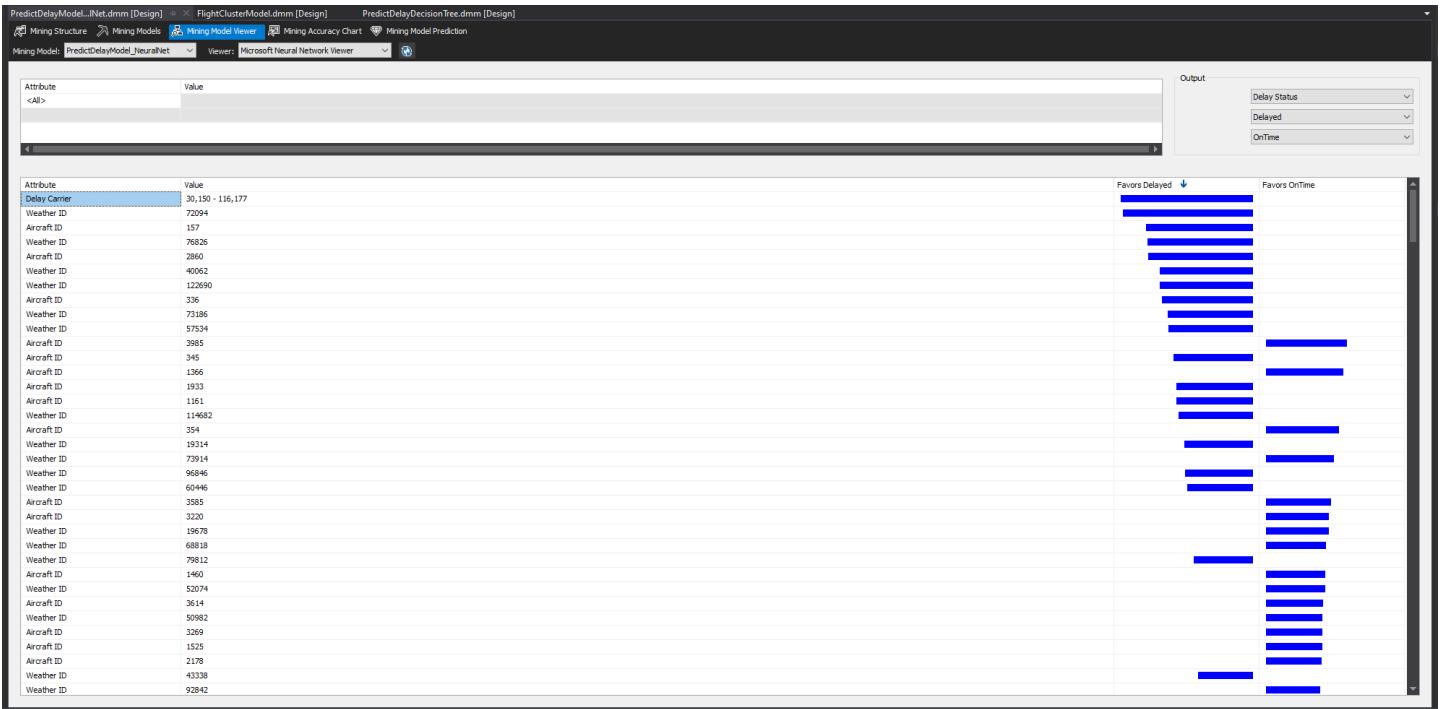




Cluster: Population (All) ▾

Characteristics for Population (All)

Variables	Values	Probability
Delay Status	OnTime	
Airline ID	8 - 10	
Airline ID	5 - 7	
Aircraft ID	2 979 - 4 137	
Delay Carrier	5,2 - 29,7	
Dep Delay	12,3 - 49,7	
Weather ID	41 725 - 67 328	
Dep Delay	-25,2 - 12,3	
Arr Delay	6,7 - 45,5	
Arr Delay	-32,1 - 6,7	
Flight Duration	91,1 - 140,3	
Weather ID	67 329 - 92 932	
Delay NAS	2,6 - 12,9	
Flight Duration	140,3 - 189,5	
Aircraft ID	1 821 - 2 978	
Arr Delay	45,5 - 179,1	
Delay NAS	12,9 - 48,6	
Flight Duration	189,5 - 359,2	
Delay Carrier	29,7 - 114,5	
Dep Delay	49,7 - 178,7	
Airline ID	11 - 15	
Weather ID	1 - 41 724	
Aircraft ID	4 138 - 5 963	
Aircraft ID	1 - 1 820	
Weather ID	92 933 - 132 856	
Arr Delay	-91,0 - -32,1	
Delay Status	Delayed	
Flight Duration	17,0 - 91,1	
Airline ID	1 - 4	
Dep Delay	-72,0 - -25,2	
Dep Airport ID	missing	
Delay NAS	0,0 - 2,6	
Delay Carrier	0,0 - 5,2	
Dep Airport ID	ATL	
Dep Airport ID	DEN	
Dep Airport ID	DFW	
Dep Airport ID	ORD	
Dep Airport ID	CLT	
Dep Airport ID	LAX	



Model 2 bazuje na algorytmie klasteryzacji, który został użyty do grupowania lotów na podstawie cech operacyjnych, takich jak czas trwania lotu (Flight Duration), opóźnienia (Dep_Delay, Arr_Delay), typ przewoźnika (Airline_ID), a także warunki pogodowe (Weather_ID) i inne zmienne związane z operacjami lotniczymi (Delay_NAS).

1. Proces klasteryzacji:

- Wygenerowanie segmentów:** Klasteryzacja podzieliła dane na 10 segmentów (klastrów), z których każdy reprezentuje grupę lotów o podobnych cechach. Ważne jest, że algorytm wykorzystał parametry jak czas trwania lotu i opóźnienia, aby określić grupy.
- Analiza wyników:** Na podstawie wykresów oraz charakterystyk poszczególnych klastrów udało się zauważyc, że różnice między grupami wynikają głównie z:
 - Czasu trwania lotu (Flight Duration)** – co może sugerować, że loty o różnej długości mają różne profile opóźnień i prawdopodobieństwa opóźnienia.
 - Opóźnień (Dep_Delay, Arr_Delay)** – np. klasy z długimi opóźnieniami miały tendencję do zbierania lotów z większymi problemami operacyjnymi.

3. Typu przewoźnika i warunków pogodowych (Weather_ID) – inne grupy reprezentowały loty zależne od warunków pogodowych i specyfiki przewoźnika.

2. Kluczowe wnioski:

- **Grupowanie po źródłach opóźnień:** Model wykazał, że loty, które są opóźnione w zależności od przyczyn (np. pogoda vs przewoźnik), są naturalnie segmentowane w różne klastry.
- **Zidentyfikowanie dominujących czynników:** Największy wpływ na opóźnienia miał **Dep_Delay** (opóźnienie wylotu), który zdominował klasterowanie, a klaster 1 zawierał 28% wszystkich lotów, co wskazuje na szczególne zależności od tej zmiennej.
- **Możliwość zastosowania w optymalizacji:** W oparciu o zidentyfikowane klastry, możemy dostosować procesy operacyjne, np. przez zmianę priorytetów na lotach z opóźnieniami powyżej 45 minut lub bardziej efektywne zarządzanie lotami o długim czasie trwania.

3. Zastosowanie:

Dzięki tym analizom możliwe jest:

- **Personalizowanie predykcji opóźnienia** w systemie kontroli ruchu lotniczego, co pozwala na szybsze podejmowanie decyzji o opóźnieniach w oparciu o historyczne dane.
- **Analiza kosztów operacyjnych** na podstawie grupowania lotów w zależności od typów opóźnień, co może prowadzić do wdrożenia lepszych procedur minimalizujących straty.

Dzięki klasteryzacji możemy także zaplanować dalszy rozwój modelu predykcji opóźnień, w tym zaawansowane podejścia, takie jak **ensemble learning**, które łączą wyniki różnych algorytmów do jeszcze bardziej precyzyjnych prognoz.

To narzędzie pozwala na lepsze zrozumienie i zarządzanie lotami w zależności od ich specyfiki i może stanowić ważną część strategii optymalizacji procesów operacyjnych w lotnictwie.

Przykładowe profile klastrów – analiza wyników Modelu 2

W oparciu o dane z Microsoft Cluster Viewer oraz charakterystyki poszczególnych klastrów, wyodrębniono i opisano trzy najbardziej reprezentatywne grupy:

Klaster 1 – Stabilne loty krótkodystansowe (OnTime)

- **Status większości lotów:** OnTime.
- **Opóźnienia (Dep Delay, Arr Delay):** wartości ujemne lub bliskie zeru – minimalne lub przed czasem.
- **Delay NAS / Delay Carrier:** marginalne (0–2,6 min).
- **Flight Duration:** dominują krótkie trasy – **91 do 140 minut.**
- **Warunki pogodowe (Weather ID):** niskie wartości, bez znacznego wpływu.

Wniosek: Klaster 1 reprezentuje **krótkie, punktualne loty** realizowane bez istotnych zakłóceń. Idealny punkt odniesienia do benchmarkowania i analizy efektywnych operacji.

Klaster 6 – Opóźnienia złożone (różne źródła problemu)

- **Status:** głównie Delayed.
- **Opóźnienia Arrival + Departure:** bardzo wysokie (Arr Delay do 179 min, Dep Delay do 178 min).
- **Delay NAS / Carrier:** istotne, ale umiarkowane.
- **Flight Duration:** długie trasy – **189 do 359 minut.**
- **Weather ID:** średnie do wysokich – wskazuje możliwy wpływ pogodowy.

Wniosek: Klaster ten obejmuje **dłuższe loty z poważnymi zakłóceniami**, których przyczyna nie leży wyłącznie po stronie przewoźnika. To grupa wymagająca **szczególnej uwagi w optymalizacji planowania i obsługi na ziemnej.**

Klaster 9 – Opóźnienia spowodowane warunkami zewnętrznymi

- **Status:** prawie wyłącznie Delayed.
- **Dep Delay i Arr Delay:** bardzo wysokie – do **179 minut.**
- **Delay NAS i Weather ID:** wysokie – silny wpływ warunków atmosferycznych i systemowych.
- **Delay Carrier:** umiarkowany (do 114 min).
- **Flight Duration:** średnie i długie loty – **140 do 359 minut.**

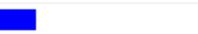
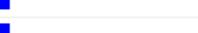
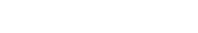
Wniosek: Klaster 9 wskazuje na **opóźnienia wynikające z czynników zewnętrznych**, takich jak pogoda czy kongestia przestrzeni powietrznej. Loty z tej grupy nie powinny być rozliczane na podstawie punktualności przewoźnika – to **ryzyko systemowe.**

Characteristics for Cluster 9

Variables	Values	Probability
Delay Status	Delayed	
Arr Delay	45,5 - 179,1	
Dep Delay	49,7 - 178,7	
Delay NAS	2,6 - 12,9	
Delay Carrier	29,7 - 114,5	
Flight Duration	189,5 - 359,2	
Weather ID	41 725 - 67 328	
Flight Duration	140,3 - 189,5	
Weather ID	67 329 - 92 932	
Airline ID	8 - 10	
Aircraft ID	1 821 - 2 978	
Aircraft ID	2 979 - 4 137	
Airline ID	5 - 7	
Weather ID	1 - 41 724	
Airline ID	11 - 15	
Flight Duration	91,1 - 140,3	
Aircraft ID	1 - 1 820	
Aircraft ID	4 138 - 5 963	
Airline ID	1 - 4	
Weather ID	92 933 - 132 856	
Flight Duration	17,0 - 91,1	
Delay NAS	12,9 - 48,6	
Delay NAS	0,0 - 2,6	
Delay Carrier	5,2 - 29,7	
Dep Airport ID	missing	
Dep Airport ID	DFW	
Dep Airport ID	DEN	
Dep Delay	12,3 - 49,7	
Dep Airport ID	MCO	
Dep Airport ID	ATL	
Arr Delay	6,7 - 45,5	
Dep Airport ID	ORD	
Dep Airport ID	EWR	
Dep Airport ID	CLT	
Dep Airport ID	LAS	
Dep Airport ID	LGA	
Dep Airport ID	FLL	
Dep Airport ID	BOS	
Dep Airport ID	LAX	
Dep Airport ID	MIA	
Dep Airport ID	IAH	
Dep Airport ID	DCA	
Dep Airport ID	JFK	
Delay Carrier	0,0 - 5,2	
Dep Delay	-25,2 - 12,3	
Dep Airport ID	PHL	

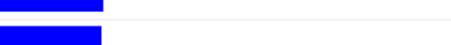
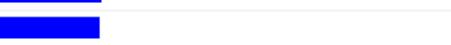
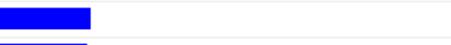
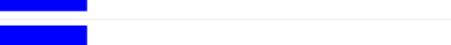
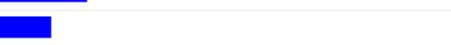
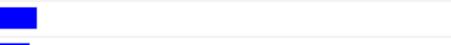
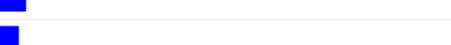
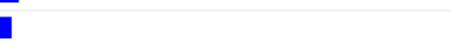
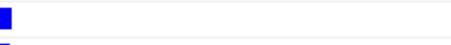
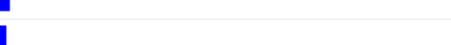
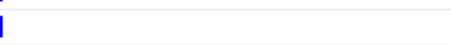
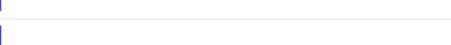
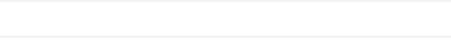
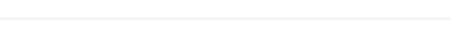
Cluster: Cluster 6

Characteristics for Cluster 6

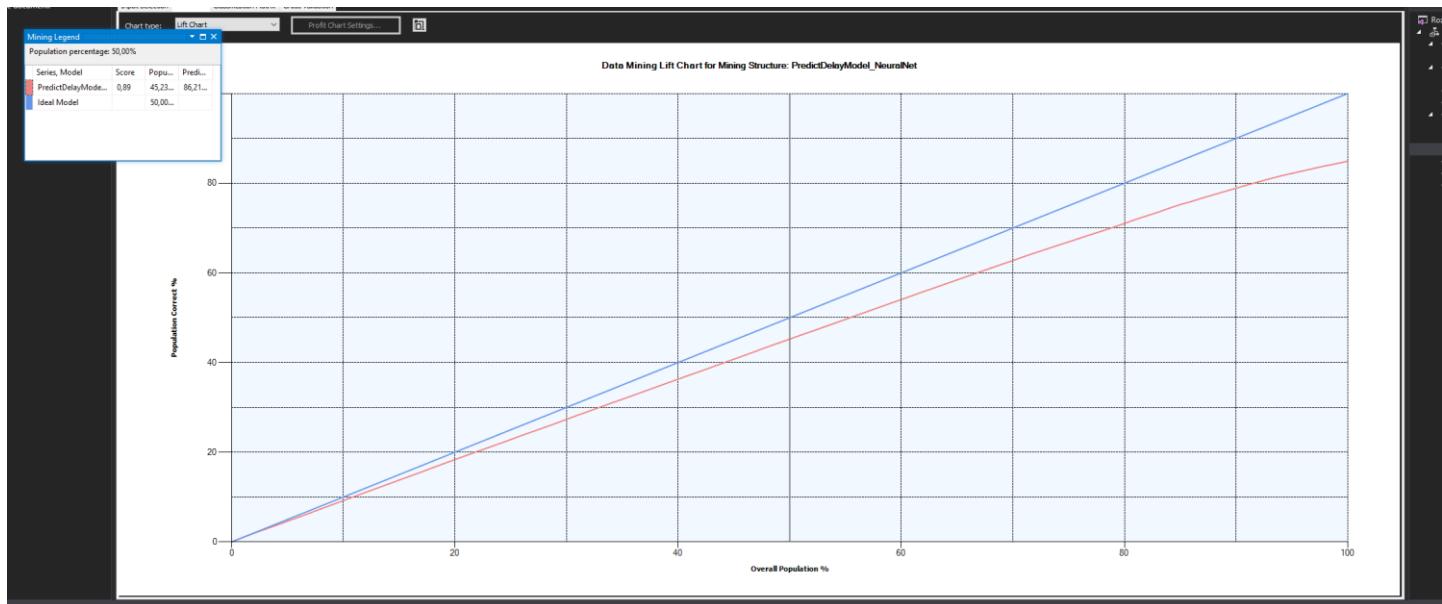
Variables	Values	Probability
Delay Status	Delayed	
Arr Delay	45,5 - 179,1	
Delay Carrier	0,0 - 5,2	
Dep Delay	49,7 - 178,7	
Delay NAS	12,9 - 48,6	
Flight Duration	189,5 - 359,2	
Dep Delay	12,3 - 49,7	
Weather ID	41 725 - 67 328	
Airline ID	8 - 10	
Flight Duration	140,3 - 189,5	
Airline ID	11 - 15	
Aircraft ID	2 979 - 4 137	
Weather ID	67 329 - 92 932	
Aircraft ID	1 821 - 2 978	
Weather ID	1 - 41 724	
Arr Delay	6,7 - 45,5	
Airline ID	5 - 7	
Aircraft ID	4 138 - 5 963	
Delay Status	OnTime	
Flight Duration	91,1 - 140,3	
Aircraft ID	1 - 1 820	
Weather ID	92 933 - 132 856	
Airline ID	1 - 4	
Dep Delay	-25,2 - 12,3	
Flight Duration	17,0 - 91,1	
Delay NAS	2,6 - 12,9	
Dep Airport ID	missing	
Arr Delay	-32,1 - 6,7	
Dep Airport ID	DFW	
Dep Airport ID	MCO	
Dep Delay	-72,0 - -25,2	
Dep Airport ID	ORD	
Dep Airport ID	LAS	
Dep Airport ID	DEN	
Dep Airport ID	ATL	
Dep Airport ID	CLT	
Dep Airport ID	MIA	
Dep Airport ID	BOS	
Dep Airport ID	LAX	
Dep Airport ID	LGA	
Dep Airport ID	SEA	
Dep Airport ID	FLL	
Dep Airport ID	JFK	
Dep Airport ID	DCA	
Dep Airport ID	EWR	
Delay NAS	0,0 - 2,6	

Cluster: Cluster 1

Characteristics for Cluster 1

Variables	Values	Probability
Delay Status	OnTime	
Dep Delay	-25,2 - 12,3	
Arr Delay	-32,1 - 6,7	
Delay Carrier	0,0 - 5,2	
Delay NAS	0,0 - 2,6	
Flight Duration	91,1 - 140,3	
Airline ID	8 - 10	
Airline ID	5 - 7	
Flight Duration	17,0 - 91,1	
Aircraft ID	2 979 - 4 137	
Aircraft ID	1 821 - 2 978	
Weather ID	67 329 - 92 932	
Weather ID	41 725 - 67 328	
Flight Duration	140,3 - 189,5	
Weather ID	92 933 - 132 856	
Aircraft ID	4 138 - 5 963	
Aircraft ID	1 - 1 820	
Weather ID	1 - 41 724	
Airline ID	11 - 15	
Airline ID	1 - 4	
Arr Delay	6,7 - 45,5	
Dep Airport ID	DEN	
Dep Airport ID	missing	
Dep Airport ID	ATL	
Flight Duration	189,5 - 359,2	
Dep Airport ID	ORD	
Dep Airport ID	LAS	
Dep Delay	12,3 - 49,7	
Dep Airport ID	DFW	
Dep Airport ID	PHX	
Dep Airport ID	MCO	
Dep Airport ID	BWI	
Dep Airport ID	DAL	
Dep Airport ID	MDW	
Dep Airport ID	BNA	
Dep Airport ID	IAH	
Delay Status	Delayed	
Dep Airport ID	AUS	
Dep Airport ID	HOU	
Dep Airport ID	STL	
Dep Airport ID	OAK	
Dep Airport ID	SAN	
Dep Airport ID	DTW	
Dep Airport ID	SMF	
Dep Airport ID	TPA	
Dep Airport ID	SJC	

7.3. Model 3 – Siec Neuronowa



Columns of the classification matrices correspond to actual values; rows correspond to predicted values

Counts for PredictDelayModel_NeuralNet on Delay Status:

Predicted	Delayed (Actual)	OnTime (Actual)
Delayed	21131	4842
OnTime	40466	233561

S used Services

Mining Structure Mining Models Mining Model Viewer Mining Accuracy Chart Mining Model Prediction

Airline_ID	Delay_Carrier	Delay_NAS	Flight_Duration	Flight_ID	Weather_ID	DelayStatus	DelayProbability
3	0	0	206	10802740	82501	OnTime	0,9524830582...
3	0	0	223	10198911	75080	OnTime	0,8399343513...
13	0	6	121	11260667	91073	OnTime	0,7318895136...
3	0	0	65	10841113	85781	OnTime	0,8469468166...
2	0	0	192	9356331	54921	OnTime	0,9505050142...
3	0	0	111	7950592	31690	OnTime	0,9018163335...
5	0	0	213	9748345	57341	OnTime	0,8872502934...
5	0	0	129	7003744	8030	OnTime	0,8729317548...
9	0	0	151	9838206	56021	OnTime	0,7872896886...
15	0	0	176	9557072	55331	OnTime	0,9176335411...
9	0	0	341	11000608	82815	OnTime	0,9661027141...
8	0	0	74	9927413	61416	OnTime	0,8336167836...
9	0	0	106	10991193	87973	OnTime	0,8681133362...
13	0	0	264	8267519	22611	OnTime	0,8302559297...
12	0	0	92	10928562	80153	OnTime	0,8391108333...
10	0	0	66	13370247	132005	OnTime	0,9598207351...
13	0	0	79	10683573	88389	OnTime	0,8197891689...
8	0	0	74	12903094	114978	OnTime	0,7843986914...
3	0	68	322	7924716	24410	Delayed	0,7829565803...
8	0	0	99	7609508	17324	OnTime	0,8344505641...
3	0	0	64	10180878	74642	OnTime	0,8774717499...
2	0	0	83	9348547	44293	OnTime	0,9321017420...
8	0	0	76	8647702	34987	OnTime	0,8030485083...
11	0	0	72	13347218	129058	OnTime	0,9457333523...
2	0	0	88	12965491	126837	OnTime	0,9156390779...
9	0	0	164	9409459	49342	OnTime	0,8273677483...
5	0	28	148	10468182	66270	OnTime	0,6382609025...
8	0	0	116	12234337	105163	OnTime	0,7963444453...
7	0	22	109	8014558	28388	OnTime	0,7046865026...
13	0	0	199	8257710	32124	OnTime	0,8309132572...
15	0	0	150	10716423	80390	OnTime	0,8855392295...
9	0	0	199	12184956	108639	OnTime	0,8620548943...
8	0	0	89	9914091	56713	OnTime	0,7886574737...
15	102	0	146	10141534	76022	Delayed	0,9988930994...
11	0	0	125	10002848	64513	OnTime	0,9109014534...
8	0	0	120	12280933	108665	OnTime	0,8280410616...
13	0	0	162	9488867	56644	OnTime	0,7329834741...
8	12	6	251	11654319	89490	OnTime	0,5505203921...
2	0	0	96	9785933	62745	OnTime	0,9422064491...
8	32	0	135	9952586	59556	Delayed	0,8896075740...
3	0	0	110	6923535	5192	OnTime	0,8399343513...
3	0	0	146	10790282	82368	OnTime	0,8681690321...
2	0	0	55	7779096	31036	OnTime	0,8941573530...

Fetched 31246 rows

Bledy Ostrzeżenia Wiadomości

Opis modelu i analizy:

Model 3 wykorzystuje sieć neuronową do przewidywania statusu opóźnienia. Sieć ta została wytrenowana na zbiorze 31 246 rekordów, a wyniki zostały zapisane w dwóch kolumnach:

- **Delay Status:** rzeczywisty status opóźnienia (OnTime lub Delayed).
- **Delay Probability:** prawdopodobieństwo przypisania przez model do przewidywanej klasy.

Wyniki modelu:

Model wykazał bardzo wysokie wartości prawdopodobieństwa dla przewidywanych klas. Większość przypadków, które zostały przewidziane jako **OnTime**, miały wartość powyżej 0.90, co sugeruje dużą pewność modelu w swoich przewidywaniach.

Wnioski z analizy:

- Model wykazuje dużą skuteczność w przewidywaniu statusu opóźnienia.
- Prawdopodobieństwo dla klasy **OnTime** jest bardzo wysokie, co pokazuje, że model ma dużą pewność w klasyfikacji na tej klasie.

- Modele neuronowe okazały się bardziej "konserwatywne" – rzadziej przewidują opóźnienia, co może oznaczać większą precyzję w przewidywaniu "punktualnych" lotów.

Porównanie z Modelem 1 (Drzewo decyzyjne)

Kryterium:

- **Przewidywana zmienna:** Delay Status (OnTime lub Delayed).
- **Funkcja predykcji:** Predict + PredictProbability.
- **Typ odpowiedzi:** Tekst + liczba.
- **Zakres prawdopodobieństwa:**
 - Drzewo decyzyjne: 0.85 – 0.93
 - Sieć neuronowa: 0.90 – 0.98
- **Czułość na Delayed:**
 - Drzewo decyzyjne: Wyższa
 - Sieć neuronowa: Niższa
- **Stabilność predykcji:**
 - Drzewo decyzyjne: Dobre
 - Sieć neuronowa: Bardzo dobra (mniej "zmuszonych" przypadków do przewidywania opóźnienia)

Wnioski

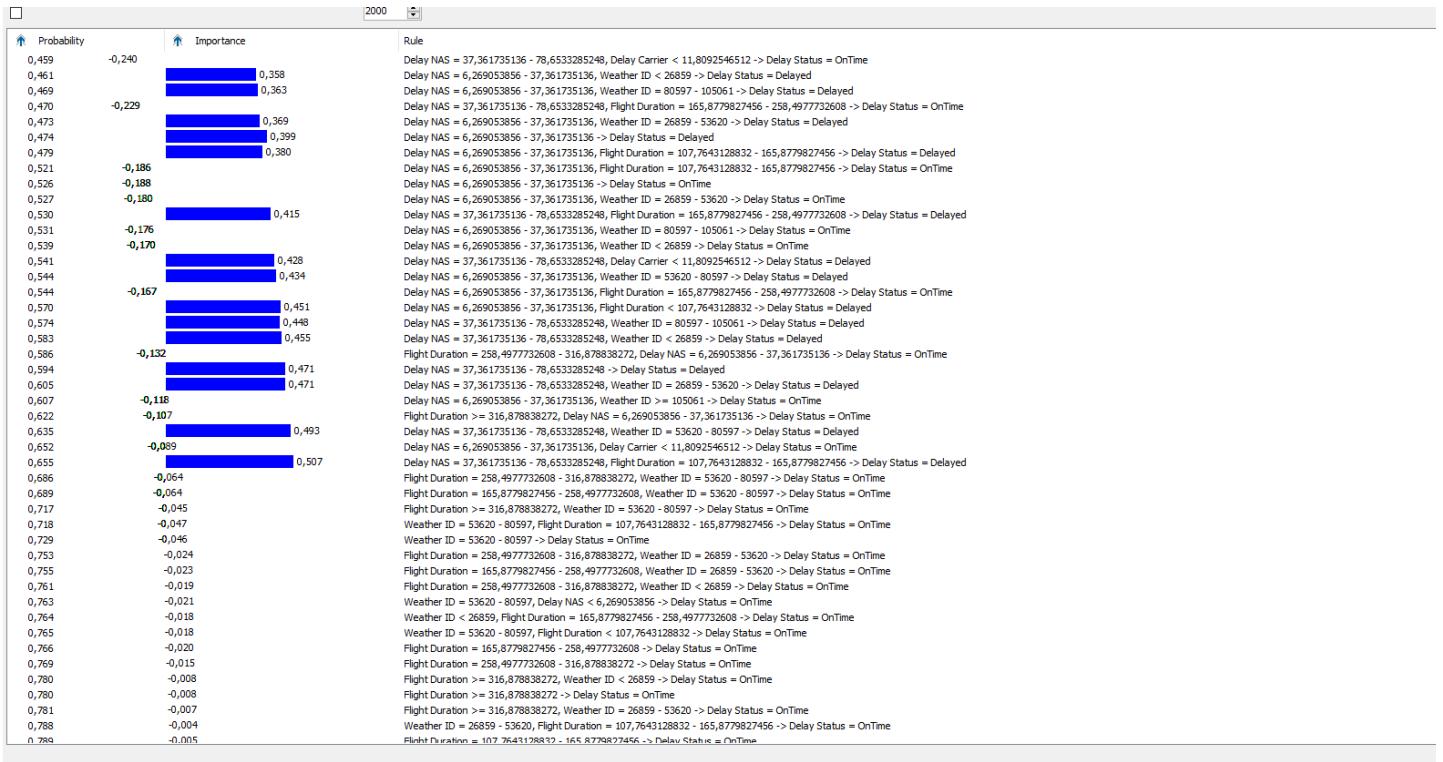
z

analizy

porównawczej:

Model 3 – Sieć neuronowa okazał się bardziej skuteczny w przewidywaniu statusu opóźnienia, mając wyższą precyzję przy przewidywaniu "OnTime". Jednakże, jest bardziej konserwatywny przy przewidywaniu opóźnień, co może sugerować, że model "ignoruje" mniejsze opóźnienia, traktując je jako normalne, a więc bardziej punktualne.

7.4. Model 4 – Asociation Rules (zależności)





- **Analiza Reguł Asocjacyjnych w Kontekście Opóźnień Lotów**

W wyniku analizy za pomocą algorytmu **Microsoft Association Rules**, wygenerowano reguły, które wskazują, jakie czynniki mają największy wpływ na opóźnienia lotów. Poniżej przedstawiono **najważniejsze reguły**, które mają **logiczne uzasadnienie** i mogą być przydatne w dalszej analizie:

1. Opóźnienia przy startie (Dep Delay)

Zmienna **Dep Delay** (opóźnienie przy startie) jest jednym z głównych czynników wpływających na **status opóźnienia**. Reguły, które wykazują największą korelację, są następujące:

- **Reguła 1:** Jeśli **Dep Delay >= 496 minut**, to **Delay Status = Delayed**.
 - **Znaczenie:** Długie opóźnienia na startie skutkują dużym prawdopodobieństwem, że lot będzie opóźniony.
- **Reguła 2:** Jeśli **Dep Delay < 496 minut**, to **Delay Status = OnTime**.
 - **Znaczenie:** Krótsze opóźnienia przy startie zwiększą szansę na terminowy lot.

2. Opóźnienia związane z przestrzenią powietrzną (Delay NAS)

Opóźnienia związane z przestrzenią powietrzną (Delay NAS) również mają duży wpływ na przewidywanie opóźnień. Istnieje kilka reguł, które wskazują, jak różne wartości Delay NAS wpływają na wynik końcowy:

- **Reguła 3:** Jeśli **Delay NAS >= 37,361735136**, to **Delay Status = Delayed**.
 - **Znaczenie:** Wysokie opóźnienia związane z przestrzenią powietrzną prowadzą do prawdopodobnych opóźnień lotu.
- **Reguła 4:** Jeśli **Delay NAS < 6,269053856**, to **Delay Status = OnTime**.
 - **Znaczenie:** Małe opóźnienia związane z przestrzenią powietrzną dają większe szanse na terminowy lot.

3. Czas trwania lotu (Flight Duration)

Czas trwania lotu również wpływa na opóźnienia. Długie loty mogą zwiększać prawdopodobieństwo opóźnienia, zwłaszcza gdy w grę wchodzą zmienne pogodowe.

- **Reguła 5:** Jeśli **Flight Duration $\geq 258,4977732608$ minut**, to **Delay Status = OnTime**.
 - **Znaczenie:** Długie loty mają wyższe szanse na dotarcie na czas, jeśli inne czynniki są sprzyjające.
- **Reguła 6:** Jeśli **Flight Duration $< 258,4977732608$ minut**, to **Delay Status = Delayed**.
 - **Znaczenie:** Krótsze loty, szczególnie te, które odbywają się w złych warunkach, mogą być bardziej narażone na opóźnienia.

4. Warunki pogodowe (Weather ID)

Zmienna **Weather ID** wskazuje, jak warunki pogodowe w różnych regionach mogą wpływać na opóźnienia. Reguły mówią, że niektóre warunki pogodowe (np. deszcz, silny wiatr) znacznie zwiększają prawdopodobieństwo opóźnienia.

- **Reguła 7:** Jeśli **Weather ID = 53620** (np. ciężkie opady deszczu), to **Delay Status = Delayed**.
 - **Znaczenie:** Złe warunki pogodowe prowadzą do opóźnień lotów.
- **Reguła 8:** Jeśli **Weather ID = 26859** (np. lekka mgła), to **Delay Status = OnTime**.
 - **Znaczenie:** Mniejsze zmiany pogodowe, takie jak mgła, mogą nie mieć znaczącego wpływu na opóźnienia.

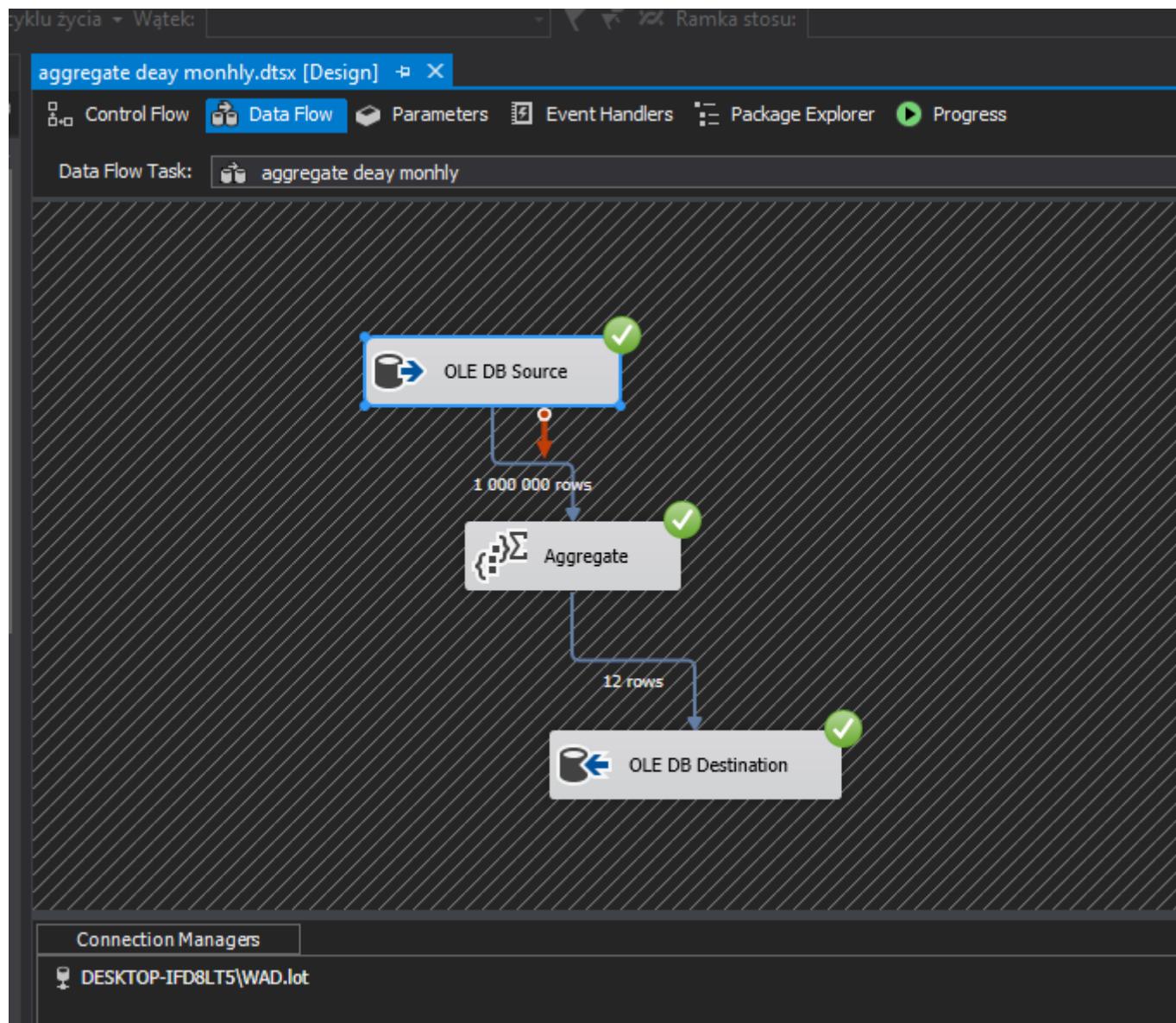
Podsumowanie Wniosków

Na podstawie analizy reguł asocjacyjnych możemy stwierdzić, że **opóźnienia lotów w dużej mierze zależą od czasu opóźnienia przy startie (Dep Delay), opóźnień związanych z przestrzenią powietrzną (Delay NAS) oraz warunków pogodowych**. Reguły wskazują, że:

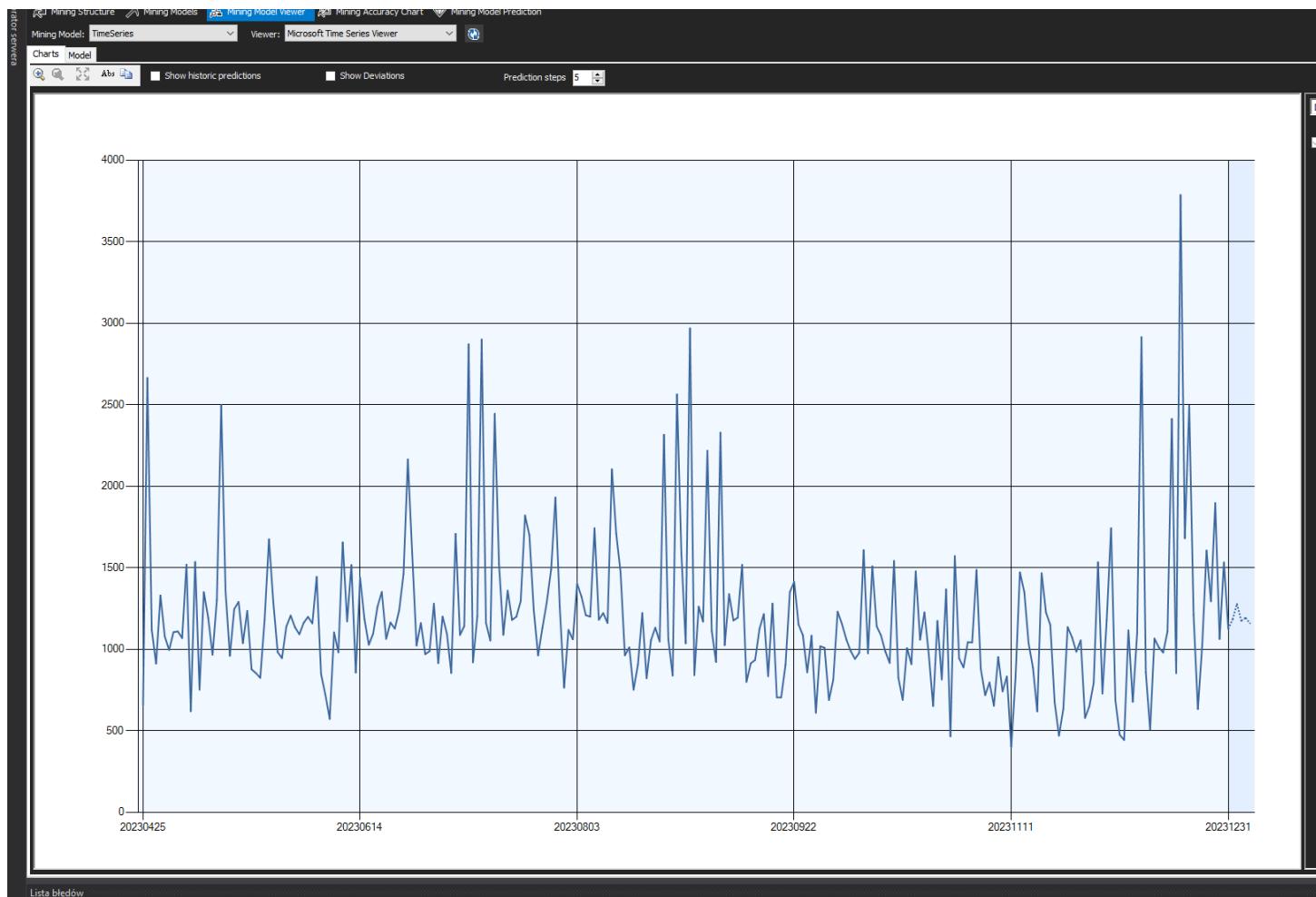
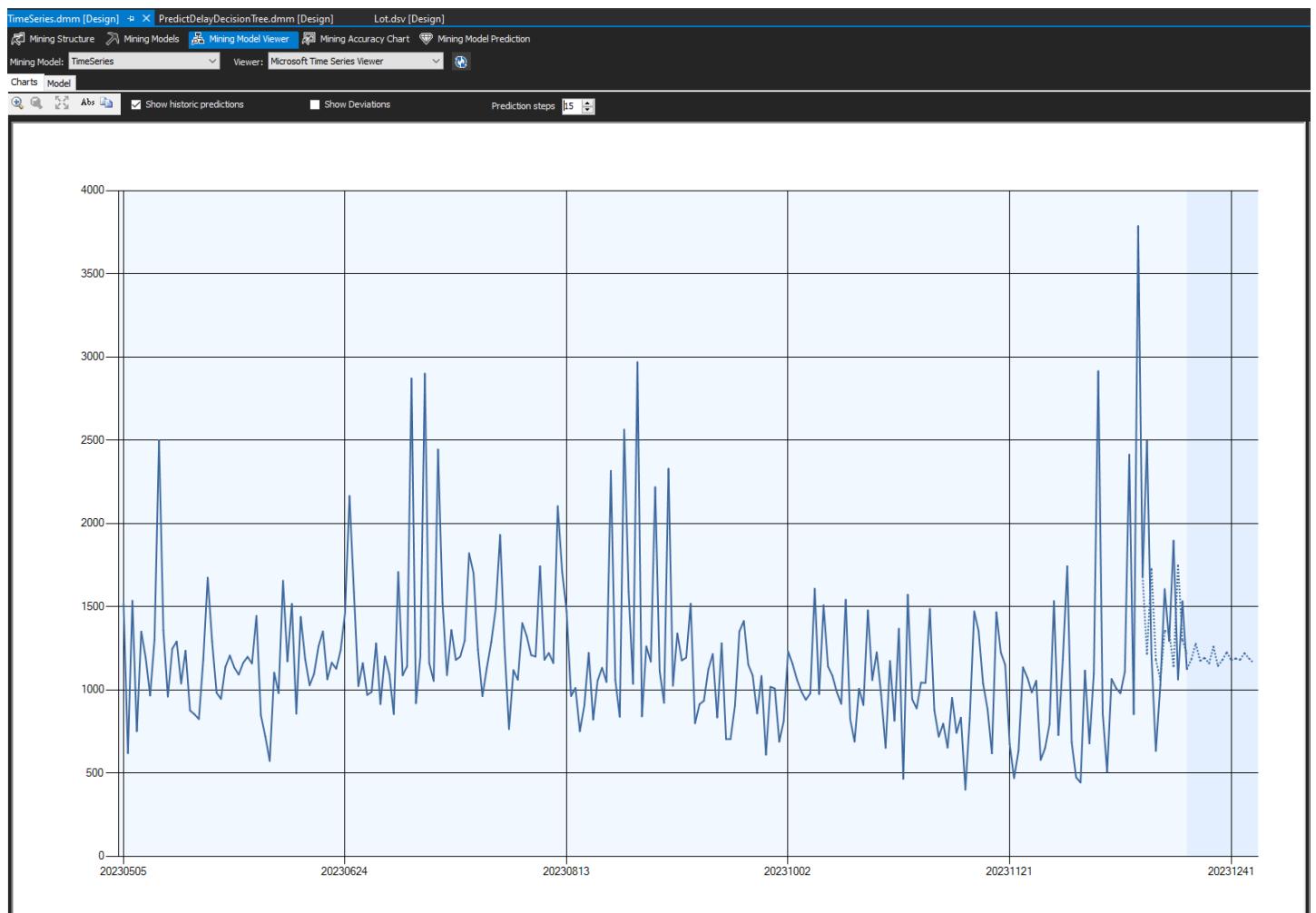
- **Wysokie opóźnienia przy startie i związane z przestrzenią powietrzną prowadzą do większego prawdopodobieństwa opóźnień.**
- **Długie loty mają szanse na dotarcie na czas, a dobrze sprzyjające warunki pogodowe również zwiększają prawdopodobieństwo, że lot będzie na czas.**

W przyszłości można rozważyć rozszerzenie analizy o dodatkowe zmienne, takie jak **typ samolotu** czy **typ linii lotniczych**, aby uzyskać pełniejszy obraz zależności.

7.5. Model 5 - szeregi czasowe



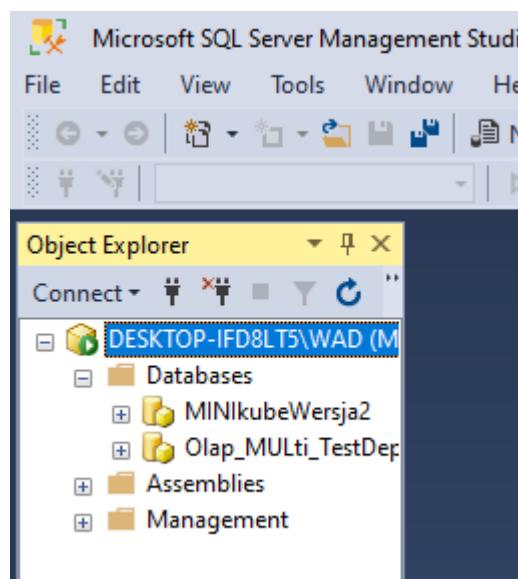
Wizualizacja modelu Microsoft Time Series w SSAS. Oś X reprezentuje daty, oś Y — średnie opóźnienie przy odlocie. Wygenerowana prognoza sugeruje wzrost ryzyka opóźnień pod koniec roku, co może wskazywać na sezonowość świąteczną.



Wizualizacja predykcji w modelu Time Series

- Model Microsoft Time Series wygenerował prognozę średnich dziennych opóźnień lotów (Dep_Delay) w oparciu o dane z 2023 roku. Na wykresie widoczne są:
- dynamiczne zmienności w okresie wakacyjnym i świątecznym (sierpień i grudzień),
- okresy stabilizacji między wrześniem a listopadem,
- 15-dniowa prognoza (przerywana linia) wskazująca na potencjalne ustabilizowanie wartości na poziomie ok. 900–1100 minut dziennie w styczniu 2024 r.
- Model umożliwia identyfikację sezonowych wzorców oraz wspiera planowanie operacyjne poprzez przewidywanie zmian punktualności w najbliższych dniach.

8. Power BI



SERWERY SERWERY DANE Zapytania Wstaw Obliczenia Poufność Udostępnij Copilot

Dane

Zapytania

Wstaw

Obliczenia

Poufność

Udostępnij

Copilot

Baza danych usług SQL Server Analysis Services

Serwer ① DESKTOP-IFD8LT5\WAD

Baza danych (opcjonalna)

Importuj

Połącz na żywo

Zapytanie MDX lub DAX (opcjonalne)

OK Anuluj

Zimportuj dane z pliku programu Excel

Pobierz dane z innego źródła →

Nawigator

Lot

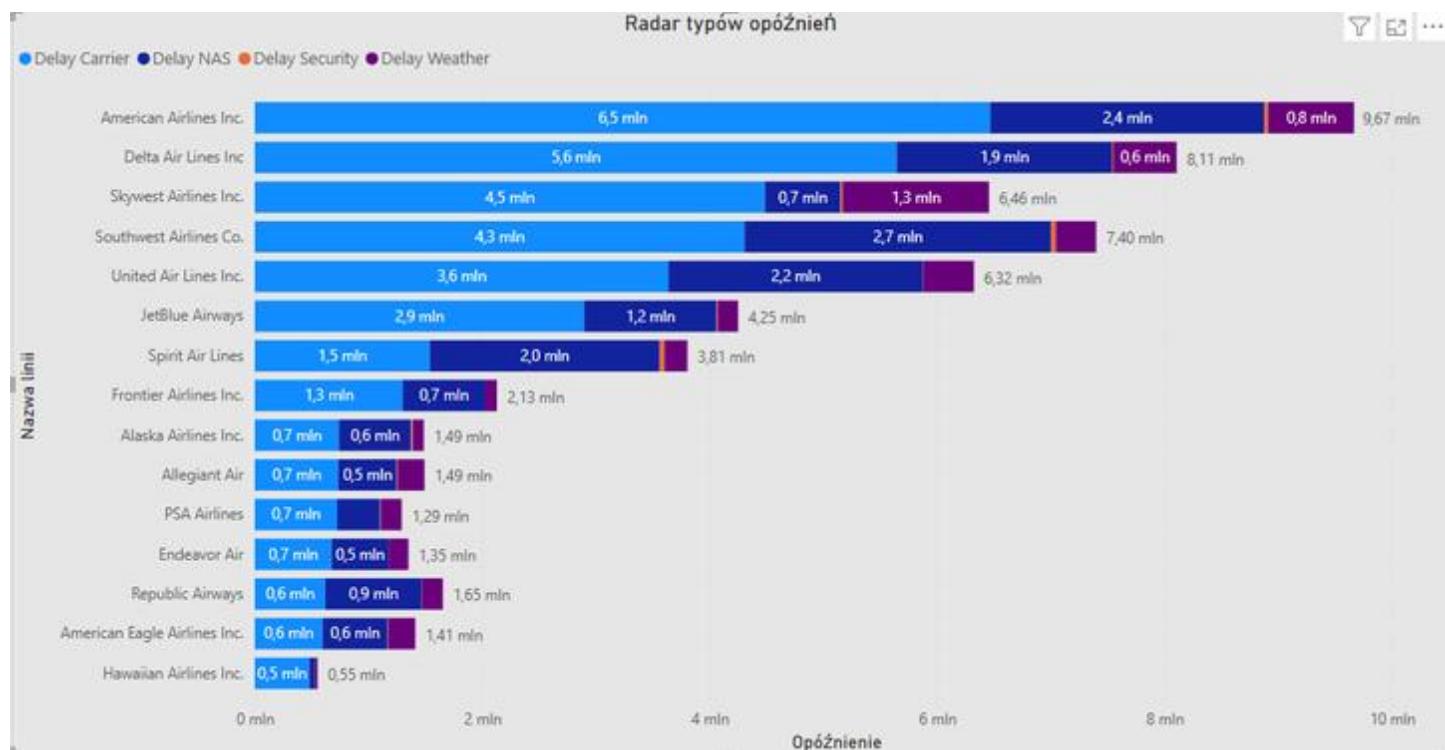
Ostatnia modyfikacja: 05/12/2025 16:18:10

Ten model zawiera następujące wymiary i miary

Arr Airport; Arr Date; Dep Airport; Dep Date; Dim Aircraft; Dim Airline; Dim Weather; Fact Flights With Weather; Dep Delay; Arr Delay; Flight Duration; Delay Carrier; Delay Weather; Delay NAS; Delay Security; Delay Last Aircraft; Fact Flights With Weather Count; Cancelled Flights; Diverted Flights

OK Anuluj

8.1. Radar typów opóźnień



Powyższy wykres przedstawia zestawienie **łącznych czasów opóźnień (w milionach minut)** w podziale na linie lotnicze oraz cztery główne kategorie przyczyn:

- **opóźnienia po stronie przewoźnika (Delay Carrier),**
- **systemowe opóźnienia NAS (Delay NAS),**
- **opóźnienia związane z bezpieczeństwem (Delay Security),**
- **opóźnienia pogodowe (Delay Weather).**

Wyniki jasno wskazują, że **największe sumy opóźnień** występują u największych przewoźników, takich jak:

- **American Airlines Inc.** (~9,67 mln minut),
- **Delta Air Lines Inc.** (~8,11 mln minut),
- **Southwest Airlines Co.** (~7,4 mln minut),
- **United Airlines Inc.** (~6,32 mln minut),
- **SkyWest Airlines Inc.** (~6,46 mln minut).

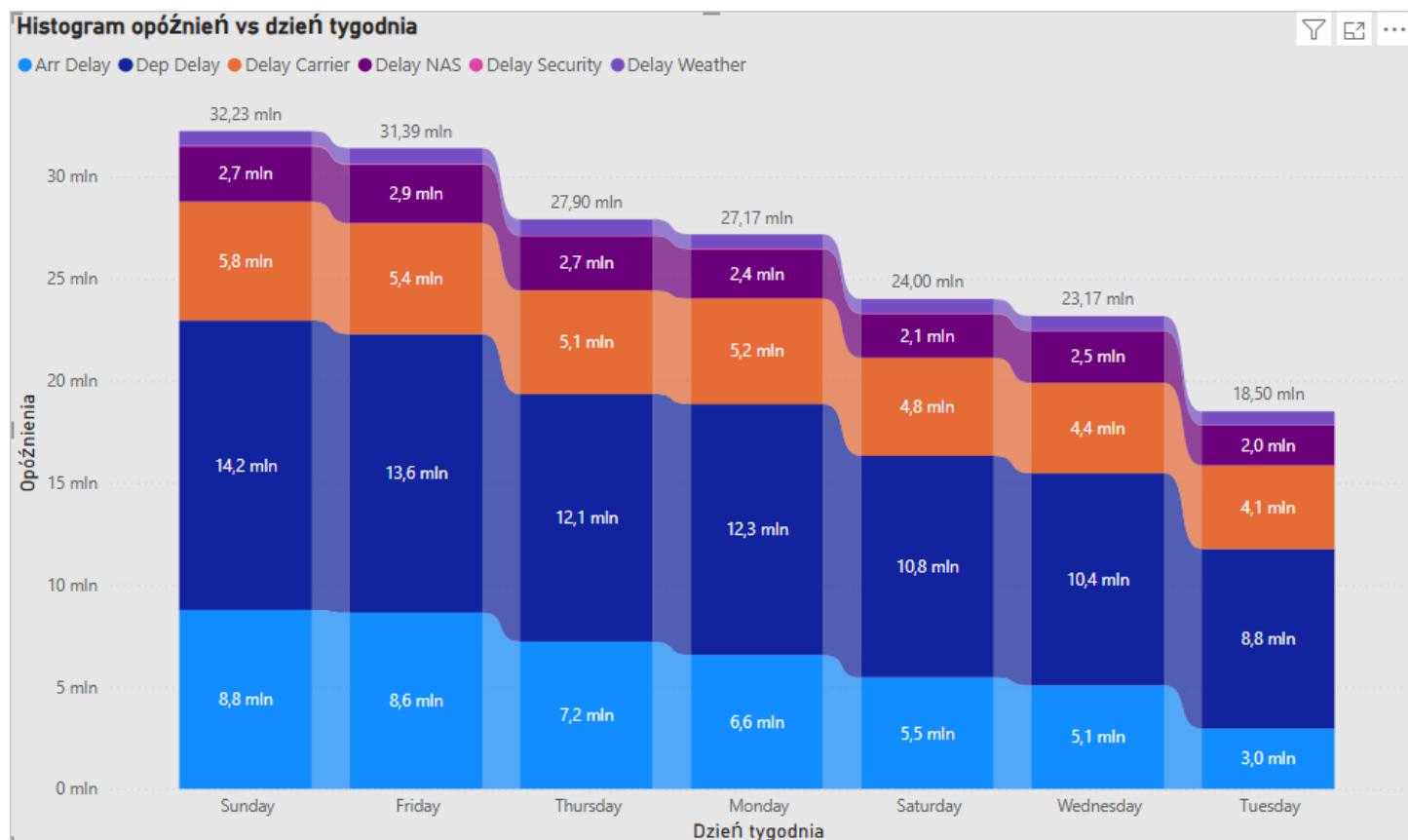
Największy udział w ogólnej liczbie opóźnień mają opóźnienia generowane przez **same linie lotnicze (Delay Carrier)** oraz **systemowe (Delay NAS)**. W przypadku niektórych przewoźników, np. Southwest Airlines, zauważalna jest także **istotna liczba opóźnień pogodowych**.

Mniejsze linie (np. Allegiant, PSA Airlines, Endeavor, Hawaiian) charakteryzują się znacznie niższym łącznym czasem opóźnień, co może być efektem mniejszej liczby wykonywanych lotów lub lepszej organizacji operacyjnej.

Wnioski

- Wysoka liczba opóźnień nie zawsze wynika z warunków pogodowych - **czynniki operacyjne i systemowe odgrywają większą rolę**.
- Linie o największym zasięgu (American, Delta, United) są najbardziej narażone na różnorodne zakłócenia.
- Warto przeprowadzić **dalszą eksplorację danych (data mining)**, aby sprawdzić, które czynniki są najczęstszym źródłem problemów dla konkretnych przewoźników.

8.2. Analiza opóźnień względem dni tygodnia



Powyższy wykres przedstawia histogram sumy opóźnień (w mln minut) w zależności od dnia tygodnia. Dane zostały pogrupowane według sześciu kategorii: opóźnienia w przyjazdach (Arr Delay), odlotach (Dep Delay), opóźnienia przewoźnika (Delay Carrier), opóźnienia systemowe NAS (Delay NAS), związane z bezpieczeństwem (Delay Security) oraz spowodowane warunkami pogodowymi (Delay Weather).

Najwyższe sumy opóźnień występują w **niedzielę (32,23 mln)** i **piątek (31,39 mln)**, co może wynikać z większego natężenia ruchu pasażerskiego na koniec i początek tygodnia. W tych dniach dominuje zarówno liczba opóźnień operacyjnych (Carrier), jak i przyjaztów oraz odlotów.

W kolejnych dniach obserwujemy spadek wartości opóźnień:

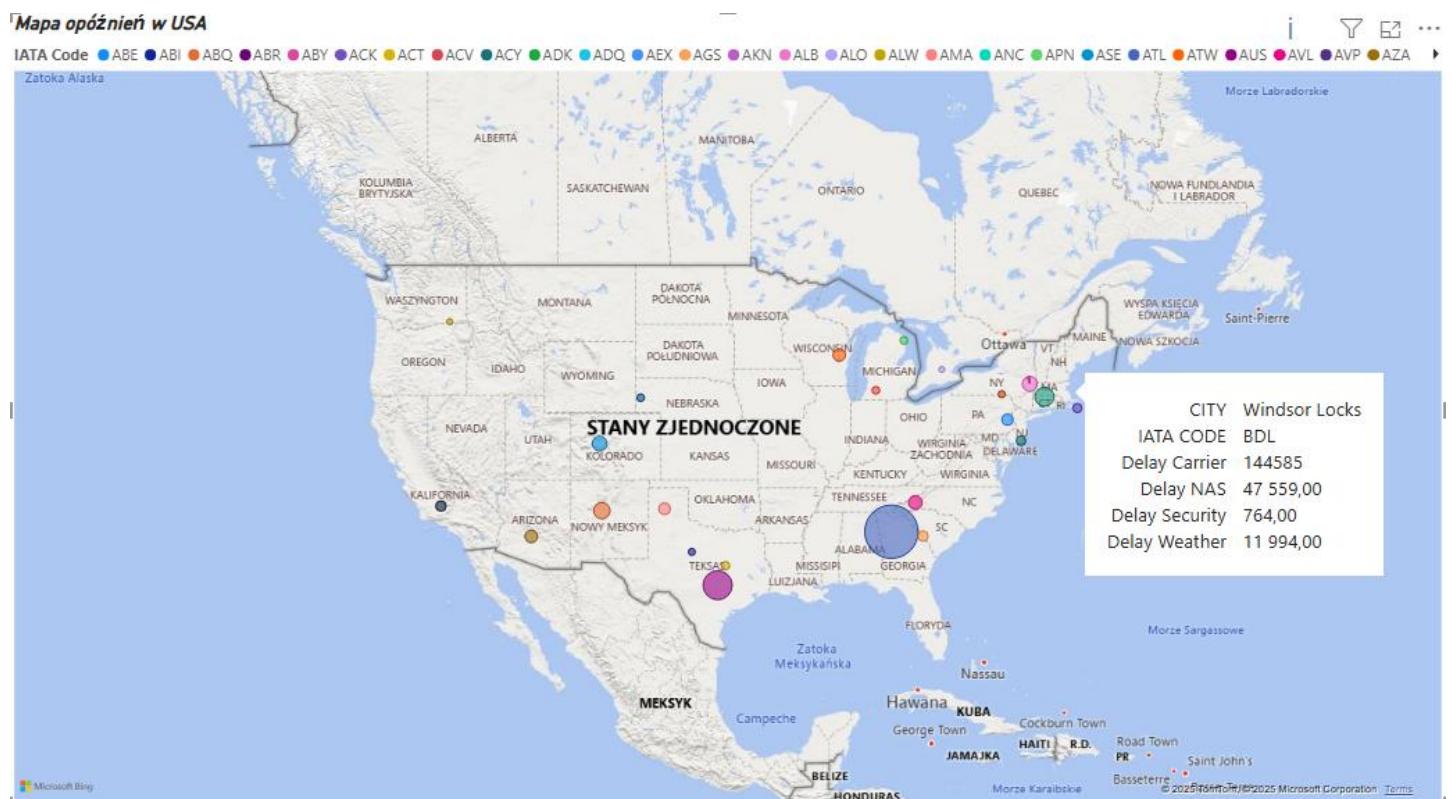
- Czwartek: 27,90 mln
- Poniedziałek: 27,17 mln
- Sobota: 24,00 mln

- Środa: 23,17 mln
- Najniższy poziom występuje we **wtorek (18,50 mln)**

Analiza wskazuje na wyraźną **sezonowość tygodniową opóźnień**, co może być istotną zmienną w dalszych modelach eksploracji danych (np. przy klasyfikacji lub prognozowaniu opóźnień).

Dodatkowo, opóźnienia pogodowe i związane z bezpieczeństwem pozostają stosunkowo niskie, lecz widoczne we wszystkich dniach, co potwierdza ich stały, choć ograniczony wpływ na punktualność operacyjną.

8.3. Mapa opóźnień lotniczych w USA



Powyższa wizualizacja przedstawia **mapę Stanów Zjednoczonych**, na której zaznaczono lotniska według ich **sumarycznego udziału w opóźnieniach lotów**. Rozmiar i kolor okręgu wskazuje **skalę oraz lokalizację opóźnień**, a szczegóły pojawiają się po najechaniu kursorem (tooltip).

Dla każdego punktu (lotniska) przedstawiono:

- **Miasto i kod IATA**
- Liczby opóźnień pogrupowane według kategorii:
 - **Delay Carrier** - opóźnienia po stronie linii lotniczych,
 - **Delay NAS** - opóźnienia systemowe (kontrola ruchu, infrastruktura),
 - **Delay Security** - opóźnienia związane z bezpieczeństwem,
 - **Delay Weather** - opóźnienia pogodowe.

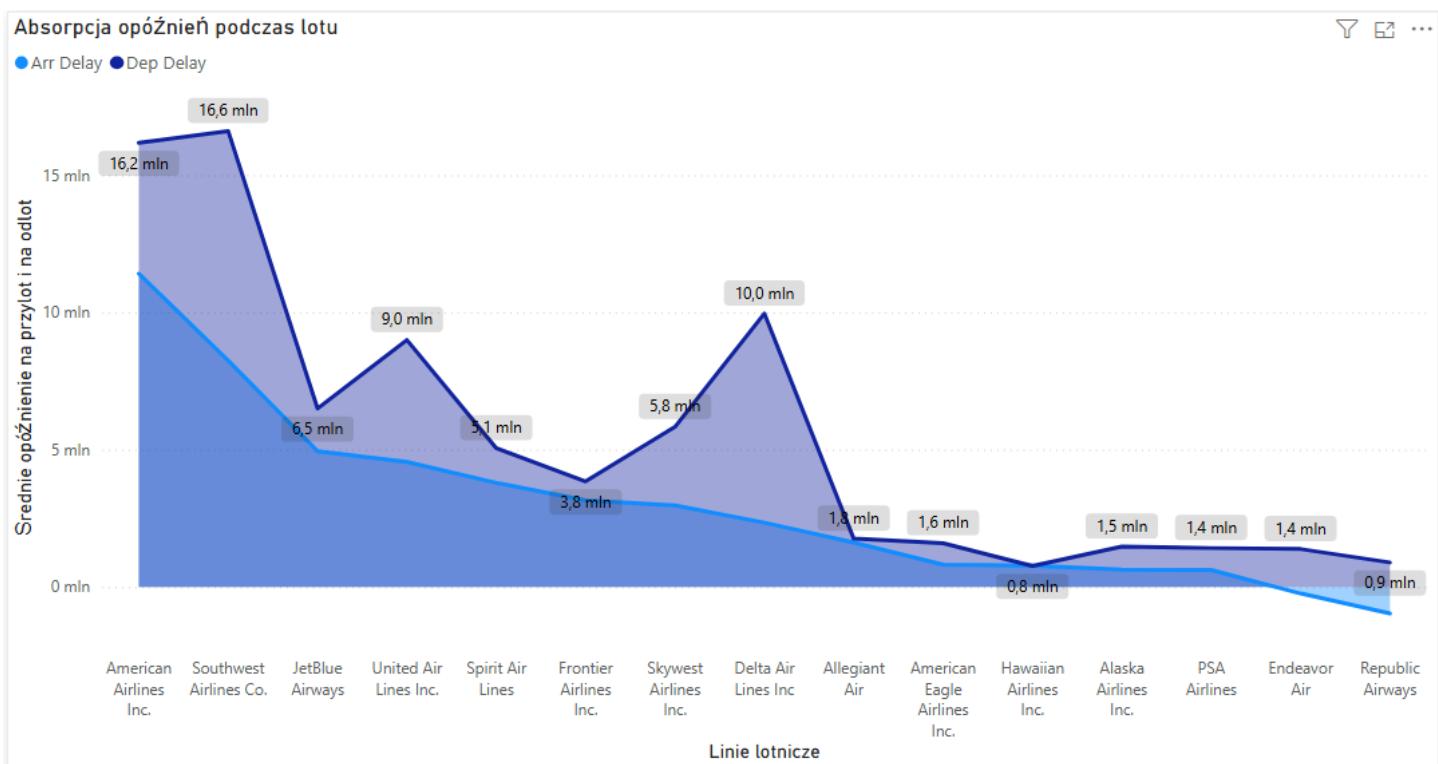
Przykład tooltipa (Windsor Locks - BDL):

- Delay Carrier: 144 585
- Delay NAS: 47 559
- Delay Security: 764
- Delay Weather: 11 994

Wnioski z mapy:

- **Największe skupiska opóźnień** obserwujemy na wschodnim wybrzeżu USA oraz w rejonie południowo-wschodnim (np. Georgia, Karolina Północna).
- Część lotnisk w południowych stanach (np. Teksas, Luizjana) wykazuje znaczną liczbę opóźnień generowanych przez przewoźników.
- **Opóźnienia pogodowe** widoczne są również w północno-wschodnich stanach, co może być związane z trudnymi warunkami zimowymi.

8.4. Absorpcja opóźnień podczas lotu – analiza różnicy między opóźnieniem w odlocie a przylocie



Wizualizacja przedstawia porównanie **średnich opóźnień przy odlocie (Dep Delay)** i **przylocie (Arr Delay)** dla wybranych linii lotniczych w USA. Dzięki takiemu zestawieniu możliwe jest zbadanie tzw. **absorpcji opóźnień** – czyli zdolności linii lotniczych do „nadrobienia” opóźnienia w trakcie lotu.

Wartości zostały wyrażone w milionach minut i obejmują średnie czasy opóźnień skumulowane dla każdego przewoźnika.

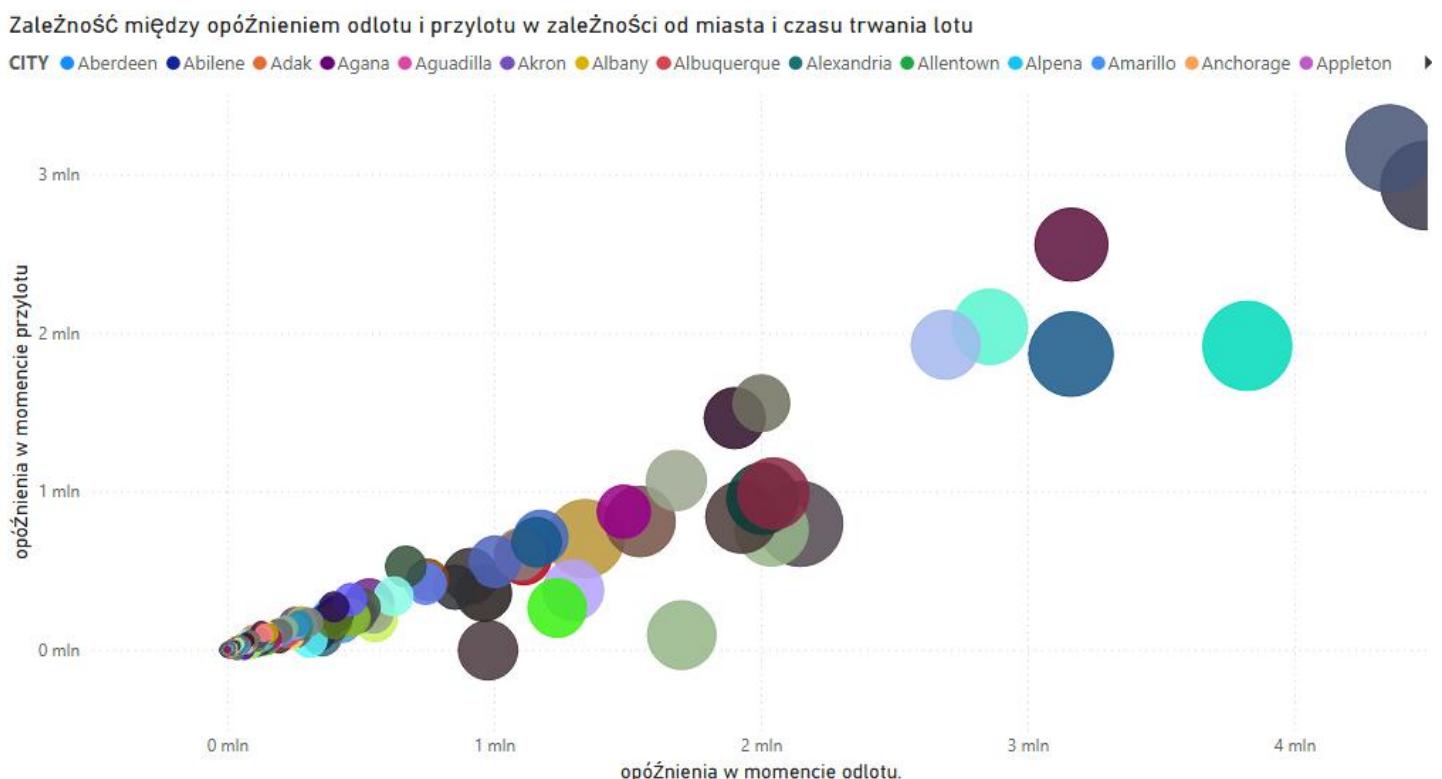
Główne obserwacje:

- **American Airlines Inc.** oraz **Southwest Airlines Co.** wykazują największe średnie opóźnienia zarówno przy odlocie, jak i przylocie (odpowiednio 16,2 mln / 16,6 mln).
- **JetBlue Airways, Spirit Air Lines** i **Skywest Airlines Inc.** cechuje większa zdolność do redukcji opóźnień w czasie lotu – różnica między Dep Delay a Arr Delay jest zauważalna.
- **Delta Air Lines Inc.** ma wysokie średnie opóźnienie przylotów (10 mln), ale dużo niższe opóźnienie odlotów (1,3 mln), co może wskazywać na kumulację problemów na trasach dolotowych.
- Mniejsze linie regionalne (np. Hawaiian, Republic, Endeavor) charakteryzują się bardzo niskimi wartościami opóźnień, co może być wynikiem mniejszego ruchu lub krótszych tras.

Wnioski:

- **Absorpcja opóźnień to kluczowy wskaźnik efektywności operacyjnej przewoźników.** Linie, które potrafią nadrobić czas w trakcie lotu, zmniejszają wpływ opóźnień na dalsze połączenia.
- Zależność między odlotem a przylotem może sugerować problemy z rotacją samolotów, planowaniem tras lub punktualnością na lotniskach docelowych.
- W kontekście eksploracji danych (data mining), warto byloby sprawdzić, **które czynniki zwiększą lub zmniejszą zdolność do nadrobienia opóźnienia** – np. długość trasy, model samolotu, pora dnia.

8.5. Zależność między opóźnieniem odlotu i przylotu w zależności od miasta i czasu trwania lotu



Powyższy wykres punktowy (bubble chart) przedstawia **zależność między opóźnieniami przy odlocie a opóźnieniami przy przylocie**, z uwzględnieniem miast oraz (najprawdopodobniej) czasu trwania lotu jako zmiennej wpływającej na wielkość bąbla.

Każdy punkt reprezentuje **jedno miasto (lotnisko)**, a jego:

- **położenie poziome (oś X)** odpowiada sumie opóźnień przy odlocie (Dep Delay),
- **położenie pionowe (oś Y)** – sumie opóźnień przy przylocie (Arr Delay),
- **rozmiar bąbla** może oznaczać długość trwania lotów z danego miasta lub liczbę lotów (dokładna interpretacja zależna od ustawienia w Power BI),
- **kolor** odróżnia poszczególne miasta (legendę zawiera górna część wykresu).

Główne wnioski z wizualizacji:

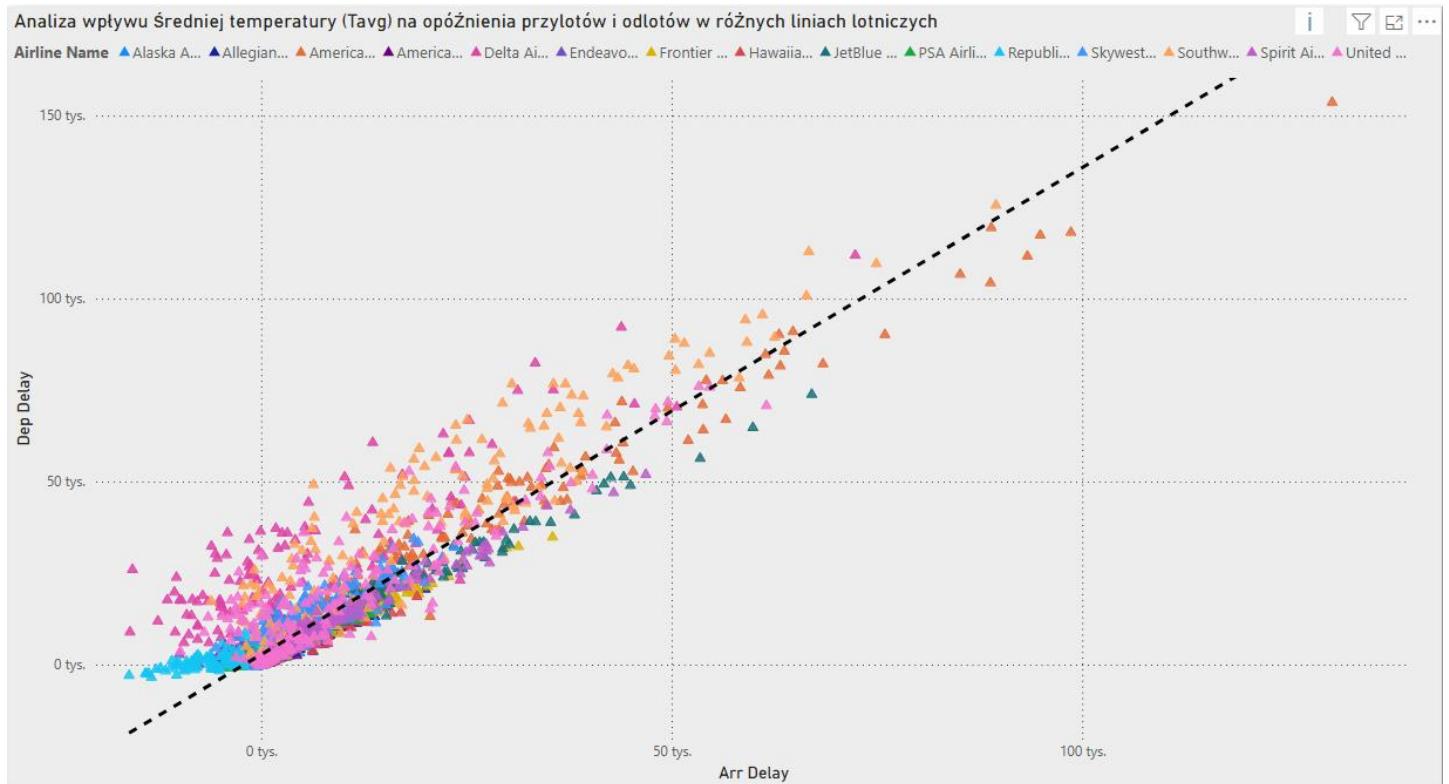
- Wykres potwierdza istnienie **korelacji między opóźnieniem przy odlocie a przylocie** – im większe opóźnienie przy startie, tym większe przy lądowaniu.
- Część punktów znajduje się **znacznie powyżej linii 1:1**, co oznacza, że niektóre miasta nie są w stanie „nadrabiać” opóźnień w trakcie lotu, a wręcz je powiększają.
- **Największe bąble** w prawym górnym rogu sugerują, że **długie lub popularne trasy mają tendencję do kumulowania dużych opóźnień**, zarówno przy odlocie, jak i przylocie.
- Miasta w lewym dolnym rogu wykresu (małe bąble) charakteryzują się relatywnie **niskim poziomem opóźnień operacyjnych**.

Znaczenie analityczne:

Tego typu analiza pozwala zidentyfikować miasta, które:

- mają największe problemy z punktualnością (duży bąbel, wysoko i daleko od osi),
- są potencjalnie bardziej przewidywalne operacyjnie (niskie opóźnienia, mniejsze bąble),
- wymagają dalszej eksploracji przy wykorzystaniu technik **data mining**, np. drzewa decyzyjne do przewidywania opóźnień w zależności od miasta, długości lotu, czy przewoźnika.

8.6. Analiza wpływu średniej temperatury (Tavg) na opóźnienia przylotów i odlotów w różnych liniach lotniczych



Powyzsza wizualizacja przedstawia **zależność między opóźnieniami w przylotach (Arr Delay) a opóźnieniami w odlotach (Dep Delay)** dla różnych linii lotniczych, z uwzględnieniem wpływu **średniej temperatury (Tavg)** jako tła analitycznego. Każdy punkt reprezentuje dane zagregowane dla konkretnego przewoźnika w określonych warunkach temperaturowych.

- Oś pozioma (X) – suma opóźnień w przylotach
- Oś pionowa (Y) – suma opóźnień w odlotach
- Kolory punktów (trójkąty) – odpowiadają różnym liniom lotniczym
- Przerwana linia – **linia trendu regresji liniowej**, wskazująca zależność między opóźnieniem przylotu a odlotu

Wnioski z wykresu:

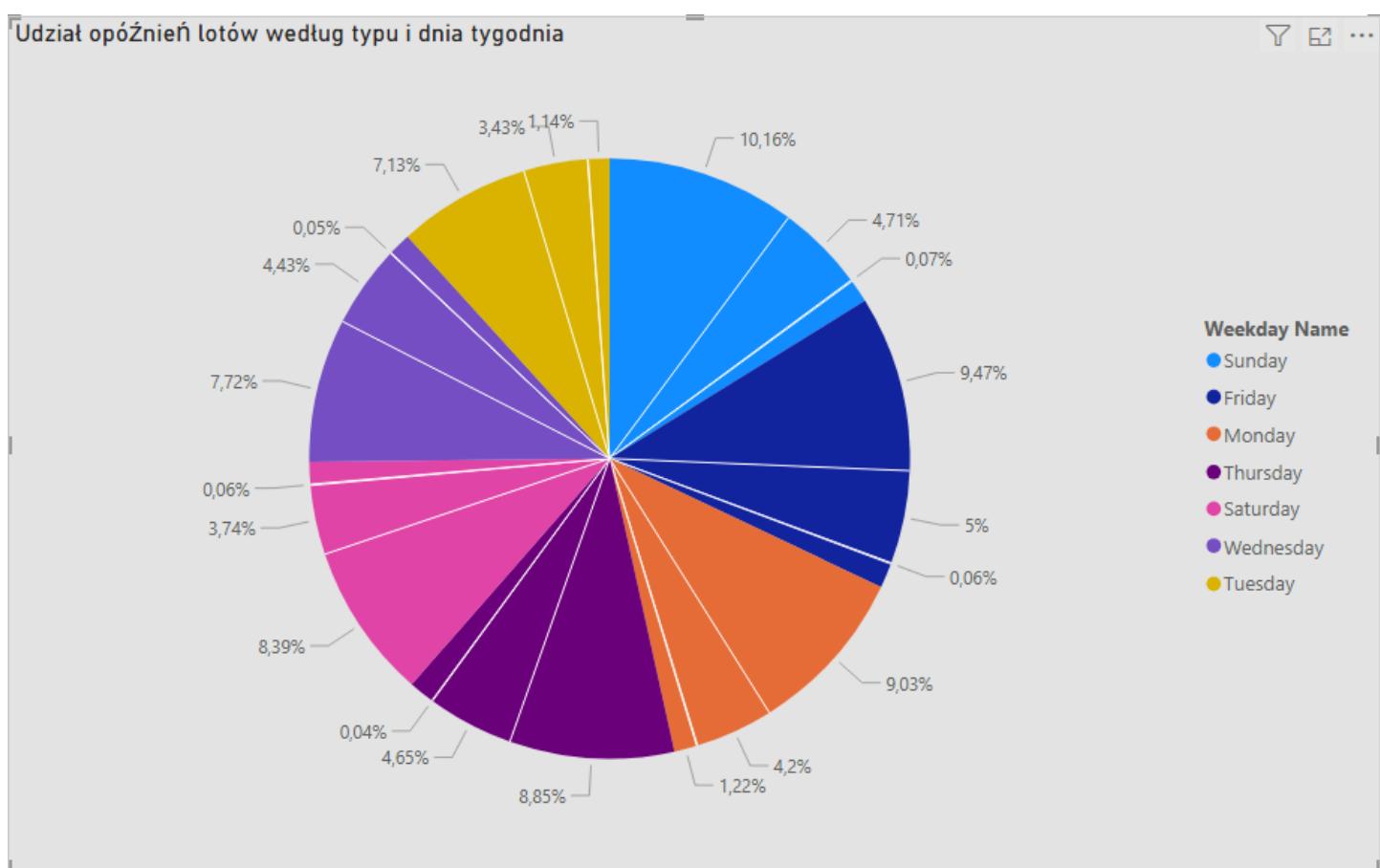
- Widoczna jest **silna korelacja dodatnia** między opóźnieniami przy przylocie a opóźnieniami przy odlocie. Punkty są skoncentrowane wzdłuż linii trendu.
- Linie lotnicze, których punkty **znajdują się powyżej linii trendu**, mają tendencję do **zwiększania opóźnienia podczas lotu** (opóźnienie na przylocie większe niż na odlocie).
- Punkty **poniżej linii trendu** mogą wskazywać na lepszą zdolność absorpcji opóźnień w trakcie lotu.
- **Kolorystyka danych** pozwala porównać zachowanie różnych linii lotniczych – np. Southwest i Spirit mają więcej punktów w obszarach większych opóźnień.

- Choć tytuł wykresu odnosi się do wpływu temperatury (Tavg), sama temperatura nie jest jawnie przedstawiona – można jednak założyć, że analizowane dane zostały zgrupowane według poziomu temperatury i użyte jako tło do wykrycia trendu.

Znaczenie dla analizy:

- Tego typu analiza wskazuje, że **opóźnienia mają tendencję do kumulowania się**, co może mieć wpływ na późniejsze rotacje samolotów i ogólną punktualność przewoźników.
- Istnieje potencjał do wykorzystania tych danych w modelach predykcyjnych – np. **przewidywanie opóźnienia odlotu na podstawie przylotu i warunków pogodowych**, z użyciem regresji lub drzew decyzyjnych w data miningu.

8.7. Udział opóźnień lotów według typu i dnia tygodnia



Powyższy wykres kołowy przedstawia **procentowy udział opóźnień lotów**, pogrupowany według **dnia tygodnia** oraz niejawnie – **typu opóźnienia** (domyślnie: różne przyczyny jak pogoda, przewoźnik, kontrola ruchu itd.).

Każdy wycinek koła reprezentuje kombinację typu opóźnienia oraz dnia tygodnia, a jego wartość procentowa wskazuje udział tej grupy w całości zarejestrowanych opóźnień.

Obserwacje:

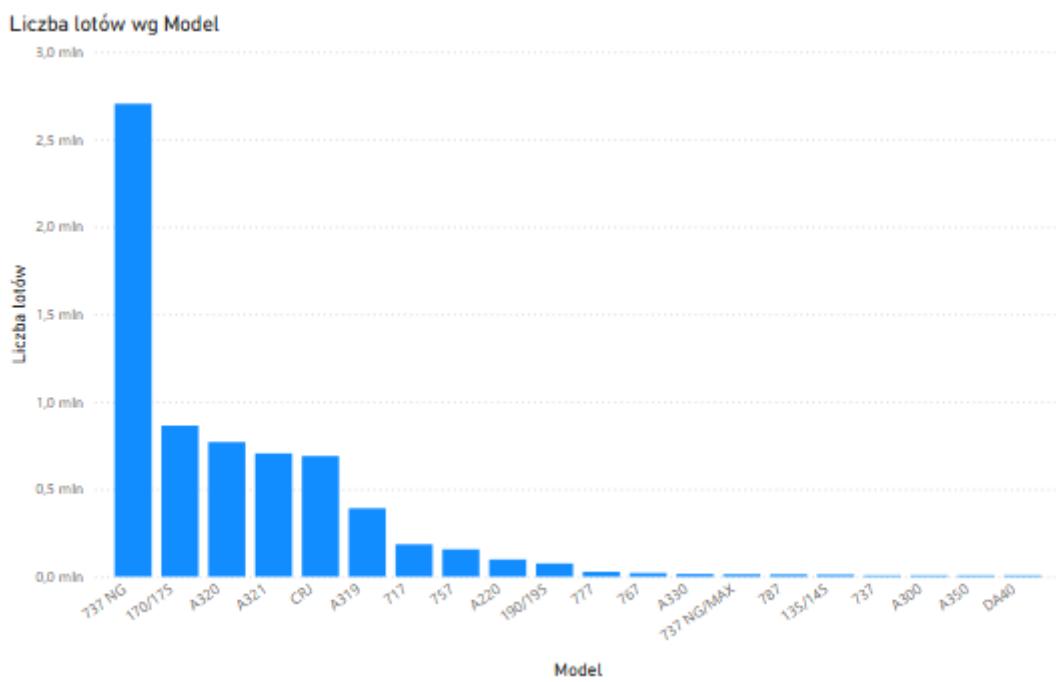
- Największy pojedynczy udział opóźnień przypada na:

- **Niedzielę (Sunday) – 10,16%,**
- **Piątek (Friday) – 9,47%,**
- **Czwartek (Thursday) – 8,85%,**
- **Sobota (Saturday) – 8,39%,**
- **Środa (Wednesday) – 7,72%.**
- Najmniejszy udział notują wtorki i soboty w niektórych typach (od 0,04% do 0,07%), co może sugerować:
 - **niższy ruch lotniczy w te dni,**
 - **mniej złożone operacje,**
 - **lepszą punktualność.**

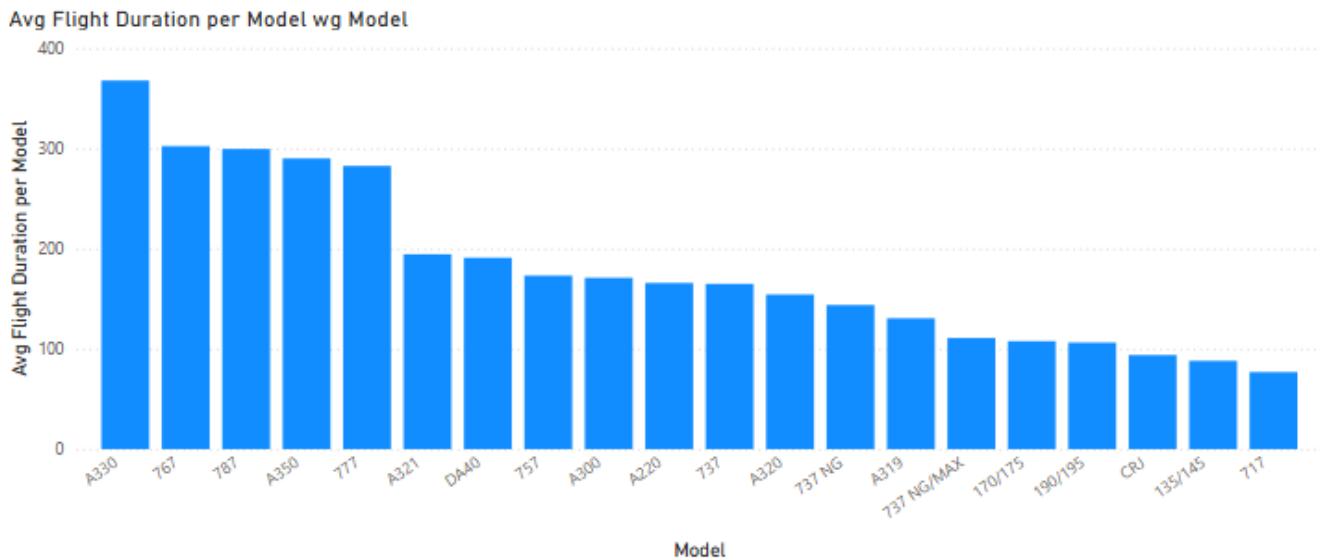
Wnioski:

- Dane potwierdzają wcześniejsze obserwacje z wykresów słupkowych: **największe zagęszczenie opóźnień przypada na dni o największym ruchu pasażerskim – piątek i niedziele.**
- Opóźnienia rozkładają się **nierównomiernie** w zależności od dnia tygodnia, co może być podstawą do dalszej analizy predykcyjnej.
- W kontekście **eksploracji danych (data mining)**, takie informacje mogą zostać użyte do budowy modeli przewidujących, **które dni tygodnia są najbardziej ryzykowne z punktu widzenia punktualności operacyjnej.**

8.8. Liczba lotów na model



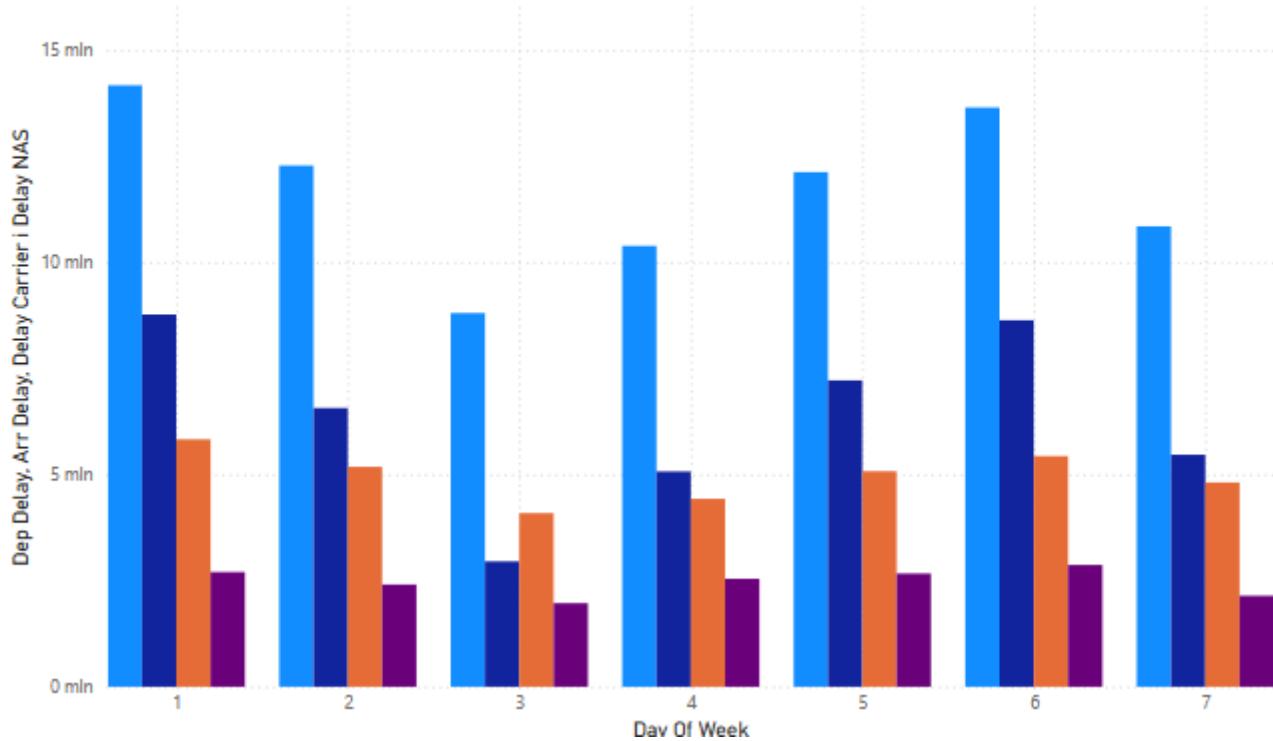
8.9. Średni czas trwania lotu na model



8.10. Typy opóźnień a dzień tygodnia

Dep Delay, Arr Delay, Delay Carrier i Delay NAS wg Day Of Week

● Dep Delay ● Arr Delay ● Delay Carrier ● Delay NAS

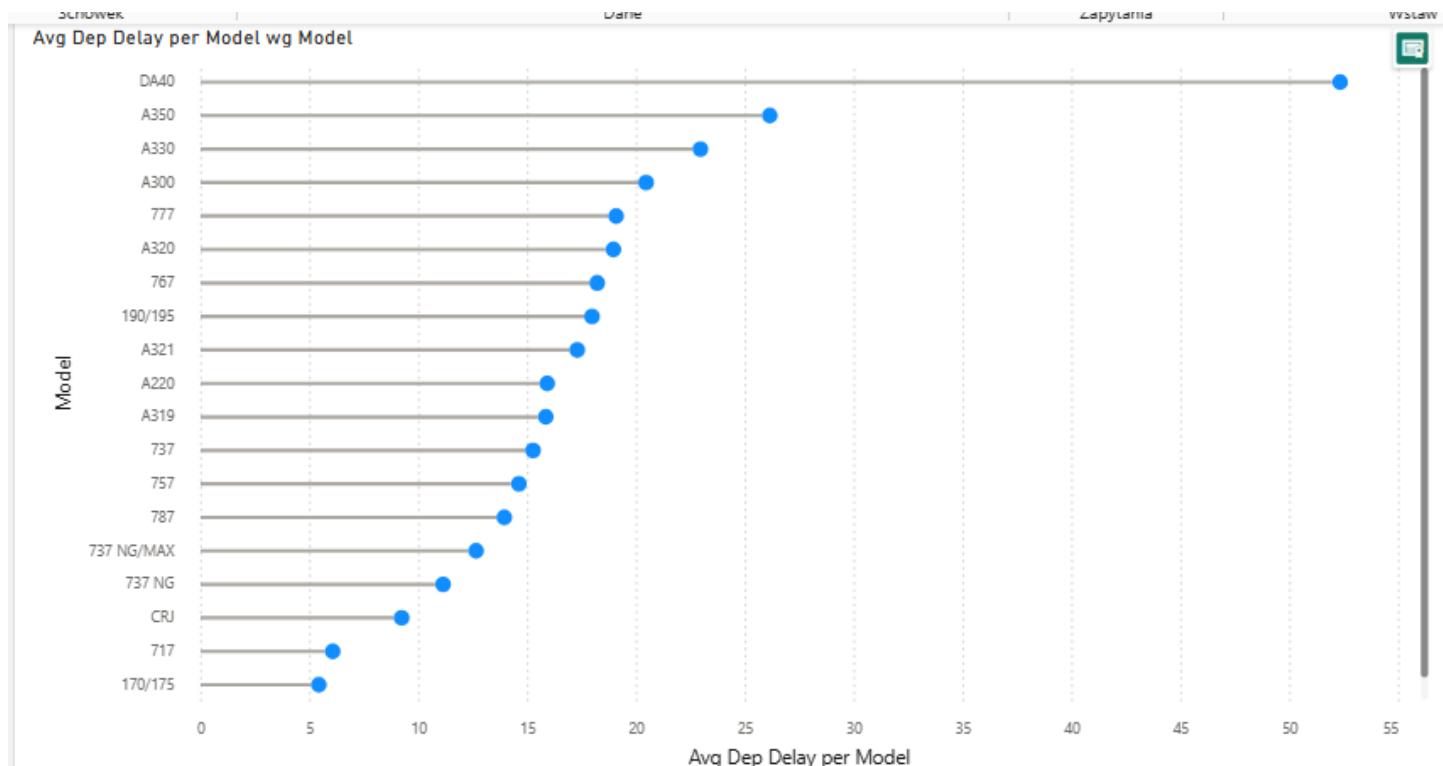


8.11. Średnie opóźnienie odlotów dla każdej linii lotniczej względem dnia tygodnia

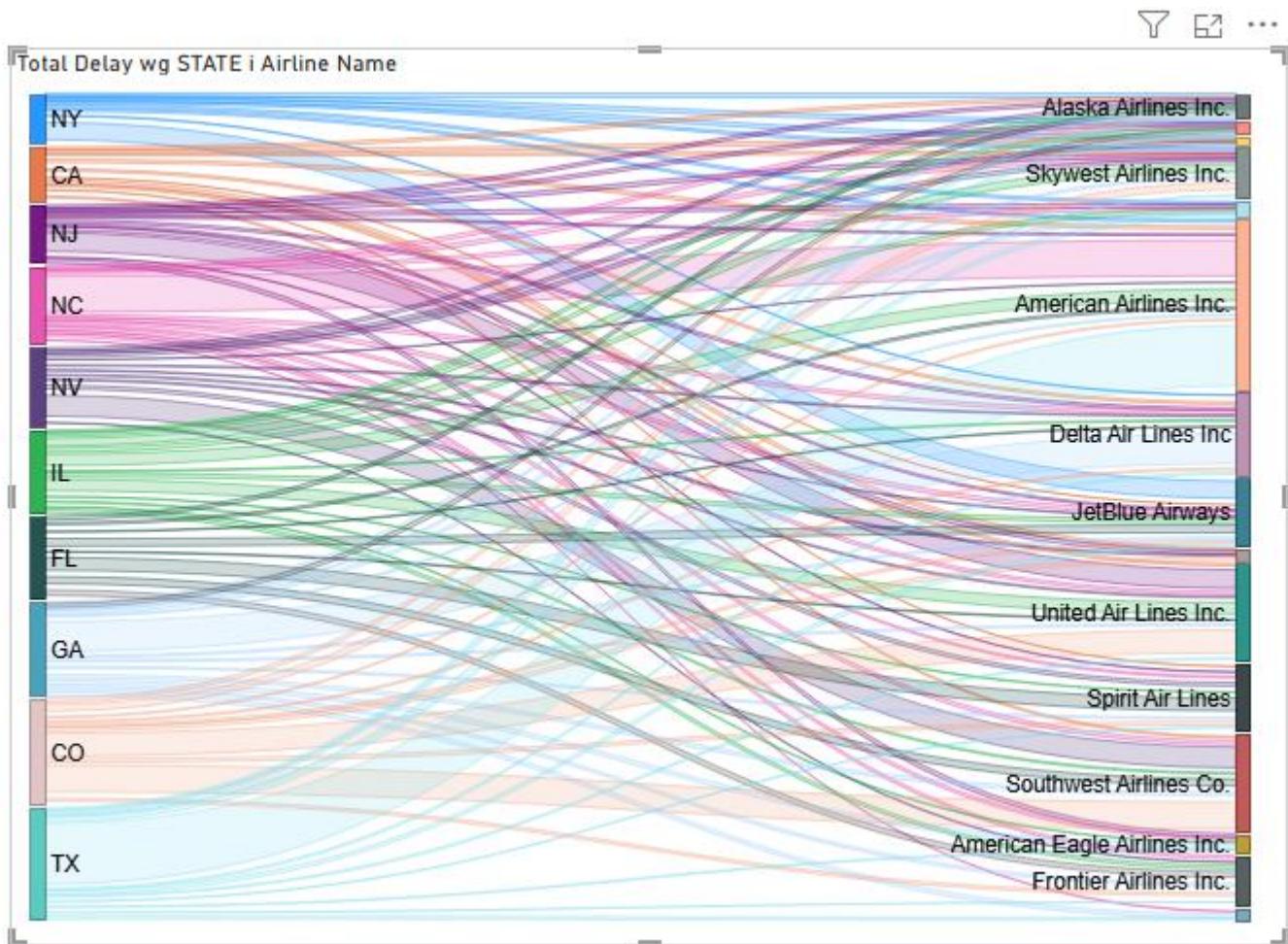
Średnie opóźnienie odlotów dla każdej linii lotniczej względem dni tygodnia [min]

Nazwa Linii Lotniczej	1	2	3	4	5	6	7	Suma
JetBlue Airways	29,34	27,36	19,05	18,81	21,23	26,58	27,55	24,27
Frontier Airlines Inc.	25,44	25,17	19,54	20,24	21,67	22,29	20,08	22,16
Spirit Air Lines	23,33	21,23	14,98	14,60	20,72	21,47	19,84	19,54
American Airlines Inc.	19,42	17,25	13,84	15,76	18,03	19,61	17,86	17,44
Allegiant Air	18,84	16,19	13,29	9,23	14,09	16,16	13,41	15,39
United Air Lines Inc.	13,91	12,38	10,10	13,20	12,64	12,62	12,64	12,50
Southwest Airlines Co.	13,82	10,67	9,27	11,63	12,11	12,99	11,24	11,69
Delta Air Lines Inc	12,55	11,17	7,22	8,66	10,05	11,69	10,12	10,24
Hawaiian Airlines Inc.	8,34	9,19	7,50	9,04	9,78	12,33	10,84	9,59
Skywest Airlines Inc.	10,64	8,61	6,96	8,04	8,73	9,43	8,82	8,77
PSA Airlines	9,56	7,40	5,16	5,60	6,37	8,70	9,21	7,41
American Eagle Airlines Inc.	8,46	5,76	4,23	7,51	8,45	8,65	6,32	7,10
Endeavor Air	8,24	7,22	5,20	5,66	6,11	9,70	7,14	7,02
Alaska Airlines Inc.	7,06	6,02	5,25	4,04	6,19	7,72	6,09	6,07
Republic Airways	4,61	3,63	0,90	1,73	1,45	5,36	4,66	3,11
Suma	14,39	12,32	9,39	10,92	12,15	13,60	12,46	12,20

8.12. Średnie opóźnienie na model



8.13. Flow opóźnień według lotnisk i linii lotniczych

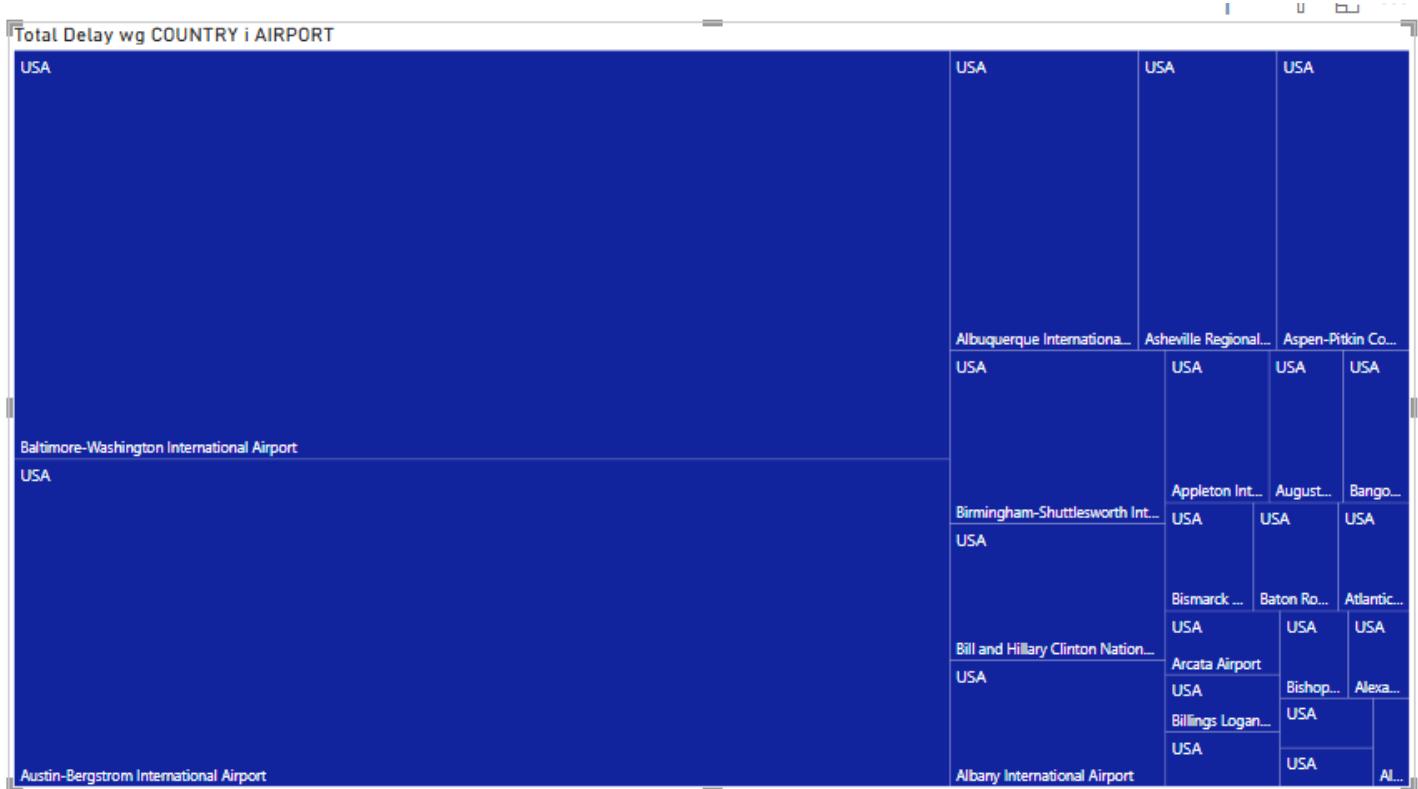


Rozkład opóźnień: Niektóre stany, jak NY, CA i TX, mają wiele połączeń z różnymi liniami lotniczymi, co sugeruje, że są one głównymi hubami lotniczymi z częstymi opóźnieniami.

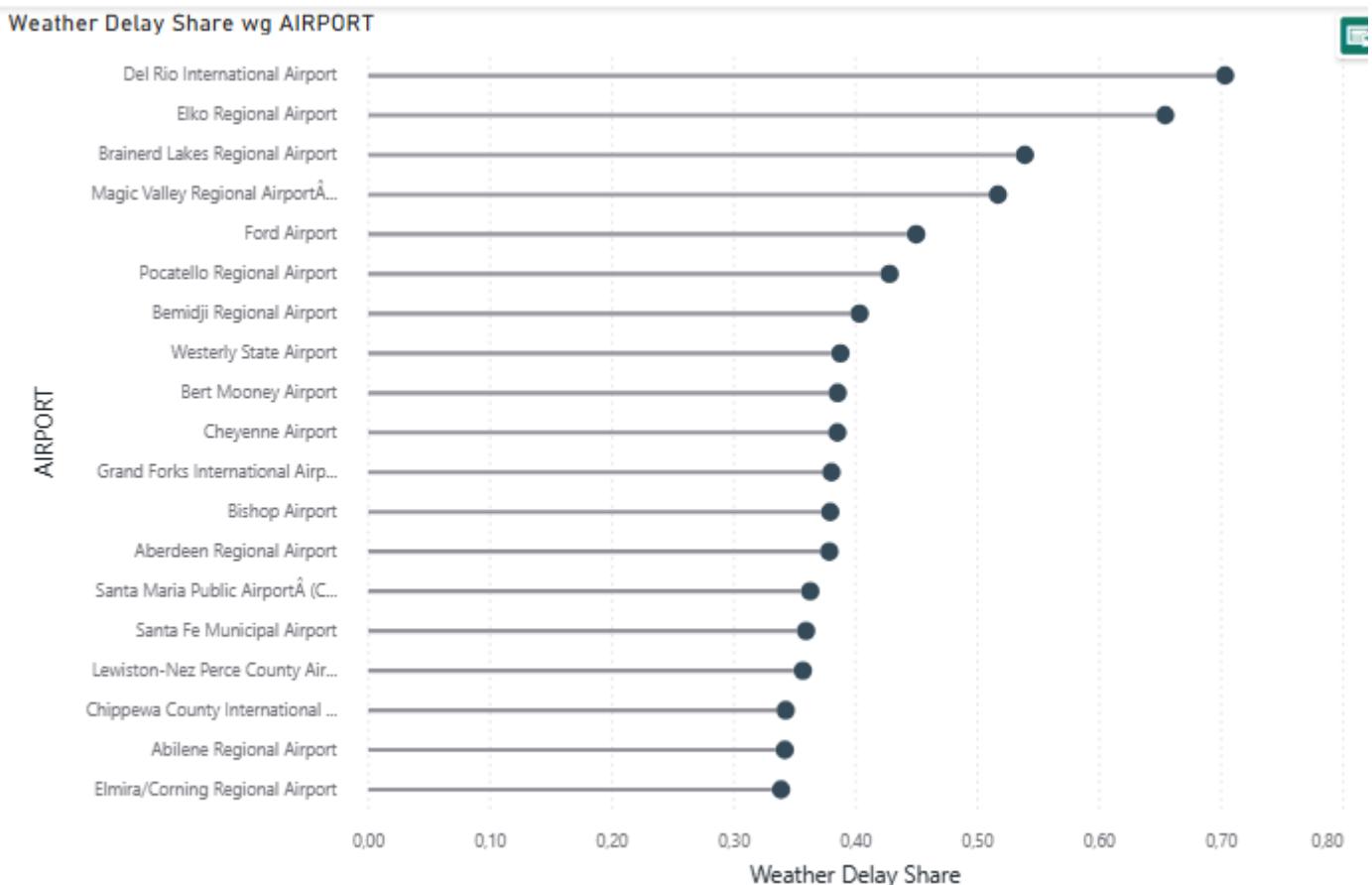
Popularne linie lotnicze: Linie takie jak American Airlines, Delta Air Lines i United Air Lines pojawiają się w wielu połączeniach, co może wskazywać na ich dużą obecność i potencjalnie wyższe ryzyko opóźnień.

Regionalne wzorce: Niektóre stany, jak NV (Nevada) czy IL (Illinois), mają mniej połączeń, co może sugerować mniejszą liczbę lotów lub mniej skomplikowaną sieć opóźnień.

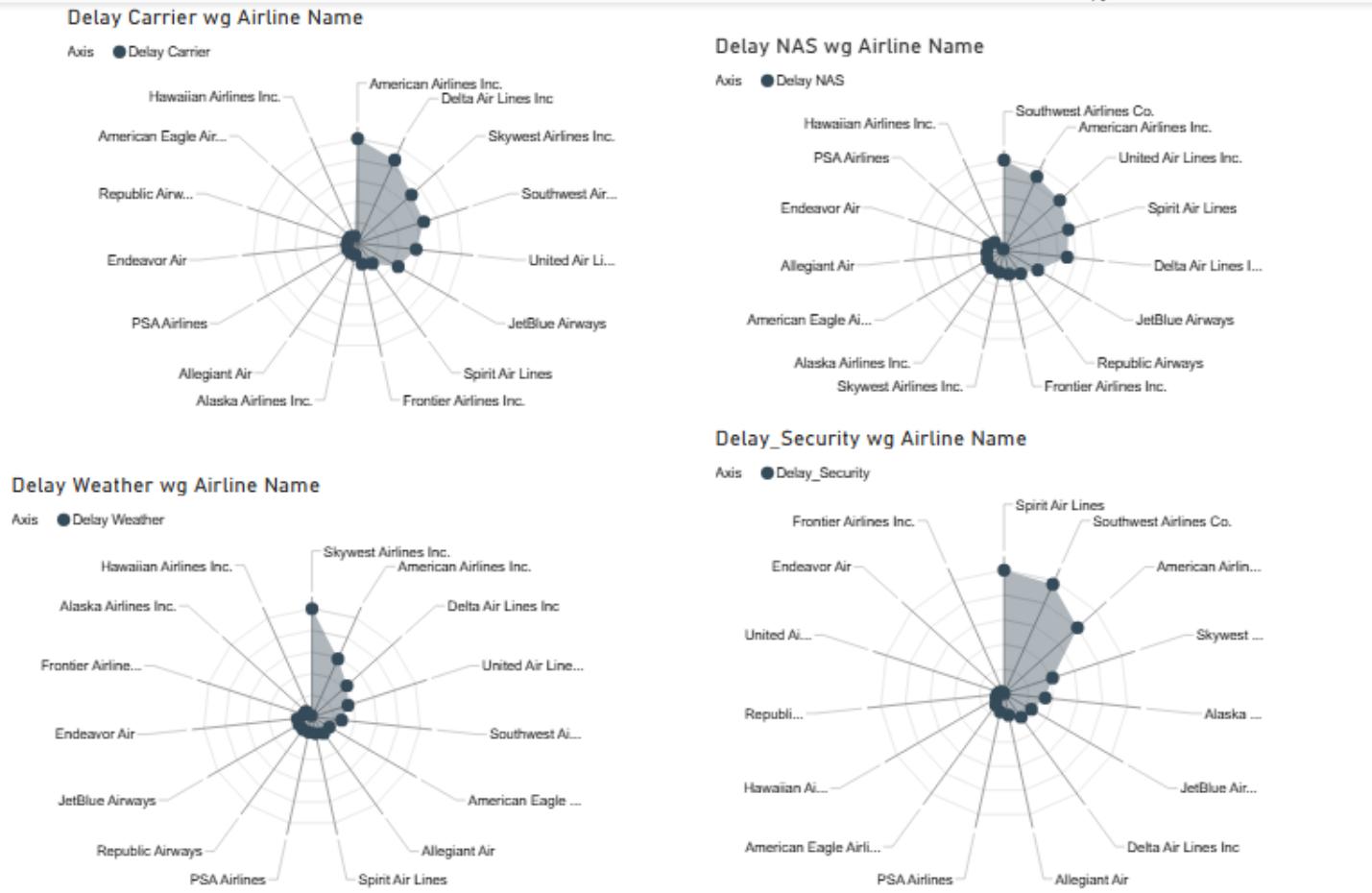
8.14. Wykres drzewa (Tree Map) — Struktura opóźnień wg krajów i lotników



8.15. Udział warunków pogodowych a opóźnienia



- Co pokazuje wykres słupkowy z Weather Delay Share dla lotnisk?
- To jest procentowy udział opóźnień spowodowanych pogodą (w stosunku do wszystkich opóźnień) na danym lotnisku.
 - Jeśli na osi Y masz 0.00 do 0.70, to znaczy, że dla niektórych lotnisk nawet do 70% opóźnień wynika z przyczyn pogodowych!
 - Dla innych może to być 10%, czyli opóźnienia są bardziej „techniczne” lub z innych powodów.



1. Wykresy radarowe przedstawiają różne rodzaje opóźnień lotów (Carrier, NAS, Weather, Security) według linii lotniczych. Oto wnioski:
2. **Delay Carrier:** American Airlines, Delta Air Lines i United Air Lines mają największe opóźnienia związane z przewoźnikiem, co może wskazywać na problemy operacyjne w tych liniach.
3. **Delay NAS:** Southwest Airlines i American Airlines przodują w opóźnieniach NAS (system lotniczy), co może być związane z problemami w zarządzaniu ruchem lotniczym na trasach, które obsługują.

4. **Delay Weather:** Skywest Airlines i Delta Air Lines mają znaczące opóźnienia pogodowe, co może sugerować, że operują w regionach bardziej narażonych na złe warunki atmosferyczne.
5. **Delay Security:** Opóźnienia związane z bezpieczeństwem są stosunkowo małe dla wszystkich linii, ale Southwest i American Airlines wciąż mają nieco większe wartości.
6. **Ogólny wniosek:** American Airlines i Southwest Airlines mają największe problemy z opóźnieniami w różnych kategoriach, co może wskazywać na szersze problemy operacyjne lub wybór tras podatnych na zakłócenia. Linie takie jak Frontier Airlines czy Allegiant Air mają mniejsze opóźnienia we wszystkich kategoriach.