Міністерство освіти і науки України

Національний університет “Львівська політехніка”

Інститут прикладної математики та фундаментальних наук

Кафедра прикладної математики

**Курсова робота**

з курсу “Робота з великими базами даних”

на тему:

“Аеропорт”

Варіант 9

**Виконав:**

студент групи ПМ-42

Петик Володимир

**Прийняв:**

Любінський Богдан Богданович

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(дата) (підпис викладача)

Львів — 2024

**Анотація**

Курсова робота присвячена розробці та аналізу інформаційної системи для предметної області "Аеропорт". Основною метою роботи є побудова бази даних і сховища даних, які дозволяють ефективно зберігати, обробляти та аналізувати інформацію про літаки, пілотів, стюардес та вильоти.

Робота охоплює такі етапи:

1. Розробка концептуальної моделі бази даних, яка описує ключові об’єкти предметної області та їх взаємозв’язки.
2. Генерація реалістичних даних для тестування системи.
3. Реалізація процесу ETL (витягування, трансформація, завантаження) з використанням SQL Server Integration Services (SSIS).
4. Побудова багатовимірного куба для аналітики за допомогою SQL Server Analysis Services (SSAS).
5. Розробка інтерактивних звітів із використанням SQL Server Reporting Services (SSRS).

Результати роботи демонструють можливість аналізу даних щодо розподілу пілотів за доступними літаками, що сприяє підтримці прийняття управлінських рішень у сфері авіації.

**Зміст**

Анотація

1. Теоретична частина  
   1.1. Основи багатовимірного аналізу даних  
   1.2. Багатовимірний підхід у зберіганні даних  
   1.3. Інструменти для багатовимірного аналізу  
   1.4. Операції з багатовимірними даними  
   1.5. Практичне значення багатовимірного аналізу
2. Практична частина  
   2.1. Структура бази даних і сховища  
   2.2. Процес ETL за допомогою SQL Server Integration Services (SSIS)  
   2.3. Побудова багатовимірного куба за допомогою SQL Server Analysis Services (SSAS)  
   2.4. Формування аналітичних звітів за допомогою SQL Server Reporting Services (SSRS)  
   - Звіт №1: Продаж в середньому білетів на один літак   
   - Звіт №2: Скільки в середньому годин проводить пілот у повітрі за тиждень

Висновки

Список використаної літератури

### ****Основна частина****

### ****1. Основи багатовимірного аналізу даних****

Багатовимірний аналіз даних — це процес, який дозволяє досліджувати інформацію з кількох ракурсів одночасно, використовуючи багатовимірні структури. Одним із ключових підходів для реалізації такого аналізу є використання технології OLAP (Online Analytical Processing). Вона дає змогу працювати з великими обсягами даних, виконувати складні запити та отримувати результати, які підтримують процес прийняття рішень.

OLAP відрізняється від традиційних баз даних тим, що дані організовані у вигляді багатовимірних структур (кубів), які забезпечують швидкий і зручний доступ до необхідної інформації. Ця технологія знайшла застосування в різних галузях: управлінні бізнесом, аналізі продажів, охороні здоров’я, фінансах тощо.

**Основні переваги OLAP:**

* **Аналіз великих обсягів даних.** Технологія дозволяє працювати з даними, що охоплюють багато аспектів одночасно.
* **Швидкодія.** Попередньо обчислені агрегати забезпечують високу швидкість виконання запитів.
* **Інтерактивність.** Дані організовані так, щоб користувачі могли швидко змінювати параметри аналізу.

### ****2. Багатовимірний підхід у зберіганні даних****

Багатовимірні структури є основою OLAP-систем і базуються на концепції багатовимірного куба. Куб є моделлю даних, що дозволяє структурувати інформацію у вигляді кількох вимірів.

**Елементи багатовимірного куба:**

* **Міри.** Це числові показники, які відображають ключові результати аналізу, наприклад, суми, кількості, середні значення.
* **Виміри.** Категорії, за якими групуються дані, наприклад, час, регіони або продукти. Вони створюють контекст для аналізу мір.
* **Ієрархії.** Багаторівневі структури у вимірах, які дозволяють деталізувати або агрегувати дані, забезпечуючи гнучкість аналізу.

Куб дозволяє одночасно аналізувати дані через декілька вимірів, що робить його потужним інструментом для дослідження складних зв’язків і закономірностей у даних.

### ****3. Інструменти для багатовимірного аналізу****

Для реалізації багатовимірного аналізу даних використовуються спеціалізовані програмні засоби. Одним із найпоширеніших є SQL Server Analysis Services (SSAS). Ця платформа дозволяє створювати багатовимірні моделі даних, які можуть використовуватися для глибокого аналізу.

**Процес створення куба включає кілька етапів:**

1. **Підключення до джерела даних.** Створюється з'єднання з базою даних, з якої буде завантажуватися інформація.
2. **Формування моделі даних.** Вибираються таблиці, визначаються зв’язки між ними, налаштовуються виміри та міри.
3. **Налаштування ієрархій.** Встановлюються багаторівневі структури для вимірів, що дозволяють деталізувати або агрегувати дані.
4. **Розгортання куба.** Куб публікується на сервері, після чого стає доступним для запитів і аналітичної роботи.

### ****4. Операції з багатовимірними даними****

Багатовимірні моделі підтримують широкий спектр операцій, які дозволяють гнучко аналізувати дані:

* **Slice (зріз):** Фокусування на одному конкретному рівні даних, наприклад, аналіз лише певного періоду.
* **Dice (вибірка):** Формування підмножини даних за декількома умовами, що дозволяє виділити необхідний набір інформації.
* **Drill-down і Roll-up:** Перехід між рівнями деталізації. Drill-down дозволяє деталізувати дані, а Roll-up — агрегувати їх на вищих рівнях.
* **Pivot (обертання):** Зміна структури представлення даних шляхом перестановки вимірів для аналізу з іншого ракурсу.

Ці операції є невід’ємною частиною роботи з OLAP-системами і забезпечують інструментарій для динамічного аналізу.

### ****5. Практичне значення багатовимірного аналізу****

Багатовимірний аналіз є основою для прийняття рішень у багатьох галузях. Він дозволяє досліджувати дані з різних аспектів, знаходити приховані закономірності та прогнозувати майбутні результати. Завдяки використанню багатовимірних моделей організації можуть:

* Оцінювати ефективність своєї діяльності.
* Оптимізувати процеси та ресурси.
* Формувати детальні аналітичні звіти, які сприяють стратегічному плануванню.

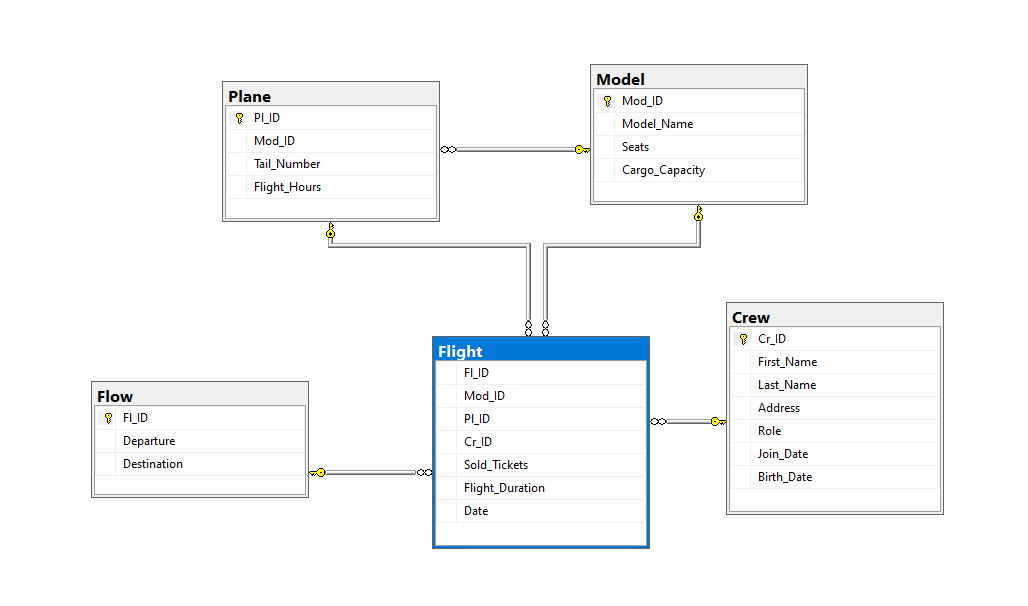
**Практична частина**

У рамках роботи було здійснено побудову та аналіз даних із використанням інструментів SQL Server, таких як SSIS, SSAS та SSRS. Представлено ключові етапи, включаючи створення бази даних, розробку багатовимірного куба для аналізу даних, а також підготовку аналітичних звітів на основі отриманих результатів.

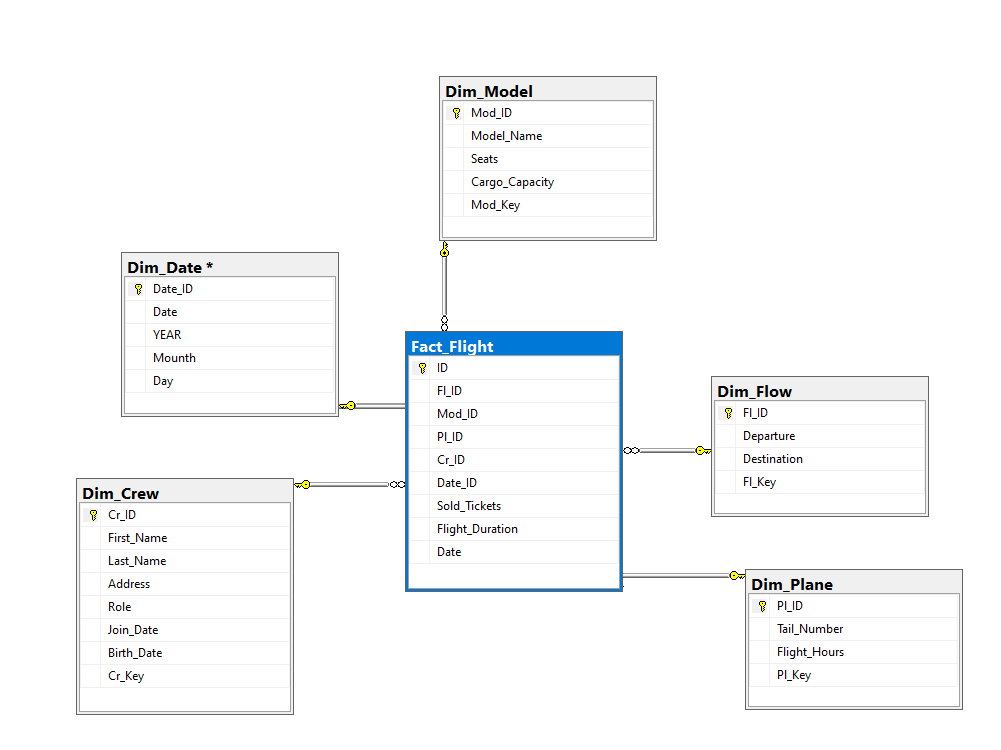
### ****1. Структура бази даних і сховища****

На першому етапі була розроблена концептуальна модель бази даних "Аеропорт", що включає таблиці, які описують літаки, моделі літаків, вильоти літаків, екіпаж та інші важливі аспекти. Структура бази даних дозволяє зберігати і обробляти дані про літаки та їх моделі, а також надавати можливість аналізу кількості членів екіпажу на літаку.

База даних (Airport)



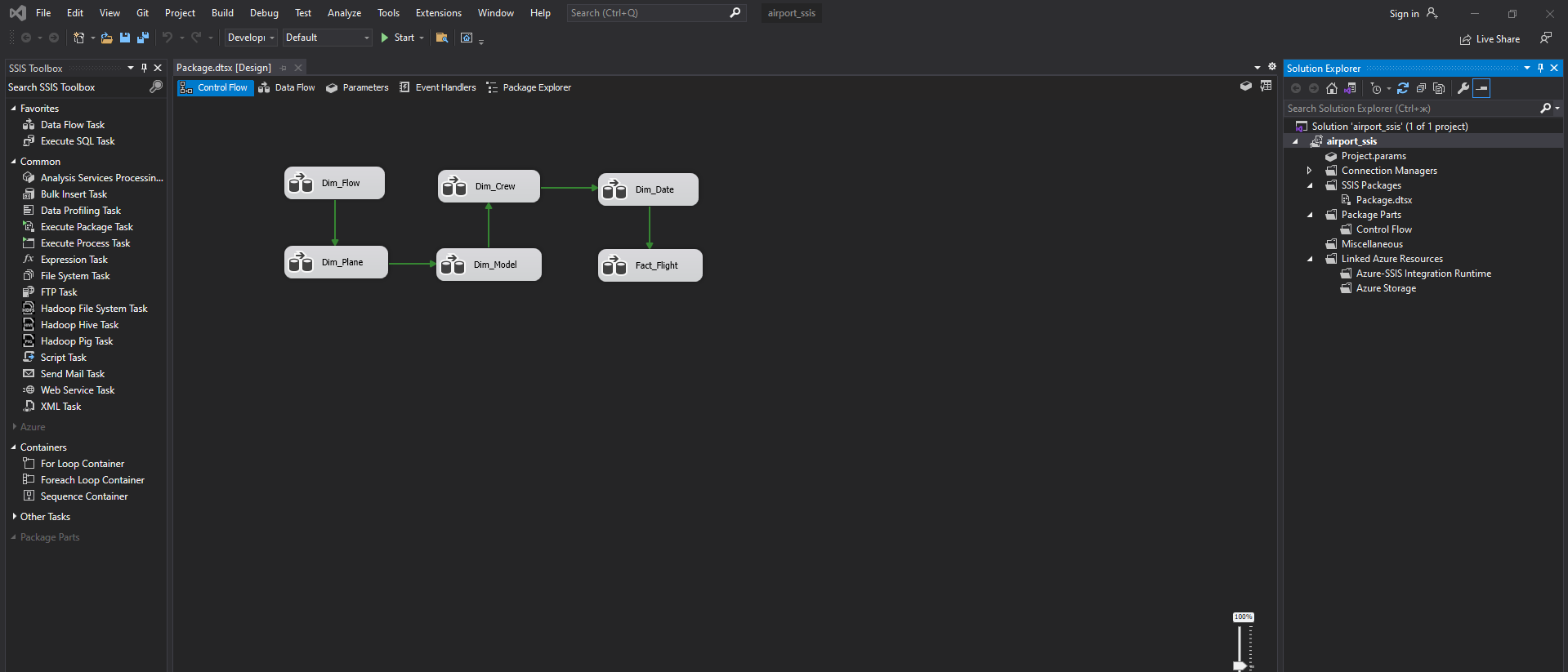
Сховище даних (AirportDW)



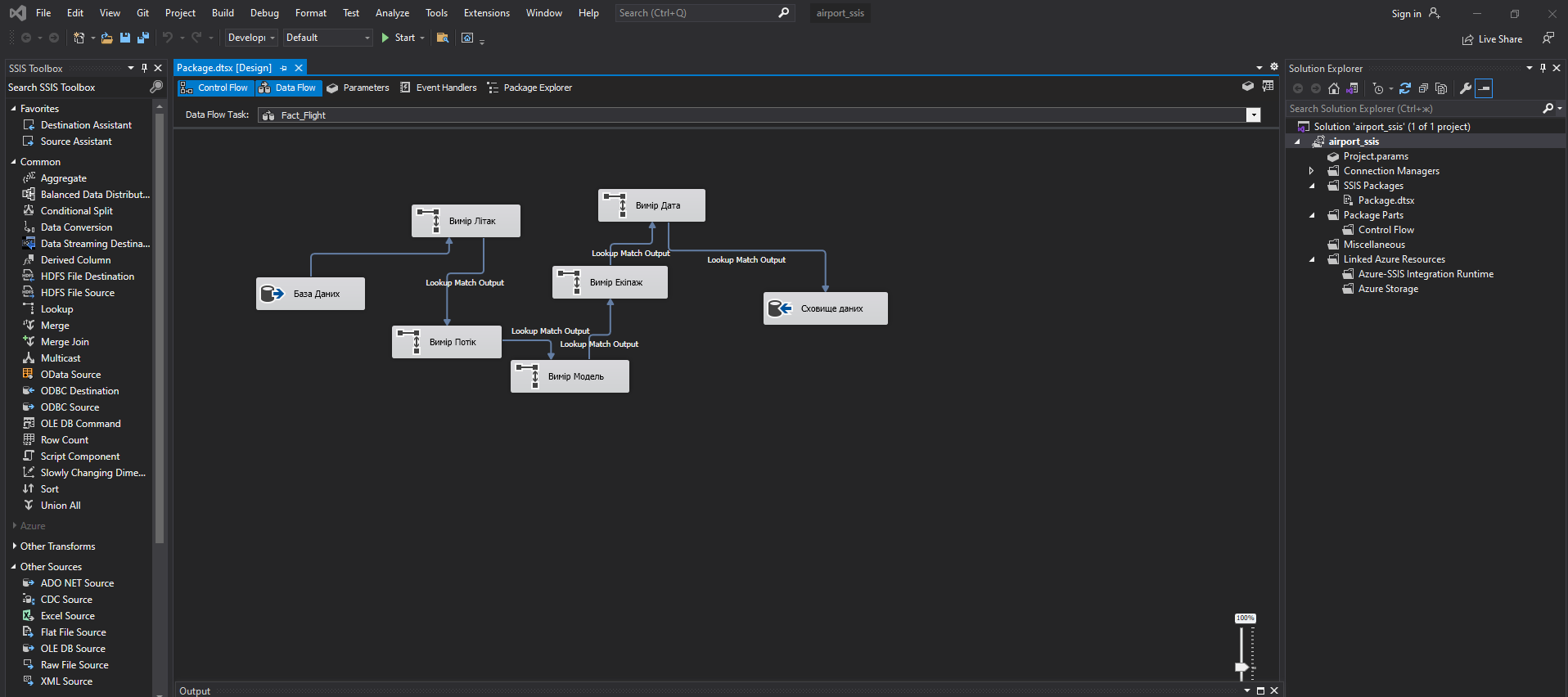
**2. Процес ETL за допомогою SQL Server Integration Services (SSIS)**

SQL Server Integration Services (SSIS) було використано для організації процесу перенесення та трансформації даних із джерела до сховища даних. На цьому етапі виконано налаштування процесів ETL (витягування, трансформація та завантаження), що охоплюють очищення даних та їх перетворення з різноманітних джерел.

Заповнення вимірів даними



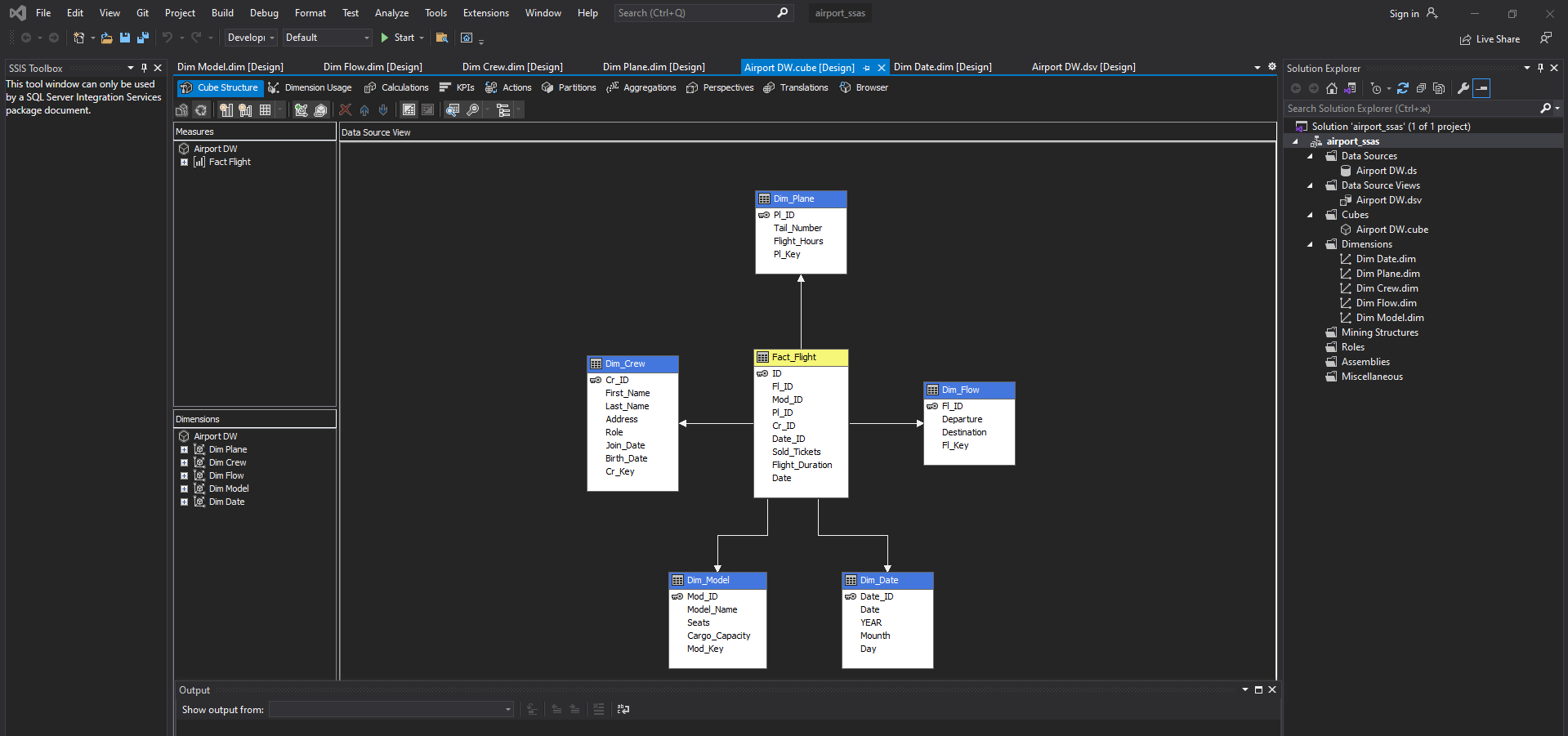
Заповнення таблиці фактів даними



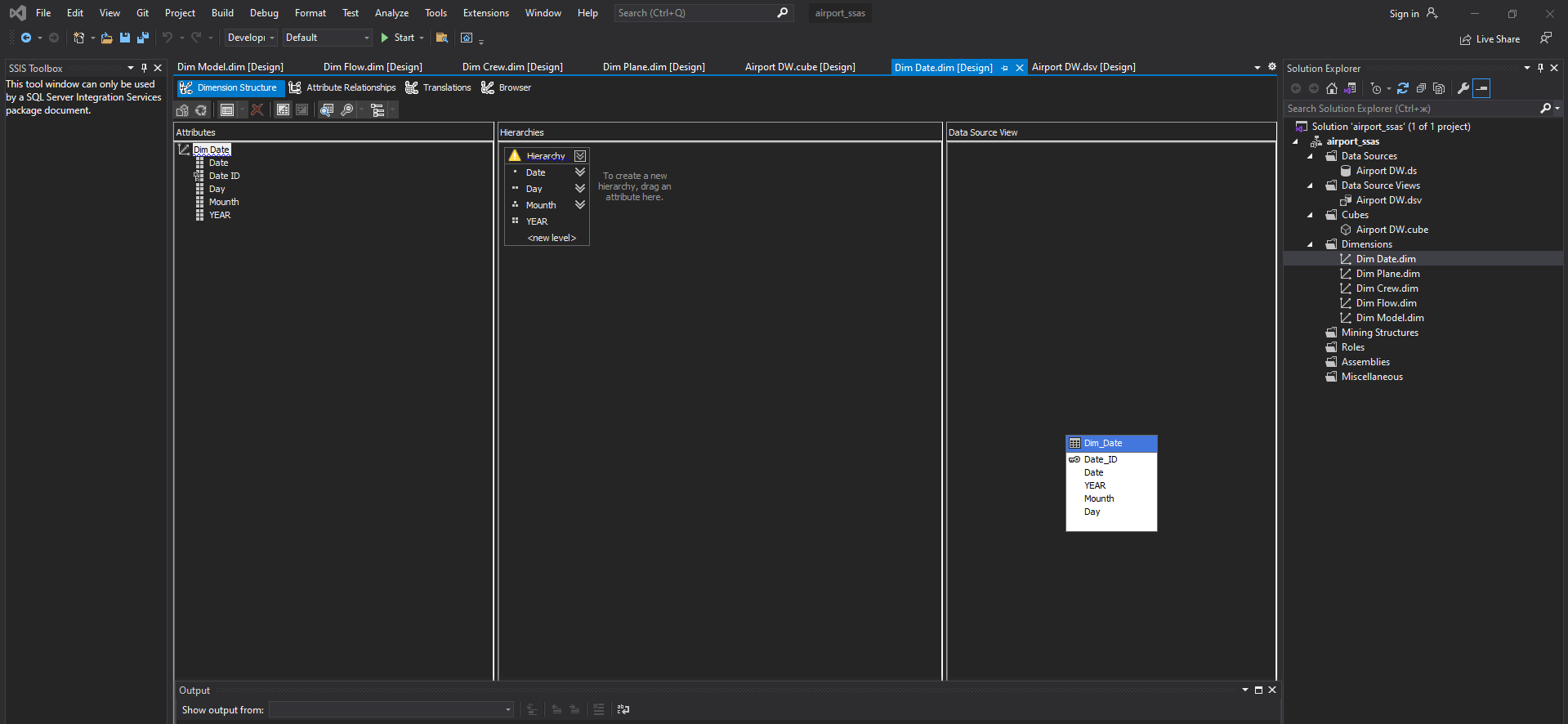
**3. Побудова багатовимірного куба за допомогою SQL Server Analysis Services (SSAS)**

Для структурування та аналізу даних у SQL Server Analysis Services (SSAS) було створено OLAP-куб. У ході роботи визначено міри та виміри куба, а також налаштовано ієрархії для розширення можливостей аналізу.

OLAP-куб



Ієрархії для виміру Dim\_Date

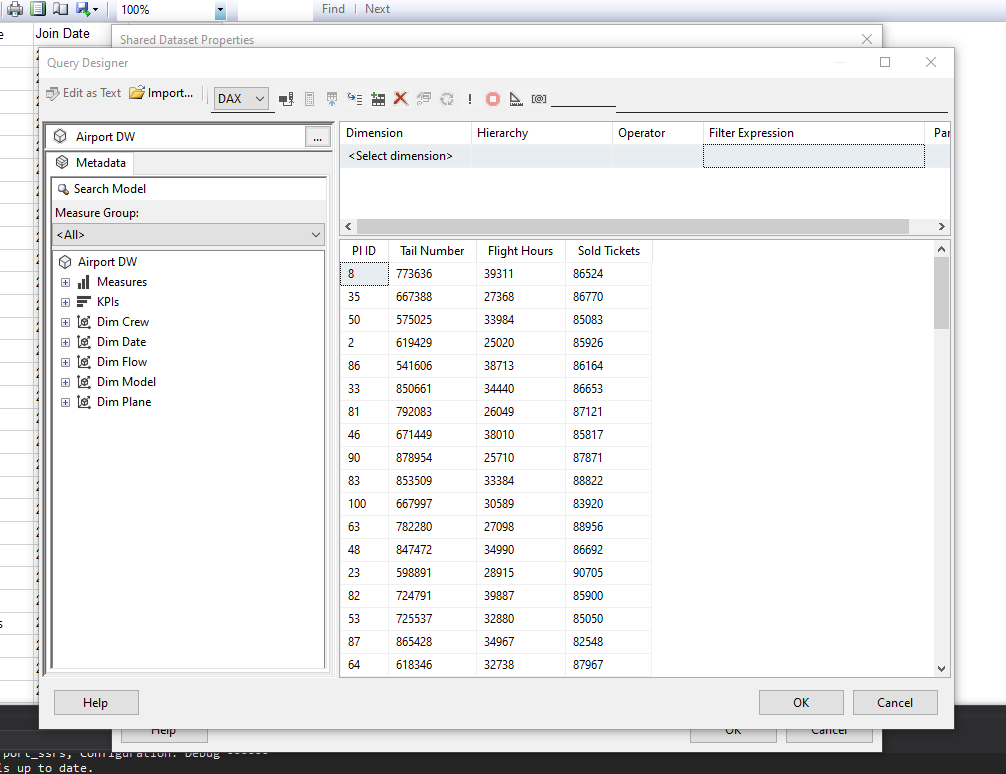


**4. Формування аналітичних звітів за допомогою SQL Server Reporting Services (SSRS)**

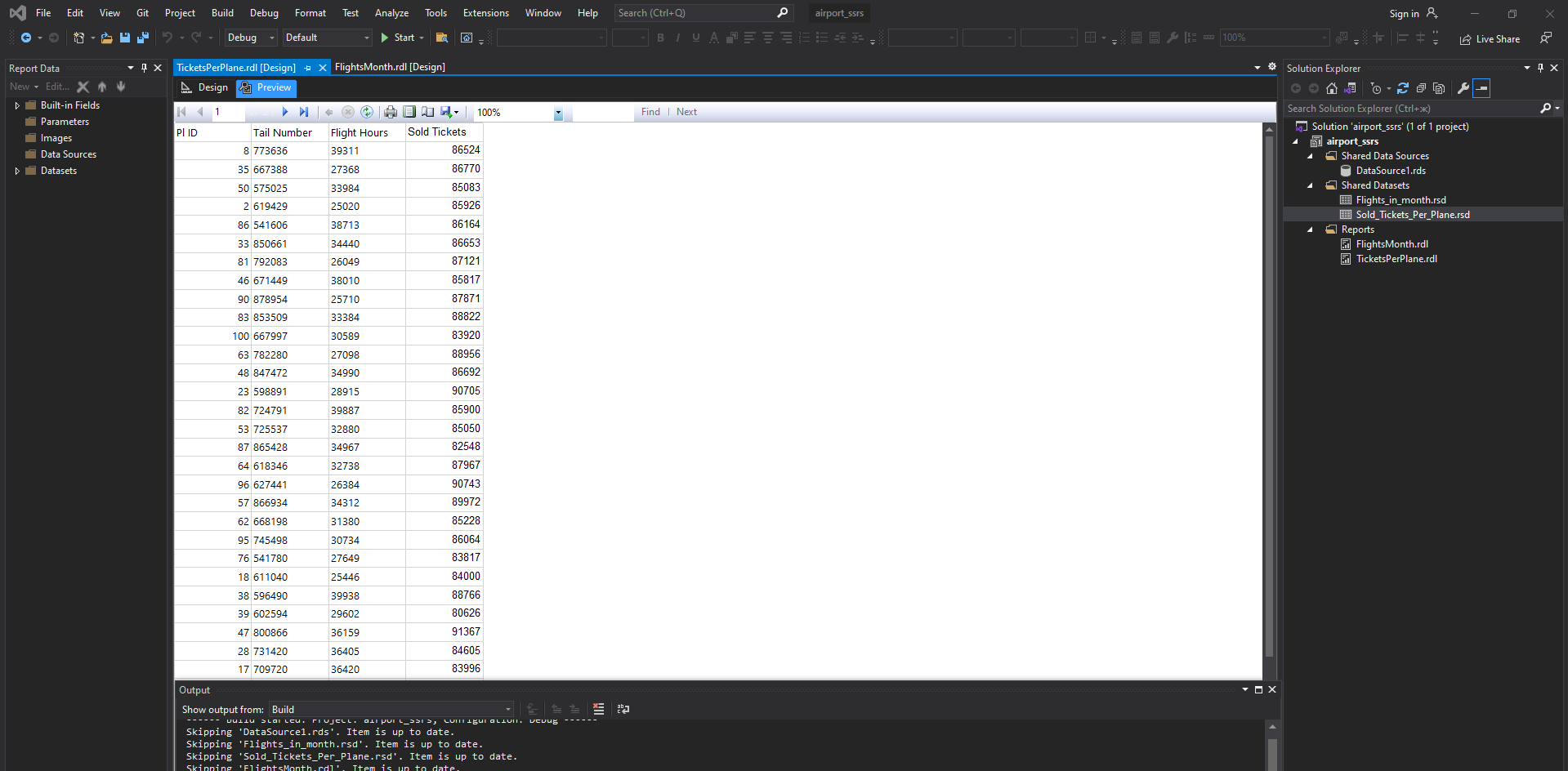
SQL Server Reporting Services (SSRS) було застосовано для візуалізації результатів аналізу та створення інтерактивних звітів. Ці звіти надають можливість користувачам працювати з даними, використовуючи фільтри для налаштування інформації відповідно до обраних параметрів.

**Звіт №1: Продаж квитків на один літак**

Цей звіт відображає статистику продажів білетів в середньому на один літак. Він містить інформацію про унікальний бортовий номер літака та проданих на нього квитків за весь. Звіт дозволяє аналізувати популярність різних типів літаків.

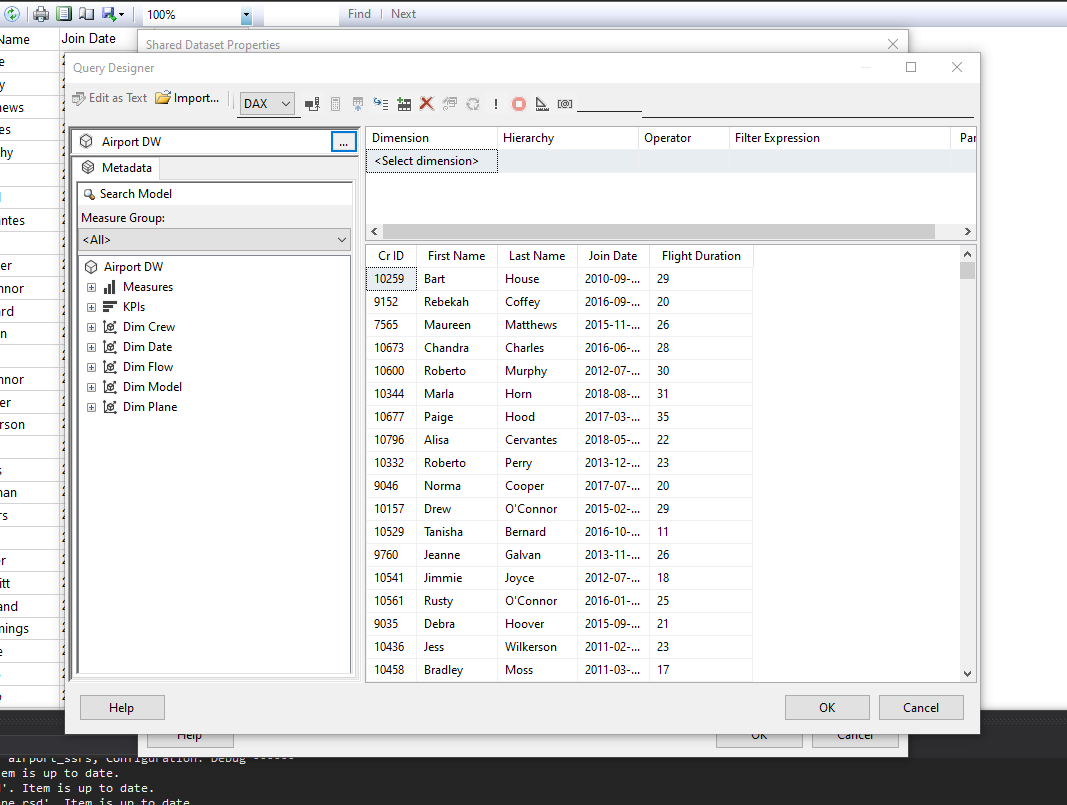


Побудова звіту №1

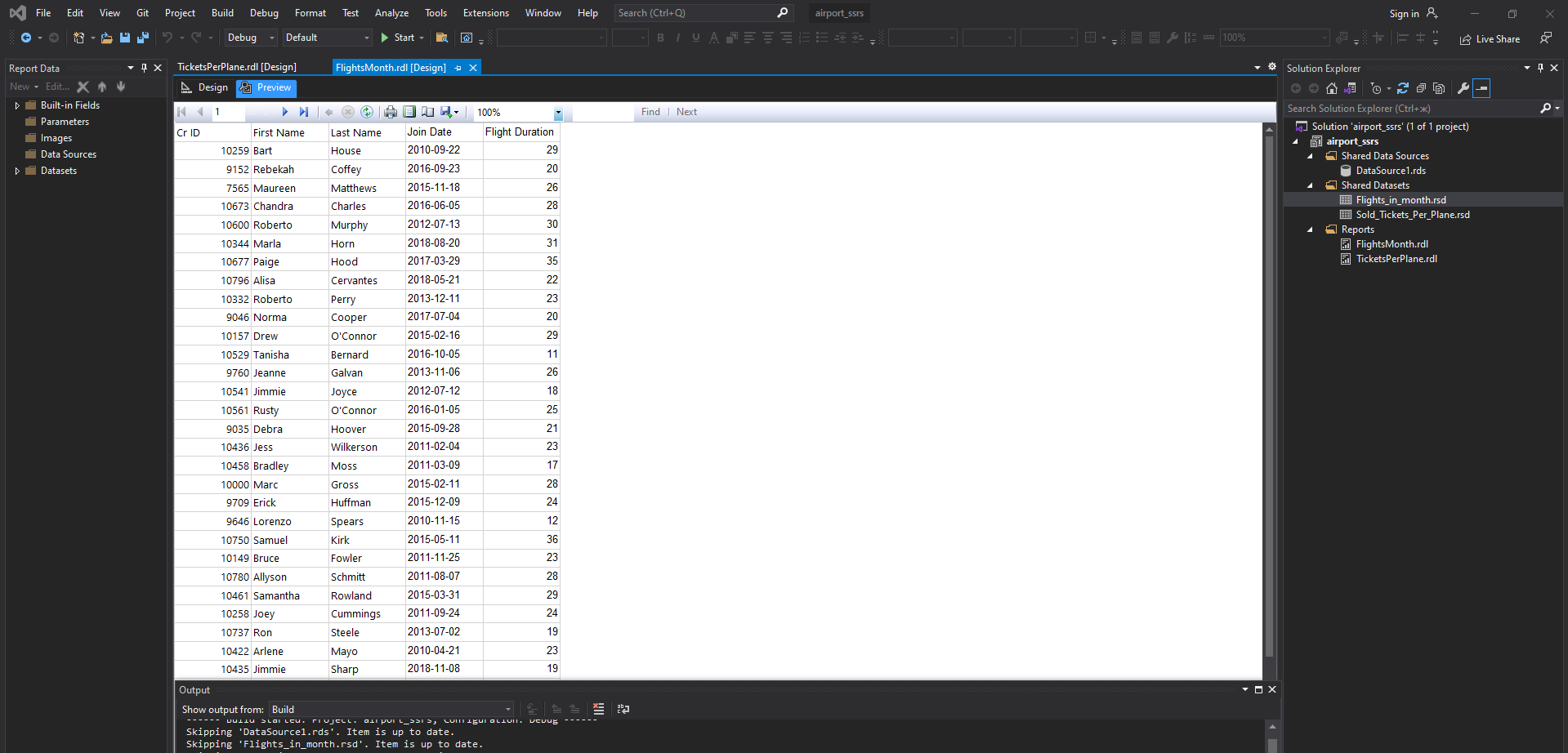


**Звіт №2: Скільки в середньому годин проводить пілот у повітрі за тиждень**

У цьому звіті представлено дані про пілотів і скільки вони проводять часу у повітрі в середньому за тиждень. Цей звіт дає змогу відслідковувати працездатність індивідуального пілота.



Побудова звіту №2



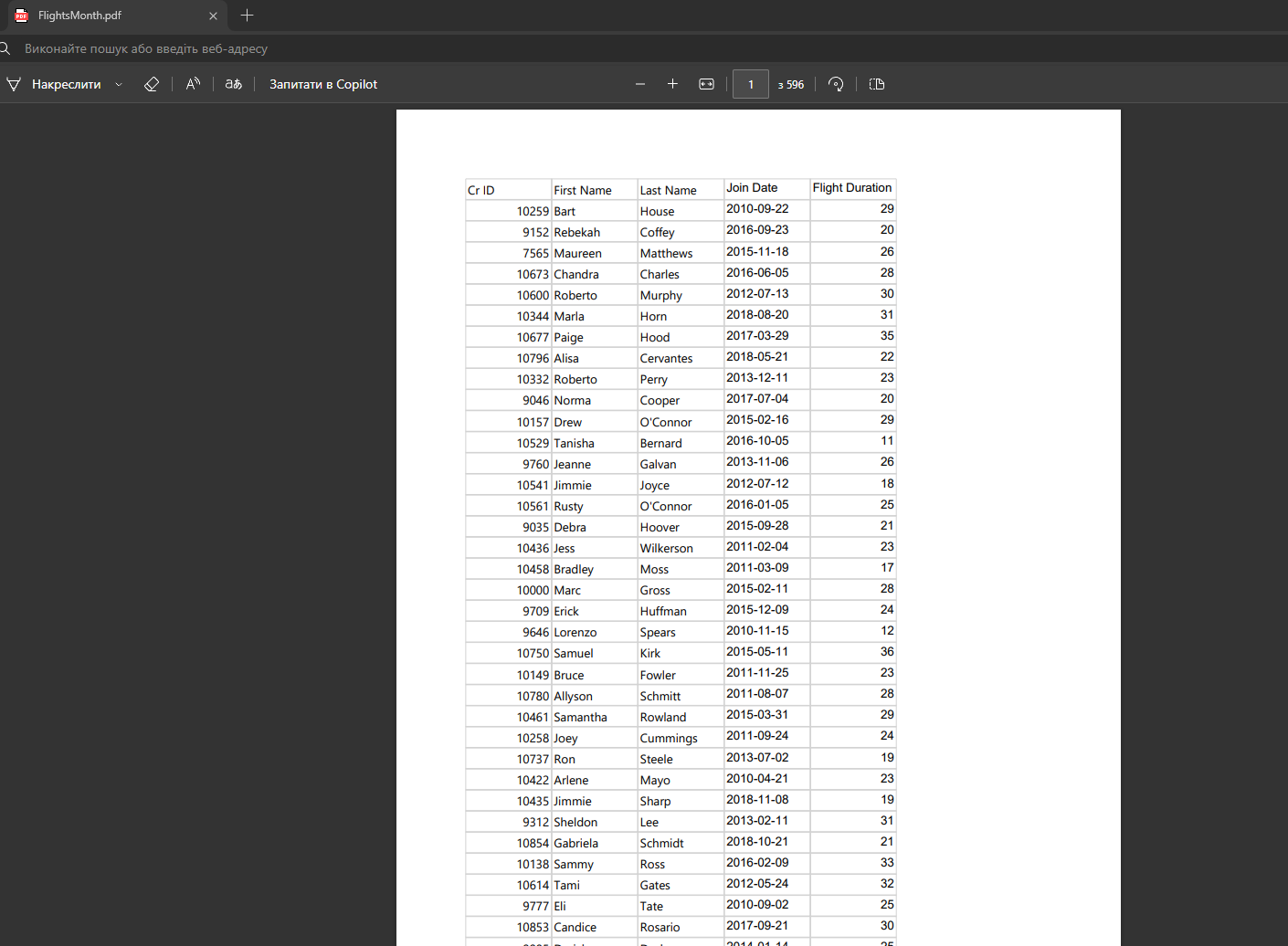
**Експорт звітів:**

У SSRS (SQL Server Reporting Services) можливість експортувати звіти є однією з ключових функцій, яка робить цей інструмент надзвичайно корисним для бізнесу та організацій. Ось кілька основних переваг експорту звітів у SSRS:

1. **Багато форматів для експорту**:
   * SSRS дозволяє експортувати звіти в різні формати, такі як Excel, PDF, Word, CSV, XML, TIFF і HTML. Це дає можливість користувачам отримувати звіти в потрібному форматі для подальшої обробки або презентації.
2. **Легке поширення та інтеграція**:
   * Експортовані звіти можуть бути легко надіслані електронною поштою, збережені на файлових серверах або інтегровані в інші системи. Це спрощує процес обміну інформацією в організації.
3. **Автоматизація процесу**:
   * SSRS дозволяє налаштувати автоматичний розсил звітів у вибраному форматі на регулярній основі. Наприклад, ви можете запланувати щоденне надсилання звіту про продажі в форматі Excel певним користувачам.
4. **Покращення аналітики**:
   * Завдяки можливості експортувати звіти в Excel або CSV, користувачі можуть легко виконувати додатковий аналіз даних за допомогою інструментів, таких як Excel або інші аналітичні платформи.
5. **Мобільність та доступність**:
   * Формати, як-от PDF чи Word, дозволяють отримувати й переглядати звіти на різних пристроях (мобільних телефонах, планшетах), що збільшує доступність звітів для користувачів на ходу.
6. **Зручність для друку**:
   * Експорт у формати PDF або TIFF дозволяє отримувати готові для друку версії звітів, що корисно для розсилки фізичних копій або для архівування звітів у високій якості.

Ці можливості роблять SSRS потужним інструментом для створення, управління та поширення бізнес-звітів у широкому діапазоні форматів, що відповідає потребам різних користувачів.

Приклад експорту у формат pdf:



**Висновки**

У процесі виконання курсової роботи було розроблено інформаційну систему для предметної області "Аеропорт". Реалізовано базу даних, сховище даних і багатовимірний аналіз за допомогою інструментів SQL Server, таких як SSIS, SSAS та SSRS.

На етапі створення бази даних було визначено ключові об’єкти предметної області та їх зв’язки. Розроблене сховище даних забезпечило можливість інтеграції, очищення та обробки даних з різних джерел.

Побудований OLAP-куб дозволив реалізувати багатовимірний аналіз, зокрема:

* Оцінку популярності літаків.
* Відстеження фінансових показників, зокрема кількості проданих квитків.
* Аналіз маршрутів літаків.

Для візуалізації даних створено інтерактивні звіти в SQL Server Reporting Services (SSRS), які дозволяють користувачам ефективно взаємодіяти з інформацією, аналізувати її за різними параметрами та приймати обґрунтовані рішення.

Результати курсової роботи продемонстрували, що розроблена система є ефективним інструментом для зберігання, обробки та аналізу даних, що сприяє покращенню організації авіаційної діяльності та підтримці управлінських рішень у цій сфері.

**Список використаної літератури**

1. **Chapple, M., & Sequeira, J. (2019).** Microsoft SQL Server 2019: A Beginner's Guide (1st ed.). McGraw-Hill Education. <https://books.google.com/books?id=xP17DwAAQBAJ&printsec=frontcover&dq=ms+sql+server&hl=uk&newbks=1&newbks_redir=1&sa=X&ved=2ahUKEwiY4L2hgpmKAxXHFBAIHSWxDjUQ6AF6BAgFEAI>
2. **Sommers, D. (2020).** *Microsoft SQL Server 2019: Всі основи для початківців* (1st ed.). Wiley. <https://books.google.com/books?id=oU3iDwAAQBAJ&printsec=frontcover&dq=ms+sql+server&hl=uk&newbks=1&newbks_redir=1&sa=X&ved=2ahUKEwiY4L2hgpmKAxXHFBAIHSWxDjUQ6AF6BAgLEAI>
3. **Inmon, W. H. (2016).** Building the Data Warehouse (4th ed.). Wiley. <https://books.google.com/books?id=QFKTmh5IFS4C&printsec=frontcover&dq=data+warehouse&hl=uk&newbks=1&newbks_redir=1&sa=X&ved=2ahUKEwjrq_66gpmKAxUsCBAIHU_MEuIQ6AF6BAgHEAI>
4. **Kimball, R., & Ross, M. (2019).** The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling (4th ed.). Wiley. <https://books.google.com/books?id=wgf9DwAAQBAJ&printsec=frontcover&dq=data+warehouse&hl=uk&newbks=1&newbks_redir=1&sa=X&ved=2ahUKEwjrq_66gpmKAxUsCBAIHU_MEuIQ6AF6BAgFEAI>
5. **Harrington, J. L. (2020).** *Data Warehouse Design: Modern Principles and Methodologies* (2nd ed.). McGraw-Hill Education. <https://books.google.com/books?id=U5x7EAAAQBAJ&printsec=frontcover&dq=data+warehouse&hl=uk&newbks=1&newbks_redir=1&sa=X&ved=2ahUKEwjrq_66gpmKAxUsCBAIHU_MEuIQ6AF6BAgEEAI>