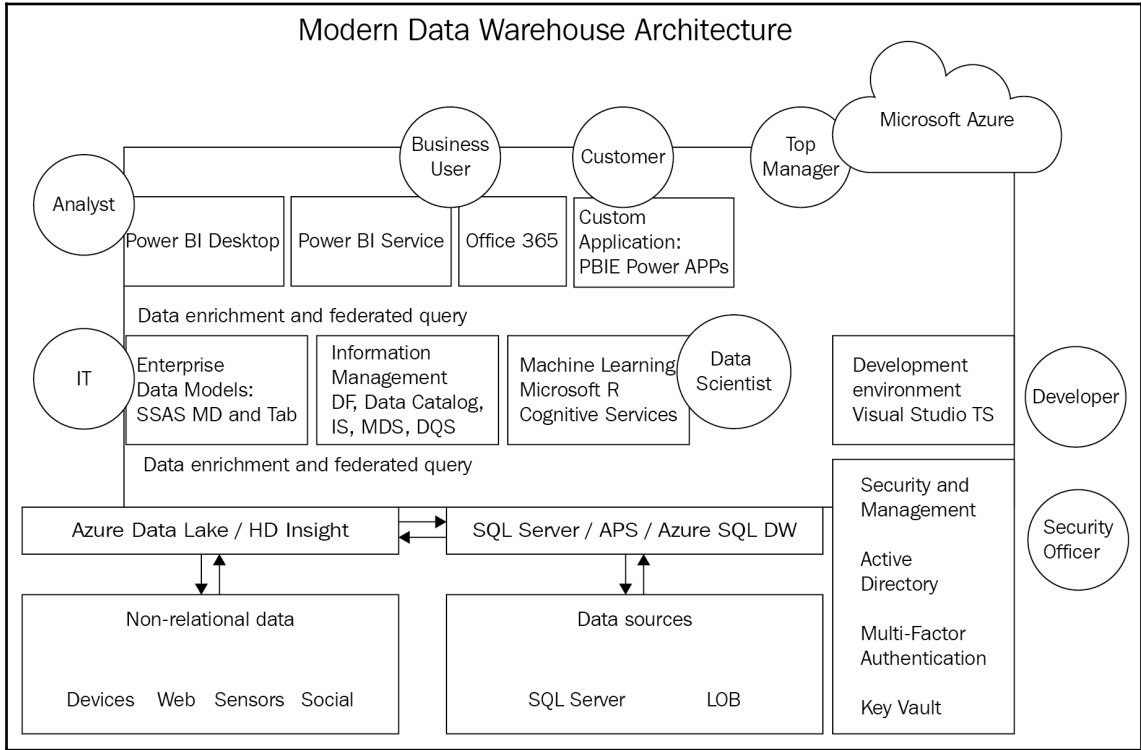


# Chapter 1: The Modern Data Warehouse



```

New data store    ... More    Add activity    Encrypt    Clone    Discard    Deploy

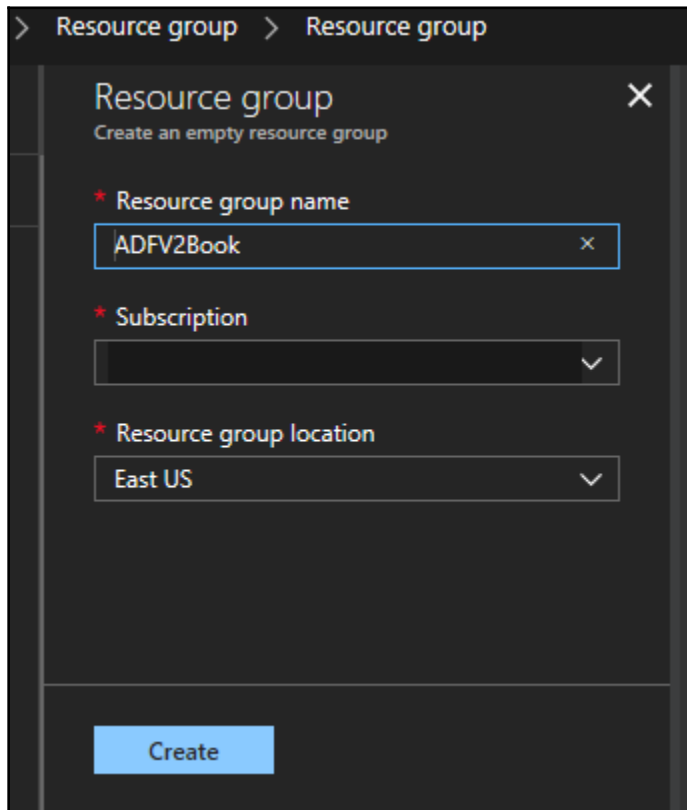
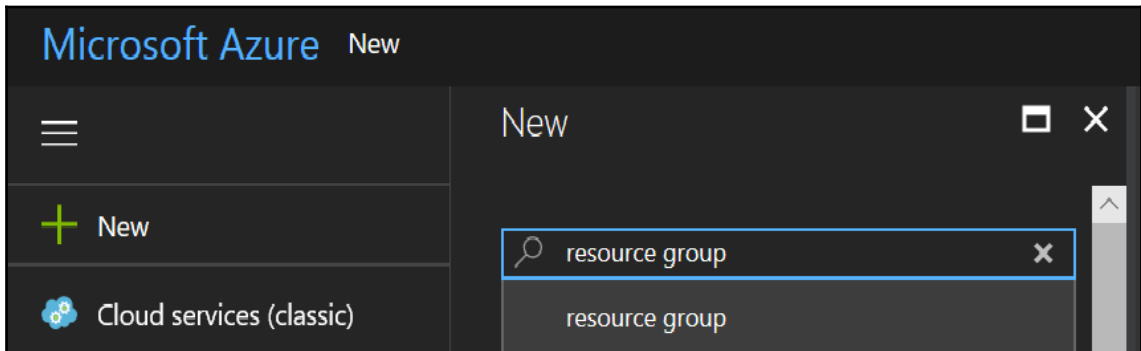
▶ Linked services
▶ Datasets
▶ Pipelines
▶ Integration runtimes (Gateways)
▼ Drafts
  Draft-1

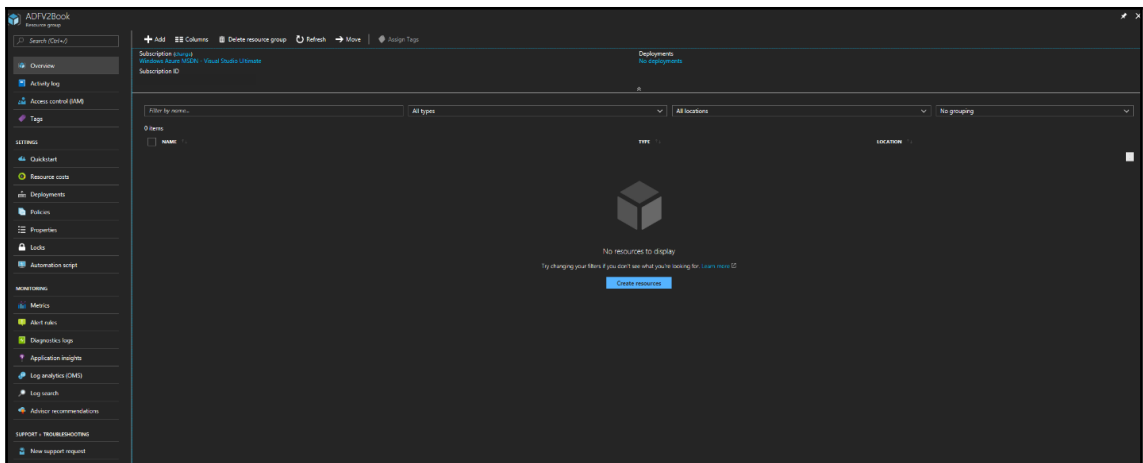
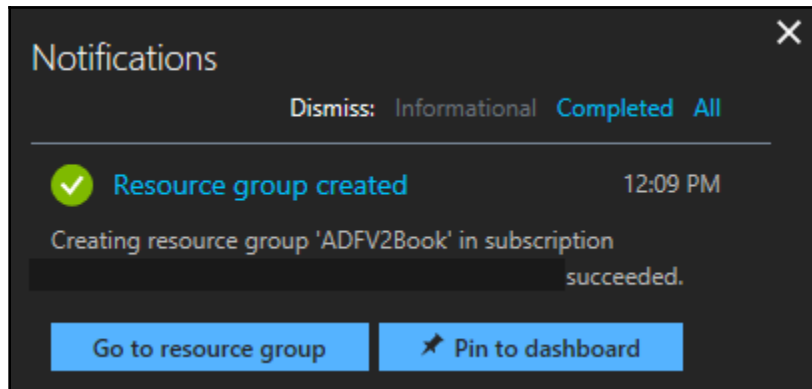
{
  "name": "PipelineTemplate",
  "properties": {
    "description": "<Enter the pipeline description here>",
    "activities": [],
    "start": "<The start date-time of the duration in which data processing will occur or the data s",
    "end": "<The end date-time of the duration in which data processing will occur or the data slice"
  }
}

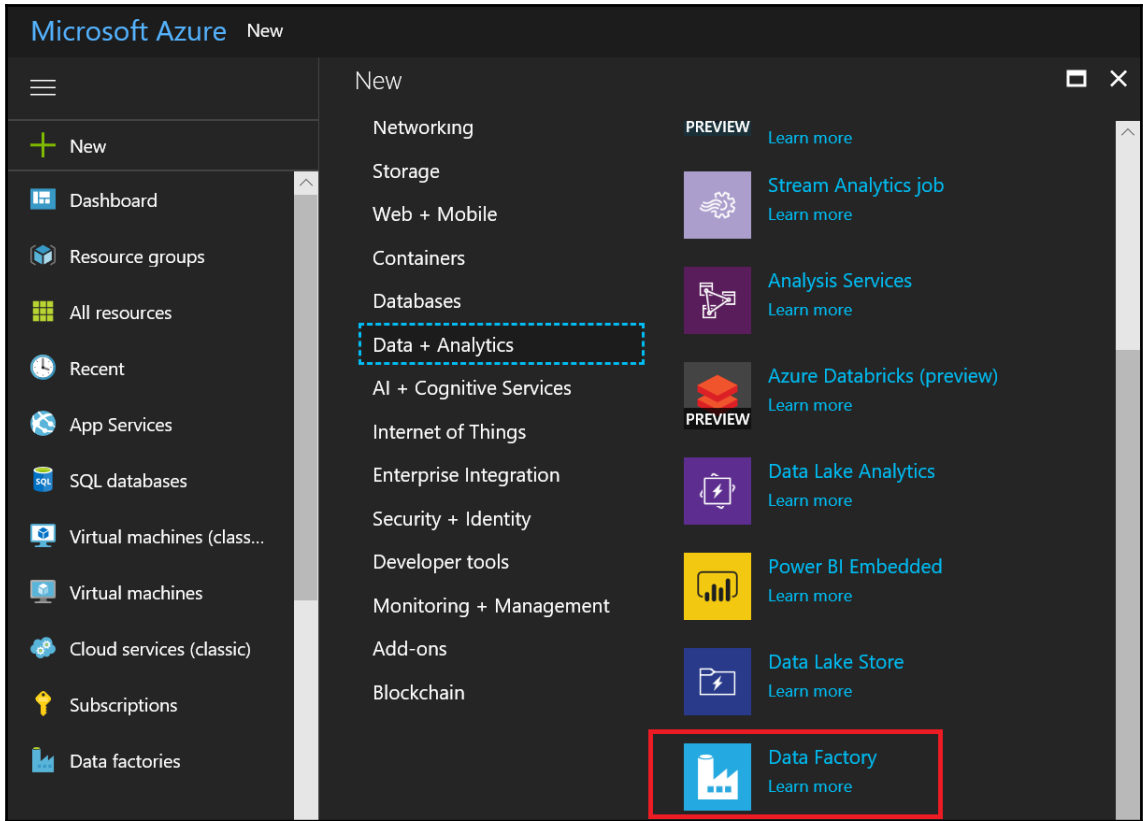
```

---

## Chapter 2: Getting Started with Our First Data Factory







---

New data factory

New data factory ✕

\* Name ⓘ  
ADFV2Book ✓

\* Subscription  
▼

\* Resource Group ⓘ  
 Create new  Use existing  
ADFV2Book ▼

Version ⓘ  
V2 (Preview) ▼

\* Location ⓘ  
East US ▼

Pin to dashboard

[Create](#) [Automation options](#)

The screenshot shows the 'Settings' page for an Azure Data Factory named 'ADFV2Book'. The interface is dark-themed and divided into several sections:

- Essentials:** Displays key information about the Data factory (V2), including its location (EastUS), provisioning state (Succeeded), and subscription details (Subscription name: Windows Azure MSDN - Visual Studio Ulti..., Subscription id: 5d9a43ae-19f6-47f6-957b-b2855a5c069c).
- Quick links:** Provides shortcuts to 'Documentation' and 'Author & Monitor'.
- Monitoring:** Shows a 'PipelineRuns' chart with a value of 100.
- Settings Panel (Right):** Offers navigation options such as 'Diagnose and solve problems', 'Activity log', 'New support request', 'Properties', 'Quick start', 'Tags', 'Locks', and 'Users'.

The screenshot displays the 'Let's get started' page for Azure Data Factory. It features a blue header with the title 'Let's get started' and four primary action buttons:

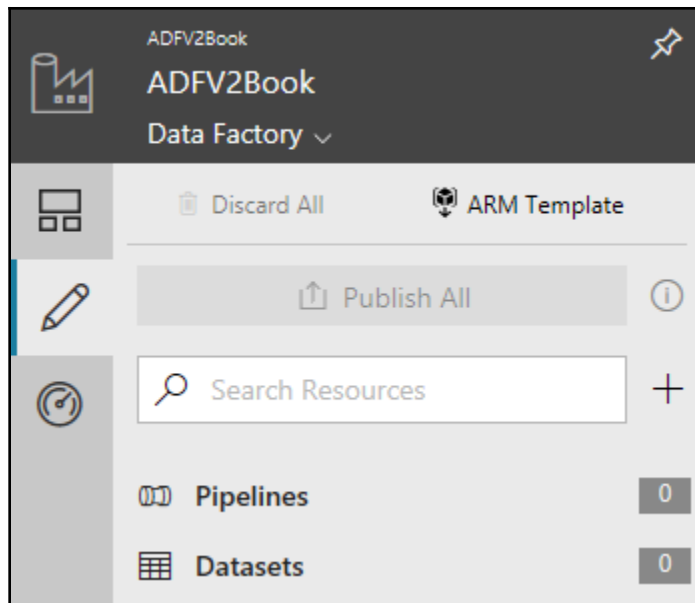
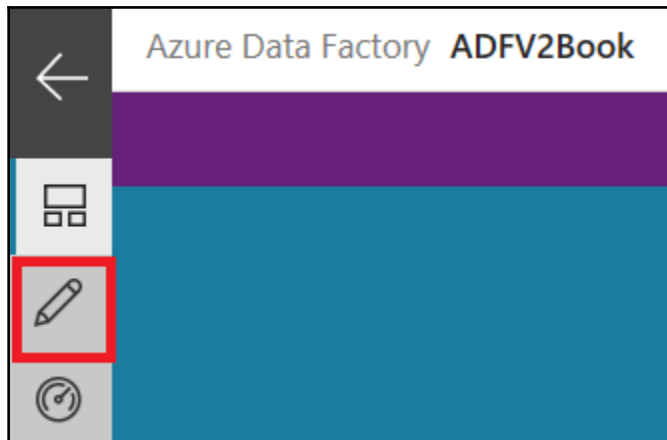
- Create pipeline
- Copy Data
- Configure SSIS Integration Runtime
- Configure Git Repository

Below these are three 'Overview' cards:

- Overview Video
- Introduction to Data Factory
- Lift & shift SSIS packages

At the bottom, there are three 'Featured Tutorials' cards:

- Copy data in cloud
- Copy on-prem data to cloud
- Transform data in cloud



←
ADFV2Book | Monitor Pipeline Runs ▾

🔄 Refresh

📅 Last 24 Hours 02/09/2018 4:14 PM - 02/10/2018 4:14 PM ▾
🌐 Time Zone (UTC-05:00) New York ▾

All Succeeded In Progress Failed Cancelled

Pipeline Name ▾	Actions	Run Start ▾	Duration	Triggered By	Status

## Storage account - blob, file, table, queue

Microsoft

🔗 🖼️ ✕

Microsoft Azure provides scalable, durable cloud storage, backup, and recovery solutions for any data, big or small. It works with the infrastructure you already have to cost-effectively enhance your existing applications and business continuity strategy, and provide the storage required by your cloud applications, including unstructured text or binary data such as video, audio, and images.

---

PUBLISHER

USEFUL LINKS

Microsoft

[Documentation](#)  
[Service overview](#)  
[Pricing](#)

Create



Create storage account ☐ ✕

The cost of your storage account depends on the usage and the options you choose below.  
[Learn more](#)

\* Name ⓘ  
adfv2book ✓  
.core.windows.net

Deployment model ⓘ  
Resource manager Classic

Account kind ⓘ  
StorageV2 (general purpose v2) ▾

Performance ⓘ  
Standard Premium

Replication ⓘ  
Locally-redundant storage (LRS) ▾

Access tier (default) ⓘ  
Cool Hot

\* Secure transfer required ⓘ  
Disabled Enabled

\* Subscription  
▾

\* Resource group  
 Create new  Use existing

ADFV2Book ▾

\* Location  
East US ▾

Virtual networks  
Configure virtual networks ⓘ  
Disabled Enabled

Pin to dashboard

Create Automation options

Open in Explorer → Move Delete storage account

Resource group (change)  
ADFV2Book

Status  
Primary: Available, Secondary: Available

Location  
East US, West US

Subscription (change)  
Windows Azure MSDN - Visual Studio Ultimate

Subscription ID  
5d9a43ae-19f6-47f6-957b-b2855a5c069c

Performance  
Standard

Replication  
Read-access geo-redundant storage (RA-GRS)

Account kind  
StorageV2 (general purpose v2)

Services

- Blobs**  
Object storage for understanding data  
[View metrics](#)  
[Configure CORS rules](#)  
[Setup custom domain](#)
- Files**  
File shares that use SMB 3.0 protocol  
[View metrics](#)  
[Configure CORS rules](#)
- Tables**  
Tabular data storage  
[View metrics](#)
- Queues**  
Scale apps depending on traffic  
[View metrics](#)

Open in Explorer → Move Delete storage account

Open 'adv2book' in Azure Storage Explorer?  
If you don't have the app yet, download it from [storageexplorer.com](http://storageexplorer.com).

Yes No

https://azure.microsoft.com/en-us/features/storage-explorer/

Microsoft Azure

SALES 1-800-867-1389 MY ACCOUNT

Why Azure Solutions Products Documentation Pricing Training Marketplace Partners Support Blog More #

# Azure Storage Explorer

Easily manage Storage anywhere from Windows, MacOS, and Linux


- ✓ Access multiple accounts and subscriptions across Azure, Azure Stack, and the sovereign Cloud
- ✓ Create, delete, view, and edit storage resources
- ✓ View and edit Blob, Queue, Table, File, and Cosmos DB storage
- ✓ Obtain shared access signature (SAS) keys
- ✓ Available for Windows, Mac, and Linux

Operating system  
Windows

Download Storage Explorer free

ACCOUNT MANAGEMENT

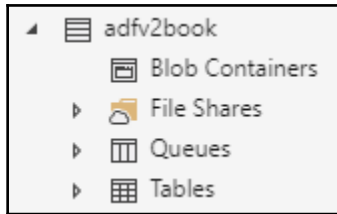
Show resources from these subscriptions:

 Microsoft account [Remove](#)  
ccote\_1@msn.com

All subscriptions:

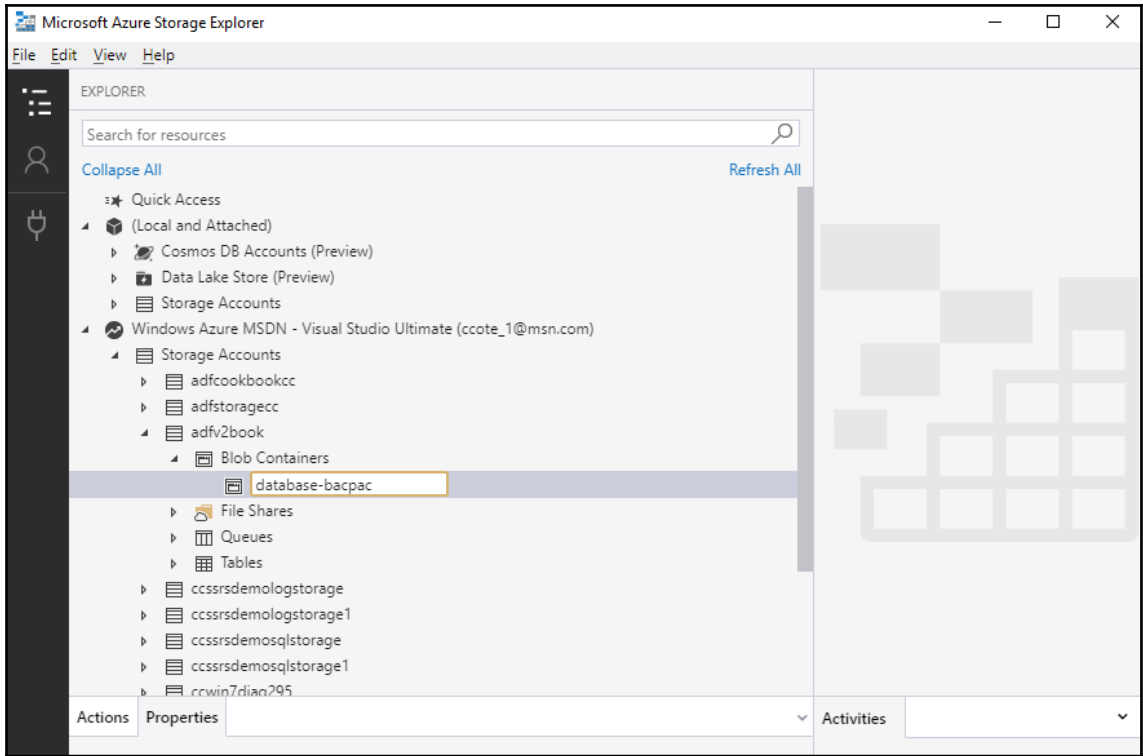
[Add an account...](#)

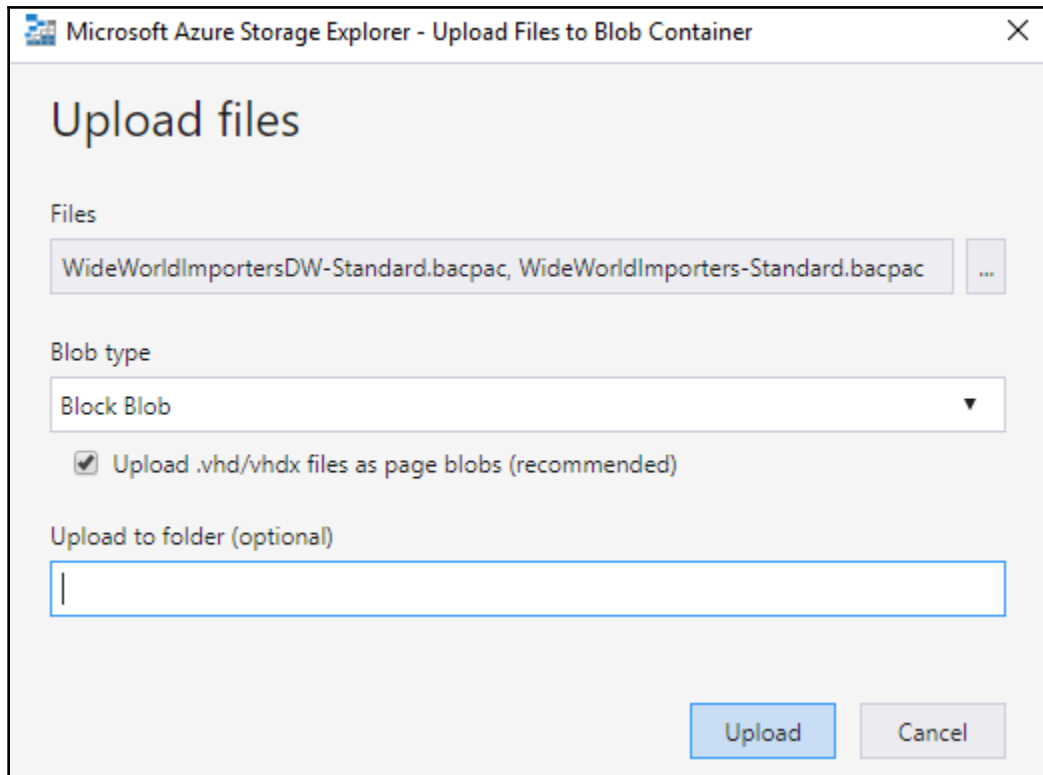
Apply Cancel

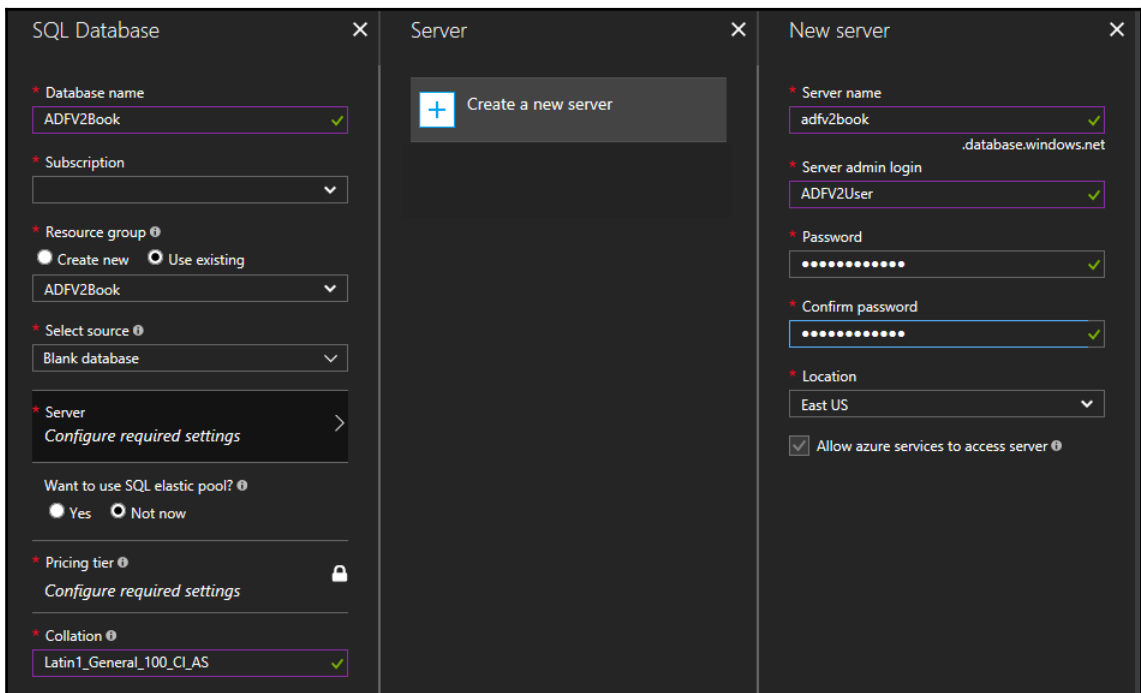
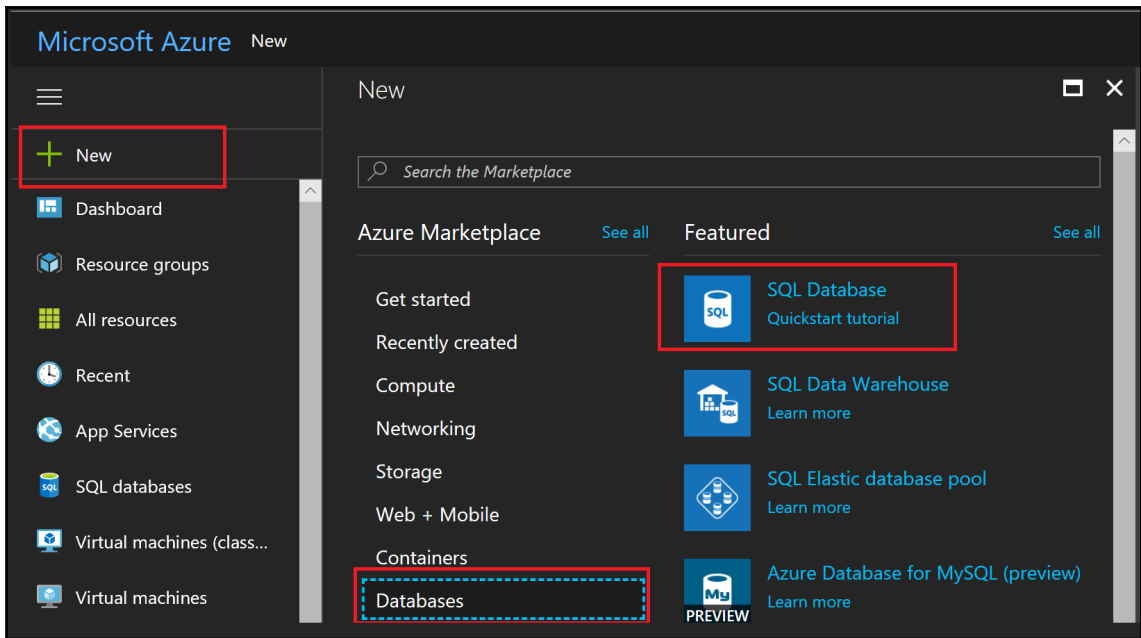


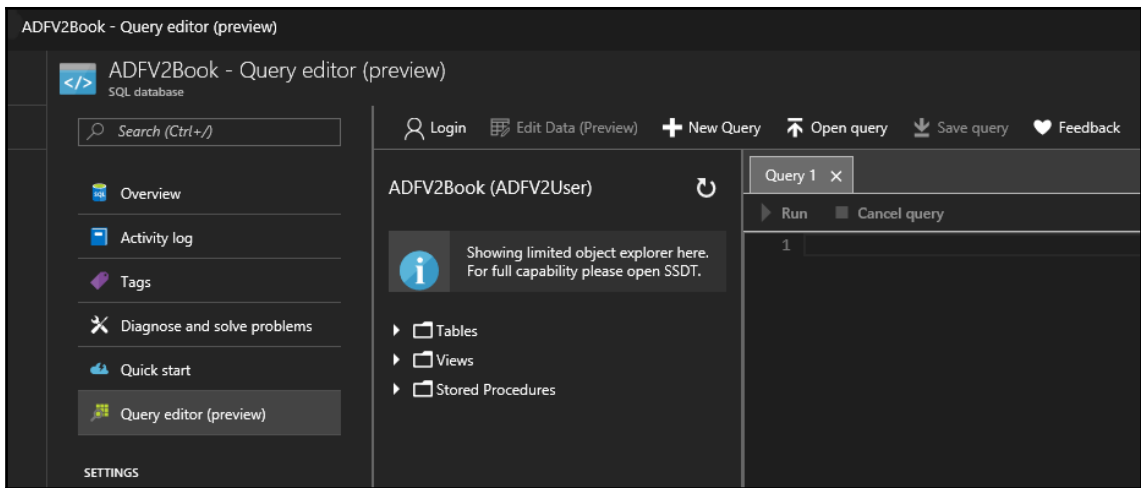
A screenshot of a GitHub release page for 'wide-world-importers-v1.0'. The page lists several database samples and their requirements. A red rectangular box highlights the 'Azure SQL Database Standard tier' section, which includes two sub-items: 'WideWorldImporters-Standard.bacpac' and 'WideWorldImportersDW-Standard.bacpac'.

- [WideWorldImporters-Full.bak](#) - full sample database for OLTP (OnLine Transaction Processing) and HTAP (hybrid transaction and analytical processing aka real-time operational analytics). For SQL Server 2016 SP1 (or higher), any edition.
  - Update (8/12/2016): updated stats on system tables to support efficient reverse engineering with Entity Framework
- [WideWorldImportersDW-Full.bak](#) - full sample database for OLAP (OnLine Analytical Processing). For SQL Server 2016 SP1 (or higher), any edition.
- Hardware requirements: installing both databases requires 10GB of available disk space, and 1.5GB of available memory for SQL Server.
- **Azure SQL Database Premium tier**
  - [WideWorldImporters-Full.bacpac](#) - full OLTP sample database in bacpac format. For Azure SQL Databases in the Premium tier.
  - [WideWorldImportersDW-Fullbacpac](#) - full OLAP sample database in bacpac format. For Azure SQL Databases in the Premium tier.
- **Azure SQL Database Standard tier**
  - [WideWorldImporters-Standard.bacpac](#) - standard edition OLTP sample database in bacpac format. For Azure SQL Databases in the Standard tier.
  - [WideWorldImportersDW-Standard.bacpac](#) - standard edition OLAP sample database in bacpac format. For Azure SQL Database in the Standard tier.

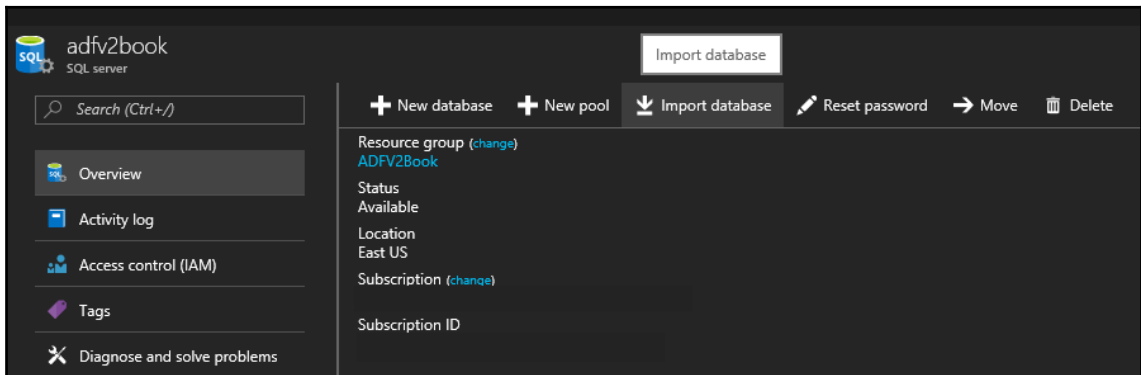
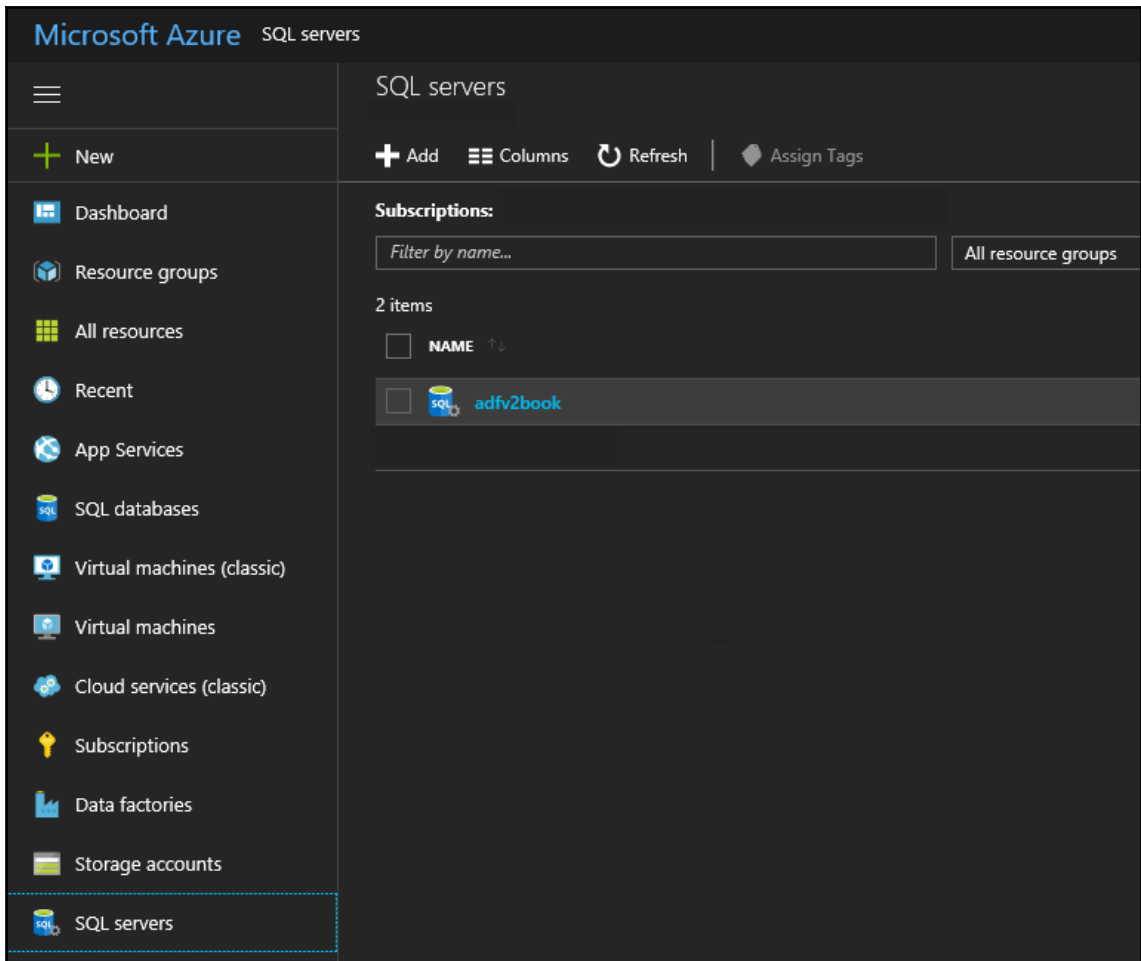












Home > adfv2book > Import database

## Import database

adfv2book

\* Subscription

---

\* Storage (Premium not supported)  
WideWorldImporters-Standard.b... >

---

\* Pricing tier ⓘ >  
Standard S3: 100 DTU, 250 GB

---

\* Database name  
WideWorldImporters-Standard ✓

\* Collation ⓘ  
Latin1\_General\_100\_CI\_AS ✓

Authentication type  
SQL Server ▾

\* Server admin login  
ADfv2User

\* Password  
●●●●●●●● ✓

Pin to dashboard

OK

Home > adfv2book - Firewall / Virtual Networks

adfv2book - Firewall / Virtual Networks  
SQL server

Search (Ctrl+ /)

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems

SETTINGS

- Quick start
- Firewall / Virtual Networks**
- Failover groups
- Long-term backup retention
- Auditing & Threat Detection
- Transparent data encryption
- Active Directory admin

Save Discard **Add client IP**

**+** Add client IP

**i** Connections from the IPs specified below provides access to all the databases in adfv2book.

Allow access to Azure services  
**ON** OFF

Client IP address 24.203.239.147

RULE NAME	START IP	END IP
		...

No firewall rules configured.

**i** Connections from the VNET/Subnet specified below provides access to all databases in adfv2book.

Virtual networks [+ Add existing virtual network](#) [+ Create new virtual network](#)

RULE NAME	RESOURCE GROUP/VNET NAME
No vnet rules for this server.	

Object Explorer

- System Databases
- ADFV2Book
- WideWorldImporters-Standard
  - Database Diagrams
  - Tables
    - System Tables
    - External Tables
    - GraphTables
    - Application.Cities (System-Versioned)
    - Application.Countries (System-Versioned)
    - Application.DeliveryMethods (System-Versioned)
    - Application.PaymentMethods (System-Versioned)
    - Application.People (System-Versioned)
    - Application.StateProvinces (System-Versioned)
    - Application.SystemParameters
    - Application.TransactionTypes (System-Versioned)
    - Purchasing.PurchaseOrderLines
    - Purchasing.PurchaseOrders
    - Purchasing.SupplierCategories (System-Versioned)
    - Purchasing.Suppliers (System-Versioned)
    - Purchasing.SupplierTransactions
    - Sales.BuyingGroups (System-Versioned)
    - Sales.CustomerCategories (System-Versioned)
    - Sales.Customers (System-Versioned)
    - Sales.CustomerTransactions
    - Sales.InvoiceLines
    - Sales.Invoices
    - Sales.OrderLines
    - Sales.Orders
    - Sales.SpecialDeals
    - Warehouse.ColdRoomTemperatures (System-Versioned)
    - Warehouse.Colors (System-Versioned)
    - Warehouse.PackageTypes (System-Versioned)
    - Warehouse.StockGroups (System-Versioned)
    - Warehouse.StockItemHoldings
    - Warehouse.StockItems (System-Versioned)
    - Warehouse.StockItemStockGroups
    - Warehouse.StockItemTransactions

SQLQuery1.sql - ad...d (ADFV2User (110))

```

/***** Script for SelectTopRows command from SSMS *****/
SELECT TOP (1000) [PurchaseOrderID]
,[SupplierID]
,[OrderDate]
,[DeliveryMethodID]
,[ContactPersonID]
,[ExpectedDeliveryDate]
,[SupplierReference]
,[IsOrderFinalized]
,[Comments]
,[InternalComments]
,[LastEditedBy]
,[LastEditedWhen]
FROM [Purchasing].[PurchaseOrders]
  
```

Results

	PurchaseOrderID	SupplierID	OrderDate	DeliveryMethodID	ContactPersonID
1	1	2	2013-01-01	9	2
2	2	4	2013-01-01	7	2
3	3	5	2013-01-01	10	2
4	4	7	2013-01-01	2	2
5	5	10	2013-01-01	8	2
6	6	12	2013-01-01	7	2
7	7	4	2013-01-02	7	2
8	8	5	2013-01-02	10	2
9	9	7	2013-01-02	2	2
10	10	10	2013-01-02	8	2
11	11	12	2013-01-02	7	2
12	12	4	2013-01-03	7	2
13	13	5	2013-01-03	10	2
14	14	7	2013-01-03	2	2

1000 rows

Azure Data Factory

# Let's get started

Create pipeline

Copy Data

Configure SSIS  
Integration Runtime

Configure Git Repository

1 Properties

2 Source

3 Destination

4 Settings  
Fault tolerance

5 Summary

6 Deployment

## Properties

Enter name and description for the copy data task.

Task name **(required)**



CopyPipeline\_ADFV2Book

Task description

This a simple pipeline that copies data from SQL Server on Azure to an Azure Storage account

Task cadence or Task schedule

- Run once now
- Run regularly on schedule

Next

1 Properties  
One time copy







2 Source  
Connection  
Dataset

3 Destination

4 Settings  
Fault tolerance

5 Summary

6 Deployment

Azure Data Lake Store	Azure Database for MySQL	Azure Database for PostgreSQL
		
Azure SQL Data Warehouse	Azure SQL Database	Azure Table Storage
		
DB2	File System	FTP

Previous Next

**Copy Data**

**1 Properties**  
One time copy

**2 Source**

- Connection
- Dataset

**3 Destination**

**4 Settings**  
Fault tolerance

**5 Summary**

**6 Deployment**

### Select tables from which to copy the data or use a custom query.

You can select multiple tables, or you can provide single custom query

EXISTING TABLES    USE QUERY

Query ⓘ

```
[WebsiteName]
[DeliveryMethod]
[CityName]
FROM [Website],[Customers]
```

Validate Query

---

PREVIEW    SCHEMA

CustomerID	CustomerName	CustomerCategoryName	PrimaryContact	AlternateContact	PhoneNumber	FaxNumber	BuyingGroupName	WebsiteURL
841	Camille Authier	Gift Store	Camille Authier	(Null)	(252) 555-0100	(252) 555-0101	(Null)	http://www.micros...

Previous    Next

**Copy Data**

**1 Properties**  
One time copy

**2 Source**  
Azure SQL Database

**3 Destination**  
Connection  
Dataset









**4 Settings**  
Fault tolerance, Performance

**5 Summary**

**6 Deployment**

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store. Click [HERE](#) to suggest new copy destinations or give comments.

FROM EXISTING CONNECTIONS    CONNECT TO A DATA STORE

 Azure Blob Storage	 Azure Cosmos DB	 Azure Data Lake Store	 Azure File Storage
 Azure Search	 Azure SQL Data Warehouse	 Azure SQL Database	 Azure Table Storage

Previous    Next



**Copy Data**

**1 Properties**  
One time copy

**2 Source**  
Azure SQL Database

**3 Destination**

- Connection
- Dataset

**4 Settings**  
Fault tolerance, Performance

**5 Summary**

**6 Deployment**

## Specify the Azure Blob storage account

Connection name **(required)** ⓘ


Network Environment **(required)**

Account selection method **(required)** ⓘ

Azure subscription **(required)** ⓘ

Storage account name **(required)**

[Previous](#) [Next](#)

 Copy Data

**1 Properties**  
One time copy

**2 Source**  
Azure SQL Database

**3 Destination**



- Connection
- Dataset



**4 Settings**  
Fault tolerance, Performance

**5 Summary**



**6 Deployment**

## File format settings


File format   
Text format 


Column delimiter   
Pipe (|) 


Use custom delimiter

Row delimiter   
Carriage Return + Line feed (\r\n) 

Use custom delimiter

Add header to file 

 Advanced settings

Previous 

**Copy Data**

- 1 **Properties**  
One time copy
- 2 **Source**  
Azure SQL Database
- 3 **Destination**  
Azure Blob Storage
- 4 **Settings**
- 5 **Summary**
- 6 **Deployment**

## Settings

More options for data movement

^ **Fault tolerance settings**

**Error handling for incompatible rows between source and destination** ⓘ

Actions

Abort copy activity on first incompatible row (cc) ▾

^ **Performance settings**

^ **Advanced settings**

**Parallel copy** ⓘ

Cloud units

Auto ▾


Previous Next

**Copy Data**


- 1 Properties  
One time copy
- 2 Source  
Azure SQL Database
- 3 Destination  
Azure Blob Storage
- 4 Settings  
Fault tolerance, Performance
- 5 Summary
- 6 Deployment

## Summary

You are running one time pipeline to copy data from Azure SQL Database to Azure Blob Storage.

 **Azure SQL Database**  
1 table(s) ⓘ  
Region: EastUs

Copy Run Time Region: EastUs

 **Azure Blob Storage import-data**  
Region: EastUs

### Properties

[Edit](#)

Task name: CopyPipeline\_ADFV2Book  
Task description: This is a simple pipeline that copies data from SQL Server on Azure to an Azure Storage account

### Source

[Edit](#)

Connection: Server: adfv2book.database.windows.net, database: WideWorldImporters-Standard  
Connection name: Source\_SQLAzure\_zi2  
Dataset name suffix: i\_zi2  
Table: 1 table(s) ⓘ  
Region: EastUs

### Destination

[Edit](#)

[Previous](#) [Next](#)

**ADfv2Book** | Monitor Pipeline Runs ▾




Refresh

Last 24 Hours 02/20/2018 8:07 AM - 02/21/2018 8:07 AM ▾

Time Zone (UTC-05:00) Toronto ▾

1 Filter Applied

All Succeeded In Progress Failed Cancelled

Pipeline Name ▾	Actions	Run Start	Duration	Triggered By	Status	Parameters	Error	RunID
CopyPipeline_ADFV2Bo...	 	02/21/2018, 8:06:12 AM	00:00:23	Manual trigger	 Succeeded			3d0e0739-a04e-4dfb-b23c-484e578f4d1d

## Services



### [Blobs](#)

Object storage for understanding data

[View metrics](#)

[Configure CORS rules](#)

[Setup custom domain](#)

**Blob service** ★ ×  
adfv2book

+ Container ↻ Refresh 🗑️ Delete

Essentials ⌵

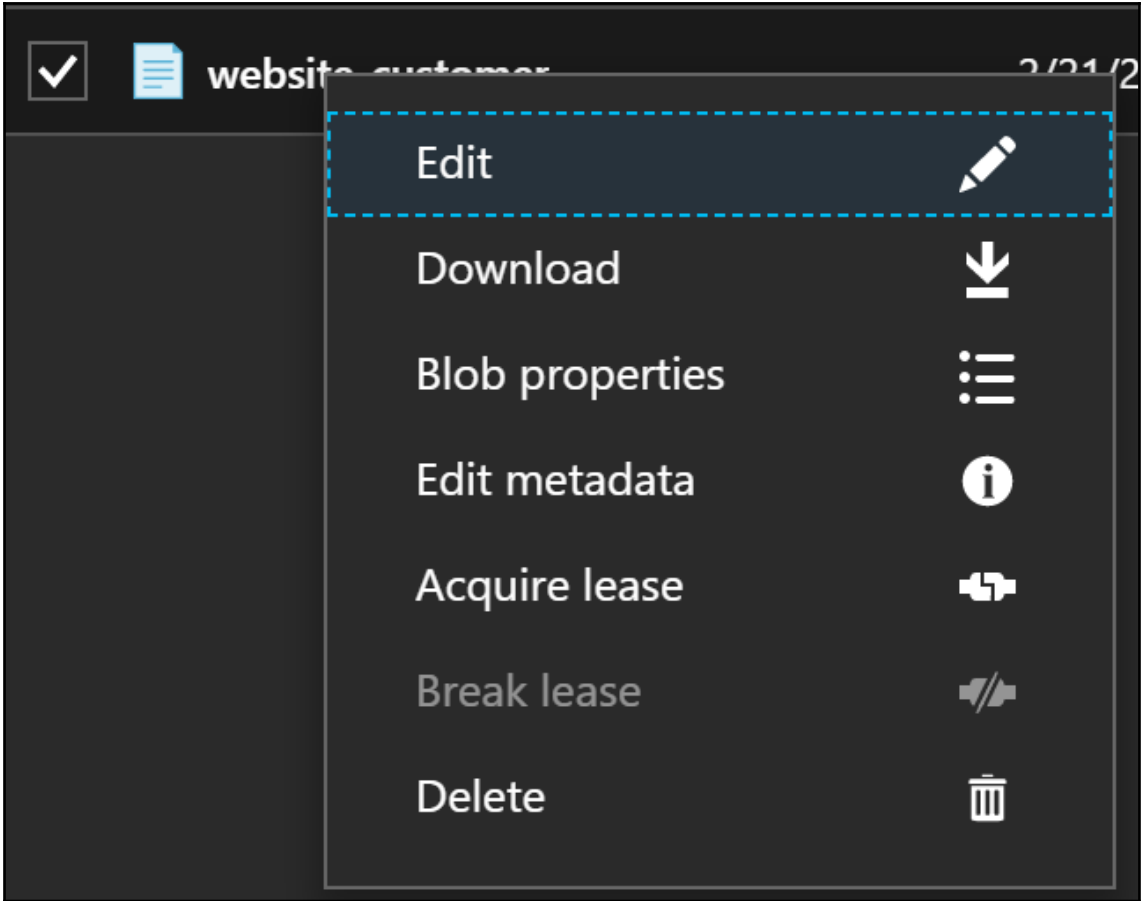
NAME	
database-bacpac	...
<input type="checkbox"/> import-data	...

**import-data**  
Container

↑ Upload ↻ Refresh 🗑️ Delete

**Location:** [import-data](#)

NAME
<input type="checkbox"/> website-customer



Home > adfv2book > Blob service > import-data > website-customer

**import-data**  
Container

Upload Refresh More

Location: import-data

Search blobs by prefix (case-sensitive)

NAME

- website-customer

**website-customer**  
import-data/

Save Discard changes Download

⚠ The file 'website-customer' may not render correctly as it contains an unrecognized extension.

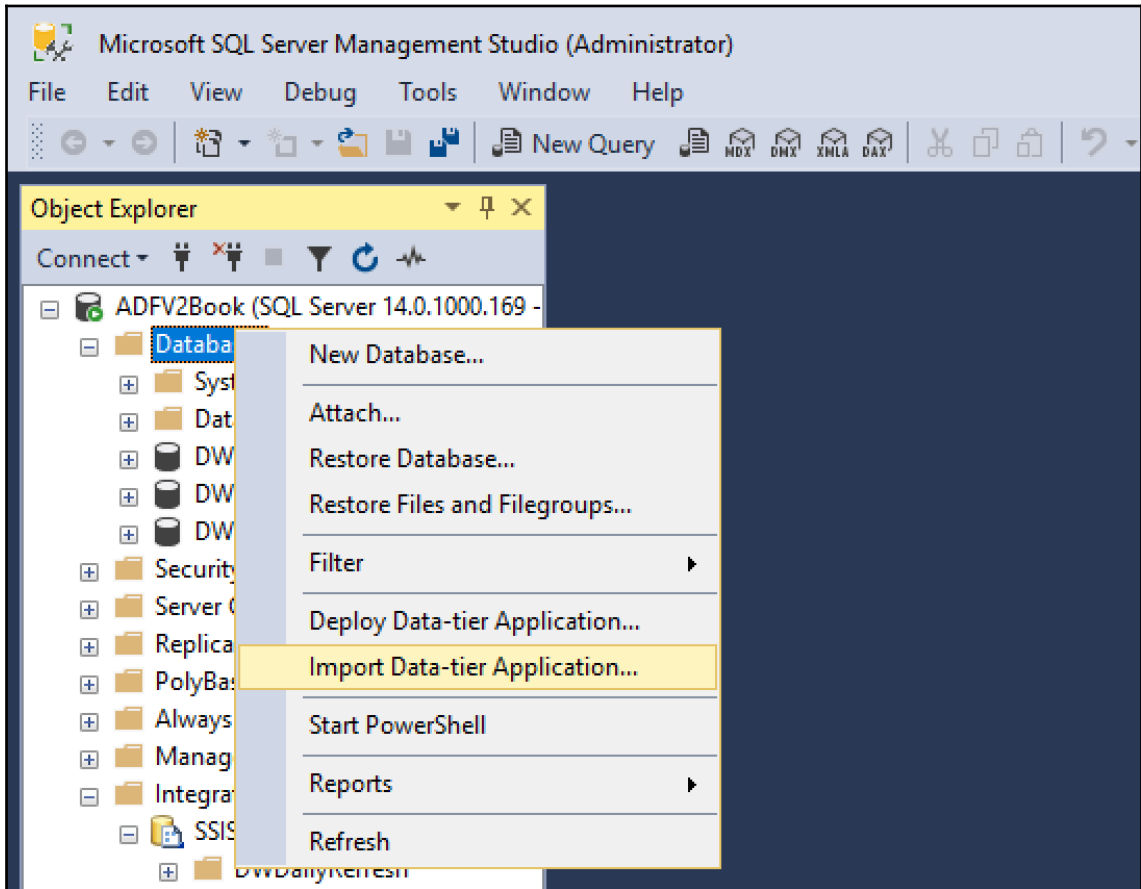
```

1 CustomerID|CustomerName|CustomerCategoryName|PrimaryContact|AlternateContact|PhoneNumber|FaxNumber
2 841|Camille Authier|Gift Store|Camille Authier|\N|(252) 555-0100|(252) 555-0101|\N|http://www.mic
3 83|Tailspin Toys (Absecon, NJ)|Novelty Shop|Sang Tran|Bela Nemeth|(201) 555-0100|(201) 555-0101|T
4 595|Wingtip Toys (Accomac, VA)|Novelty Shop|Baanuraekha Voleti|Renata Alva|(276) 555-0100|(276) 5
5 84|Tailspin Toys (Aceitunas, PR)|Novelty Shop|Eekalabya Bose|Margherita Bucco|(787) 555-0100|(787
6 38|Tailspin Toys (Airport Drive, MO)|Novelty Shop|Nils Kalnins|Rajiv Kotadia|(314) 555-0100|(314)
7 853|Caterina Pinto|Gift Store|Caterina Pinto|\N|(907) 555-0100|(907) 555-0101|\N|http://www.micro
8 537|Wingtip Toys (Akhiok, AK)|Novelty Shop|Deepesh Bose|Ambareesh Madhunapantula|(907) 555-0100|
9 589|Wingtip Toys (Alcester, SD)|Novelty Shop|Deviprasad Prabhupada|Tess Seegers|(605) 555-0100|
10 892|Bahaar Asef zade|Supermarket|Bahaar Asef zade|\N|(225) 555-0100|(225) 555-0101|\N|http://www.
11 91|Tailspin Toys (Alstead, NH)|Novelty Shop|GokCe Karaca|Maria Munteanu|(603) 555-0100|(603) 555-
12 566|Wingtip Toys (Amado, AZ)|Novelty Shop|Corneliu Vladu|Ivo Valek|(480) 555-0100|(480) 555-0101|
13 44|Tailspin Toys (Amanda Park, WA)|Novelty Shop|Adrijana Blagojevic|Lina Martins|(206) 555-0100|
14 85|Tailspin Toys (Andrix, CO)|Novelty Shop|Tommy Kilpi|Giri Bhamidipati|(303) 555-0100|(303) 555-
15 125|Tailspin Toys (Annamoriah, WV)|Novelty Shop|Blanka Stavinohova|Josefa Diaz|(304) 555-0100|(30
16 151|Tailspin Toys (Antares, AZ)|Novelty Shop|Donato Araujo|Alaine Berthelette|(480) 555-0100|(480
17 199|Tailspin Toys (Antonito, CO)|Novelty Shop|Cezar Pavel|Bahaar Yekta|(303) 555-0100|(303) 555-0
18 155|Tailspin Toys (Arbor Vitae, WI)|Novelty Shop|Michal Polaskova|Debendra Chakraborty|(262) 555-
19 889|Sointu Savonheimo|Gift Store|Sointu Savonheimo|\N|(701) 555-0100|(701) 555-0101|\N|http://www
20 42|Tailspin Toys (Arietta, NY)|Novelty Shop|Chandrakanta Raut|Ivan Castellanos|(212) 555-0100|(21
21 178|Tailspin Toys (Armstrong Creek, WI)|Novelty Shop|Hyun-Doo Song|Bahadãr Korkmaz|(262) 555-010

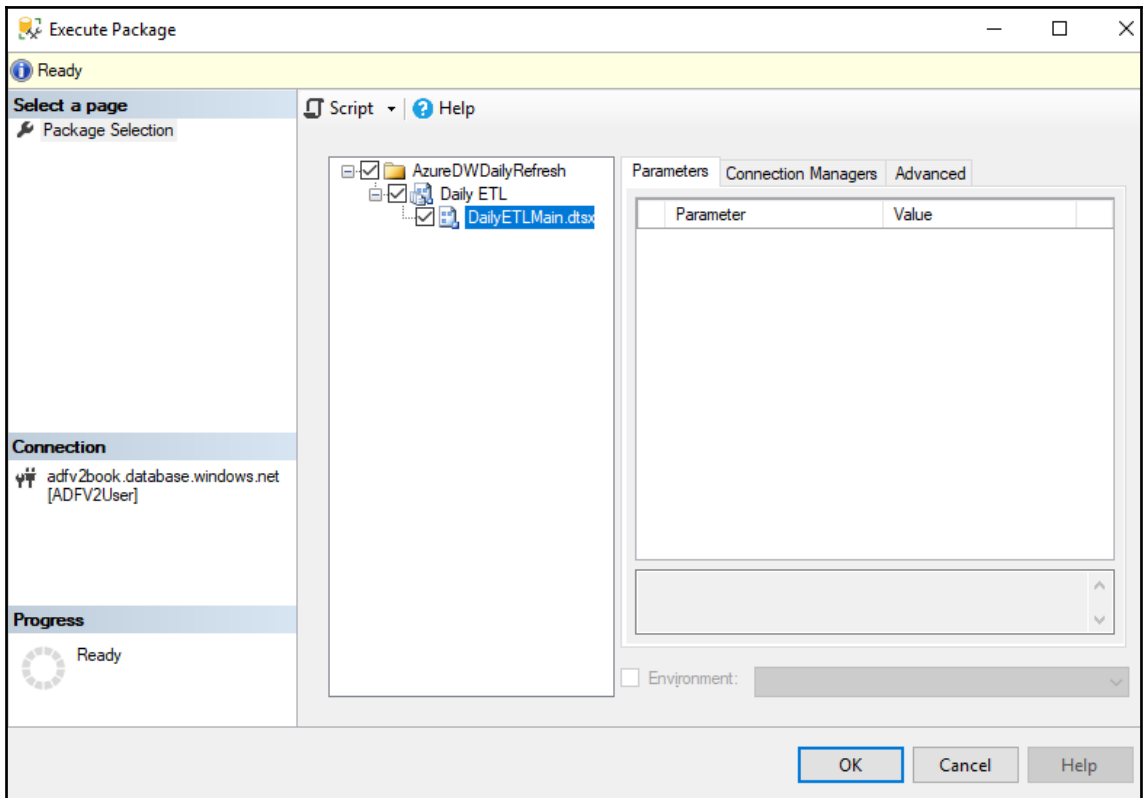
```

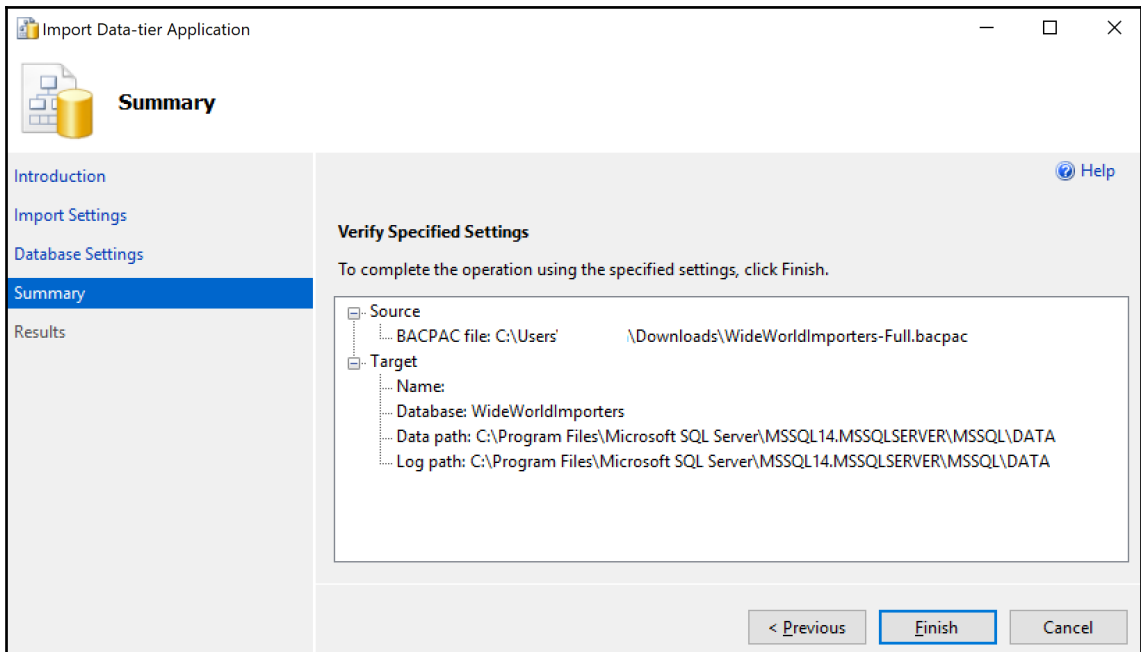
---

## Chapter 3: SSIS Lift and Shift









Import Data-tier Application

**Results**

Introduction Help

Import Settings

Database Settings

Summary

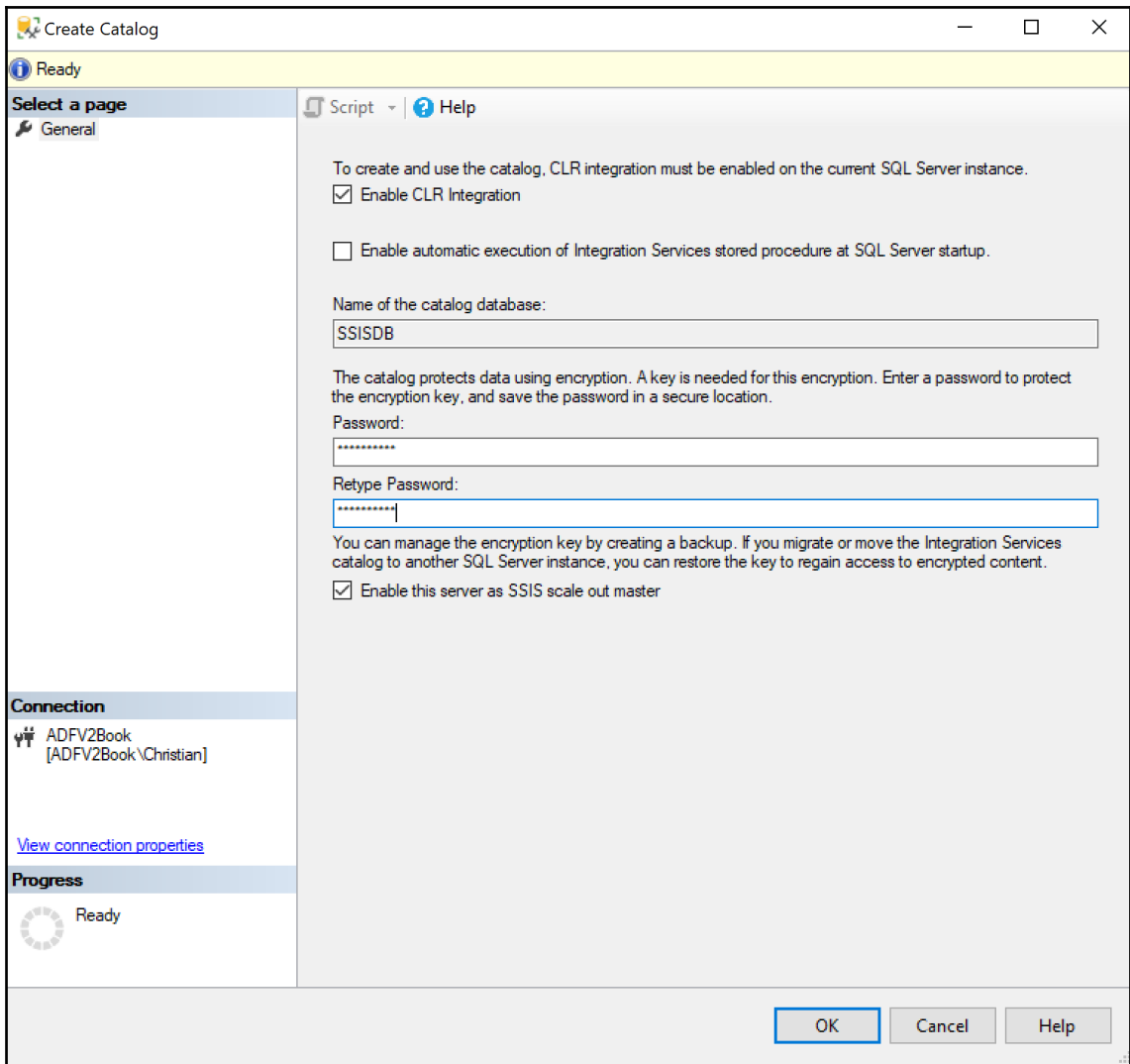
**Results**

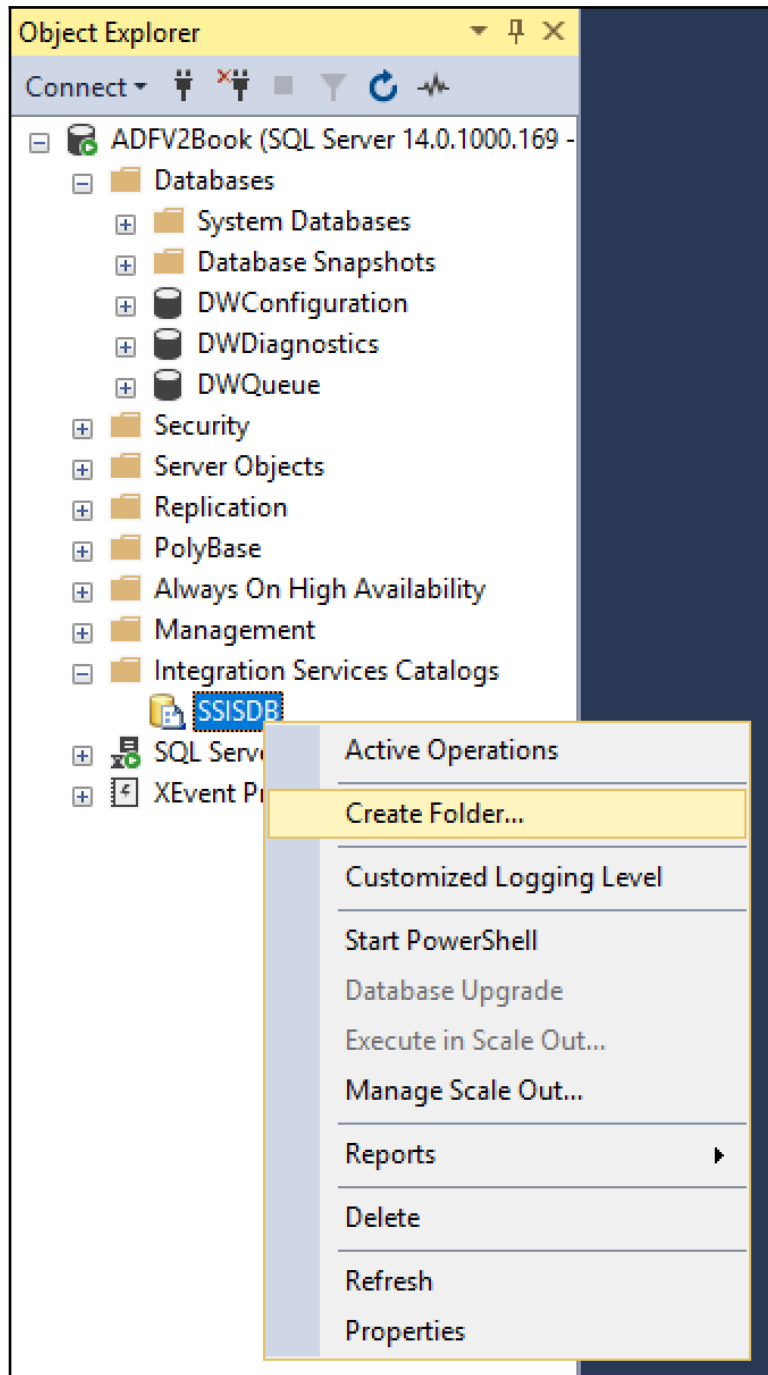
**✓ Operation Complete**

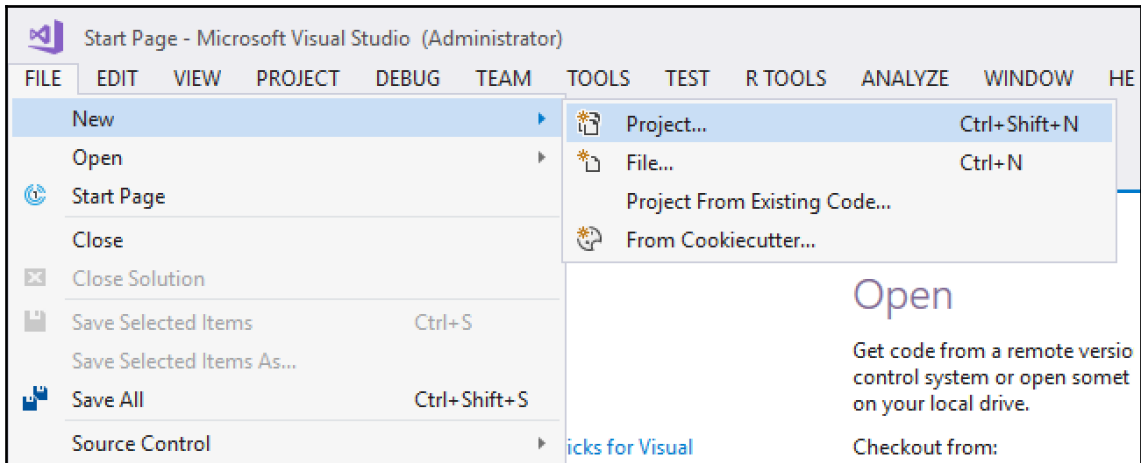
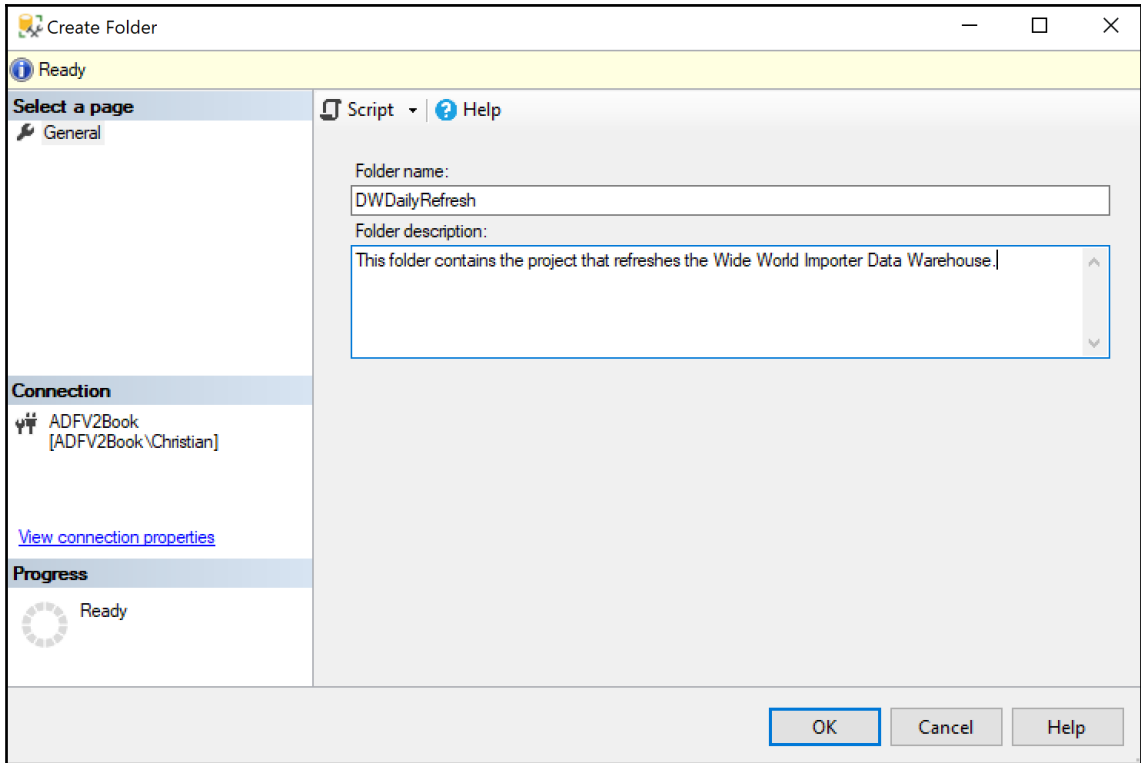
Summary:

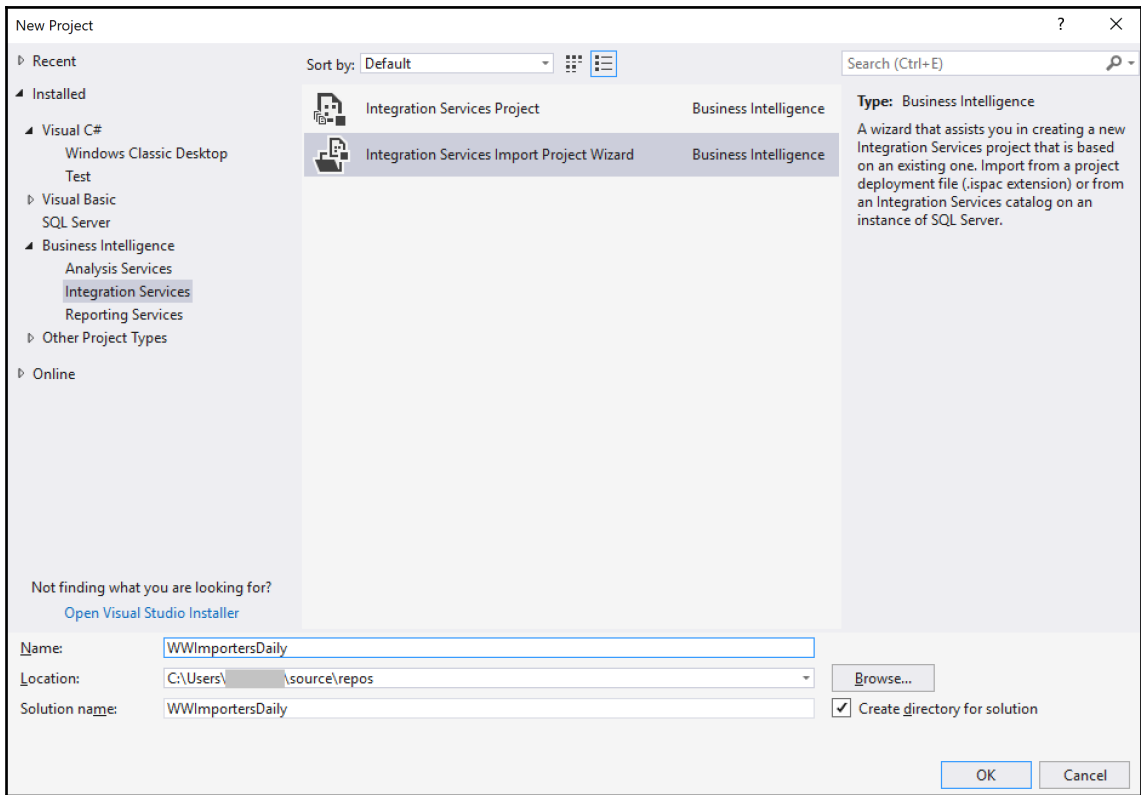
Name	Result
✓ Initializing deployment	<a href="#">Success</a>
✓ Verifying deployment plan	<a href="#">Success</a>
✓ Analyzing deployment plan	<a href="#">Success</a>
✓ Importing package schema and data into database	<a href="#">Success</a>
✓ Updating database	<a href="#">Success</a>
✓ Importing data	<a href="#">Success</a>
✓ Processing Import.	<a href="#">Success</a>
✓ Disabling indexes.	<a href="#">Success</a>
✓ Disabling index 'FK_Application_Cities_StateProvinceID'.	<a href="#">Success</a>
✓ Disabling index 'UQ_Application_Countries_CountryName'.	<a href="#">Success</a>
✓ Disabling index 'UQ_Application_Countries_FormalName'.	<a href="#">Success</a>
✓ Disabling index 'UQ_Application_DeliveryMethods_DeliveryMethodName'.	<a href="#">Success</a>
✓ Disabling index 'UQ_Application_PaymentMethods_PaymentMethodName'.	<a href="#">Success</a>
✓ Disabling index 'IX_Application_People_FullName'.	<a href="#">Success</a>
✓ Disabling index 'IX_Application_People_IsEmployee'.	<a href="#">Success</a>
✓ Disabling index 'IX_Application_People_IsSalesperson'.	<a href="#">Success</a>
✓ Disabling index 'IX_Application_People_Perf_20160301_05'.	<a href="#">Success</a>
✓ Disabling index 'FK_Application_StateProvinces_CountryID'.	<a href="#">Success</a>
✓ Disabling index 'IX_Application_StateProvinces_SalesTerritory'.	<a href="#">Success</a>
✓ Disabling index 'UQ_Application_StateProvinces_StateProvinceName'.	<a href="#">Success</a>

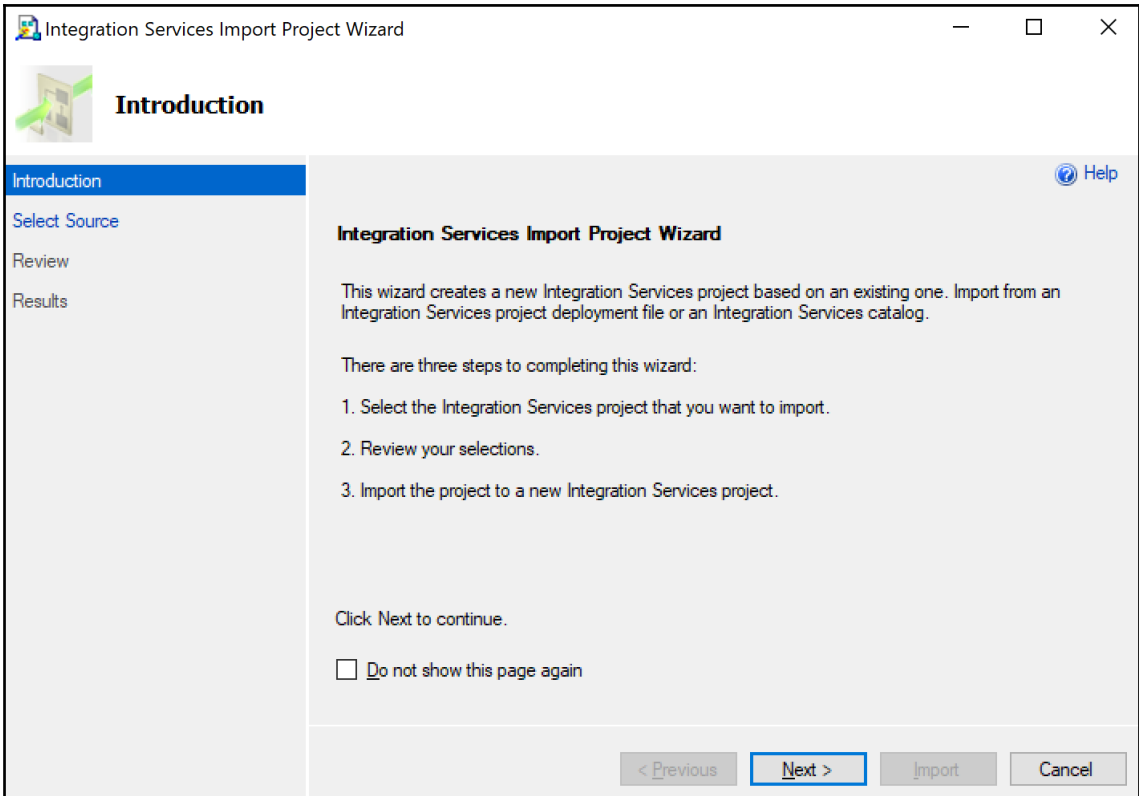
< Previous    Next >    **Close**



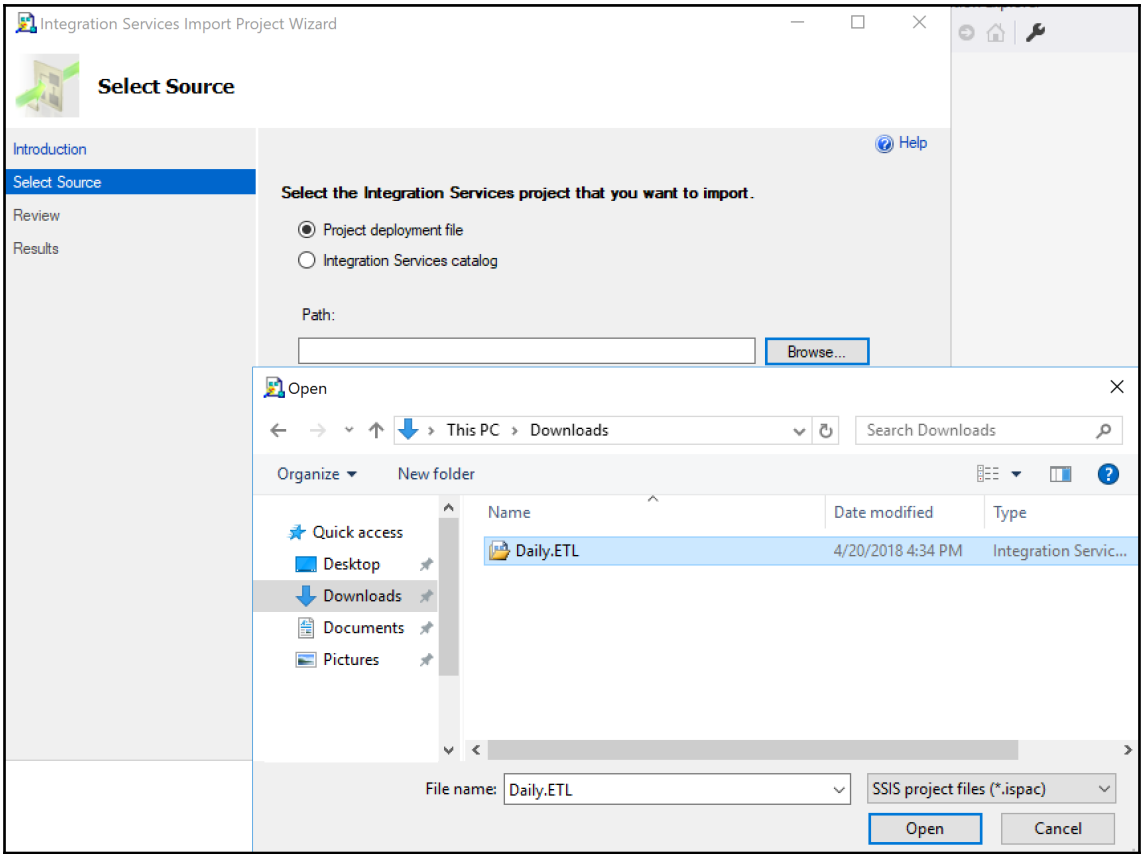


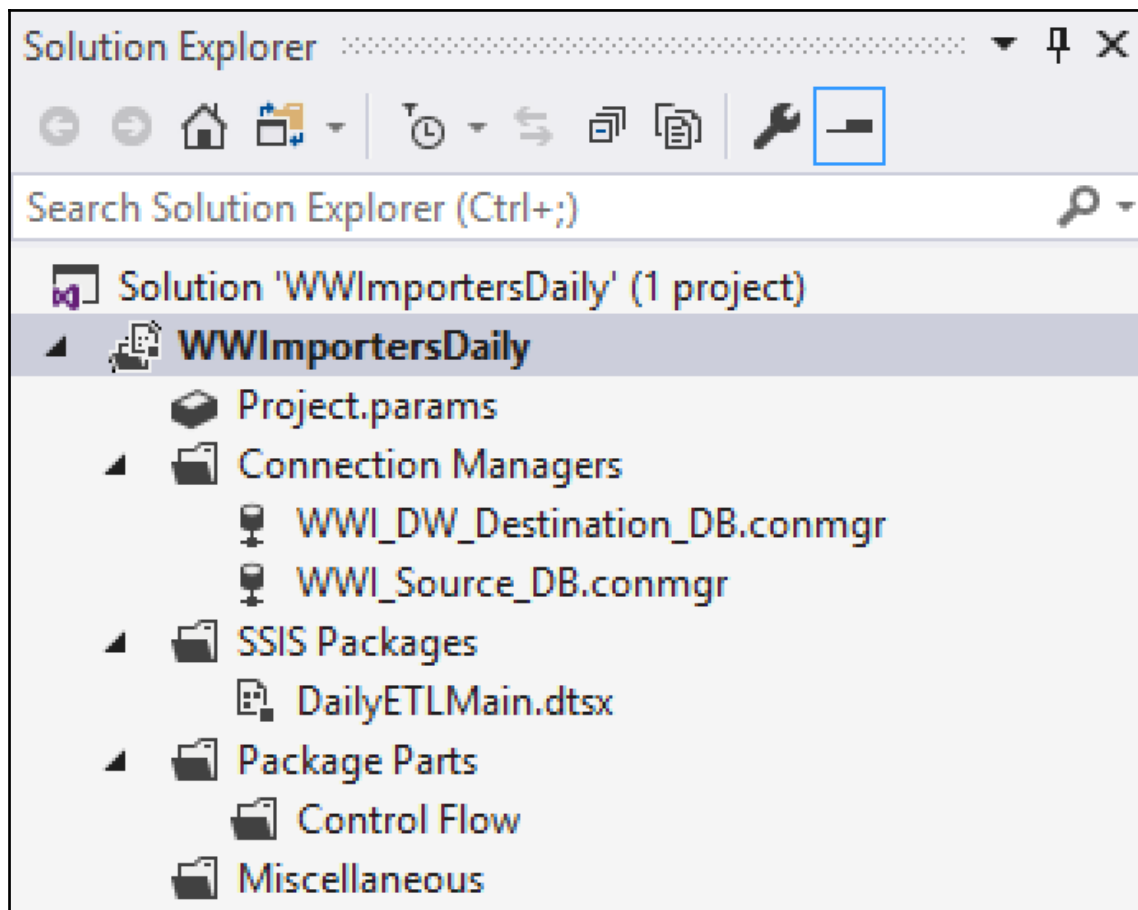


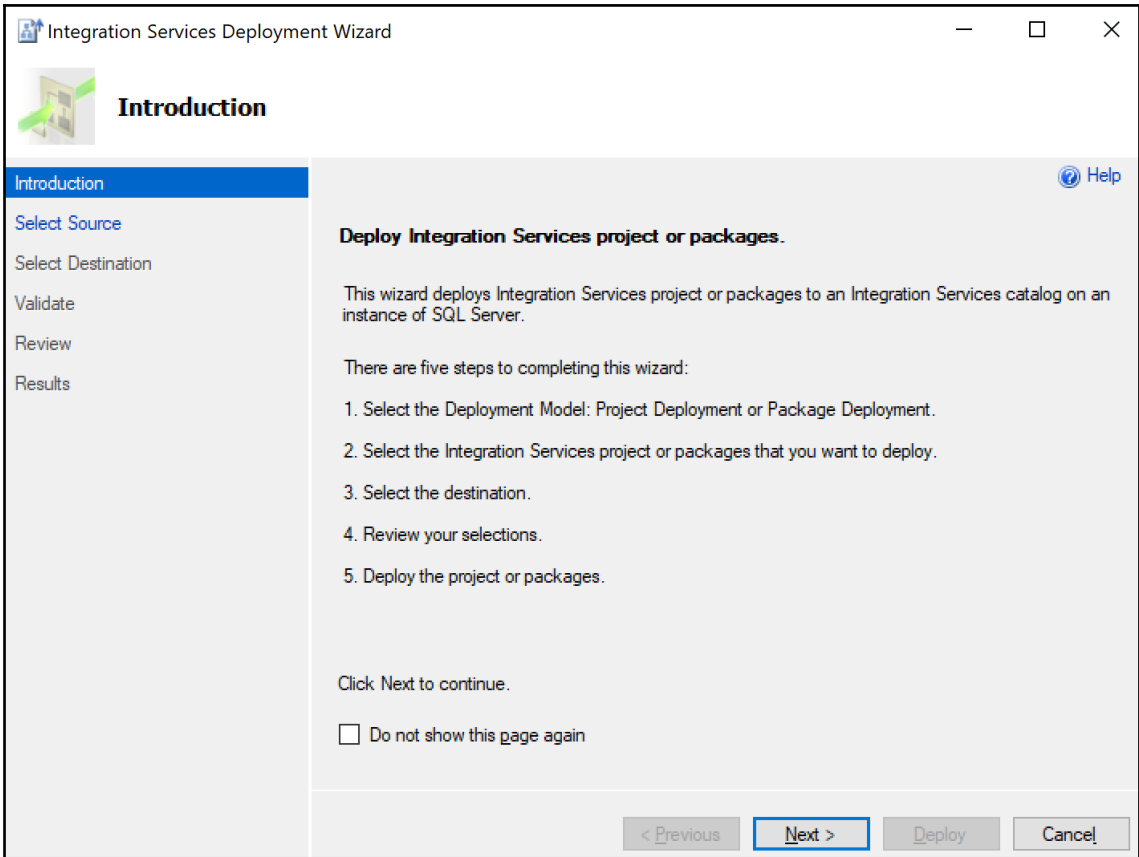


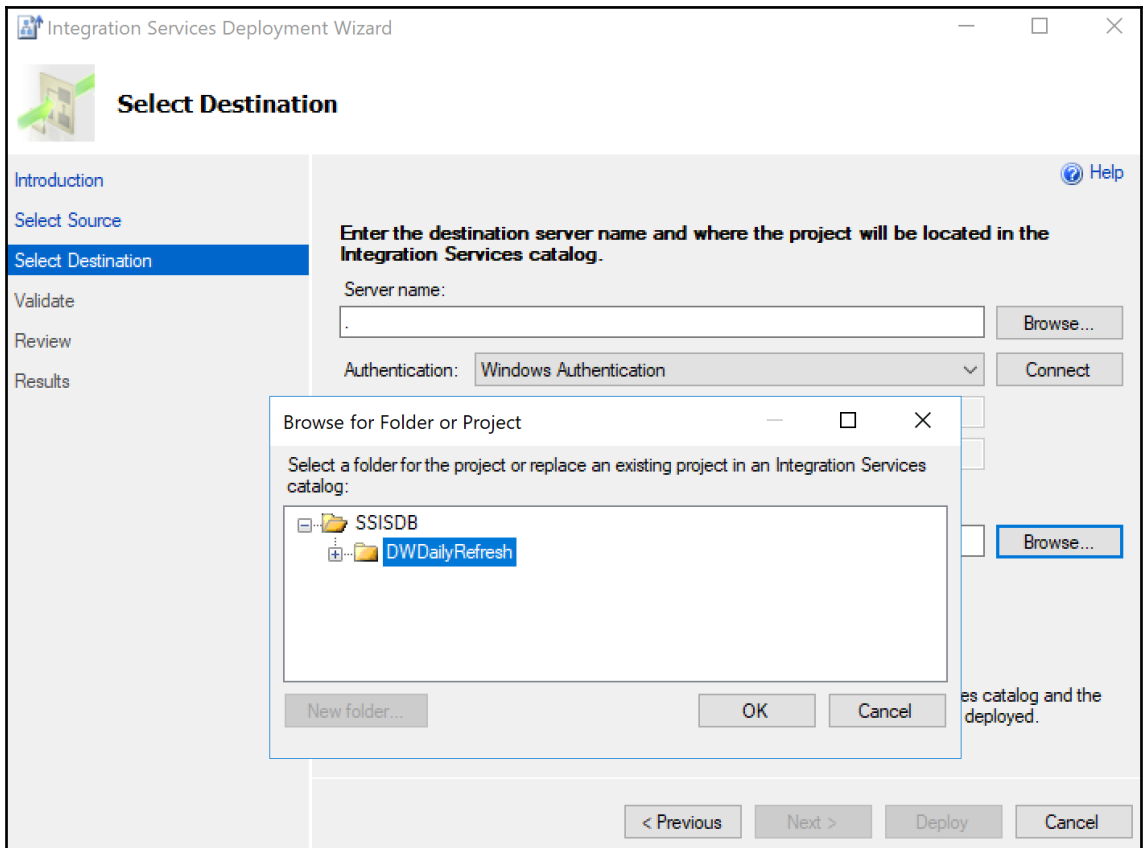


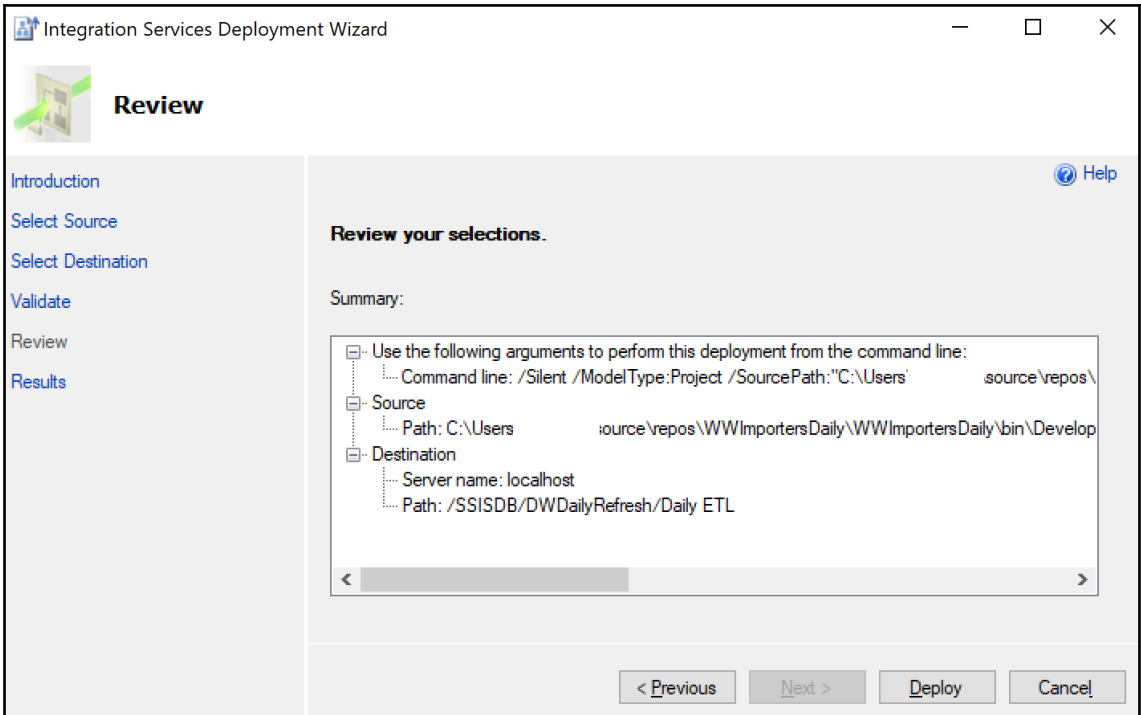












Integration Services Deployment Wizard

## Results

Introduction  
Select Source  
Select Destination  
Validate  
Review  
Results

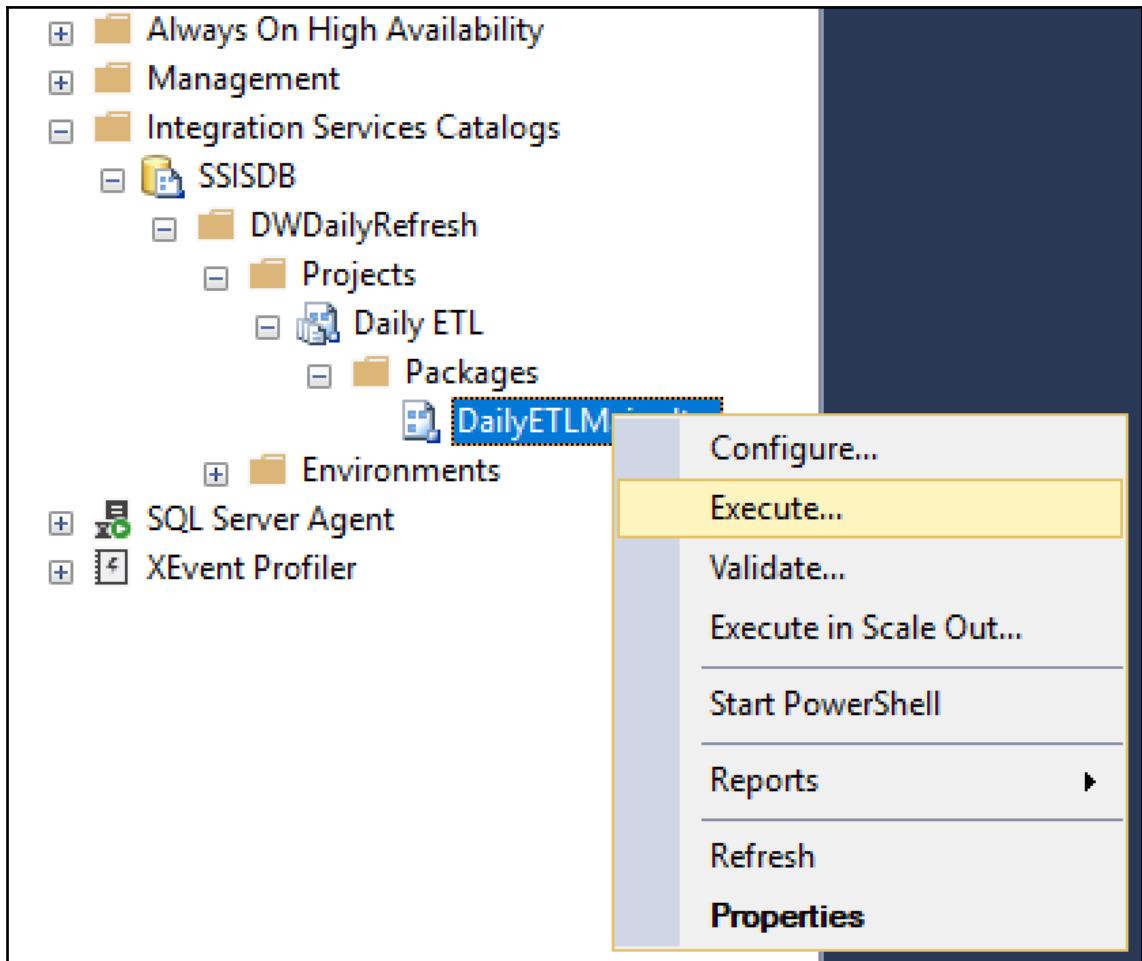
Help

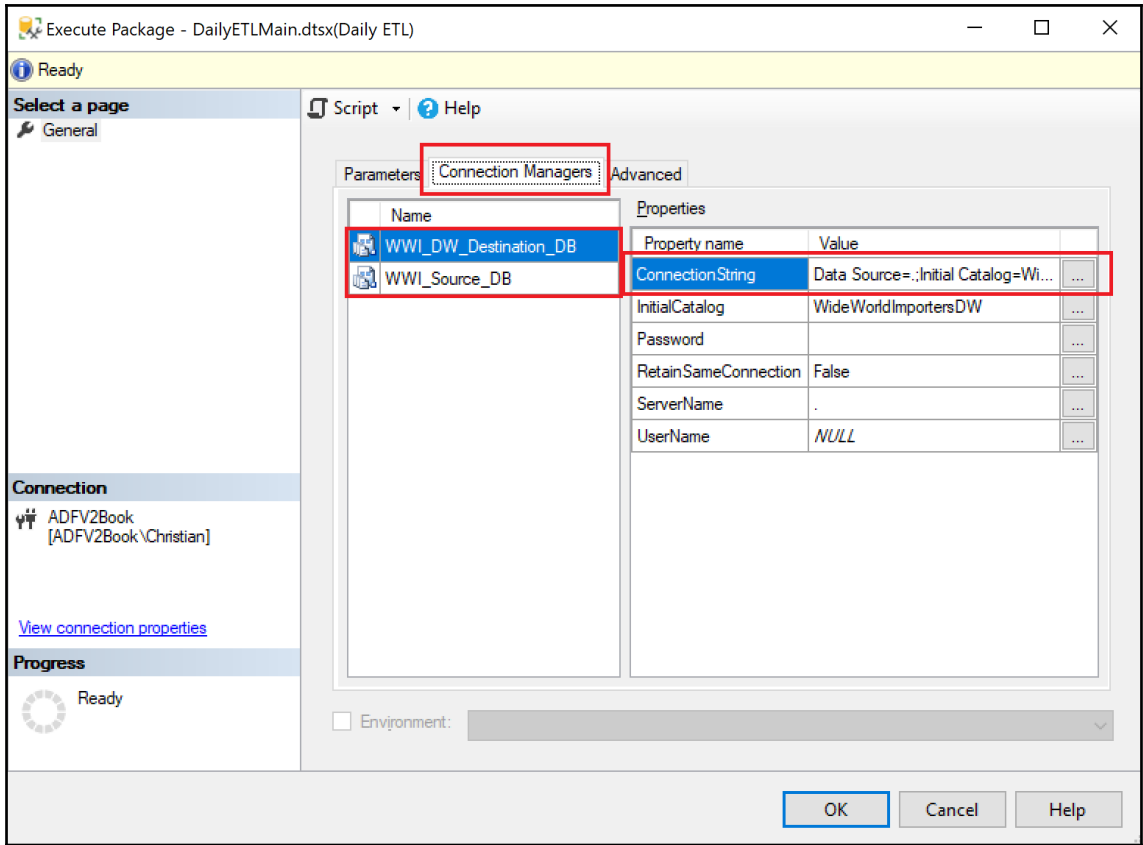
### Results

Action	Result
✓ Loading project	<a href="#">Passed</a>
✓ Connecting to destination server	<a href="#">Passed</a>
✓ Changing protection level	<a href="#">Passed</a>
✓ Deploying project	<a href="#">Passed</a>

Save Report...

< Previous   Next >   **Close**   Cancel







Operation ID	6
Package	DWDailyRefresh\Daily ETL\DailyETLMain.dtsx
Environment	-
Status	Succeeded
Machine	

Execution Overview

Filter: Result: All; (3 more)

Result	Duration (sec)	Package Name	Task Name	Execution Path
Succeeded	7.859	DailyETLMain.dtsx	DailyETLMain	<a href="#">\DailyETLMain</a>
Succeeded	0.016	DailyETLMain.dtsx	Calculate ETL Cutoff Time backup	<a href="#">\DailyETLMain\Calculate ETL Cutoff Time backup</a>
Succeeded	0.109	DailyETLMain.dtsx	Ensure Date Dimension	<a href="#">\DailyETLMain\Ensure Date</a>

The screenshot shows the Azure Data Factory portal interface. On the left sidebar, the 'Connections' tab is selected and highlighted with a red box. The main pane displays the 'Integration Runtimes' section, also highlighted with a red box. Below this, a table lists the integration runtime details:

Name	Actions	Type	Sub-type	Status	Region
defaultIntegrationRuntime		Azure	...	Running	Auto Resolve

## Integration Runtime Setup



Integration Runtime is the native compute used by ADF to execute or dispatch activities. Choose what integration runtime to create based on required capabilities.



Perform data movement and dispatch activities to external computes.



Lift-and-shift existing SSIS packages to execute in Azure.

Cancel

Next →

## Integration Runtime Setup

**General Settings**

Name \* (i)

Description (i)

Type

Location \* (i)

Node Size \* (i)

Node Number \* (i)

Edition/License \* (i)

Save Money

Save with a license you already own. Already have a SQL Server license?

By selecting "yes", I confirm I have a SQL Server license with Software Assurance to apply this [Azure Hybrid Benefit for SQL Server](#).

## Integration Runtime Setup ✕

**SQL Settings**

Subscription \* i

[Redacted]

Location i

East US

Catalog Database Server Endpoint \* i

adf2book.database.windows.net

Use AAD authentication with your ADF MSI i

Admin Username \* i

ADFV2User

Admin Password \* i

.....










Catalog Database Service Tier \* i

Basic

Allow Azure services to access i

Connection successful

Cancel ← Previous Next →

Connections X						
Linked Services		Integration Runtimes				
Name ↓	Actions	Type ↓	Sub-type	Status	Region	
SSISIR	   	Azure-SSIS	---		Starting	East US
defaultIntegrationRuntime	  	Azure	---		Running	Auto Resolve

Connect to Server X

SQL Server

Login Connection Properties Additional Connection Parameters

Server \_\_\_\_\_

Type the server name, or choose it from the drop-down list.

Server type: Database Engine ▼

Server name: adf2book.database.windows.net ▼

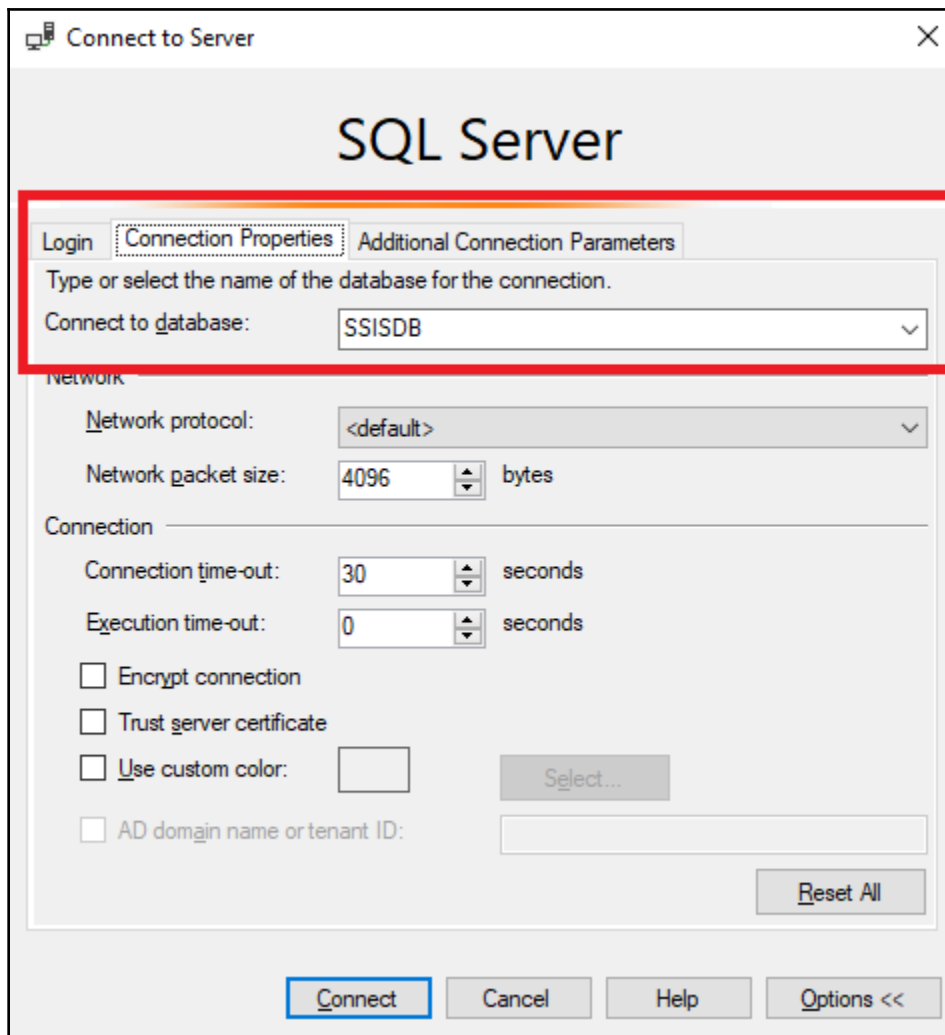
Authentication: SQL Server Authentication ▼

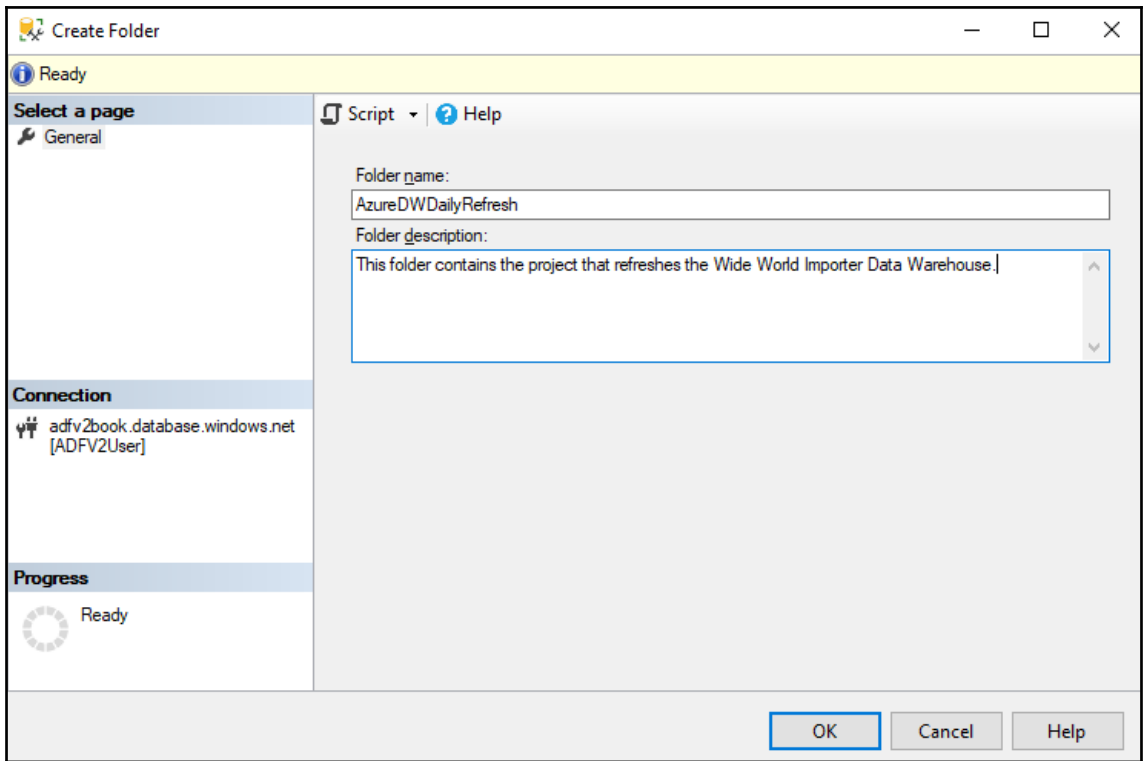
  Login: ADFV2User ▼

  Password: \*\*\*\*\*

Remember password

Connect
Cancel
Help
Options <<





Integration Services Deployment Wizard

## Select Destination

Introduction  
Select Source  
**Select Destination**  
Validate  
Review  
Results

Help

**Enter the destination server name and where the project will be located in the Integration Services catalog.**

Server name:

Authentication:

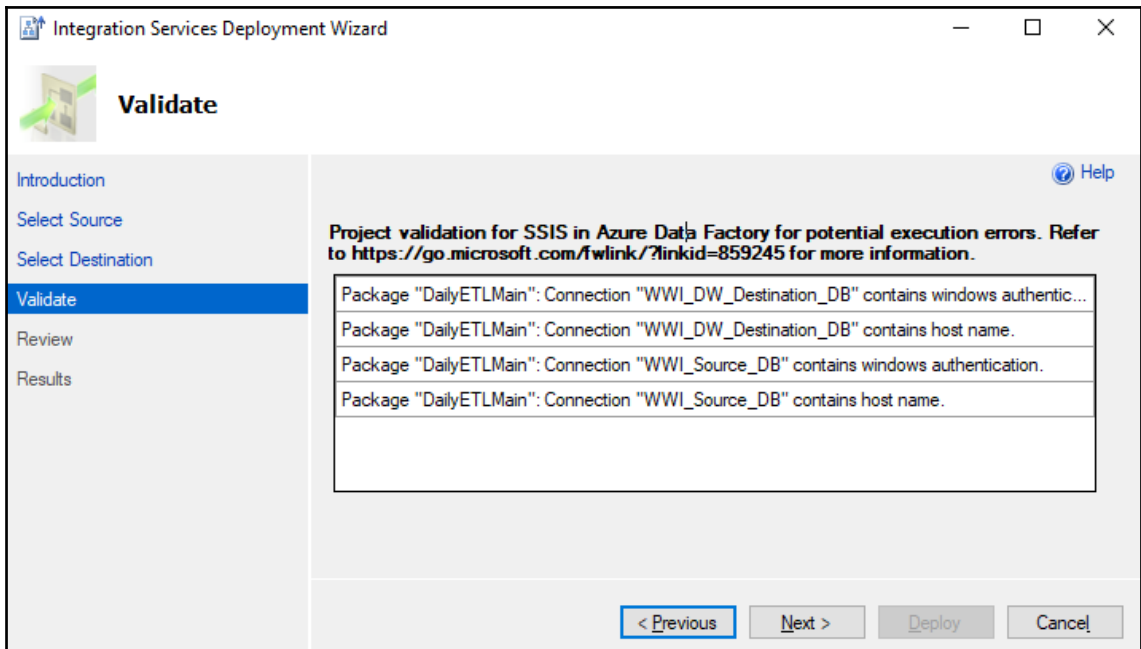
Login:

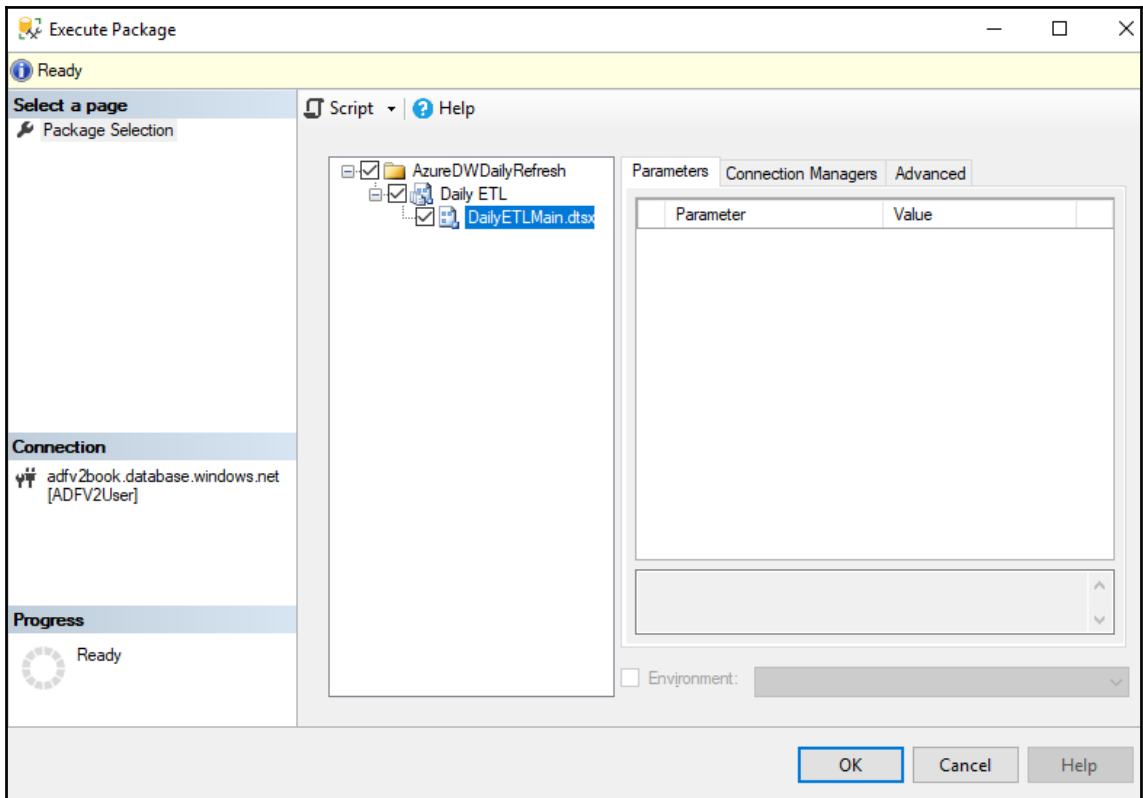
Password:

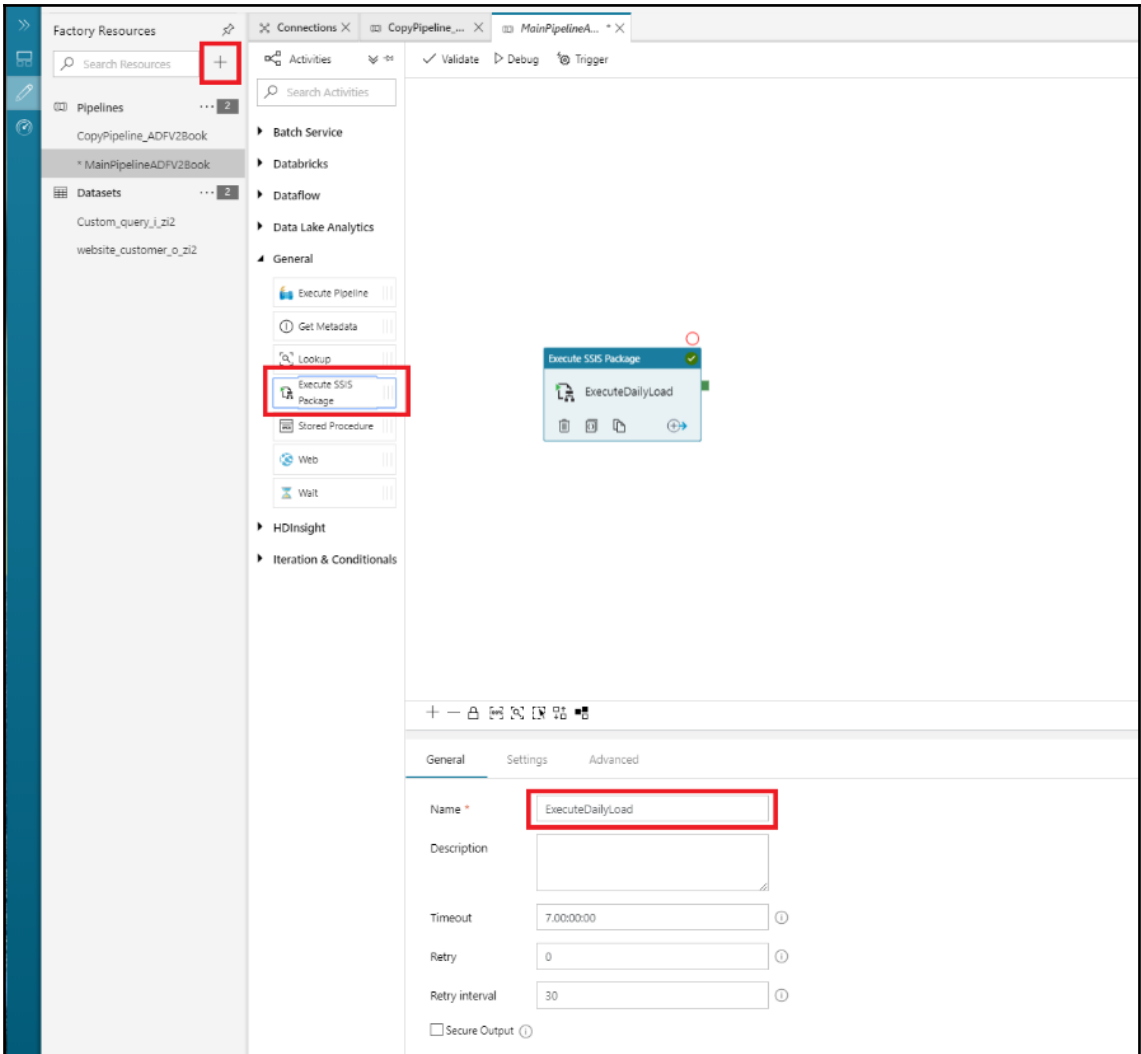
Path:

< Previous   Next >   Deploy   Cancel









General **Settings** Advanced

Azure-SSIS IR \*  ⓘ

32-Bit runtime  ⓘ

Logging level \*  ⓘ  
 Customized ⓘ

Package path \*  ⓘ

Environment path  ⓘ

General Parameters **Output**

Pipeline Run ID: **ea165ea9-4699-4d00-bd57-433ef677ce8c** ⓘ

Name	Type	Run Start	Duration	Status	Actions	RunID
ExecuteDailyLoad	ExecuteSSISPackage	04/29/2018 6:28 AM	-00:00:00	<span style="color: orange;">⚠</span> In Progress	↗	7ad7e571-8ab4-4c7e-974e-7101545c92d8

General Parameters **Output**

Pipeline Run ID: **5680b289-7129-4f3b-b556-13e71e4cf8da** ⓘ

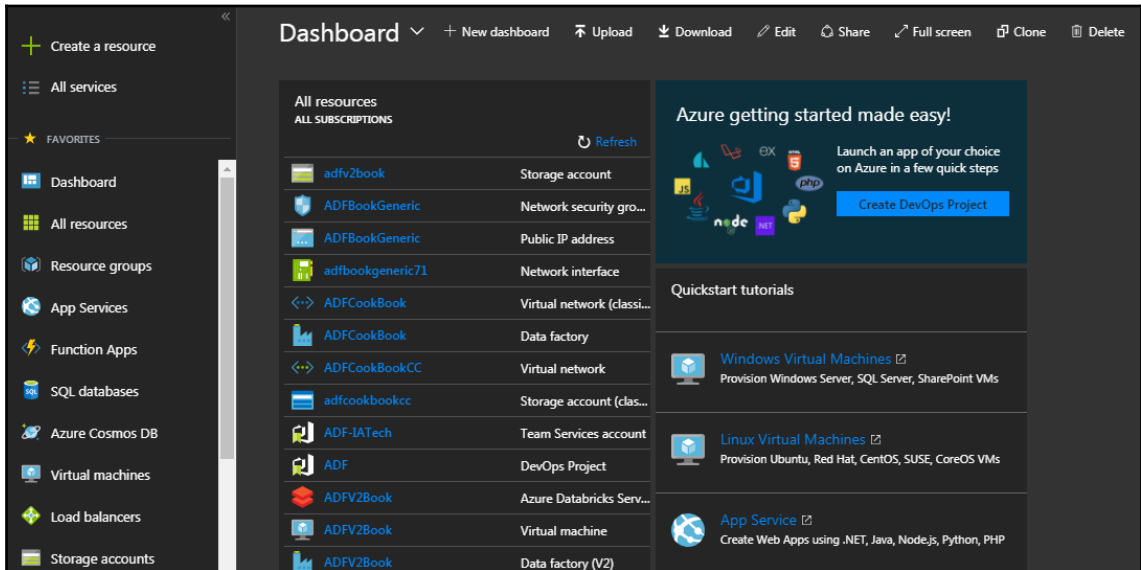
Name	Type	Run Start	Duration	Status	Actions	RunID
ExecuteDailyLoad	ExecuteSSISPackage	04/29/2018 6:42 AM	00:01:24	<span style="color: green;">✔</span> Succeeded	↗ ↘	11bf0a5b-9827-42a7-bcfd-70d330ca9f0e

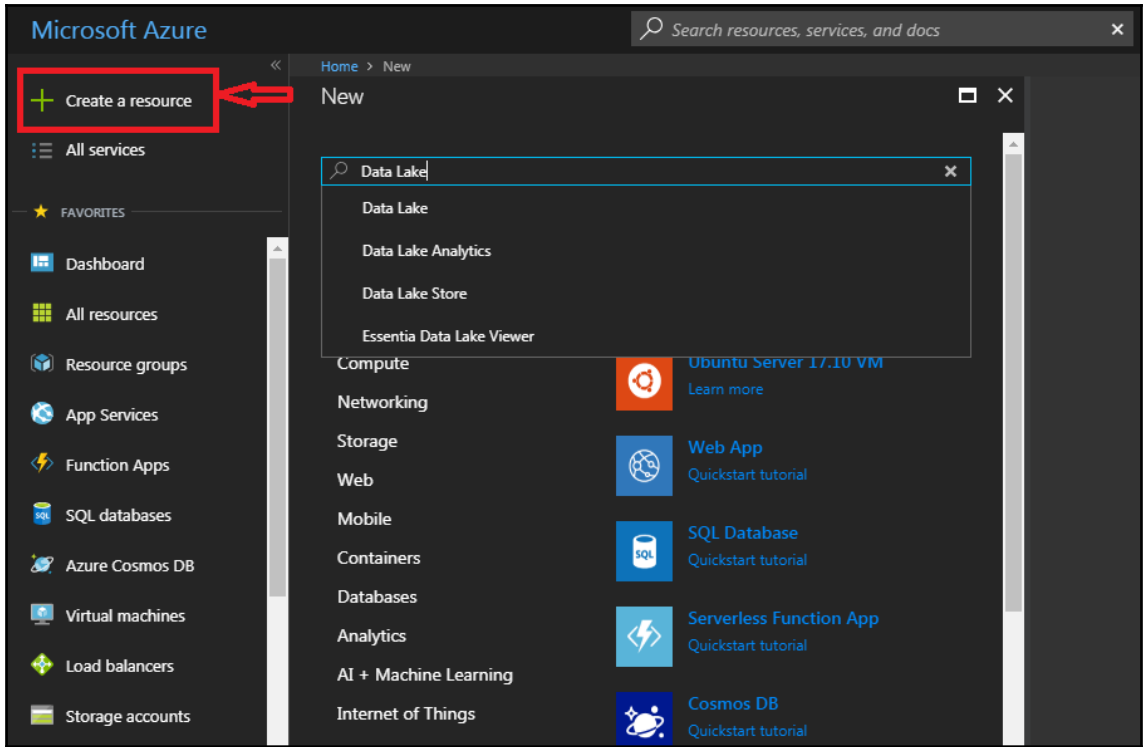
↑ Publish All ⓘ
✔ Validate All
🔄 Refresh
🗑 Discard All
🛡 ARM Template ▾

🔗 Connections ✕
🔊 MainPipelineA... \* ✕

🔗 Activities
↕ ↔
✔ Validate
▶ Debug
⚙ Trigger

# Chapter 4: Azure Data Lake





Home > New > Data Lake Store

## Data Lake Store

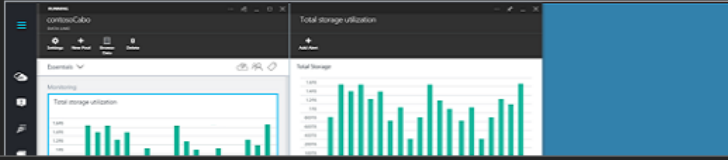
Microsoft

Azure Data Lake Store is designed to be an enterprise-wide, hyper scale repository for big data analytic workloads. In the data lake, you can easily capture data of any size, type and speed in a single place for the purposes of operational and exploratory analytics.

Highlights:

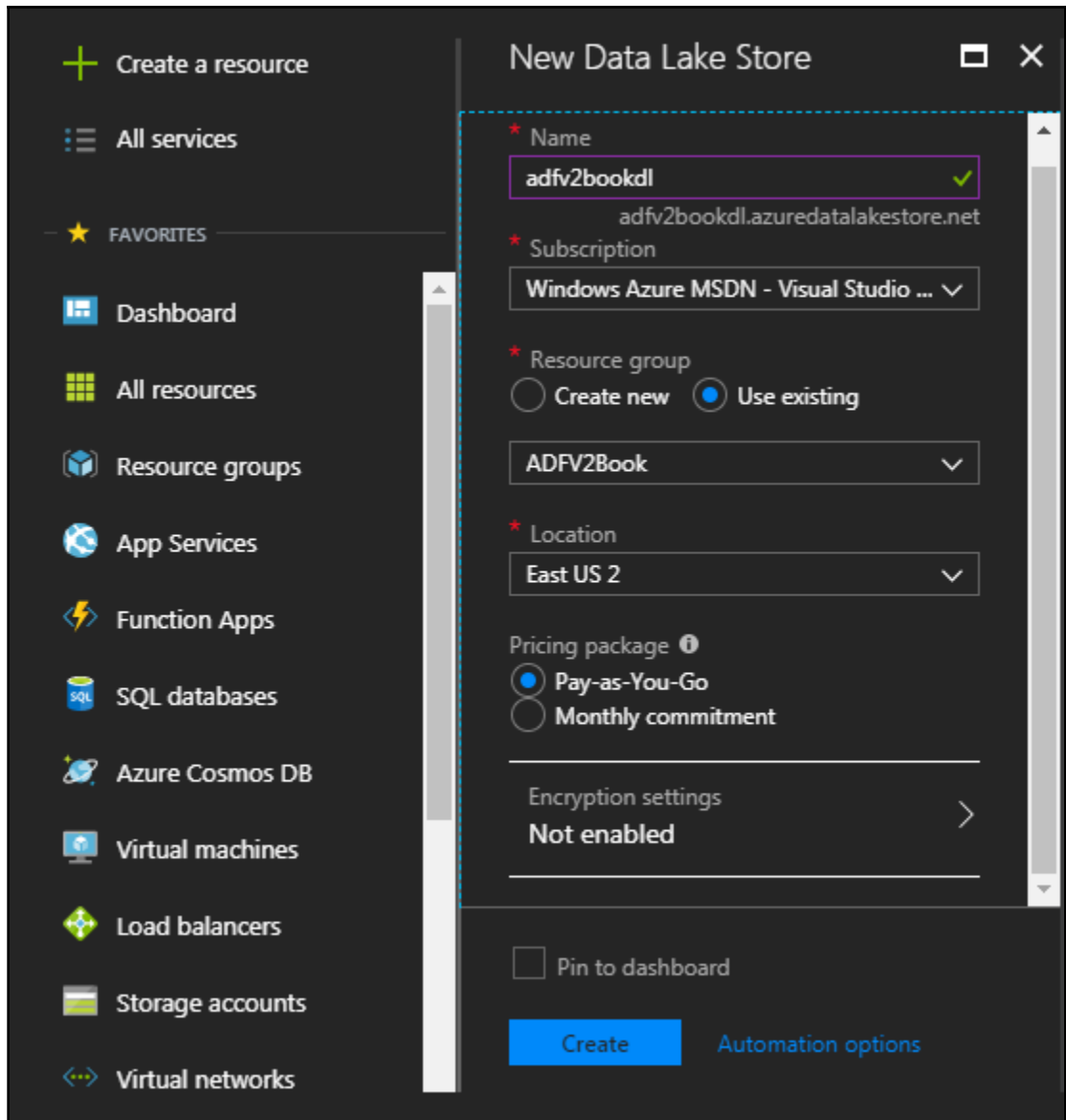
- Built for Hadoop: A Hadoop Distributed File System for the Cloud
- Unlimited storage: No fixed limits on file size, account size, or the number of files
- Performance Tuned for Big Data: Optimized for massive throughput to query and analyze any amount of data
- Enterprise Grade Security: Azure Active Directory authentication and role-based access control
- All Data: Store data in its native format without prior transformation

[Save for later](#)



[Create](#)

The screenshot shows the Azure Data Lake Store page with a left-hand navigation menu. The main content area features a description of the service, a list of highlights, a 'Save for later' button, and a preview of storage utilization charts. The charts show 'Total storage utilization' and 'Total storage' with green bars representing data points. A 'Create' button is visible at the bottom of the page.





Subscriptions: Windows Azure MSDN - Visual Studio Ultimate - Don't see a subscription? [Switch directories](#)

Filter by name... All resource groups All types All locations No grouping

73 items  Show hidden types

NAME	TYPE	RESOURCE GROUP	LOCATION	SUBSCRIPTION
adfV2book	Storage account	ADFV2Book	East US	Windows Azure MSDN - Vis...
ADFV2Book (adfV2book/ADFV2Book)	SQL database	ADFV2Book	East US	Windows Azure MSDN - Vis...
SSISDB (adfV2book/SSISDB)	SQL database	ADFV2Book	East US	Windows Azure MSDN - Vis...
WideWorldImporters-Standard (adfV2book/WideWo--	SQL database	ADFV2Book	East US	Windows Azure MSDN - Vis...
ADFV2Book_OsDisk_1_a26a268b89cb49bea2bc4281...	Disk	ADFV2BOOK	Canada East	Windows Azure MSDN - Vis...
adfV2book8	Network interface	ADFV2Book	Canada East	Windows Azure MSDN - Vis...
adfV2bookdiag376	Storage account	ADFV2Book	East US	Windows Azure MSDN - Vis...
adfV2bookdiag982	Storage account	ADFV2Book	Canada East	Windows Azure MSDN - Vis...
adfV2bookdl	Data Lake Store	ADFV2Book	East US 2	Windows Azure MSDN - Vis...
ADFV2Book-ip	Public IP address	ADFV2Book	Canada East	Windows Azure MSDN - Vis...
ADFV2Book-nsg	Network security group	ADFV2Book	Canada East	Windows Azure MSDN - Vis...

Create a resource

All services

FAVORITES

- Dashboard
- All resources
- Resource groups
- App Services
- Function Apps
- SQL databases
- Azure Cosmos DB
- Virtual machines
- Load balancers
- Storage accounts

adfV2bookdl  
Data Lake Store

Search (Ctrl+F)

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems

SETTINGS

- Encryption
- Firewall
- Pricing tier
- Properties
- Locks
- Automation script

Delete Data explorer

Essentials

adfV2bookdl

Pay-as-You-Go  Estimated cost to date

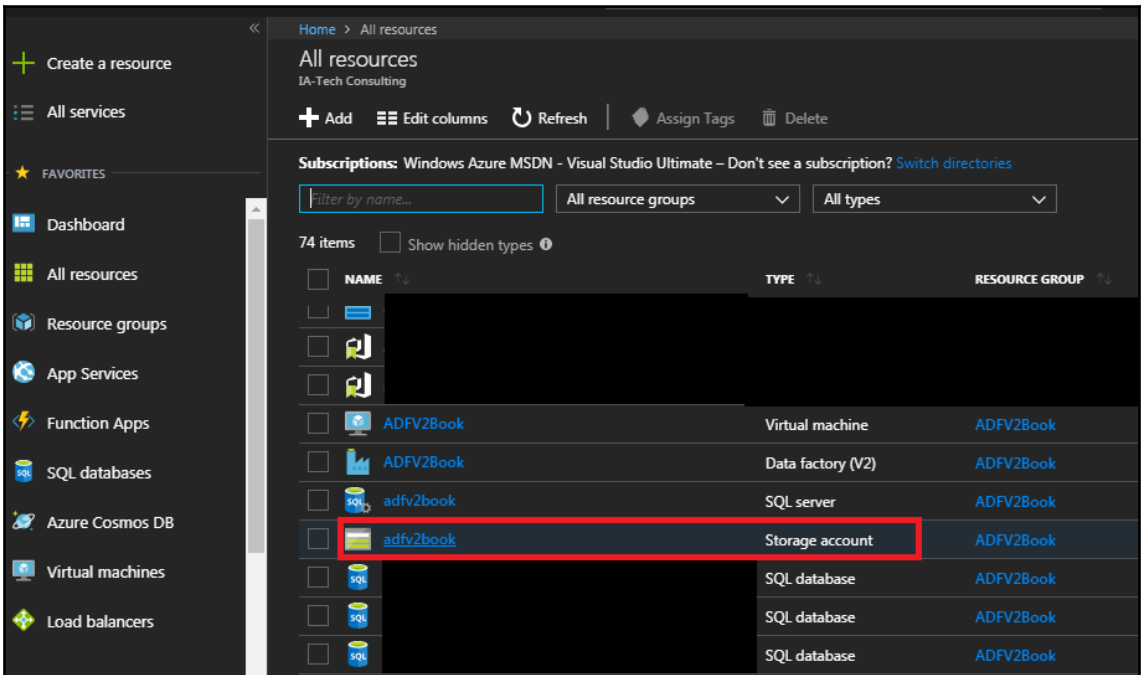
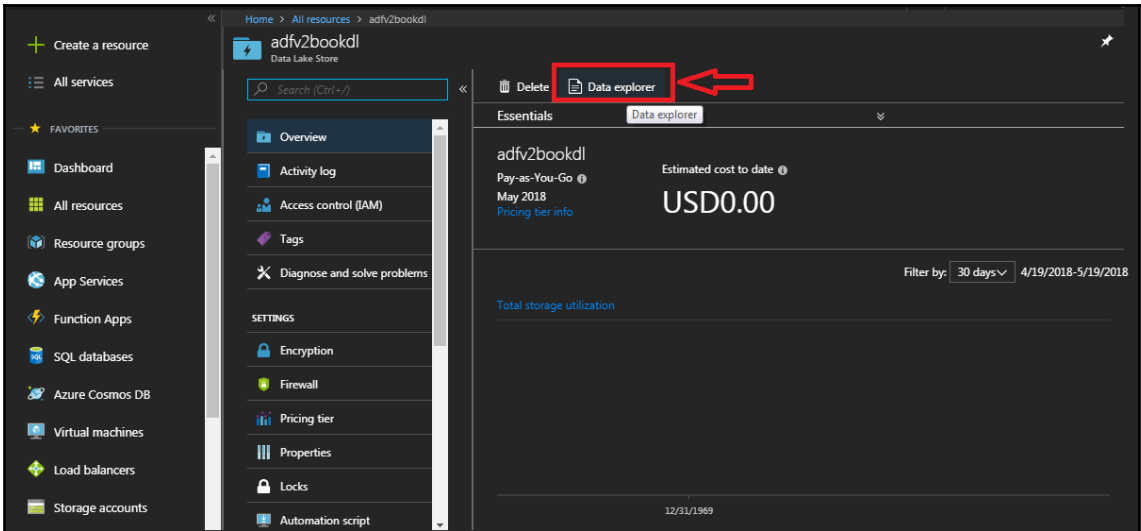
May 2018 **USD0.00**

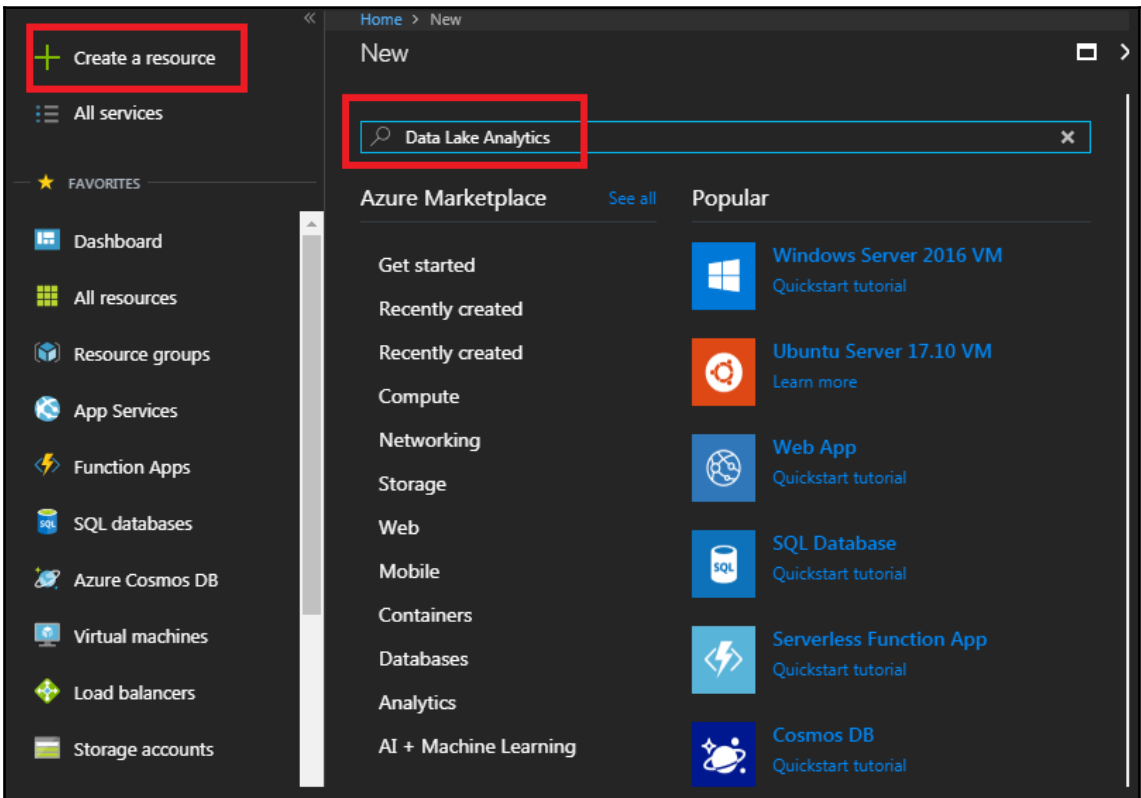
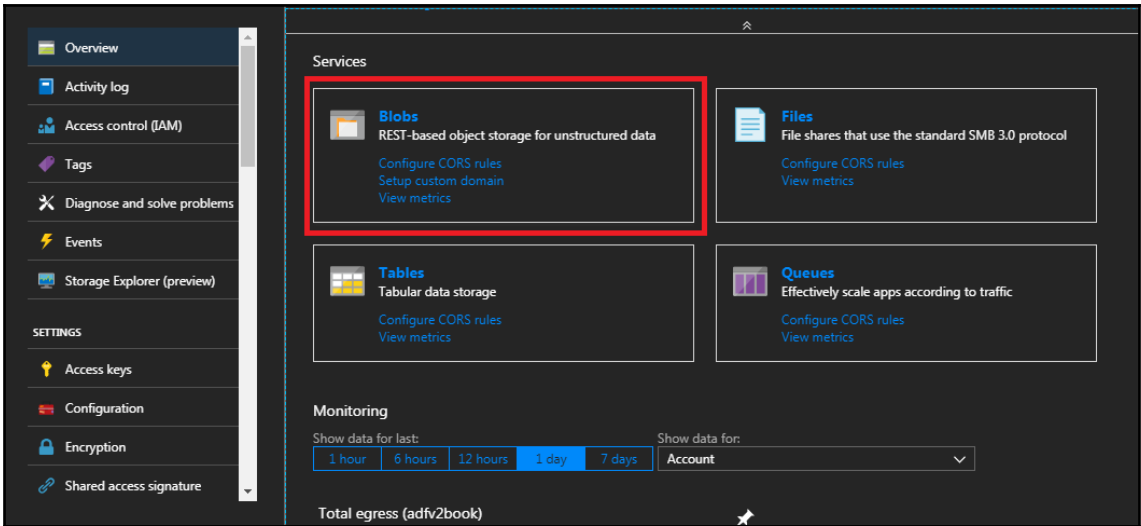
[Pricing tier info](#)

Filter by: 30 days 4/19/2018-5/19/2018

Total storage utilization

12/31/1969





**Data Lake Analytics**  
Microsoft

The Azure Data Lake Analytics service was architected from the ground up for cloud scale and performance. It takes away the complexities normally associated with big data in the cloud and ensures that Data Lake Analytics will meet your current and future business needs.

**Highlights:**

- Analyze any kind of data of any size
- Only pay for the processing power that you use
- Develop faster, debug and optimize smarter
- Introducing U-SQL: simple, familiar, and extensible
- Managed and supported with an enterprise-grade SLA
- Dynamically scales to match your business priorities
- Built on YARN, designed for the cloud
- Proven at Microsoft with more than 10,000 developers

[Save for later](#)

**Create**

**+** Create a resource

**☰** All services

**★** FAVORITES

- 🏠** Dashboard
- 🗪** All resources
- 📁** Resource groups
- 🌐** App Services
- ⚡** Function Apps
- 🗄️** SQL databases
- 🌌** Azure Cosmos DB
- 🖥️** Virtual machines
- ⚙️** Load balancers
- 📁** Storage accounts

### New Data Lake Analytics acc... ☐ ✕

Name  
 ✓  
adf2bookadla.azuredatalakeanalytics.net

\* Subscription  
 ▾

\* Resource group  
 Create new  Use existing

▾

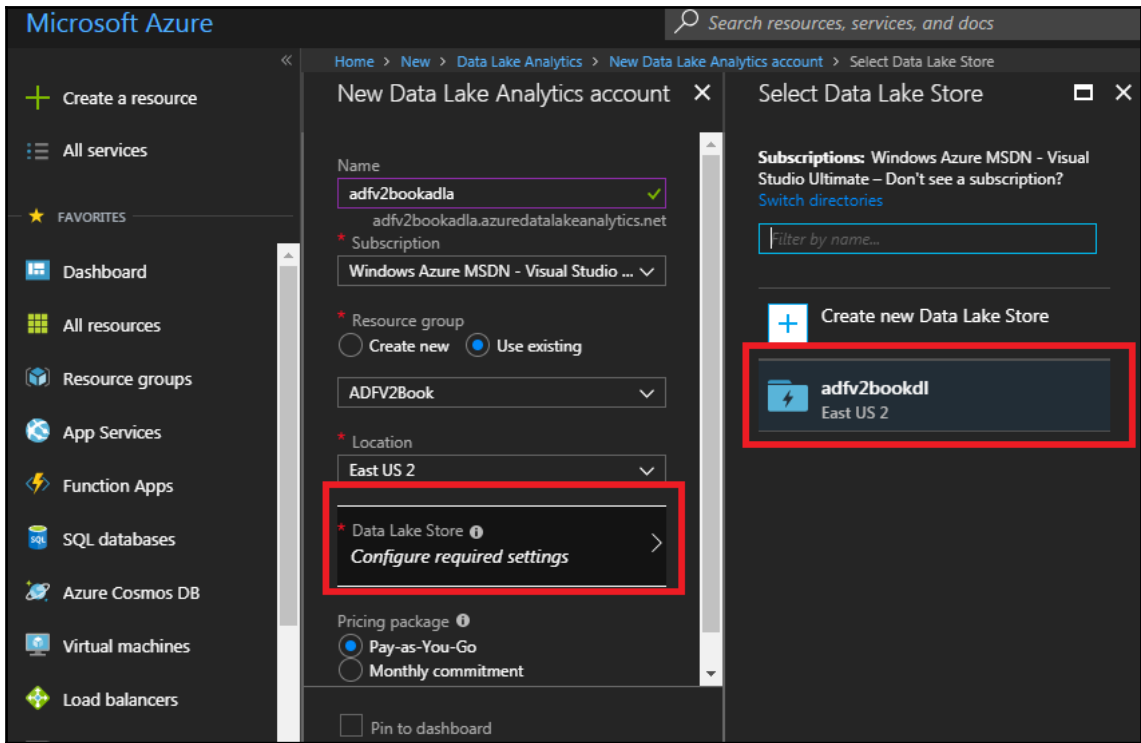
\* Location  
 ▾

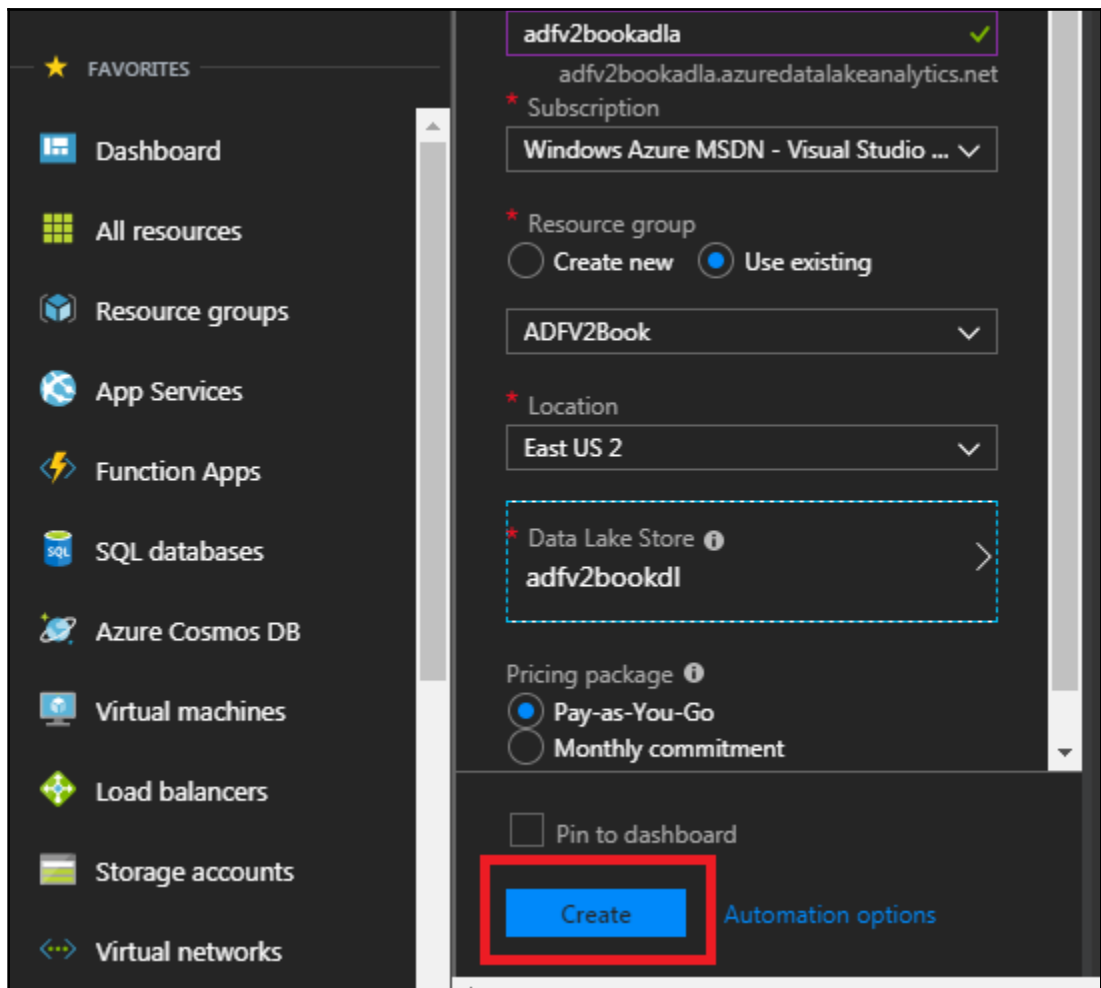
\* Data Lake Store ⓘ ➤  
*Configure required settings*

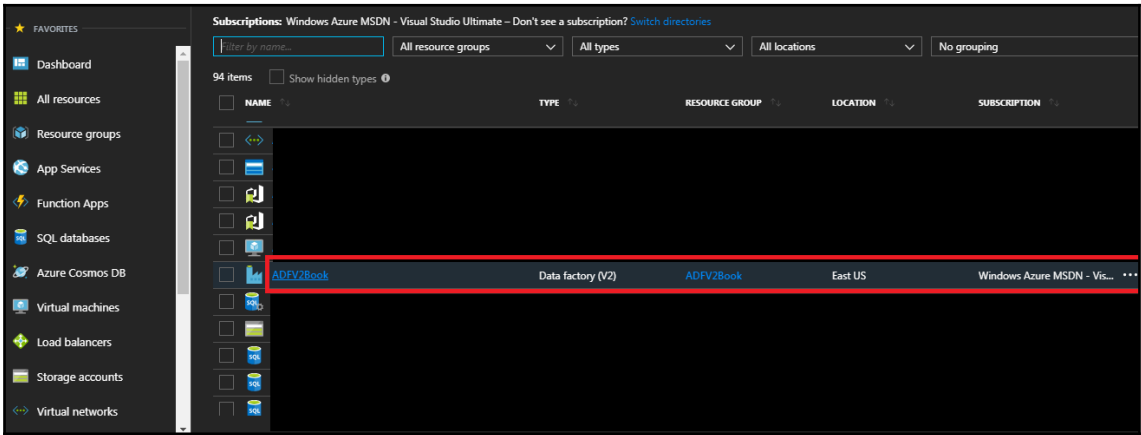
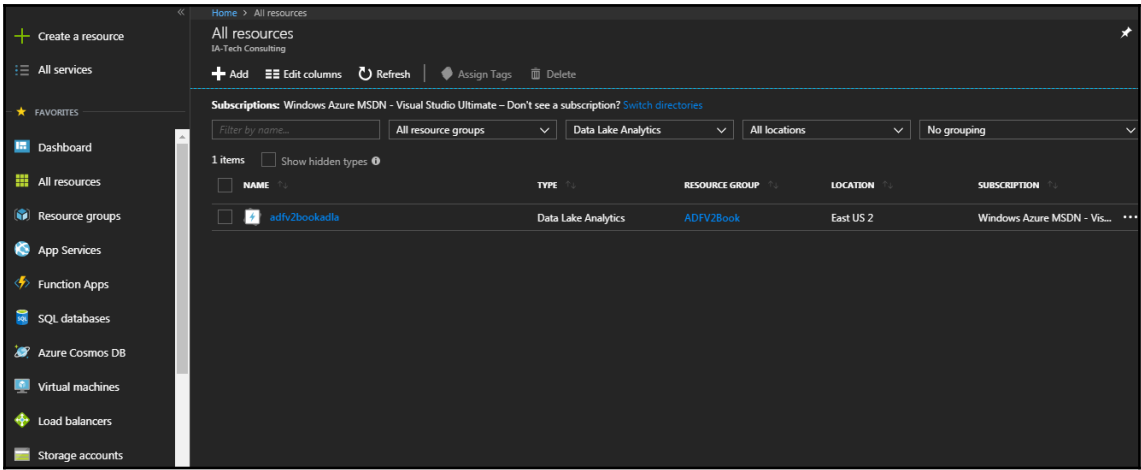
Pricing package ⓘ  
 Pay-as-You-Go  
 Monthly commitment

Pin to dashboard

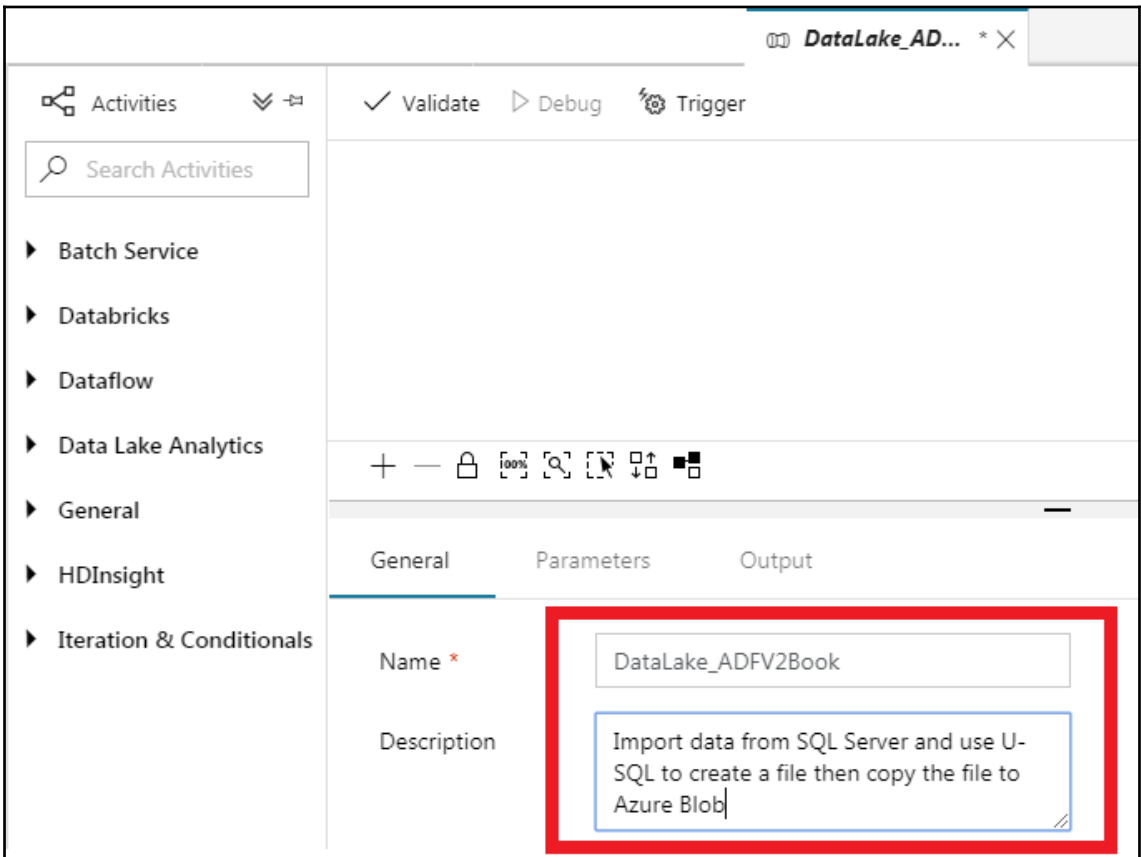
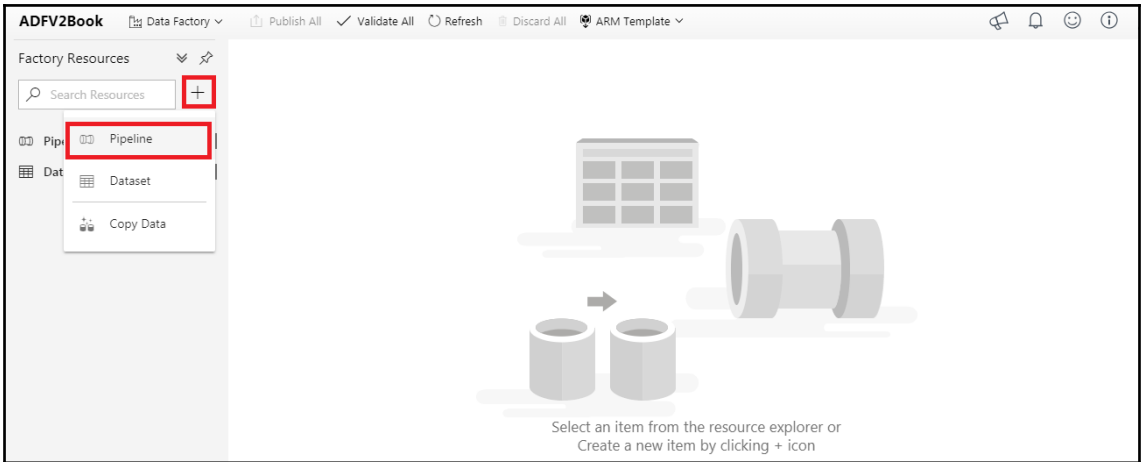
**Create** [Automation options](#)

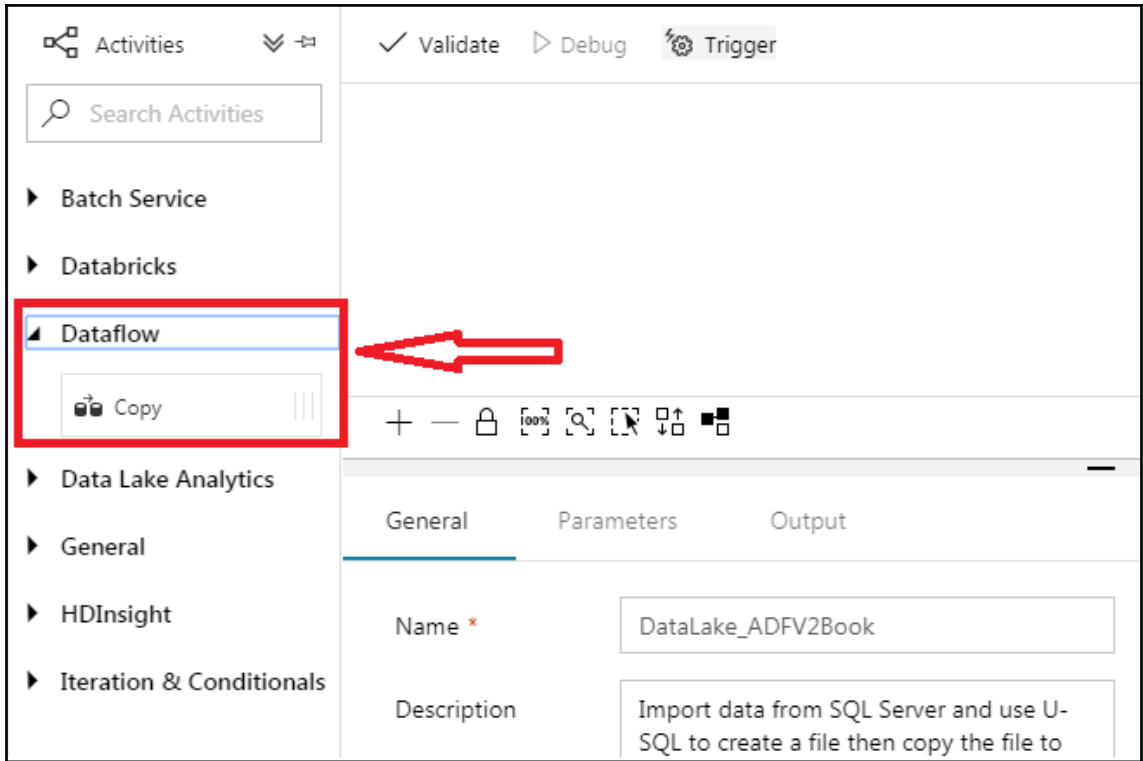


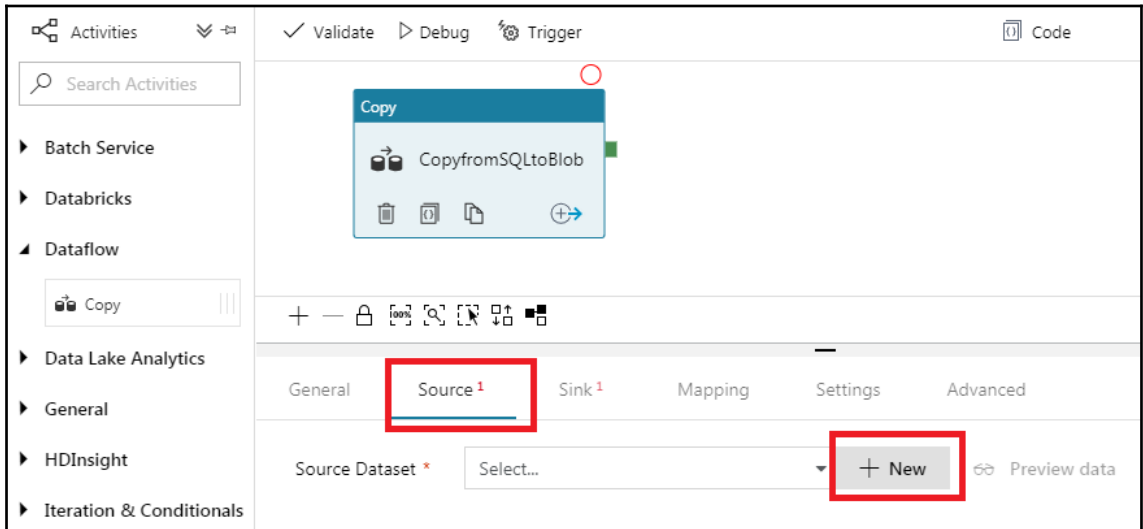
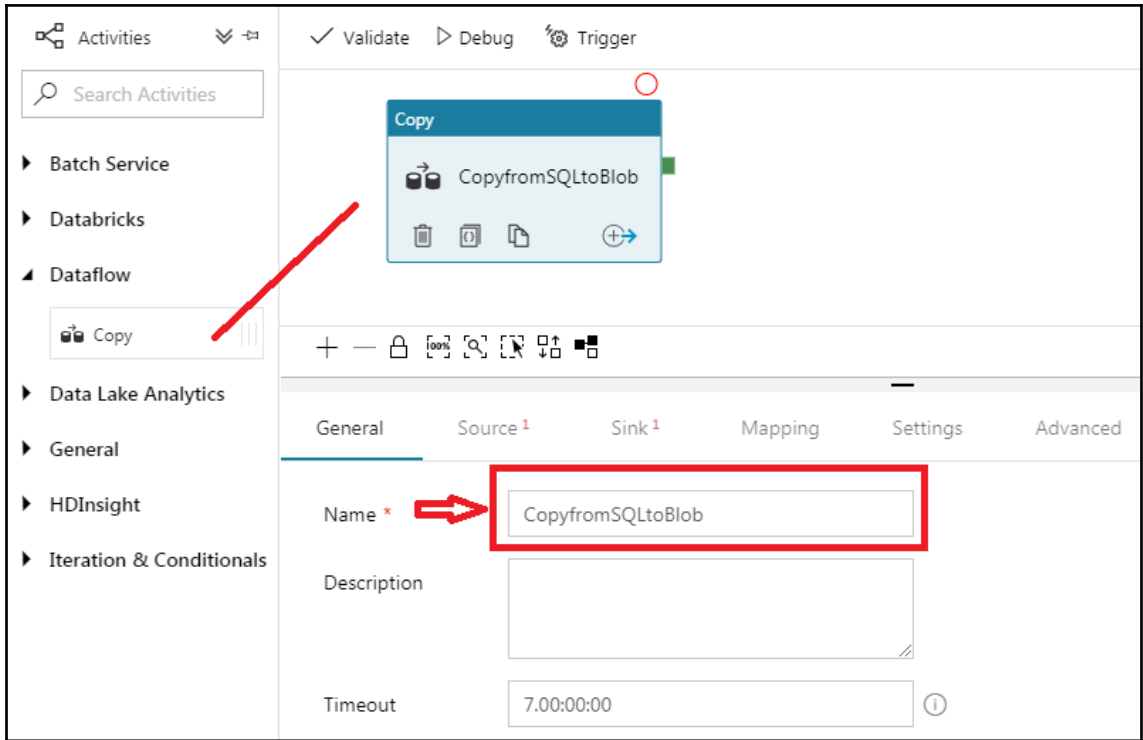


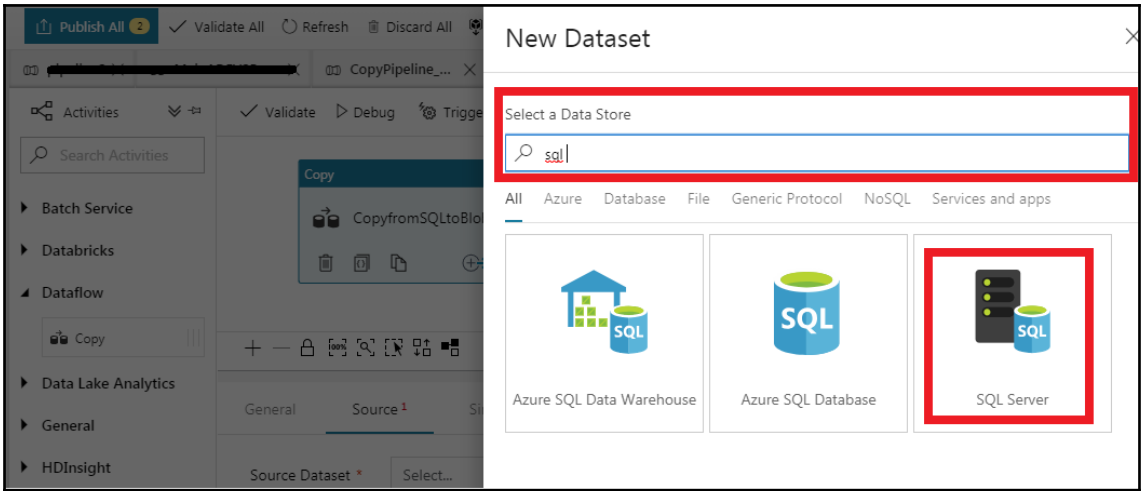


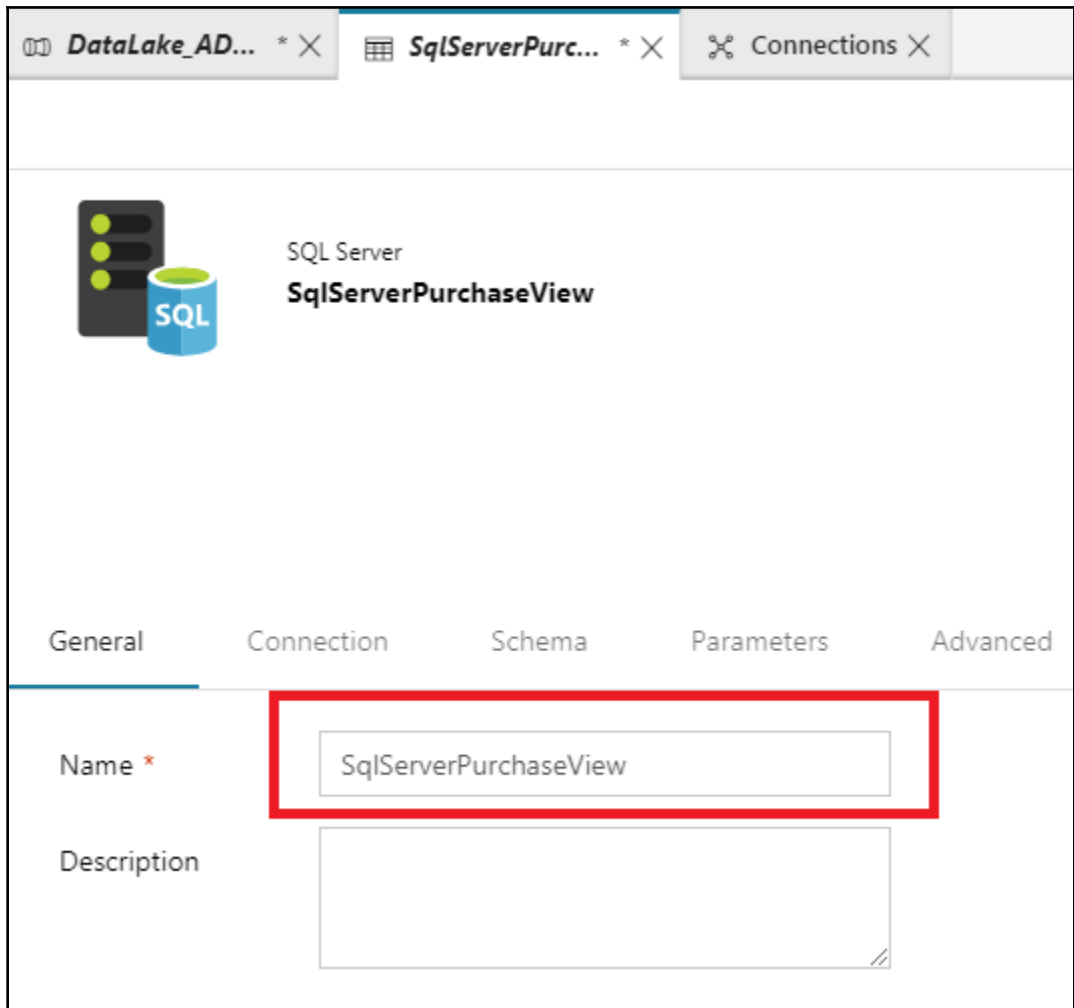


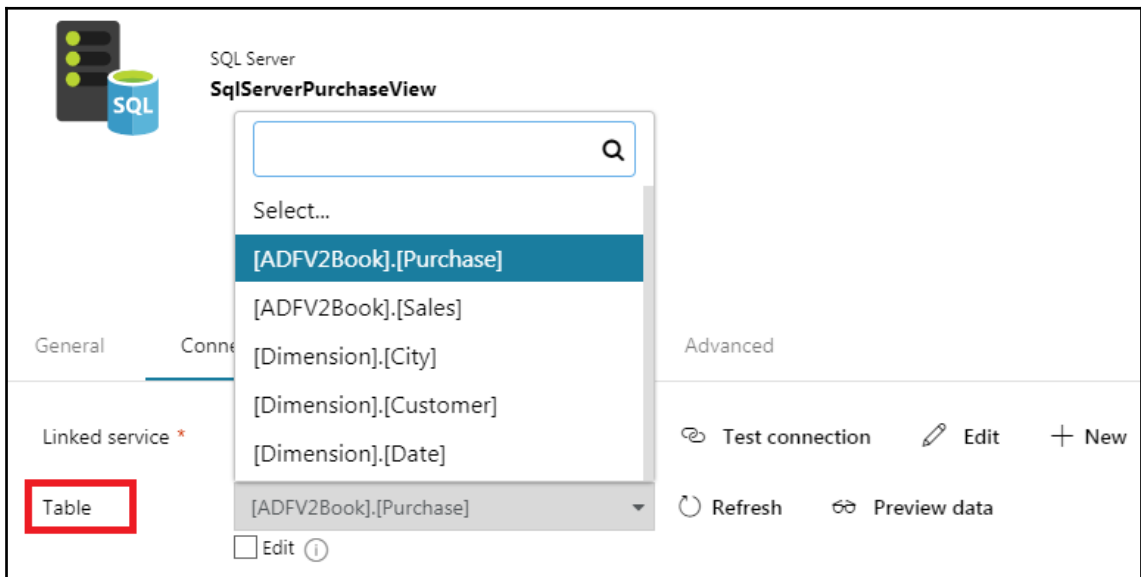
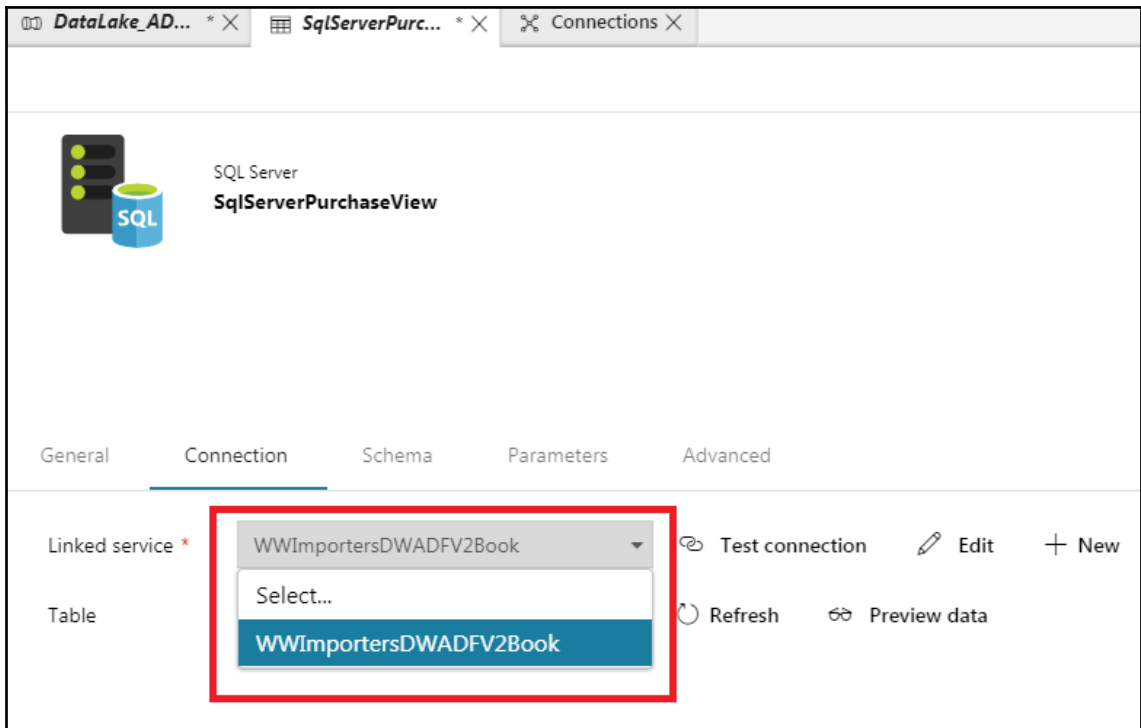


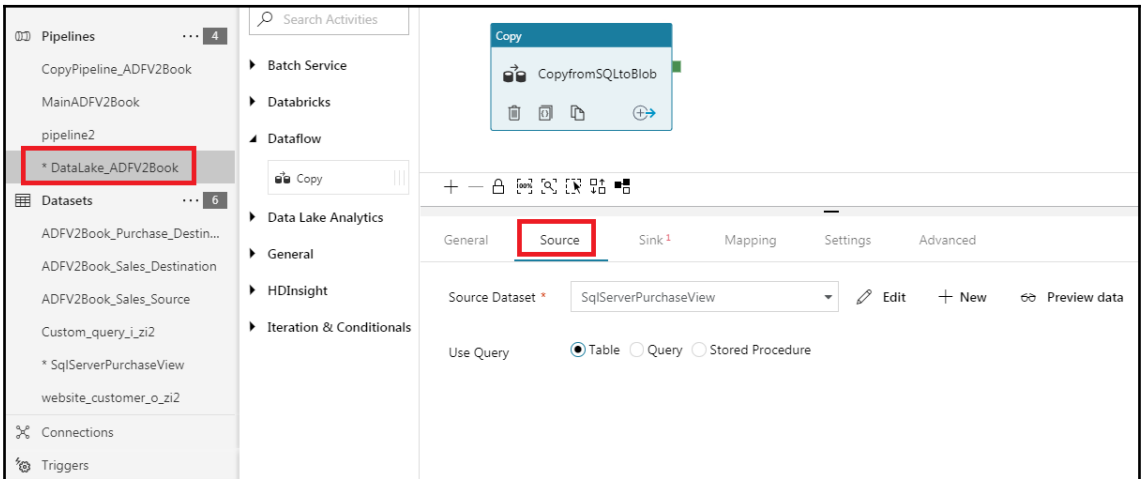
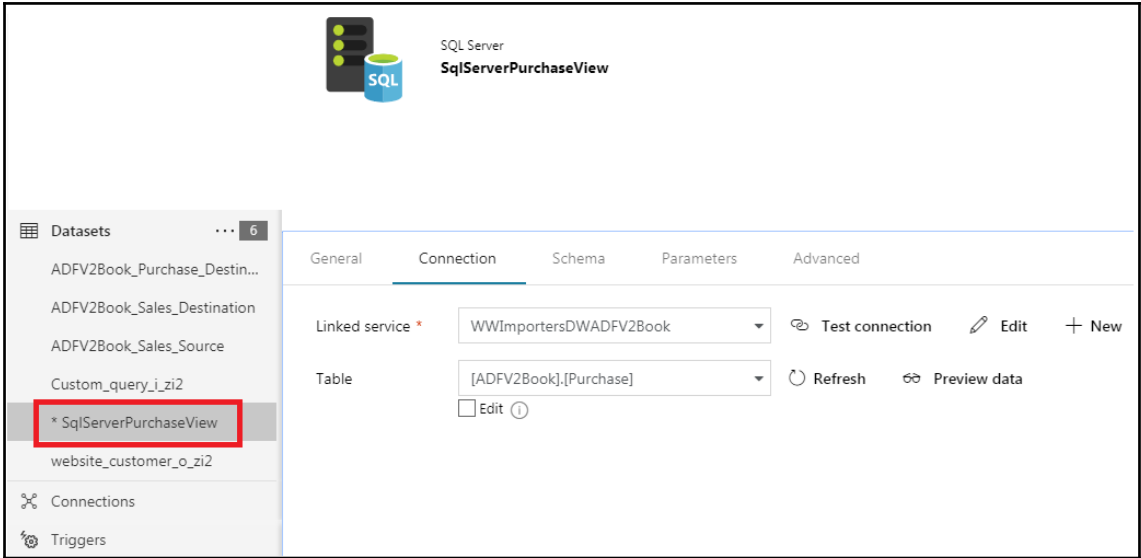












ADFV2Book Data Factory v Publish All 3 Validate All Refresh Discard All

Factory Resources Search Resources +

Pipelines 4  
CopyPipeline\_ADFV2Book  
MainADFV2Book  
pipeline2  
\* DataLake\_ADFV2Book

Datasets 6  
ADFV2Book\_Purchase\_Destin...  
ADFV2Book\_Sales\_Destination  
ADFV2Book\_Sales\_Source  
Custom\_query\_l\_zi2  
\* SqlServerPurchaseView

Activities Search Activities  
Batch Service  
Databricks  
Dataflow  
Copy  
Data Lake Analytics  
General  
HDInsight  
Iteration & Conditionals

Copy  
CopyfromSQLtoBlo

General Source Sink Dataset \* Select...


### New Dataset

Select a Data Store

Search blob

All Azure Database File Generic Protocol Services and apps

Azure Blob Storage



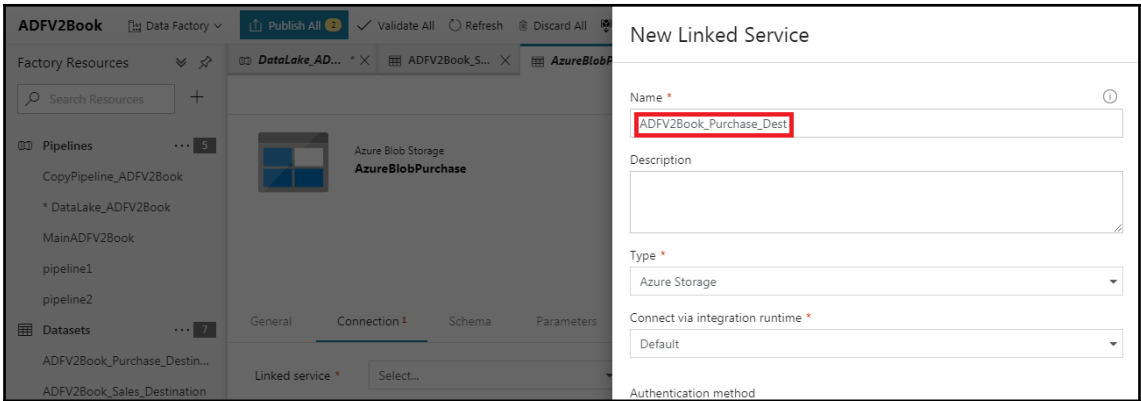
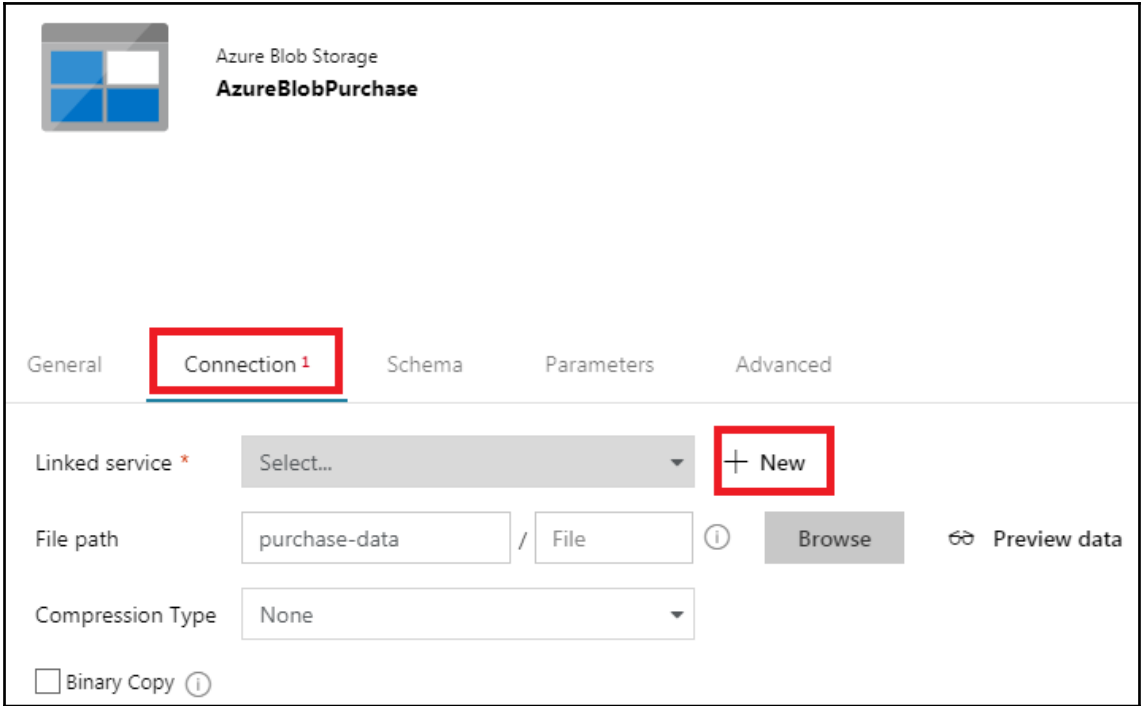
Azure Blob Storage  
**AzureBlobPurchase**

General Connection 1 Schema Parameters Advanced

Name \* AzureBlobPurchase

Description





---

Connect via integration runtime \*

Default

Authentication method

Use account key

**Connection String** Azure Key Vault

Account selection method ⓘ

From Azure subscription

**From Azure subscription**

Enter manually

b2855a5c069c)

Storage account name \*

adfv2book

Authentication method

Use account key

**Connection String** Azure Key Vault

Account selection method ⓘ

From Azure subscription

Azure subscription ⓘ

Windows Azure MSDN


Storage account name \*

adv2book

✓ Connection successful

Cancel Test connection **Finish**

---



Azure Blob Storage  
**AzureBlobPurchase**

General   **Connection**   Schema <sup>1</sup>   Parameters   Advanced

Linked service \*   ADFV2Book\_Purchase\_Destination   Test connection   Edit   New

File path   purchase-data / File   Browse   Preview data

Compression Type   None

Binary Copy

Add dynamic content [Alt+P]

---

Azure Blob Storage  
**AzureBlobPurchase**

General **Connection** Schema Parameters Advanced

Binary Copy ⓘ

▲ File Format Settings

File format Text format ⓘ Detect Text Format ⓘ

Column delimiter Comma (,)  
 Use custom delimiter

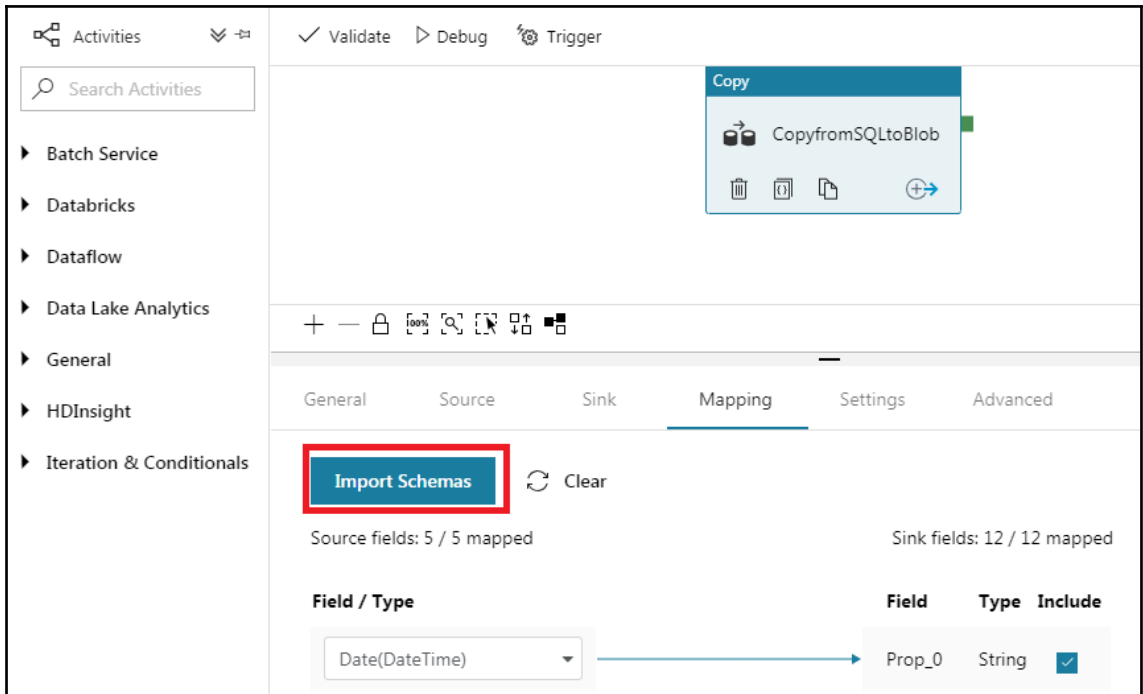
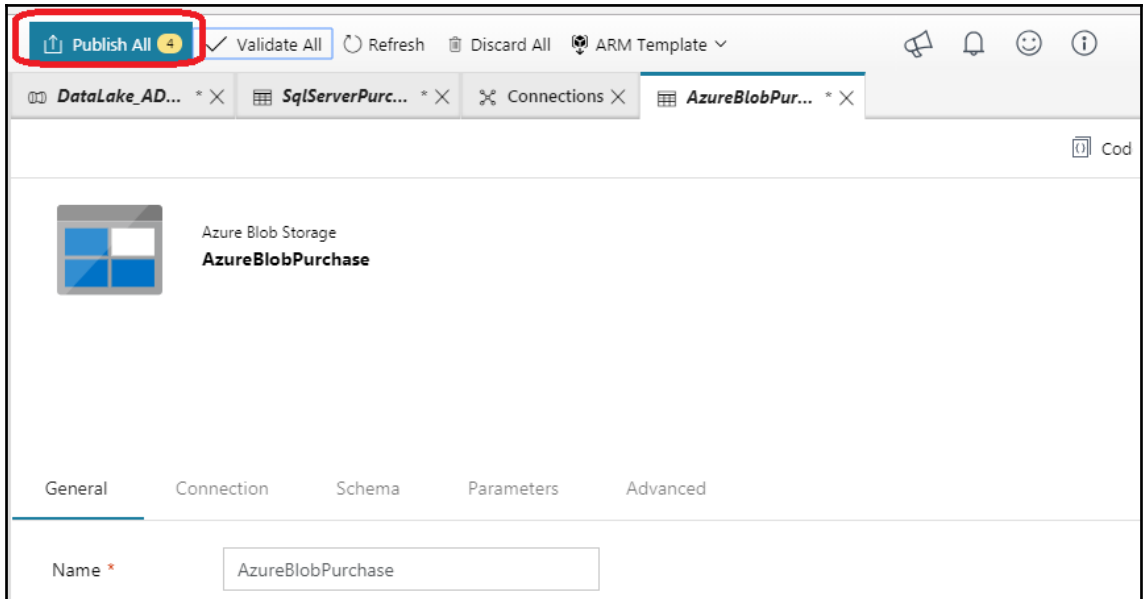
Row delimiter Carriage Return + Line feed (\r\n)  
 Use custom delimiter

Azure Blob Storage  
**AzureBlobPurchase**

General Connection **Schema** Parameters Advanced

Import Schema + New column Clear | Delete Total number of columns: 12 Filter Columns

<input type="checkbox"/>	Column	Type	Description
☰	<input type="text" value="Prop_0"/>	<span>String</span>	<input type="text"/>
☰	<input type="text" value="Prop_1"/>	<span>String</span>	<input type="text"/>
☰	<input type="text" value="Prop_2"/>	<span>String</span>	<input type="text"/>



Primary: Available  
 Location: East US  
 Subscription (name): Windows Azure MSDN - Visual Studio Ultimate  
 Subscription ID: 5d9a43ae-1916-47f6-957b-b2855a5c069c  
 Tags (name): [Click here to add tags](#)

Search containers by prefix

NAME	LAST MODIFIED	PUBLIC ACCESS ...	LEASE STATE
database-bacpac	2/17/2018, 7:01:23 PM	Private	Available ...
import-data	2/25/2018, 3:05:54 PM	Container	Available ...
<b>purchase-data</b>	5/19/2018, 11:20:04 AM	Container	Available ...
sales-data	5/19/2018, 2:29:32 PM	Container	Available ...
ssis-data	2/25/2018, 3:06:08 PM	Container	Available ...

Search blobs by prefix (case-sensitive)  Show deleted blobs

NAME	MODIFIED	ACCESS TIER	BLOB TYPE	SIZE	LEASE STATE
adla					...
[ADFV2Book].[Purchase].txt	5/24/2018, 12:53:27 PM	Hot (Inferred)	Block blob	970.33 KiB	Available ...
<b>ADFV2BookPurchaseData.csv</b>	5/24/2018, 2:51:00 PM	Hot (Inferred)	Block blob	970.33 KiB	Available ...
Purchase.txt	5/24/2018, 11:54:10 AM	Hot (Inferred)	Block blob	910.05 KiB	Leased ...

Overview Snapshots Edit blob Generate SAS

Properties

URL

<https://adv2book.blob.core.windows.net/pu...>



LAST MODIFIED

5/23/2018, 10:10:20 PM

CREATION TIME

5/22/2018, 11:23:08 AM

TYPE

Block blob

SIZE

910.01 KiB

SERVER ENCRYPTED

true

ETAG

0x8D5C11B807502F7

CONTENT-MD5

-

LEASE STATUS

Unlocked

LEASE STATE

Available

LEASE DURATION

-

COPY STATUS

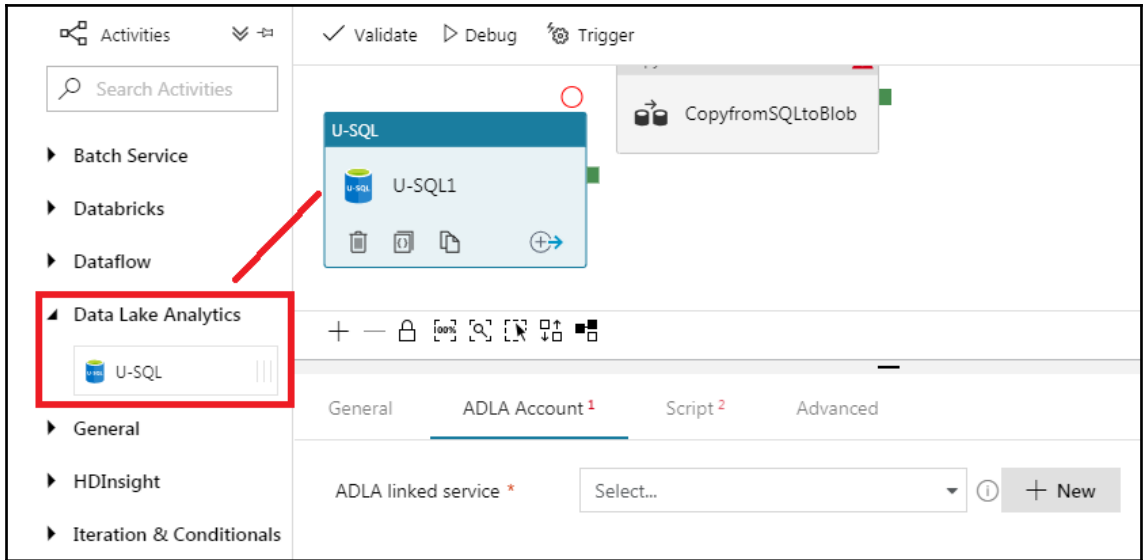
-

COPY COMPLETION TIME

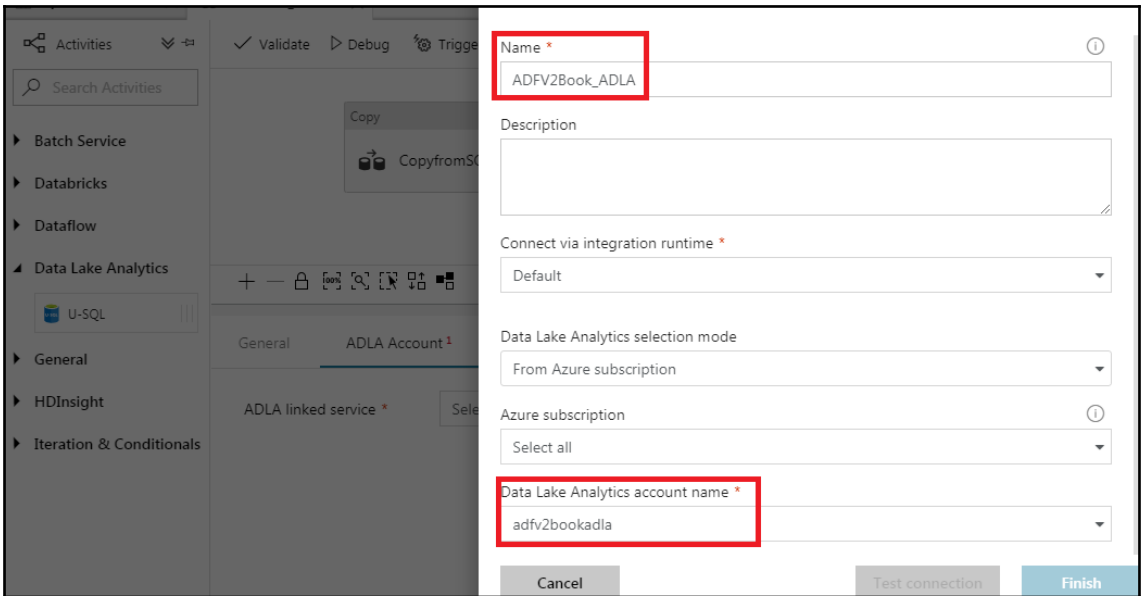
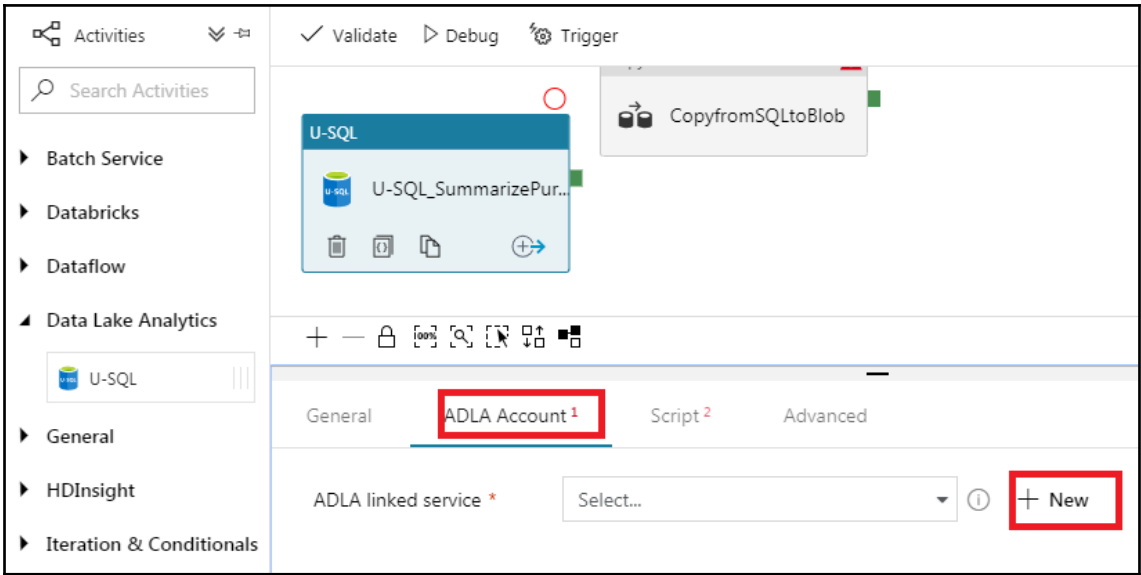
-

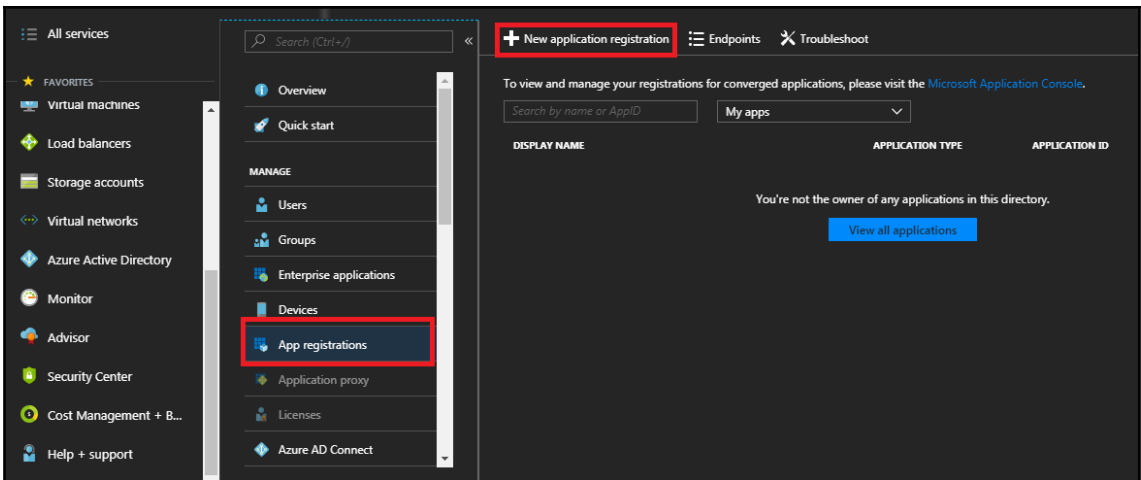
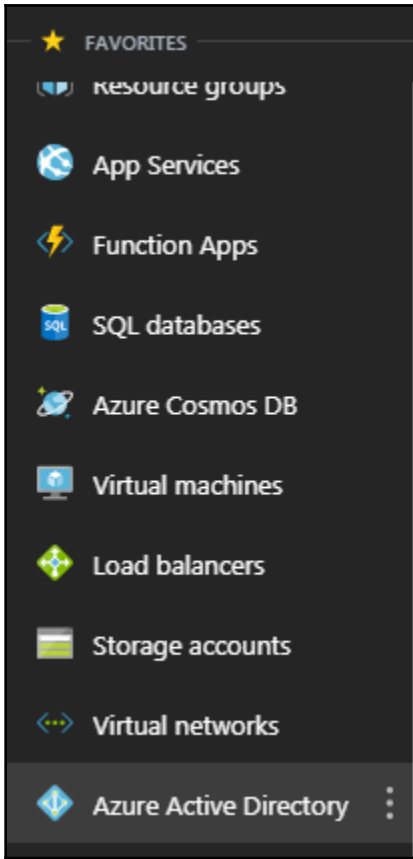
Undelete all snapshots

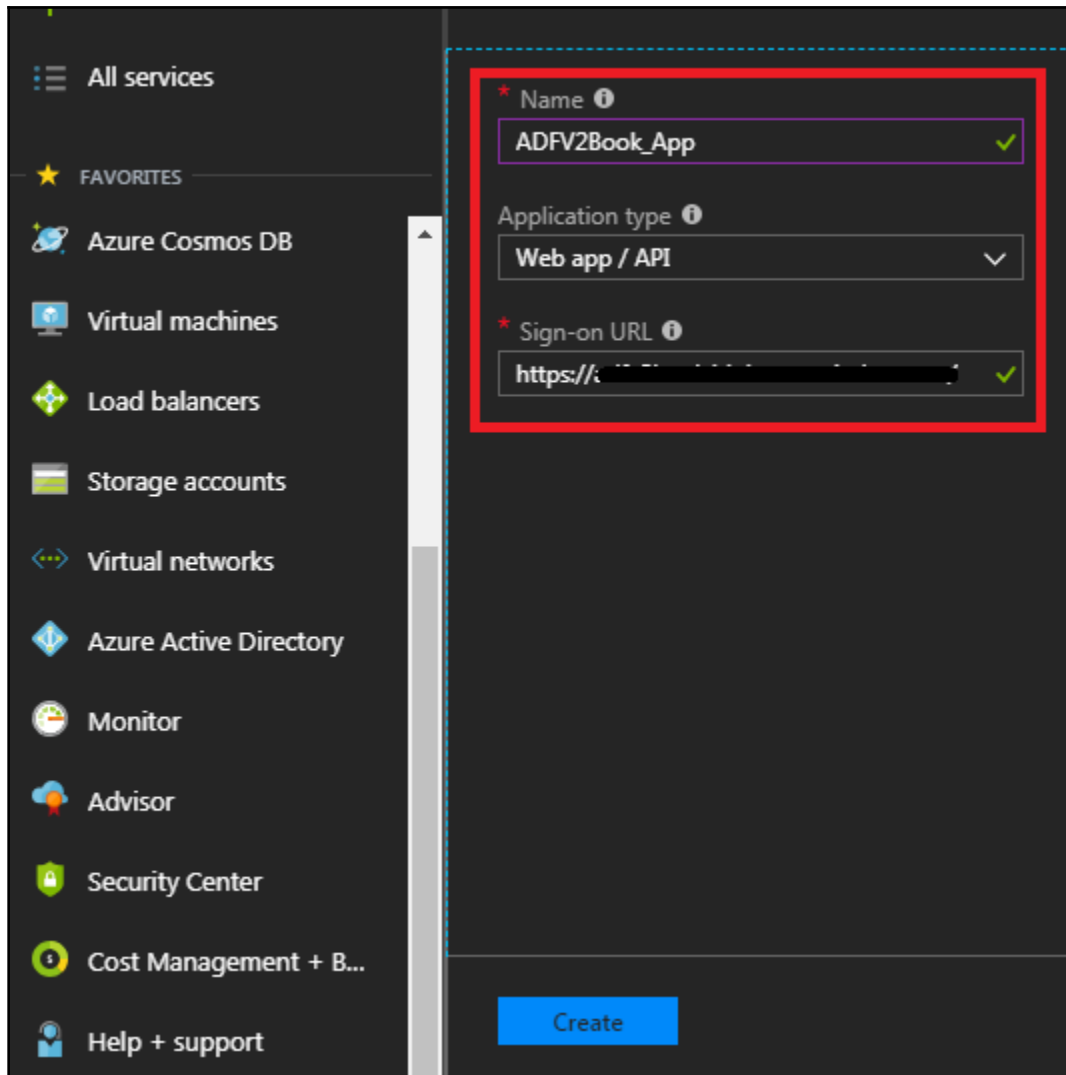


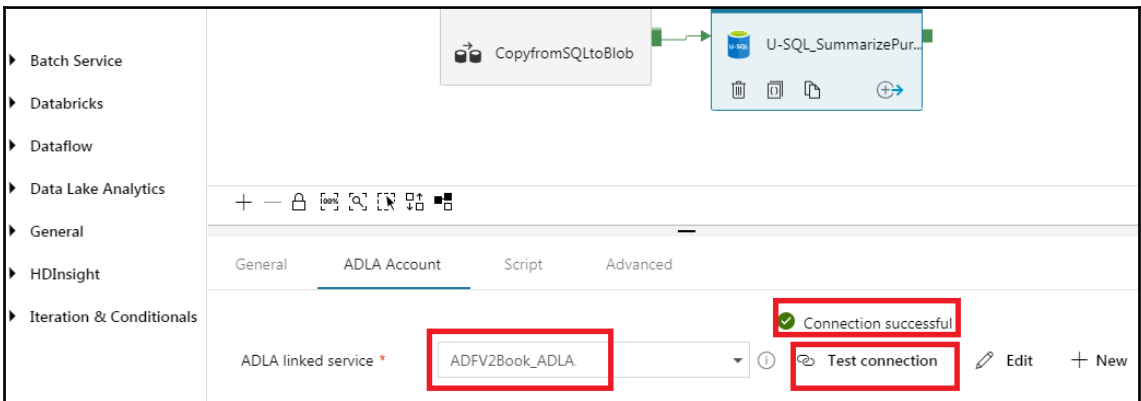
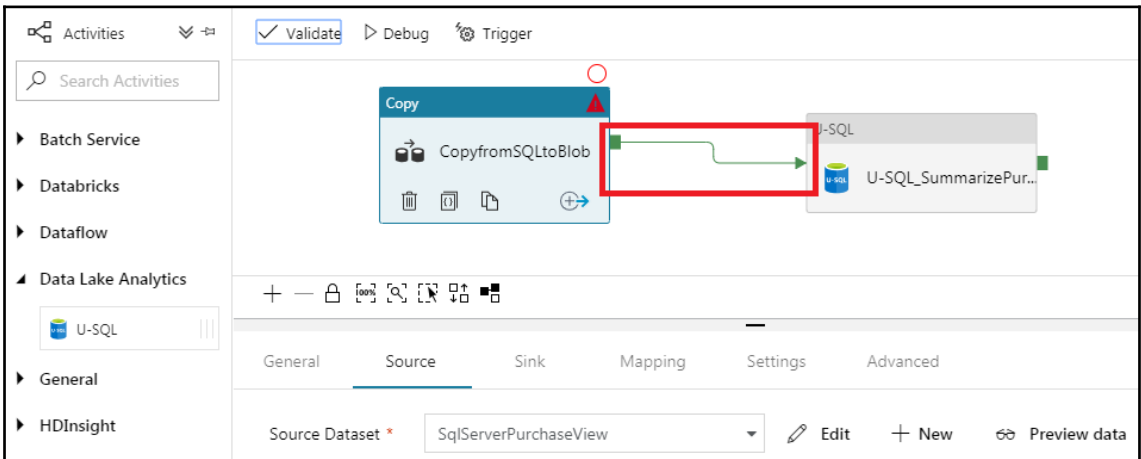
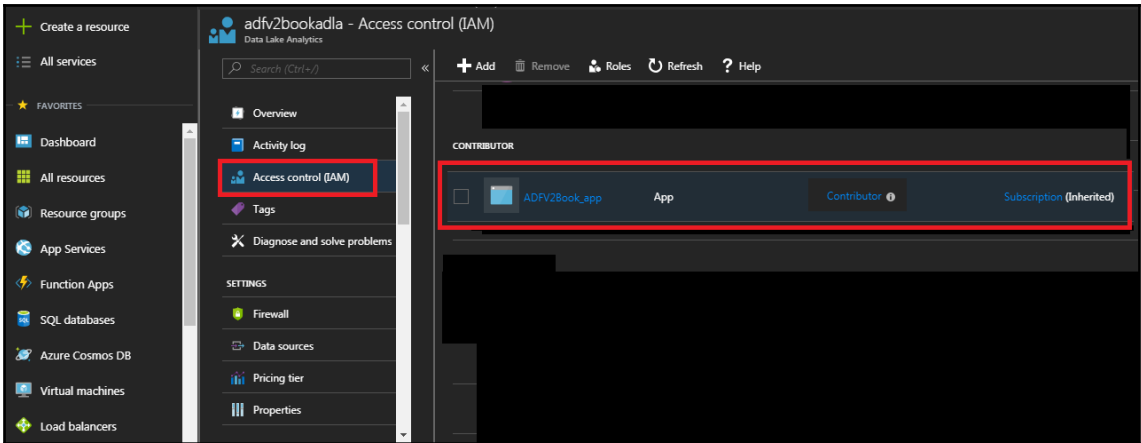


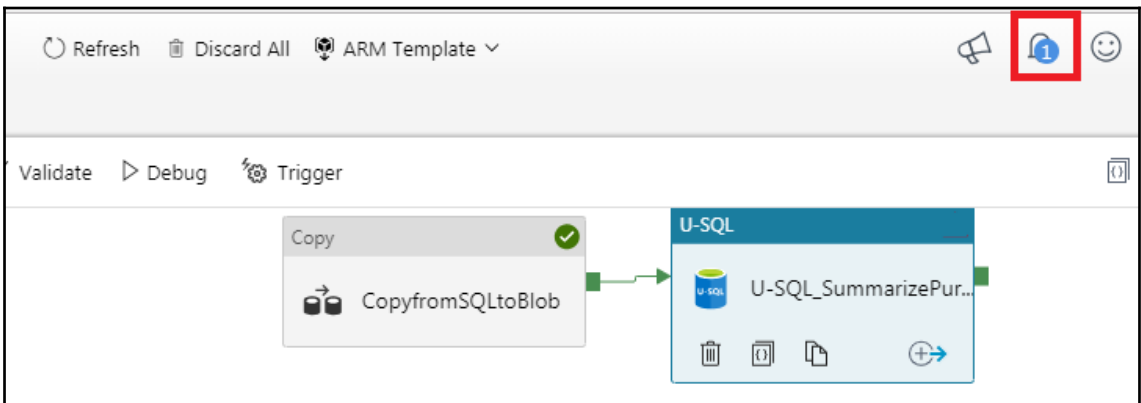
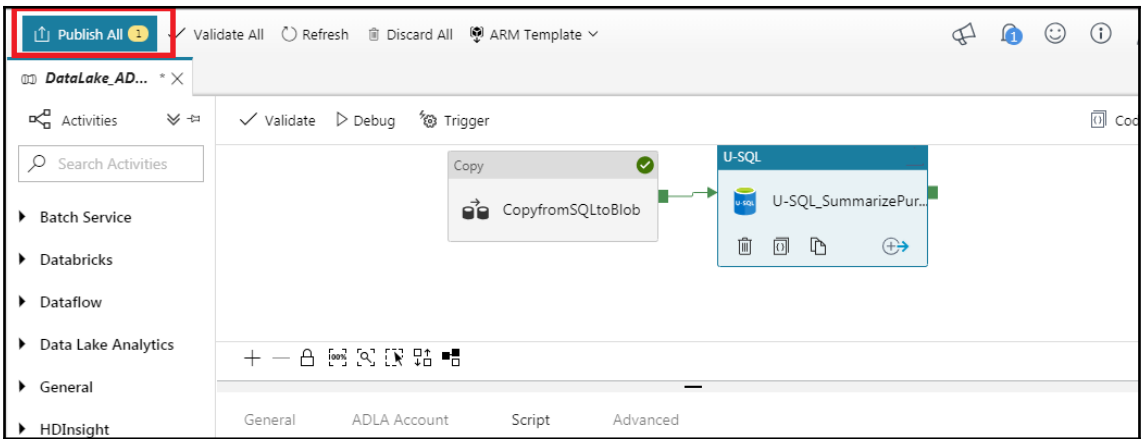
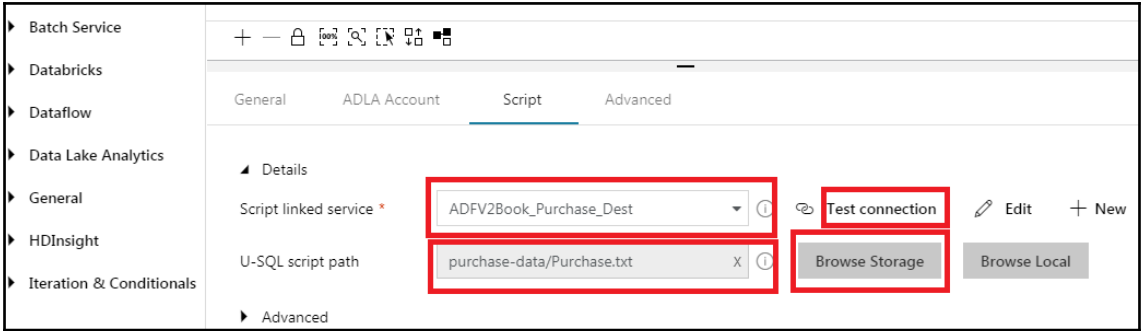
The screenshot displays the Azure Data Studio interface for configuring a U-SQL activity. On the left, the 'Activities' pane shows a search bar and a list of activity categories: Batch Service, Databricks, Dataflow, Data Lake Analytics, and U-SQL. The 'U-SQL' category is selected, and the 'General' tab is active. The main workspace shows a canvas with two activity boxes: 'U-SQL\_SummarizePur...' and 'CopyfromSQLtoBlob'. The 'U-SQL\_SummarizePur...' activity is selected, and its configuration is shown in the right-hand pane. The 'Name' field is highlighted with a red box and contains the text 'U-SQL\_SummarizePurchaseData'. Other fields include 'Description' (empty) and 'Timeout' (set to '7.00:00:00'). The top of the interface shows 'Validate', 'Debug', and 'Trigger' buttons.

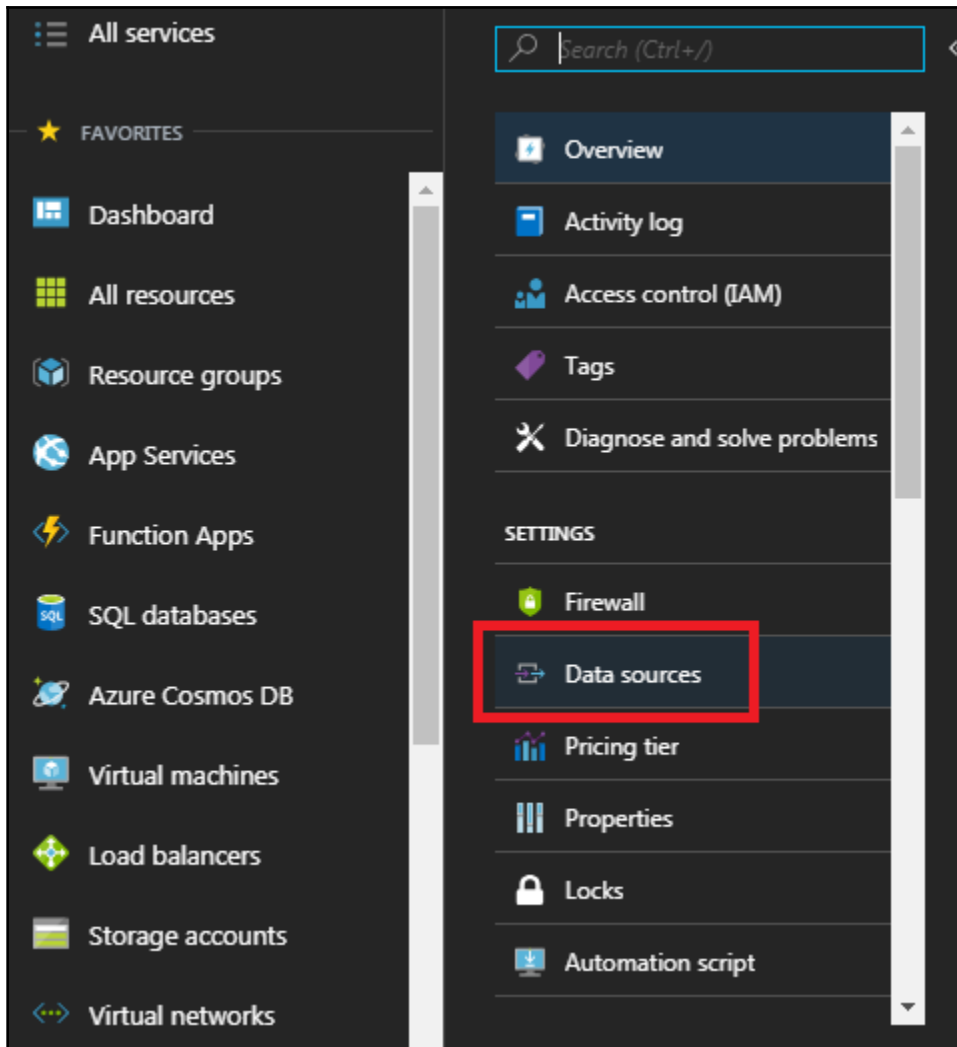














Microsoft Azure

Search resources, services, and docs

Home > adfv2bookadla - Data sources

adfv2bookadla - Data sources  
Data Lake Analytics

Search (Ctrl+V)

**Add data source**

Data sources

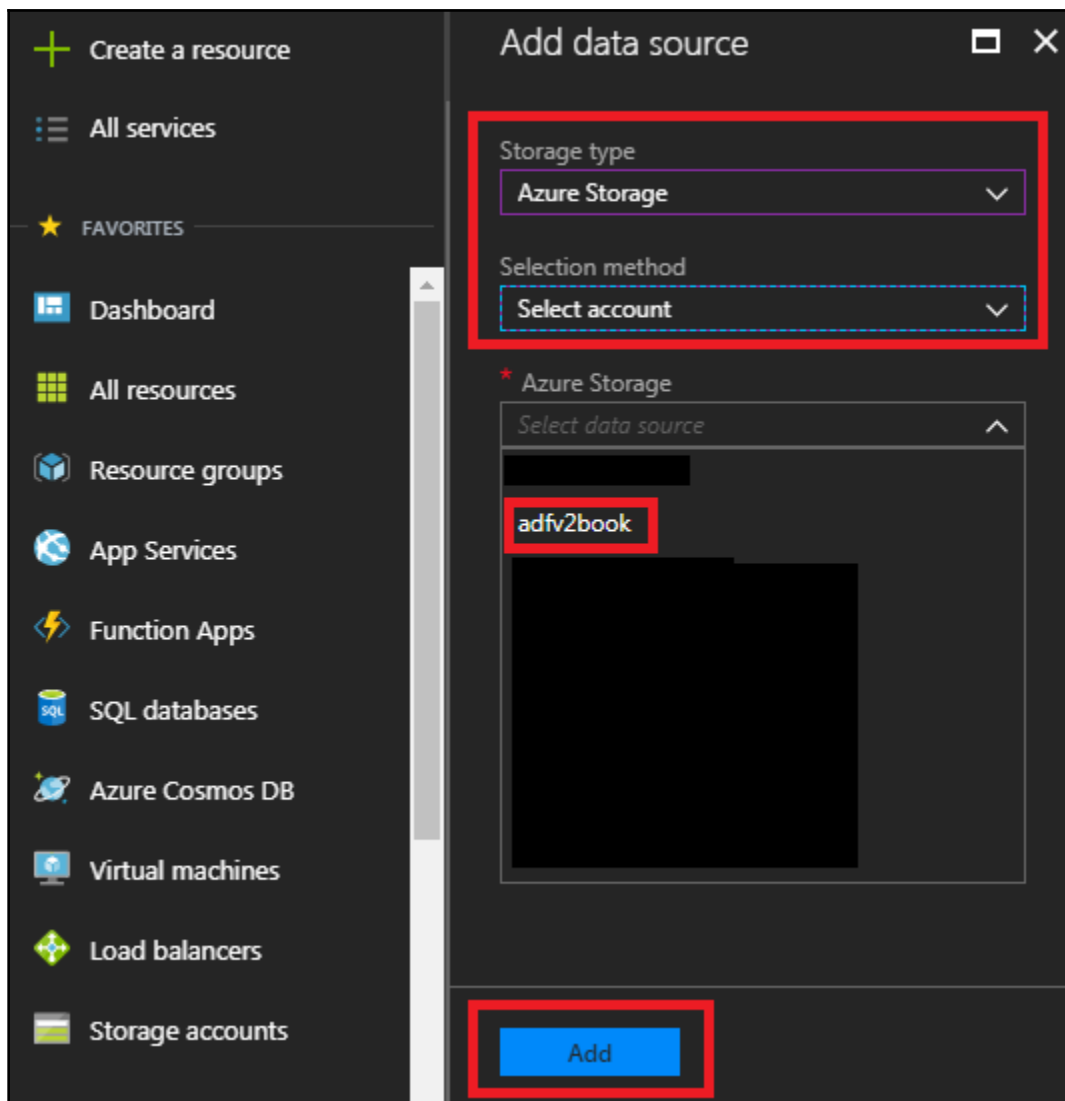
NAME	TYPE
adfv2bookdll (default)	Azure Data Lake Store

Navigation menu (left):

- Create a resource
- All services
- FAVORITES
- Dashboard
- All resources
- Resource groups
- App Services
- Function Apps
- SQL databases
- Azure Cosmos DB

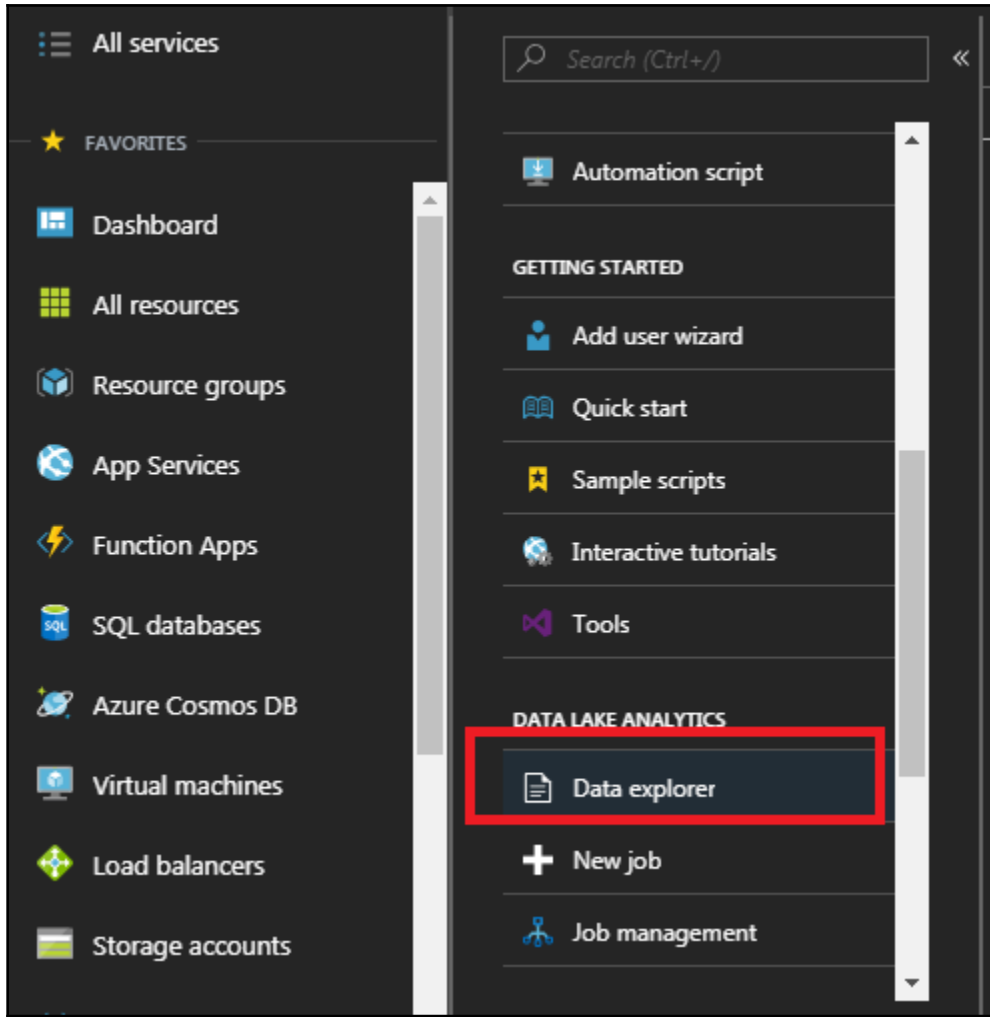
Settings menu (middle):

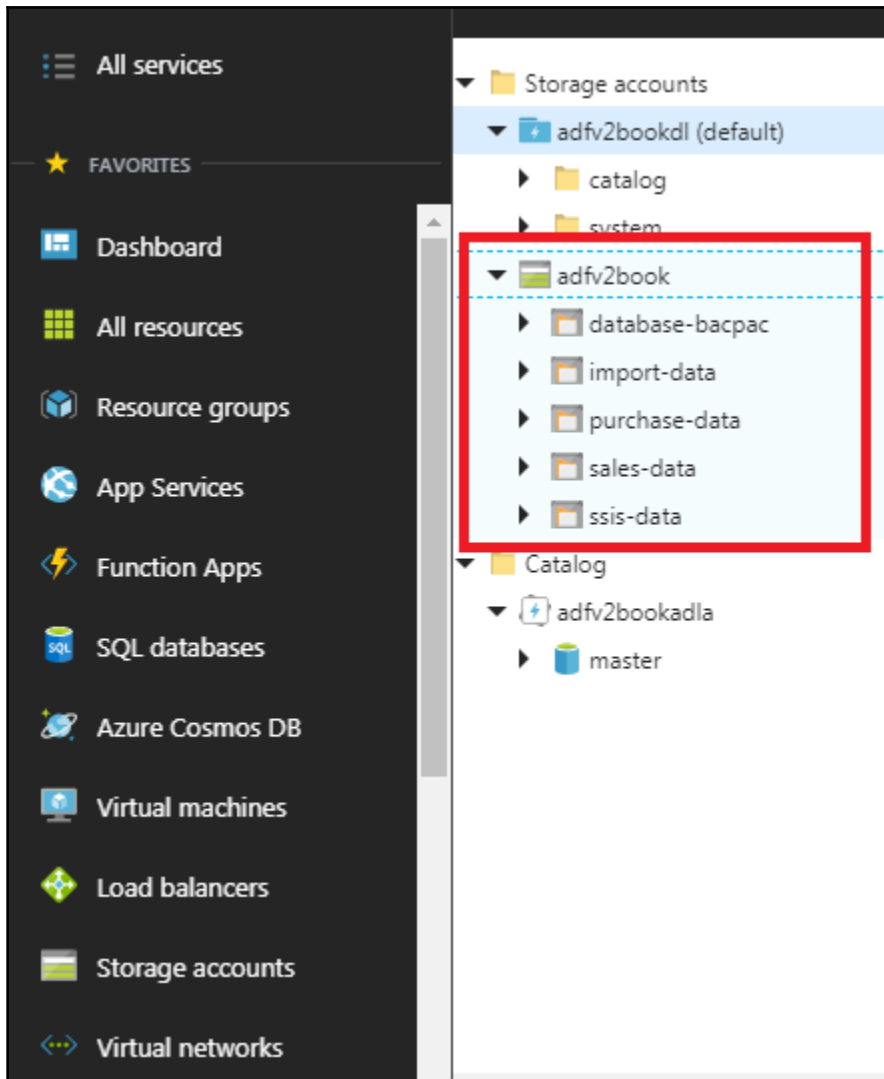
- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- SETTINGS
- Firewall
- Data sources
- Pricing tier

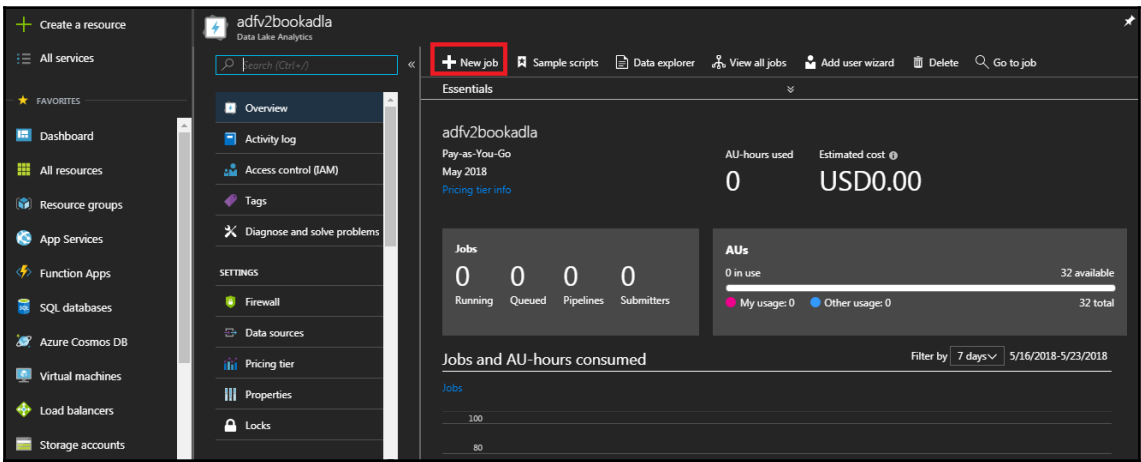
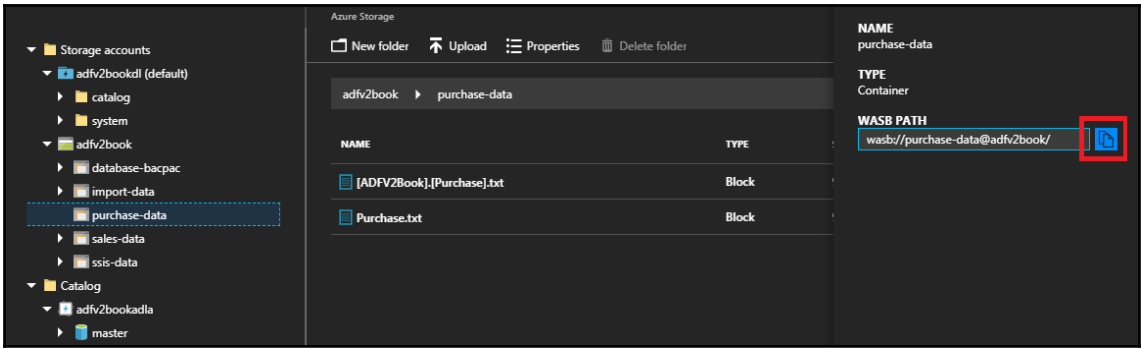


The screenshot shows the Azure portal interface for configuring data sources in a Data Lake Analytics workspace. The workspace name is 'advf2bookadla'. The 'Data sources' section is active, displaying a table of existing data sources. The 'advf2book/' data source is highlighted with a red box, indicating it is the selected or default data source.

NAME	TYPE
advf2bookdl (default)	Azure Data Lake Store
advf2book/	Azure Storage







Data explorer Open file Save as

Account: adfv2book... Job name: SummarizePurchases AUs: 1 Submitter: Estimated cost: 0.00 USD/minute Submit

```
10 ,Category string
11 ,[Stock Item] string
12 ,[Ordered Quantity] int
13
14 FROM "wasb://test@adfv2book/ADFV2BookPurchaseData.csv"
15 USING Extractors.Tsv(skipFirstNRows:1);
16
17 @sumresults =
18 SELECT Supplier, Category,
19        SUM([Ordered Quantity]) AS TotalQuantity
20 FROM @results
21 GROUP BY Supplier, Category;
22
23 OUTPUT @sumresults
24 TO "wasb://test@adfv2book/SumPurchase.csv"
25 USING Outputters.Csv();
```

### SummarizePurchases

Job details

Refresh Resubmit Reuse script

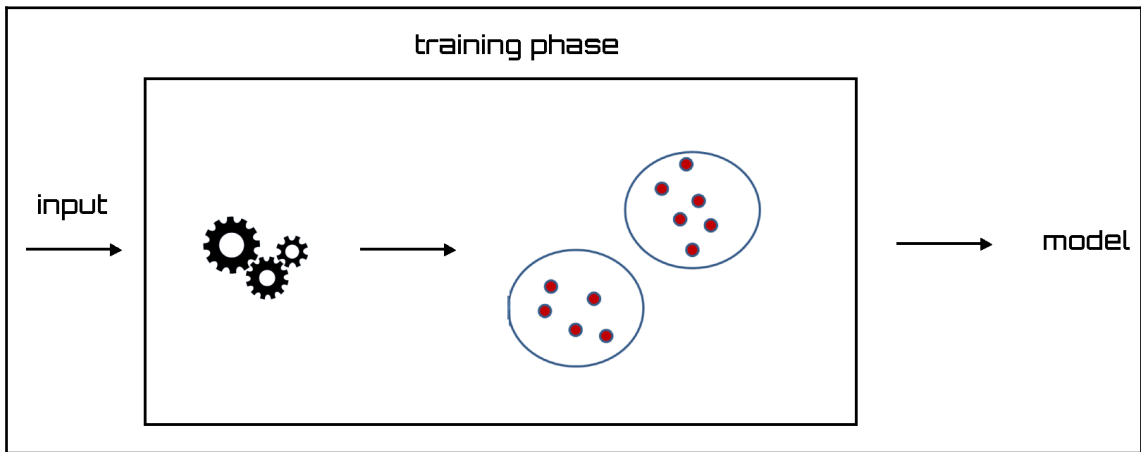
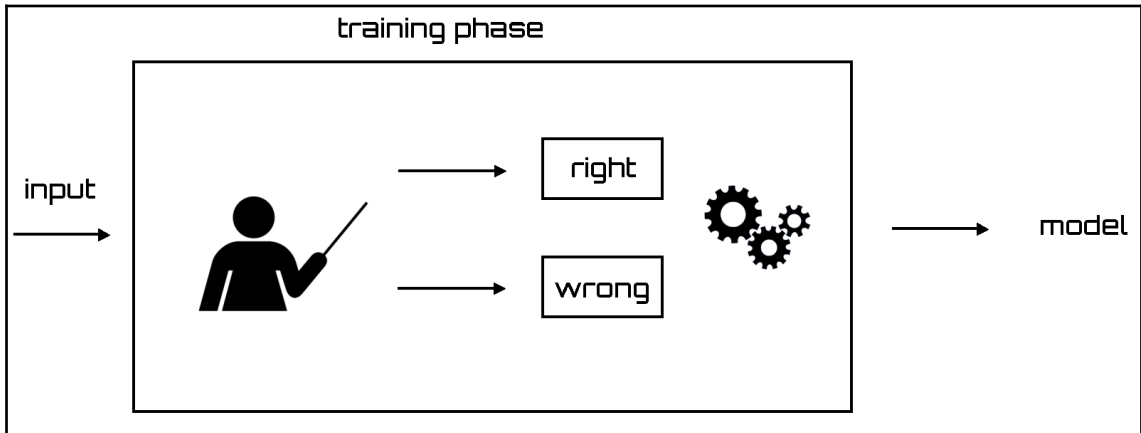
Status: **Succeeded**

Preparing	Queued	Running	Done
21s	N/A	1min 24s	

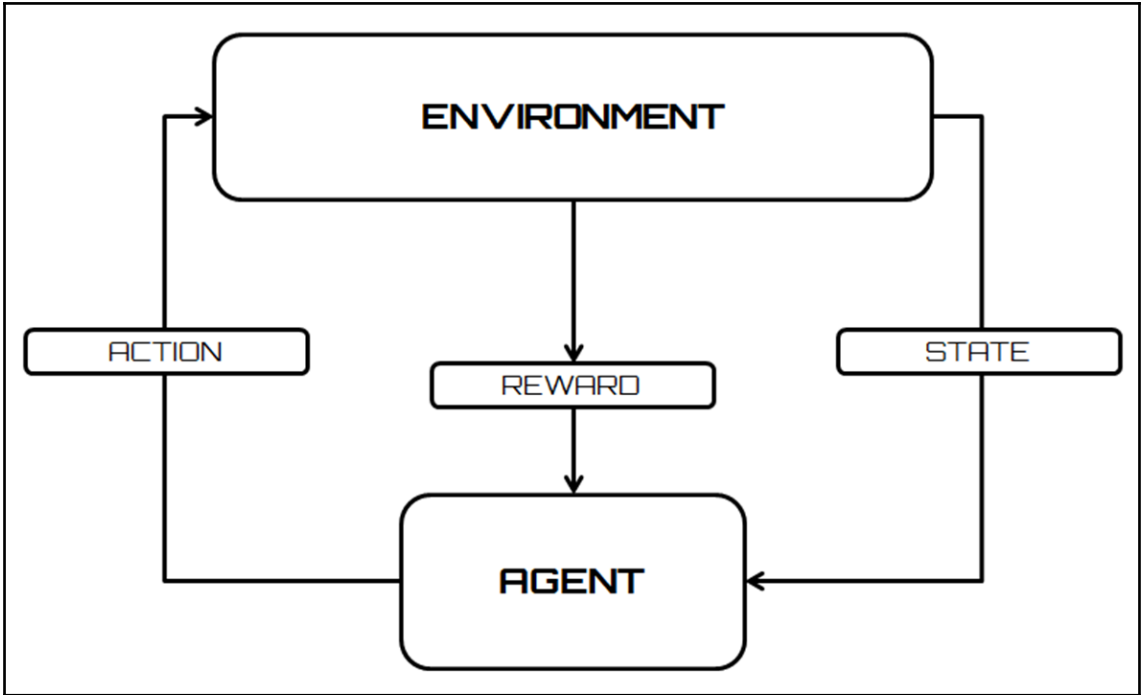
Progress 100%

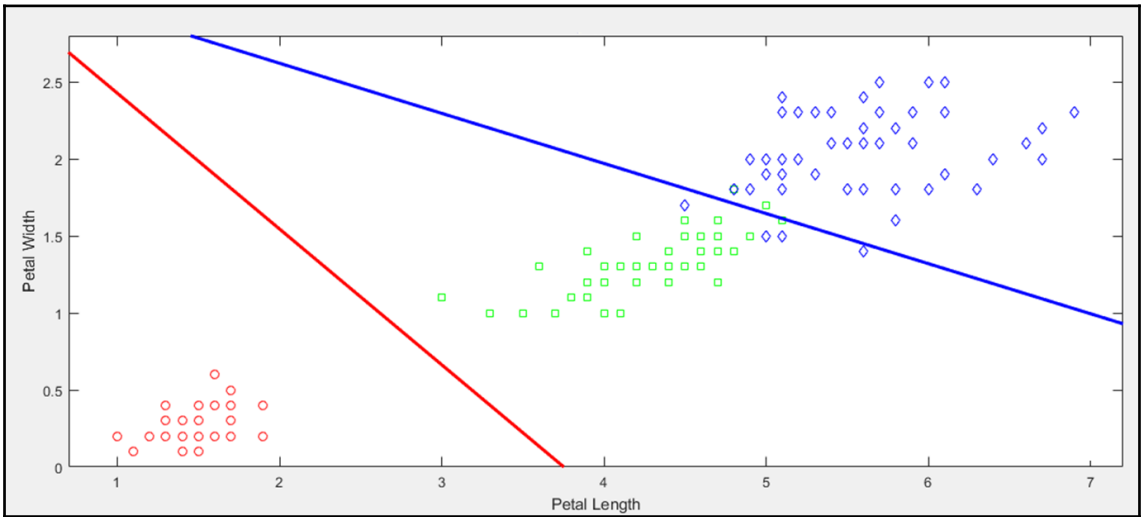
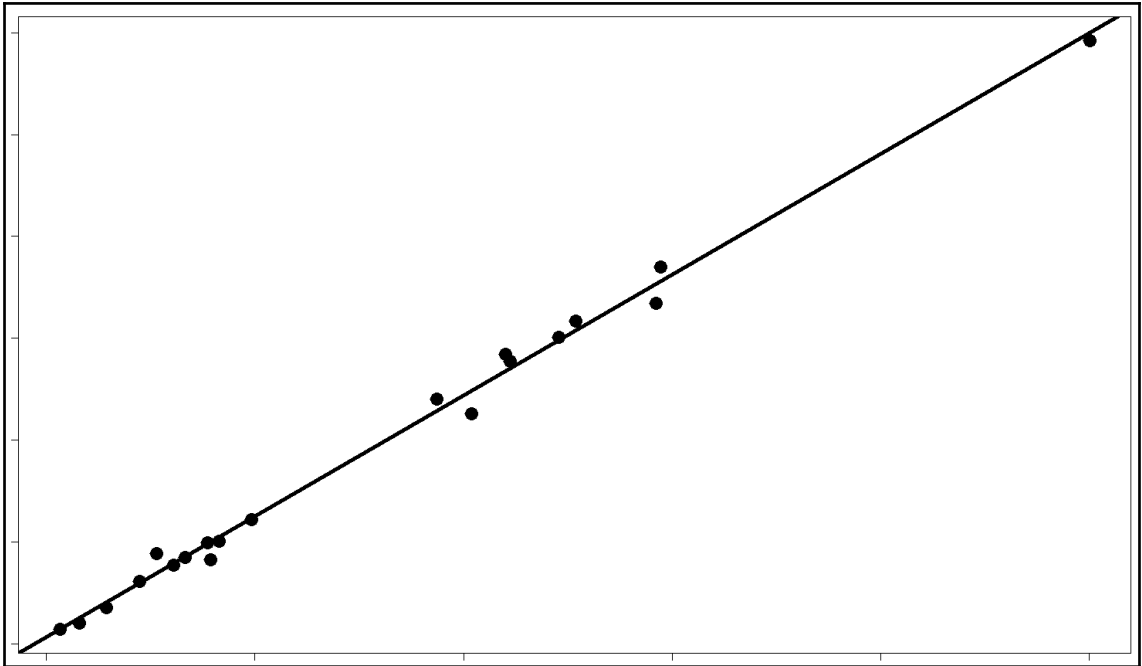
---

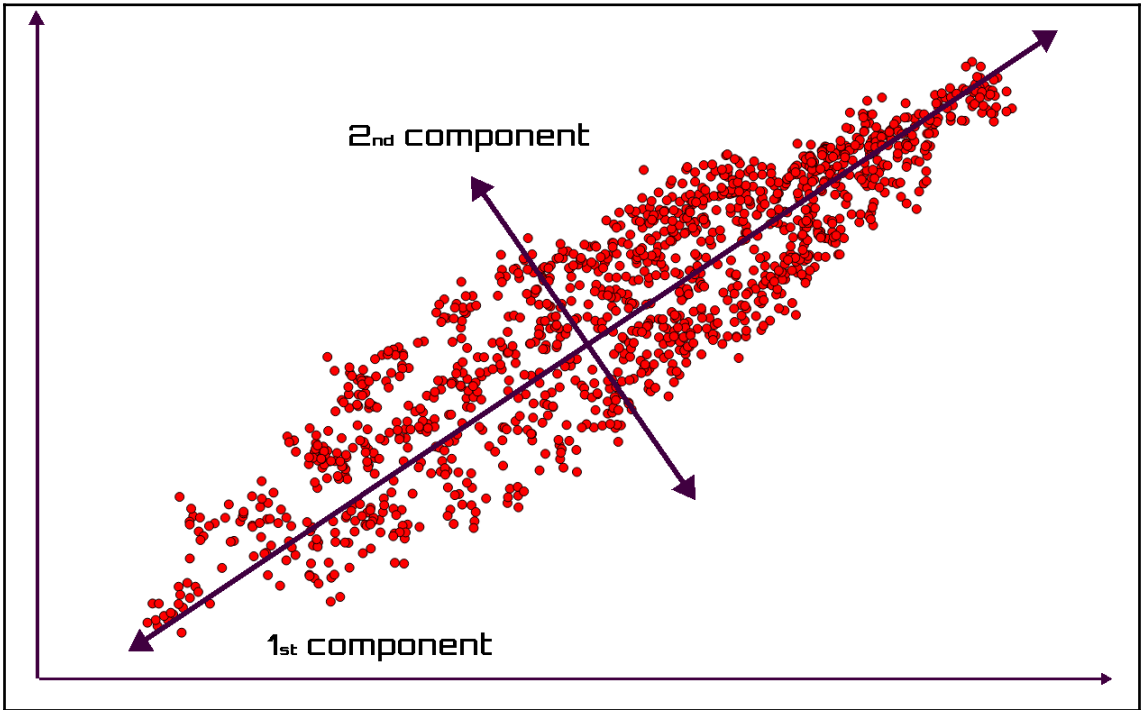
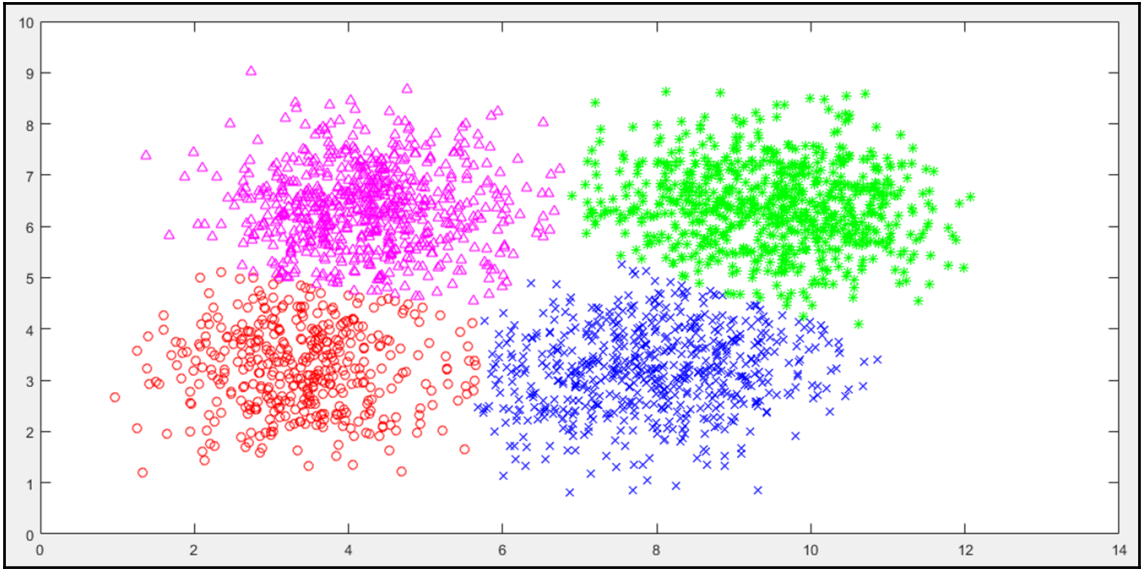
# Chapter 5: Machine Learning on the Cloud











Microsoft Azure Machine Learning Studio ? 🗨️ 😊 Sign In



## Welcome to Azure Machine Learning

Try it for free

No Azure subscription? No credit card? No problem! Choose anonymous Guest Access, or sign in with your work or school account, or a Microsoft account.

[Sign In](#)

Not an Azure ML user?  
[Sign up here](#)

[Pricing & FAQ](#)  
By using this free version, you agree to be bound by the Microsoft Azure Website Terms of Use.

**Announcements NEW!**

**Mining Campaign Funds**  
Aired on August 02, 2017

Play with 2016 Presidential Campaign finance data while learning how to prepare a large dataset for machine learning by processing and engineering features. This sample experiment works on a 2.5 GB dataset and will take about 20 minutes to run in its entirety.  
[> Learn More](#)

**Inside the Data Science VM**  
Aired on June 21, 2016

DSVM is a custom Azure Virtual Machine image that is published on the Azure marketplace and available on both Windows and Linux. It contains several popular data science and development tools both from Microsoft and from the open source community all pre-installed and pre-configured and ready to use. We will cover best practices that would show how you can use the DSVM effectively to run your next data science or analytics project.  
[> Learn More](#)

**Predictive Maintenance Modeling**  
Aired on July 05, 2016

Predictive maintenance is one of the top demanded applications of predictive modelling and is seen as a life-saver in asset-heavy industries such as manufacturing and aerospace due to its potential to provide significant cost savings by reducing downtime due to mechanical problems. We will share what data science techniques are best equipped for this sort of analytics.  
[> Learn More](#)

	FREE	STANDARD
Price	Free	\$9.99 per seat per month \$1 per studio experimentation hour
Azure subscription	Not required	Required
Max number of modules per experiment	100	Unlimited
Max experiment duration	1 hour per experiment	Up to 7 days per experiment with a maximum of 24 hours per module
Max storage space	10 GB	Unlimited - BYO
Read data from On-Premises SQL <small>Preview</small>	No	Yes
Execution/performance	Single node	Multiple nodes
Production Web API	No	Yes
SLA	No	Yes

Microsoft Azure Machine Learning Studio

GIUSEPPE CIABURRO-Free...

settings

NAME AUTHORIZATION TOKENS USERS DATA GATEWAYS

WORKSPACE NAME GIUSEPPE CIABURRO-Free-Workspace

WORKSPACE DESCRIPTION Default workspace.

WORKSPACE TYPE Free [Learn More](#)

WORKSPACE ID ed6e5f6da752429db7c27c913450ebb4

WORKSPACE STORAGE

USED AVAILABLE

0 GB

0% of 10 GIGABYTES

Want more storage? Get the standard version [learn more](#)

+ NEW

SAVE DISCARD

NEW

DATASET

MODULE

PROJECT  
PREVIEW

EXPERIMENT

NOTEBOOK  
PREVIEW

Search experiment templates

Microsoft Samples

VIEW MORE IN GALLERY

Blank Experiment

Experiment Tutorial

Sample 1: Download dataset from UCI: Adult 2 class dataset

Sample 2: Dataset Processing and Analysis: Auto Imports Regression

Sample 3: Cross Validation for Binary Classification: Adult

Sample 4: Cross Validation for Regression: Auto Imports Dataset

Two-Class Averaged Perceptron, Two-Class

Linear Regression, Boosted Decision Tree

Microsoft Azure Machine Learning Studio

GIUSEPPE CIABURRO-Free... ?

Experiment created on 17/5/2018 In draft

Search experiment items

- ▶ Saved Datasets
- ▶ Data Format Conversions
- ▶ Data Input and Output
- ▶ Data Transformation
- ▶ Feature Selection
- ▶ Machine Learning
- ▶ OpenCV Library Modules
- ▶ Python Language Modules
- ▶ R Language Modules
- ▶ Statistical Functions
- ▶ Text Analytics
- ▶ Time Series
- ▶ Web Service
- ▶ Deprecated

Drag Items Here

To create your experiment, drag and drop datasets and modules here

Mini Map

Properties Project

- ▶ Experiment Properties
  - STATUS CODE InDraft
- ▶ Summary
  - Enter a few sentences describing your experiment (up to 140 characters).
- ▶ Description
  - Enter the detailed description for your experiment.
- ▶ Quick Help

+ NEW

RUN HISTORY

SAVE

SAVE AS

DISCARD CHANGES

RUN

SET UP WEB SERVICE

PUBLISH TO GALLERY

# Upload a new dataset



SELECT THE DATA TO UPLOAD:

Browse

This is the new version of an existing dataset

ENTER A NAME FOR THE NEW DATASET:

SELECT A TYPE FOR THE NEW DATASET:

Select a dataset type...



PROVIDE AN OPTIONAL DESCRIPTION:





Microsoft Azure Machine Learning Studio

GIUSEPPE CIABURRO-Free... ?

Experiment create... In draft

Properties Project

Search experiment items

Saved Datasets

My Datasets

Breast\_Cancer.csv

Samples

Data Format Conversions

Data Input and Output

Data Transformation

Feature Selection

Machine Learning

Breast\_Cancer.csv 1

Mini Map

Breast\_Cancer.csv

SUBMITTED BY giuseppe.ci...

SIZE 20.3 KB

FORMAT GenericCSV

CREATED ON 5/17/2018 ...

View dataset

Experiment created on 17/5/2018 > Breast\_Cancer.csv > dataset

rows 699

columns 11

view as

Sample_code_number	Clump_thickness	Uniformity_of_cell...
1000025	5	1
1002945	5	4
1015425	3	1
1016277	6	8
1017023	4	1
1017122	8	10
1018099	1	1
1018561	2	1
1033078	2	1
1033078	4	2
1035283	1	1
1036172	2	1
1041801	5	3
1043999	1	1

Statistics

Mean	4.4177
Median	4
Min	1
Max	10
Standard Deviation	2.8157
Unique Values	10
Missing Values	0
Feature Type	Numeric Feature

Visualizations

Clump\_thickness

Histogram

compare to None

Clump_thickness	Frequency
1	140
2	50
3	110
4	80
5	130
6	35
7	25
8	45
9	15
10	70

Microsoft Azure Machine Learning Studio

GIUSEPPE CIABURRO-Free...

Experiment create... In draft

Breast\_Cancer.csv

Clean Missing Data

1 2

Mini Map

Breast\_Cancer.csv

Clean Missing Data

1 2

Properties Project

Clean Missing Data

Columns to be cleaned

Selected columns: All columns

Launch column selector

Minimum missing value... 0

Maximum missing value... 1

Cleaning mode: Custom substitution value

Replacement value: 0

Generate missing v...

Quick Help

Specifies how to handle the values missing from a dataset

Microsoft Azure Machine Learning Studio

GIUSEPPE CIABURRO-Free...

Experiment created on 17/5/2018 In draft

Draft saved at 09:08:40

Breast\_Cancer.csv

Clean Missing Data

Split Data

1 2

Mini Map

Breast\_Cancer.csv

Clean Missing Data

Split Data

1 2

Properties Project

Split Data

Splitting mode: Split Rows

Fraction of rows in the first set: 0.7

Randomized split: checked

Random seed: 0

Stratified split: False

Quick Help

Split the rows of a dataset into two distinct sets

Microsoft Azure Machine Learning Studio

GIUSEPPE CIABURRO-Free... ?

Experiment created on 17/5/2018

In draft

Draft saved at 10:08:55

Search experiment items

- Multiclass Logisti...
- Multiclass Neural ...
- One-vs-All Multic...
- Two-Class Averag...
- Two-Class Bayes ...
- Two-Class Booste...
- Two-Class Decisio...
- Two-Class Decisio...
- Two-Class Locally...
- Two-Class Logisti...**
- Two-Class Neural ...
- Two-Class Suppor...

Clustering

Regression

Score

Train

- Sweep Clustering
- Train Anomaly Dete...
- Train Clustering Model
- Train Matchbox Seco...
- Train Model**

**module selected**

Breast\_Cancer.csv

Clean Missing Data

Split Data

Two-Class Logistic Regression

Train Model 1

Mini Map

Properties Project

**Train Model**

Label column

Selected columns:

Column names:

Output\_Class

Launch column selector

Quick Help

Train a previously created classification or regression model (more help...)

Microsoft Azure Machine Learning Studio

GIUSEPPE CIABURRO-Free... ?

Experiment created on 17/5/2018 Finished running ✓

Search experiment items

- ▶ Saved Datasets
- ▶ Data Format Conversions
- ▶ Data Input and Output
- ▶ Data Transformation
- ▶ Feature Selection
- ▶ Machine Learning
  - ▶ Evaluate
  - ▶ Initialize Model
  - ▶ Score
    - Apply Transformation
    - Assign Data to Clusters
    - Score Matchbox Rec...
    - Score Model**
  - ▶ Train
    - OpenCV Library Modules
    - Python Language Modules
    - R Language Modules
    - Statistical Functions
    - Text Analytics

Flowchart steps:

- Breast\_Cancer .csv
- Clean Missing Data ✓
- Split Data ✓
- Two-Class Logistic Regression ✓
- Train Model ✓
- Score Model 1 ✓

Properties Project

Score Model

- Append score column...

START TIME 5/18/2018 ...

END TIME 5/18/2018 ...

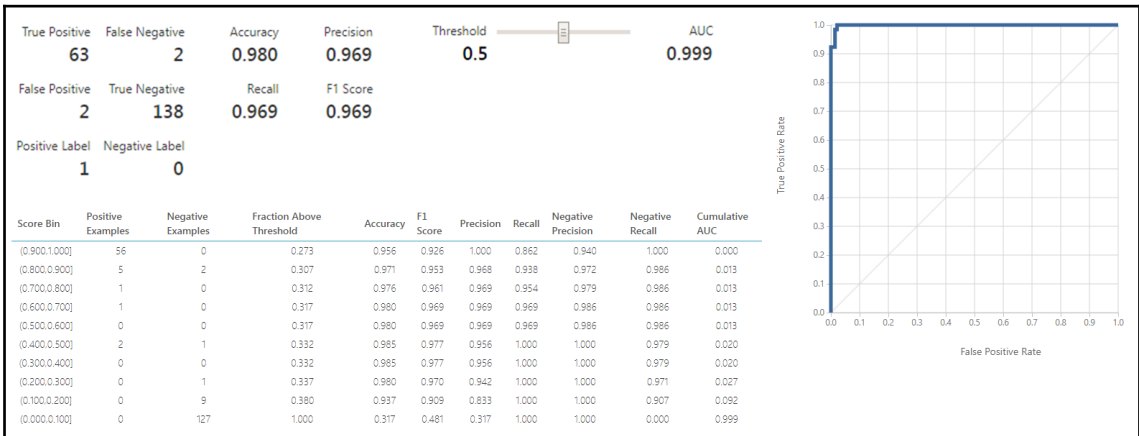
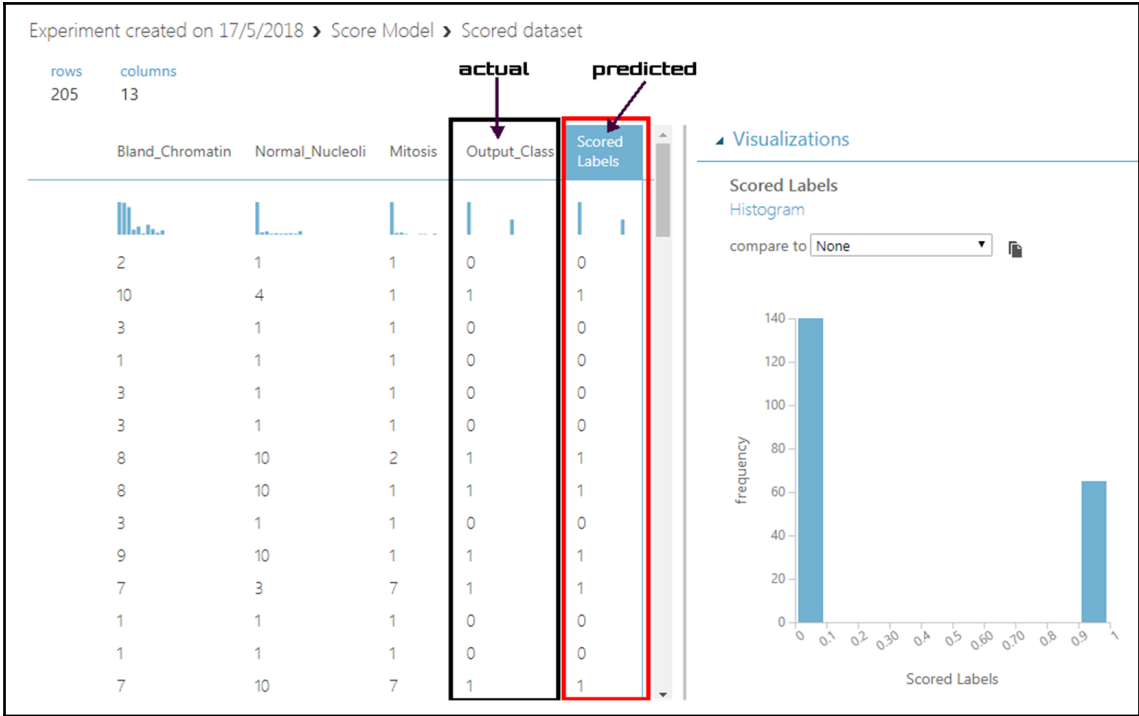
ELAPSED TIME 0:00:00.000

STATUS CODE Finished

STATUS DETAILS Task output was present in output cache

Quick Help

Score a trained classification or regression model (more help...)




---

# Chapter 6: Introduction to Azure Databricks

Databases

Analytics



Azure Databricks  
Quickstart tutorial

Home > New > Azure Databricks Service

### Azure Databricks Service

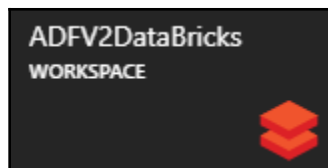
\* Workspace name  
ADFV2DataBricks ✓

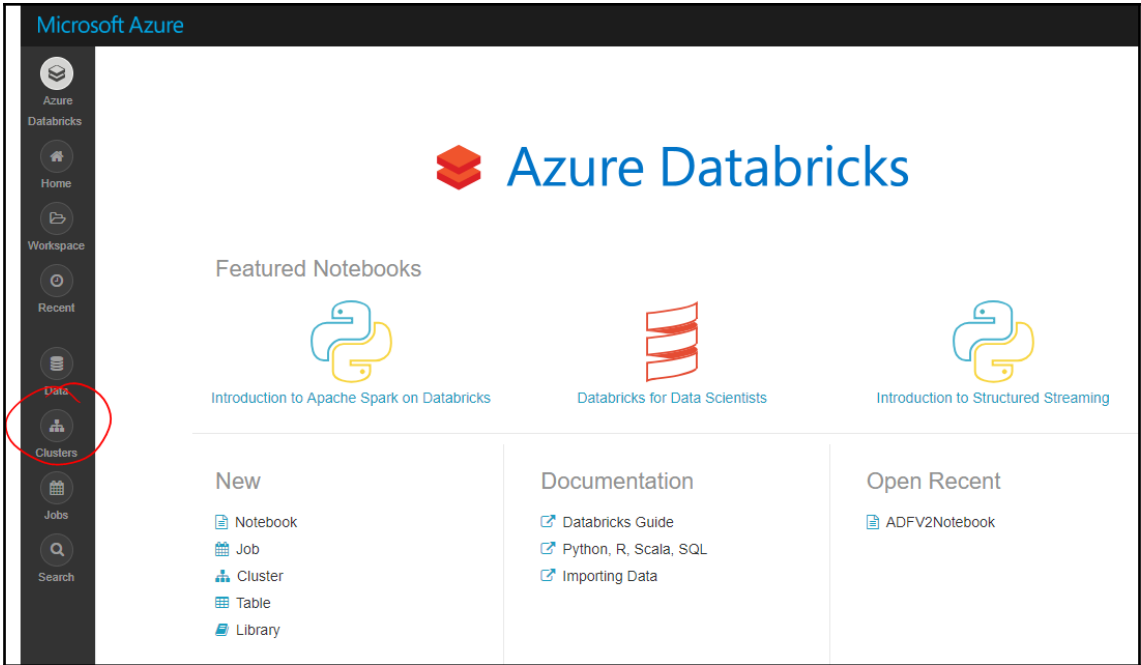
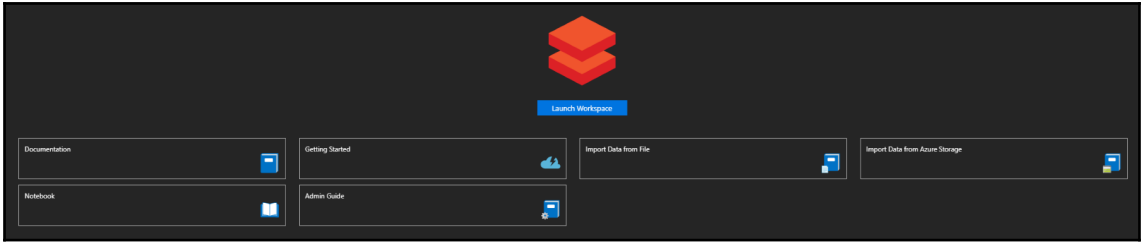
\* Subscription  
[Dropdown]

\* Resource group ⓘ  
 Create new  Use existing  
ADFV2Book [Dropdown]

\* Location  
East US [Dropdown]

\* Pricing Tier ( [View full pricing details](#) )  
Premium (+ Role-based access controls) [Dropdown]





---

Microsoft Azure

Clusters

+ Create Cluster

▼ Interactive Clusters

Name	State
ADFV2Book	Pending ⓘ

▼ Job Clusters

Table with 2 columns: Name, State

Table with 1 row: ADFV2Book, Pending ⓘ

Table with 2 columns: Name, State

Table with 0 rows

Navigation sidebar: Azure, Databricks, Home, Workspace, Recent, Data



**Create Cluster**

**New Cluster** Cancel Create Cluster 2-8 Workers: 28.0-112.0 GB Memory, 8-32 Cores, 1.5-6 DBU  
1 Driver: 14.0 GB Memory, 4 Cores, 0.75 DBU Cost \$0.40 per DBU

**Cluster Type**  
Serverless Pool (beta, R/Python/SQL) **Standard** [Learn more about Serverless Pools](#)

**Cluster Name**  
ADFV2Book

**Databricks Runtime Version**  
4.0 (includes Apache Spark 2.3.0, Scala 2.11)

**Python Version**  
2

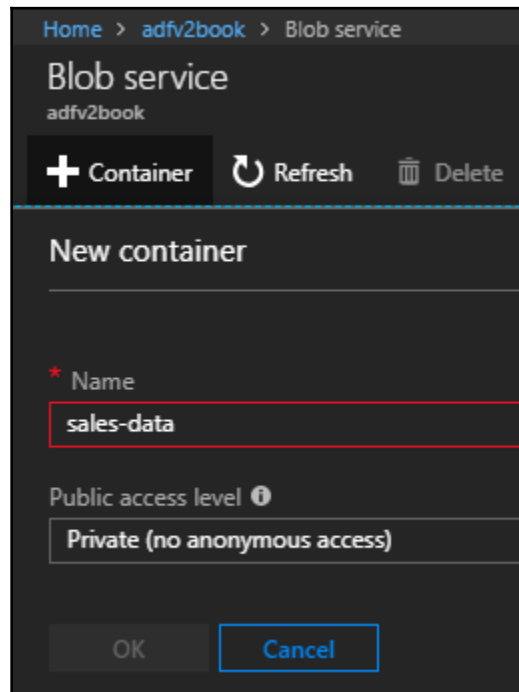
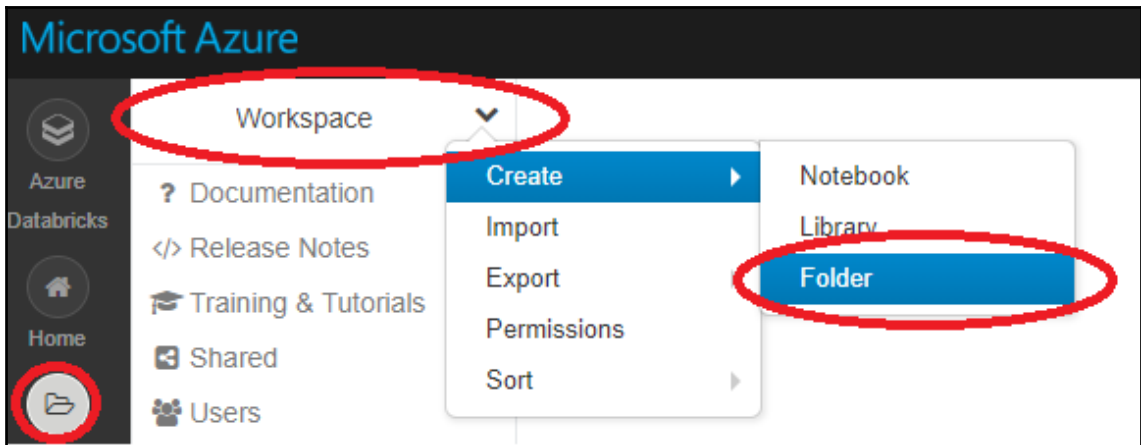
**Driver Type**  
Same as worker 14.0 GB Memory, 4 Cores, 0.75 DBU

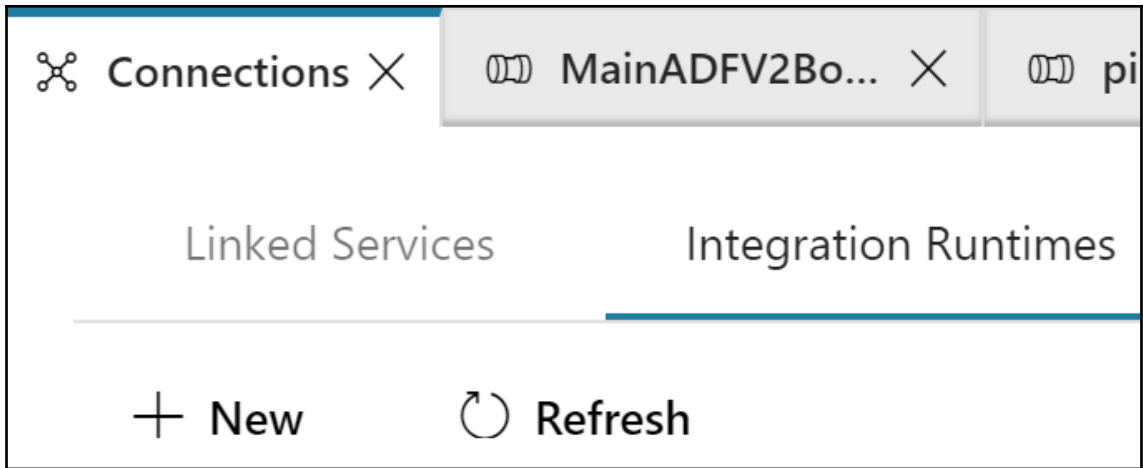
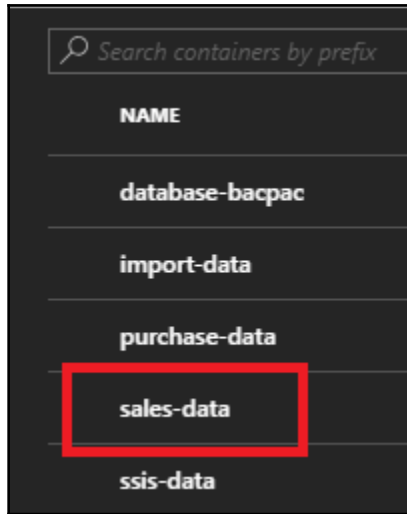
**Worker Type** **Min Workers** **Max Workers**  
Standard\_DS3\_v2 14.0 GB Memory, 4 Cores, 0.75 DBU 2 8  **Enable Autoscaling**

**Auto Termination**  
 Terminate after 120 minutes of inactivity

**Spark** [Tags](#) [Logging](#)

**Spark Config**  
Enter your Spark configuration options here. Provide only one key-value pair per line.  
Example:  
spark.speculation true  
spark.kryo.registrator my.package.MyRegistrator





---

## Integration Runtime Setup



Install integration runtime on Windows machine or add further nodes using the Authentication Key.

### ▲ Integration Runtime Settings

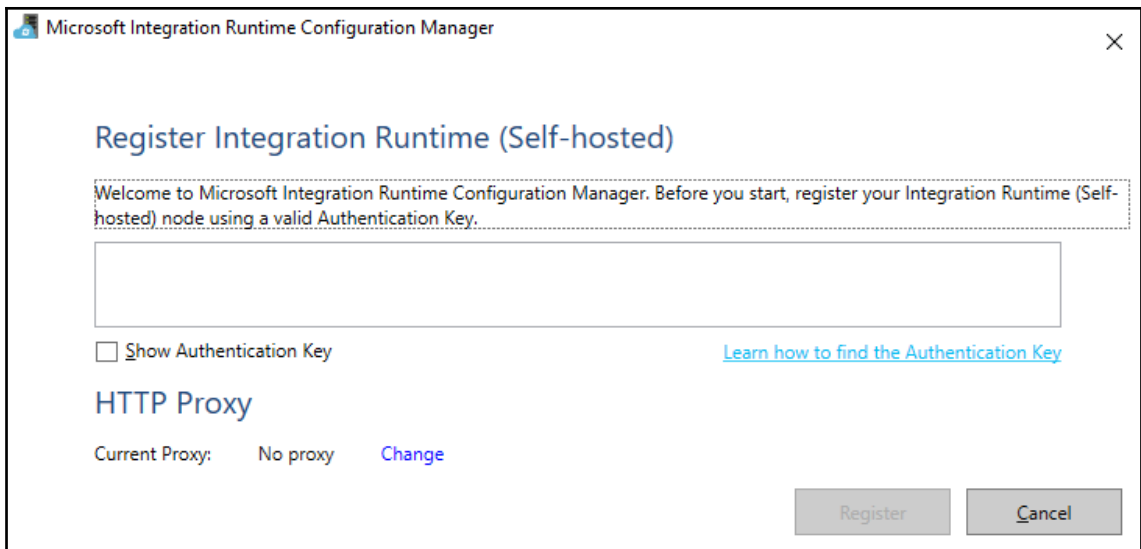
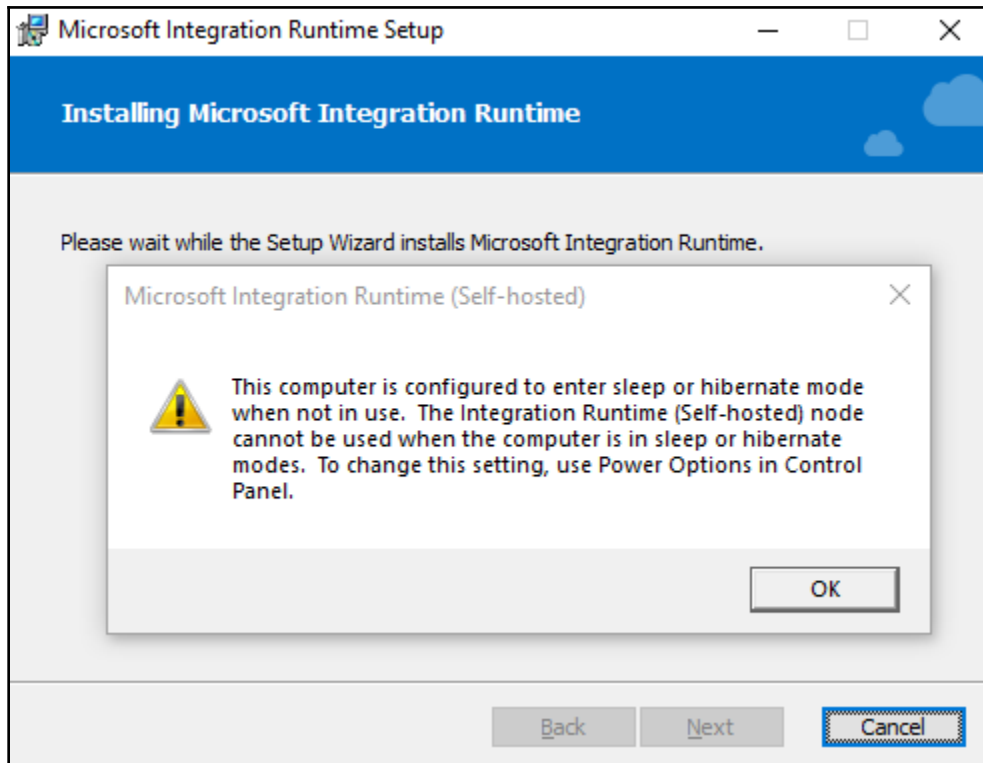
Option 1: Express setup

[Click here to launch the express setup for this computer](#)

Option 2: Manual setup

Step 1: [Download and install integration runtime](#)

Step 2: Use this key to register your integration runtime



---

## Integration Runtime Setup

Install integration runtime on Windows machine or add further nodes using the Authentication Key.

### Integration Runtime Settings





Option 1: Express setup

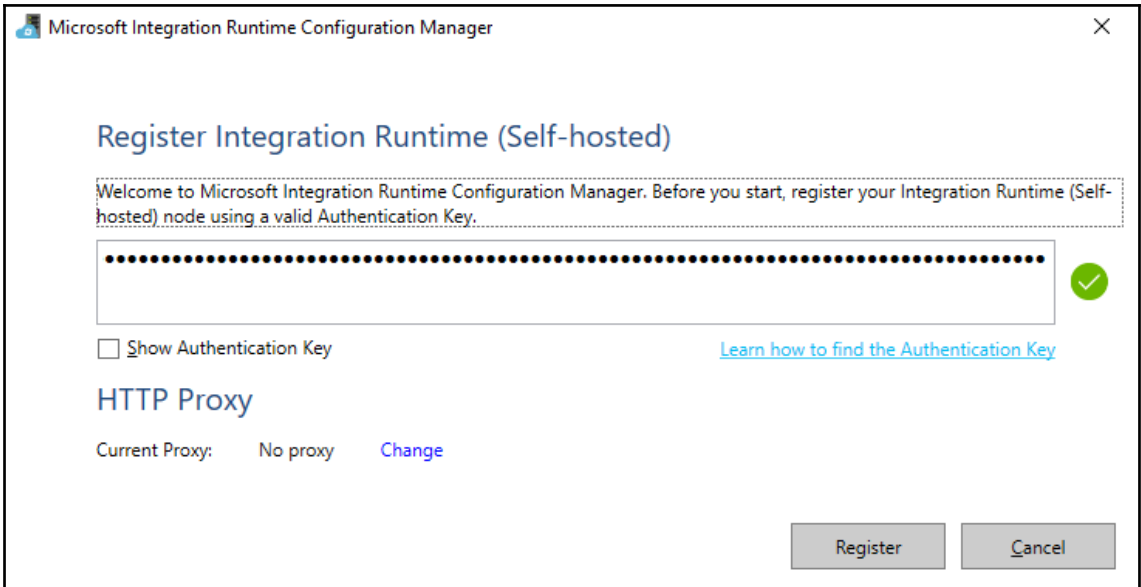
[Click here to launch the express setup for this computer](#)

Option 2: Manual setup

Step 1: [Download and install integration runtime](#)

Step 2: Use this key to register your integration runtime

Name	Authentication Key		
Key1	<input type="text" value="IR@a91dad52-b477-4d13-a436-8661d43be32b@ADFV2Book@e"/>		
Key2	<input type="text" value="IR@a91dad52-b477-4d13-a436-8661d43be32b@ADFV2Book@e"/>		



---

## New Linked Service

**Data Store** Compute

SQL Server

All Azure Database File Generic Protocol NoSQL



SQL Server



## New Linked Service ✕

Connect via integration runtime \*

WWImportersDW

[Edit Integration Runtime](#)

Connection String Azure Key Vault

Server name \*

ADFV2Book

Database name \*

WideWorldImportersDW

Authentication type

Windows Authentication

User name \*

██████████

Password Azure Key Vault

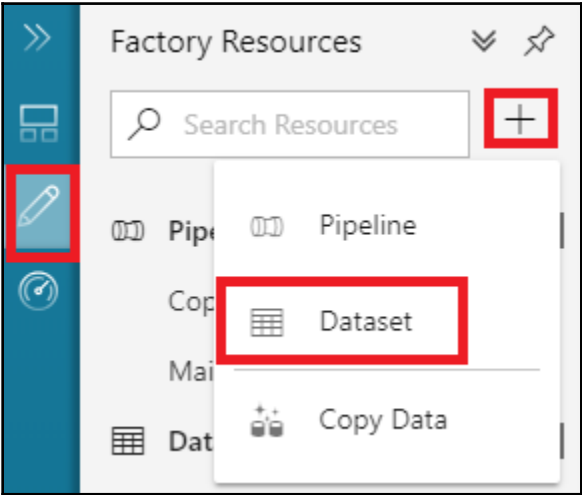
Password \*

.....

▶ Advanced ⓘ

✔ Connection successful

Cancel Test connection Finish



SQL Server  
ADFV2Book\_Sales

General Connection Schema Parameters Advanced

Linked service \* **WWImportersDWADFV2Book** Test connection Edit New

Table **[ADFV2Book].[Sales]** Refresh Preview data

**Data Preview**

Linked Service: WWImportersDWADFV2Book  
Object: [ADFV2Book].[Sales]

Date	Customer	Package	Description	Quantity	Unit Price	Tax Rate	Total Excluding Tax	Tax Amount	Profit	Total Including Tax
2013-01-01T00:00:00	Tails핀 Toys (Baraboo, WI)	Each	20 mm Anti static bubble wrap (Blue) 50m	40	102	15	4080	612	1880	4692
2013-01-01T00:00:00	Tails핀 Toys (Baraboo, WI)	Each	Developer joke mug - that's a hardware problem (Black)	4	13	15	52	7.8	34	59.8
2013-01-01T00:00:00	Tails핀 Toys (Baraboo, WI)	Each	Halloween skull mask (Gray) M	12	18	15	216	32.4	36	248.4
2013-01-01T00:00:00	Tails핀 Toys (Hahira, GA)	Each	Animal with big feet slippers (Brown) S	4	32	15	128	19.2	96	147.2

## New Linked Service



Data Store   Compute

 Blob|

All   Azure   Database   File   Generic Protocol   NoSQL   Services and apps



Azure Blob Storage

### New Linked Service ✕

Name \* i

Description

Connect via integration runtime \*

Authentication method

Connection String  Azure Key Vault


Account selection method i

Azure subscription i

Storage account name \*

▶ Advanced i

ADFV2Book\_Sa... \* X

 Azure Blob Storage  
**ADFV2Book\_Sales\_Destination**

General   Connection<sup>2</sup>   Schema   Parameters   Advanced

Name \*

Description

Azure Blob Storage  
**ADFV2Book\_Sales\_Destination**

General **Connection** Schema Parameters Advanced

Linked service \*  🔗 Test connection ✎ Edit + New

File path  /  ⓘ Browse 🔗 Preview data

Compression Type

Binary Copy ⓘ

▲ File Format Settings

File format  ⓘ Detect Text Format ⓘ

Column delimiter    
 Use custom delimiter

Row delimiter



▲ **Dataflow**

**Copy**

Copy\_ADFV2Book\_S...

General **Source** Sink Mapping Settings Advanced

Source Dataset \*  Edit + New Preview data

Use Query  Table  Query  Stored Procedure

General Source **Sink** Mapping Settings Advanced




Sink Dataset \*  Edit + New

Copy behavior


Debug

Pipeline Run ID: **00e53842-a8c8-4965-ad8e-2af05cd8ccc6**

Name	Type	Run Start	Duration	Status	Actions
Copy_ADFF...	Copy	05/21/2018 2:09 PM	00:00:17	Succeeded	


Details  Refresh  

[Learn more on copy performance details from here.](#)



SQL Server

Succeeded →

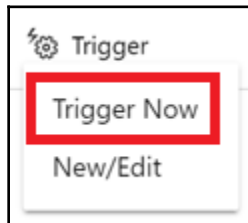


Azure Blob Storage

Data read:	51.865 MB	Data written:	33.971 MB
Rows read:	213,388	Files written:	1
		Rows written:	213,388
		Throughput:	5.628 MB/s

Copy duration 00:00:09

▶ SQL Server → Azure Blob Storage Queue 00:00:03 | Time to first byte 00:00:00 | Transfer 00:00:05





## Pipeline Run ✕

---

### Parameters

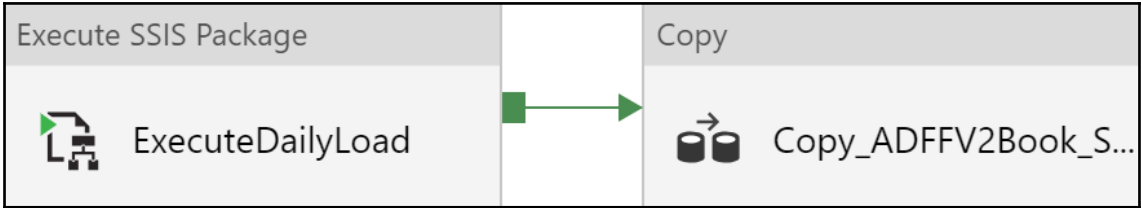
Name	Type	Value
No records found		

Cancel

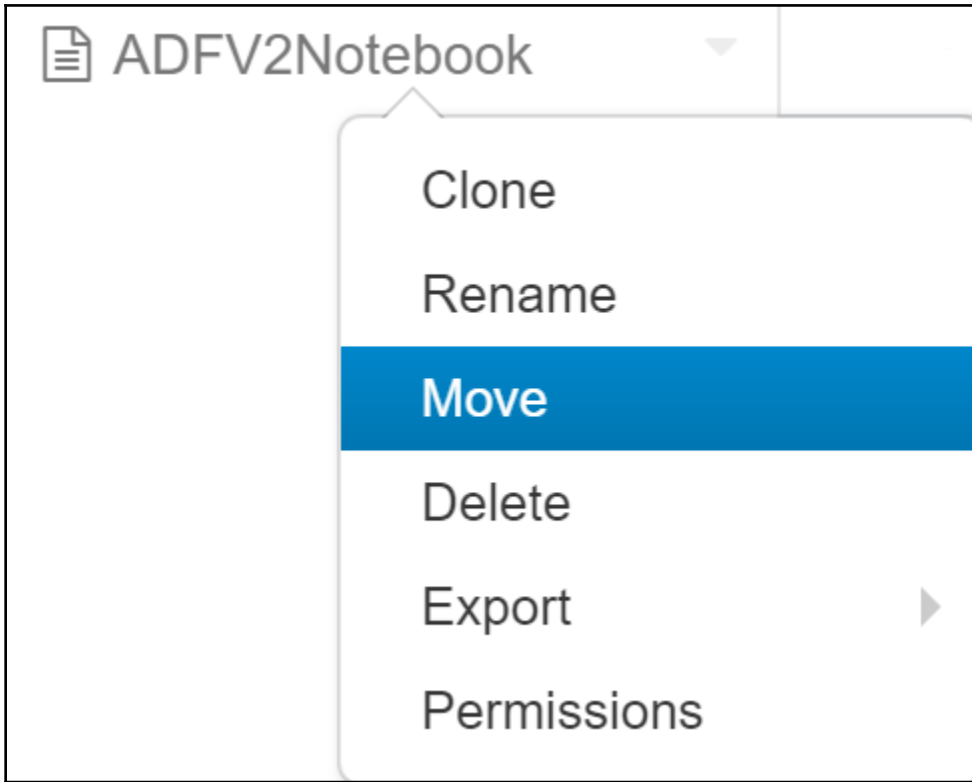
Finish

	Pipeline Name	Actions	Run Start	Duration	Triggered By	Status	Parameters	Error
<input type="checkbox"/>	pipeline1		05/21/2018, 2:24:48 PM	00:00:21	Manual trigger	<span style="color: green;">✔ Succeeded</span>		

NAME	MODIFIED	ACCESS TIER	BLOB TYPE	SIZE
Sales.txt	2018-05-21, 2:25:03 p.m.	Hot (Inferred)	Block blob	32.4 MiB



The screenshot shows the Microsoft Azure Databricks interface. On the left is a navigation sidebar with icons for "Data" (highlighted with a red box), "Recent", "Workspace", "Home", and "Databricks". The main area is divided into two columns: "Databases" and "Tables" (highlighted with a red box). The "Tables" column is currently empty. On the right, the "Create table" section is active, showing "Create New Table" options. Under "Data source", the "Spark Data Sources" button is highlighted with a red box. Under "Connector", the "Azure Blob Storage" dropdown is highlighted with a red box. At the bottom of this section, the "Create Table in Notebook" button is also highlighted with a red box.



---

## Moving 'ADFV2Notebook'

Moving to folder: /ADFCalls/

Shared	
Users	
ADFCalls	

Cancel

Select

1

**Load data from Blob storage**

2

**Transform / aggregate it**

3

**Create a table in Spark to expose it to other client tools**

# ADFV2Notebook (Python)



Detached ▼



File ▼



● ADFV2Book (Terminated, 4.0 (includes Apache Spark 2.3.0, Scala 2.11))



Attached: ● ADFV2Book ▼



File ▼



View:

Terminated

Detach

**Start Cluster**

View Spark UI

View Driver Logs

Cmd 3

```
1 dbutils.widgets.text("storage_account_name", "STORAGE_ACCOUNT_NAME", "Storage Account Name")
2 dbutils.widgets.text("storage_account_access_key", "YOUR_ACCESS_KEY", "Storage Access Key / SAS")
```

Command took 0.08 seconds -- by ccote\_1@msn.com at 5/21/2018, 9:14:38 AM on test

Cmd 4

```
1 dbutils.widgets.text("file_location", "wasbs://example/location", "Upload Location")
2 dbutils.widgets.dropdown("file_type", "csv", ["csv", 'parquet', 'json'])
```

```
Cmd 5
1 spark.conf.set(
2   "fs.azure.account.key."+dbutils.widgets.get("storage_account_name")+".blob.core.windows.net",
3   dbutils.widgets.get("storage_account_access_key"))

Command took 0.03 seconds -- by ccote_1@msn.com at 5/21/2018, 9:14:38 AM on test
```

### Step 3: Create tables

We will create a temporary table, SalesTmp from the dataframe and aggregate it's content into Sales and SalesSplit.

```
Cmd 9
This table will persist across cluster restarts and allow various users across different notebooks to query this data.

Cmd 10
1 %sql
2 DROP TABLE IF EXISTS SalesTmp

OK
Command took 1.78 seconds -- by ccote_1@msn.com at 5/21/2018, 9:14:39 AM on test

Cmd 11
1 df.write.format("parquet").saveAsTable("SalesTmp")
```

```
1 %sql
2
3 DROP TABLE IF EXISTS Sales;
4 CREATE TABLE Sales
5   AS SELECT CAST(_c0 AS DATE) Date, _c1 AS Customer, _c2 AS Description, SUM(_c3) AS Quantity FROM SalesTmp GROUP BY _c0, _c1, _c2 ORDER BY Date,
6   Customer;
7 SELECT * From Sales;
```

```
1 %sql
2
3 DROP TABLE IF EXISTS Sales;
4 CREATE TABLE Sales
5   AS SELECT CAST(_c0 AS DATE) Date, _c1 AS Customer, _c2 AS Description, SUM(_c3) AS Quantity FROM SalesTmp GROUP BY _c0, _c1, _c2 ORDER BY Date, Customer;
6 SELECT * FROM Sales;
7
```

▼ (4) Spark Jobs

- ▶ Job 6 [View](#) (Stages: 2/2)
- ▶ Job 7 [View](#) (Stages: 2/2, 1 skipped)
- ▶ Job 8 [View](#) (Stages: 1/1)
- ▶ Job 9 [View](#) (Stages: 2/2)

Showing sample based on the first 1000 rows.

Grid icon | Plot Options... | Download icon

# New Linked Service



Data Store **Compute**



Azure Batch



Azure Data Lake Analytics



Azure Databricks



Azure HDInsight



Azure ML

Cancel

Continue



Type \*

Azure Databricks

Connect via integration runtime \*

Default

Account selection method

From Azure subscription

Select cluster

New job cluster  Existing cluster

Domain/Region \*

East US

Access token \*

\*\*\*\*\*

Cluster node type \*

Standard\_DS3\_v2 (14 GB Memory, 4 Cores, 0.75 DBU)

Cluster version \*

4.0 (includes Apache Spark 2.3.0, Scala 2.11)

Workers \*

1

Enable Autoscaling

Cancel Finish

Signed in as  
ccote\_1@msn.com

**User Settings**

Admin Console

Manage Account

Log Out

Microsoft Azure

User Settings

Access Tokens   Git Integration   Notebook Settings

Personal access tokens can be used for secure authentication to the [Databricks API](#) instead of passwords.

[Generate New Token](#)

Token ID	Comme
No tokens exist.	

### Generate New Token

Comment

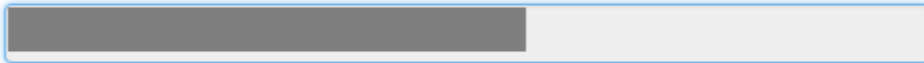
Lifetime (days)

[Cancel](#) [Generate](#)

---

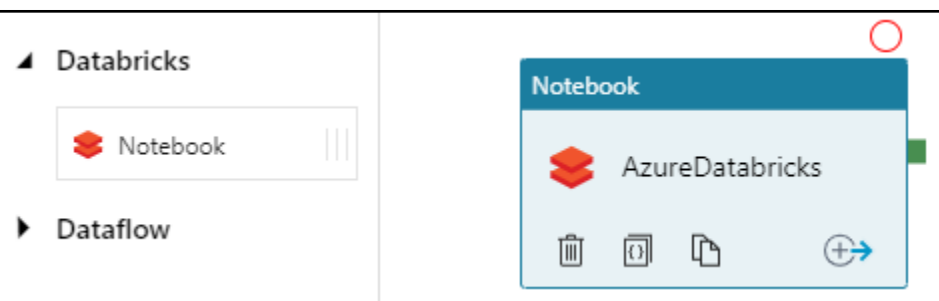
## Generate New Token

Your token has been created successfully.



⚠ Make sure to copy the token now. You won't be able to see it again.

Done





Job Clusters						
Name	State	Nodes	Driver	Worker	Runtime	
● job-11-run-1	Terminated ?	-	Standar...	Standar...	4.0 (includ...	

Untitled (Run id: 11) | ✕ Delete

Started: 2018-05-21 19:15:40 EDT  
 Duration: 6m 22s  
 Status: Succeeded  
 Run ID: 11  
 Task: Notebook at /ADFCalls/ADFV2Notebook  
 Parameters:  
 {"storage\_account\_name":"ADFV2Book","storage\_account\_access\_key":"Evy1DQh7va94QQiBlIsKzx...  
 Cluster: Driver: Standard\_DS3\_v2, Workers: Standard\_DS3\_v2, 28 GB, 4.0 (includes Apache Spark 2.3.0, Scala 2.11) - [View Spark UI / Logs](#)

### Output




This notebook shows how to create and query a table or DataFrame on Azure Blob Storage.

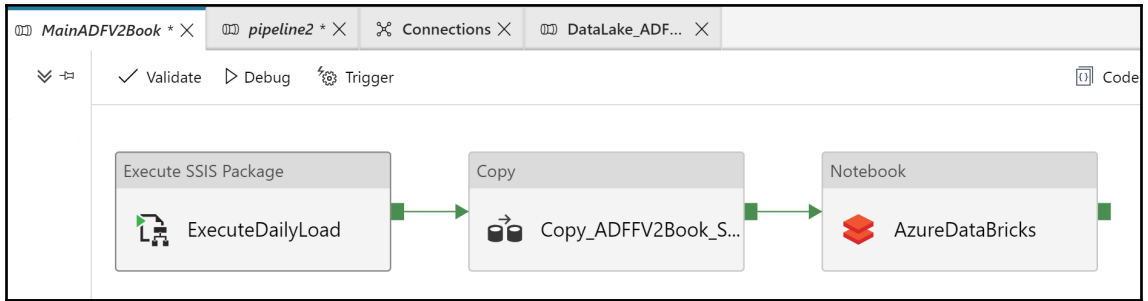
#### Step 1: Set the data location and type

There are two ways in Databricks to read from Azure Blob Storage. You can either read data using [account keys](#) or read data using shared access signatures (SAS).

To get started, we need to set the location and type of the file. We can do this using [widgets](#). Widgets allow us to parameterize the execution of this entire notebook. First we create them, then we can reference them throughout the notebook.

```
dbutils.widgets.text("storage_account_name", "STORAGE_ACCOUNT_NAME", "Storage Account Name")
dbutils.widgets.text("storage_account_access_key", "YOUR_ACCESS_KEY", "Storage Access Key / SAS")
```

General						Parameters		Output	
Pipeline Run ID: 62d9b46d-0d77-440d-9aa7-801ed380a81d									
Name	Type	Run Start	Duration	Status					
AzureDatab...	DatabricksN...	05/21/2018 7:15 PM	00:06:33	 Succeeded					



---

# **Chapter 7: Reporting on the Modern Data Warehouse**

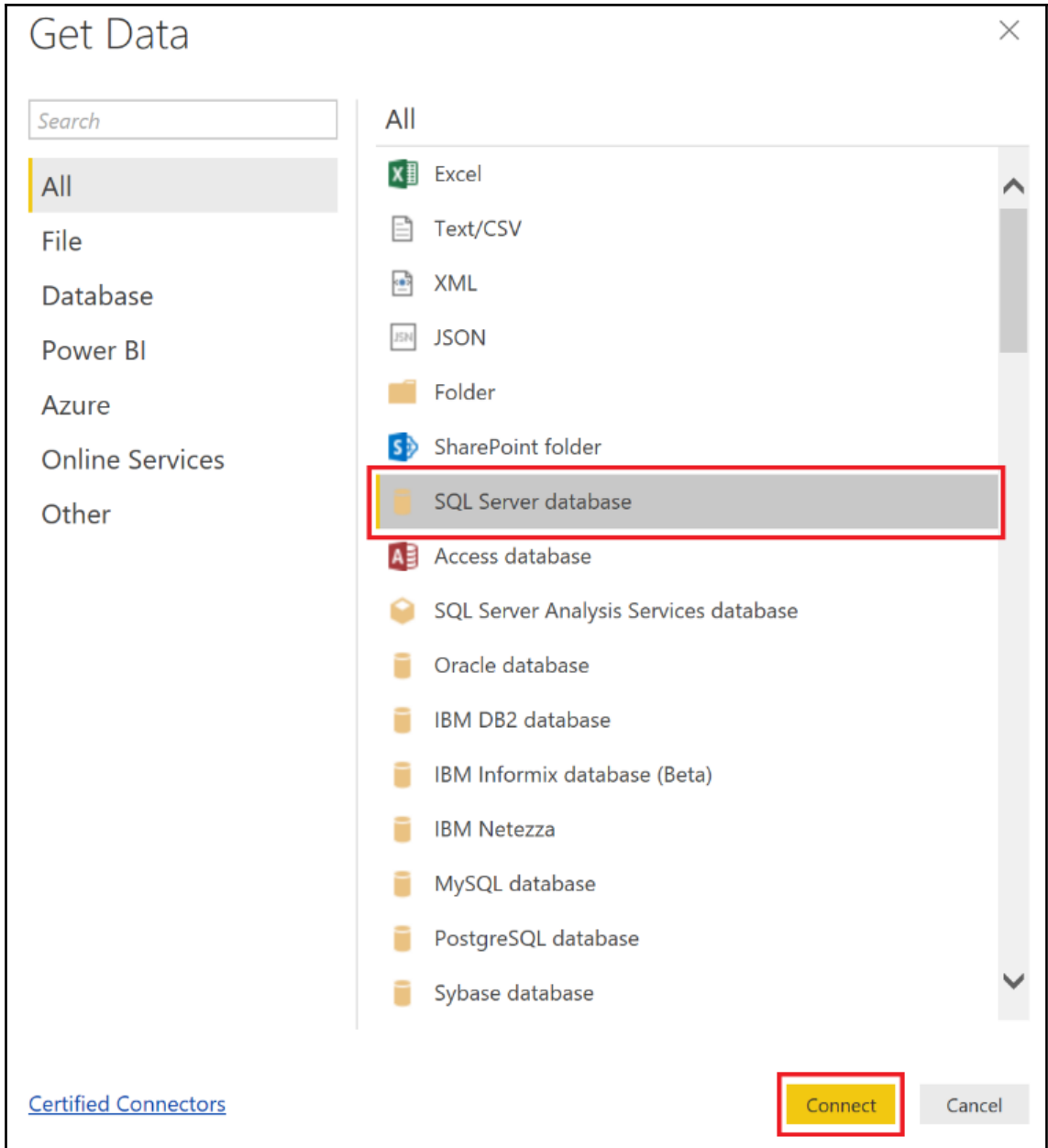
---

# Welcome to Power BI Desktop

Where can we send you the latest tips and tricks for Power BI?

Microsoft may use your contact information to provide updates and special offers about Business Intelligence and other Microsoft products and services. You can unsubscribe at any time. To learn more you can read the [privacy statement](#).





SQL Server database ✕

Server ⓘ  
ADfv2Book

Database (optional)

Data Connectivity mode ⓘ

Import

DirectQuery

> Advanced options

OK Cancel

SQL Server database ✕

Windows

Database

Microsoft account

adfv2book;WWImportersDW

Use your Windows credentials to access this database.

Use my current credentials

Use alternate credentials

User name

Password

Select which level to apply these settings to  
adfv2book

Back Connect Cancel

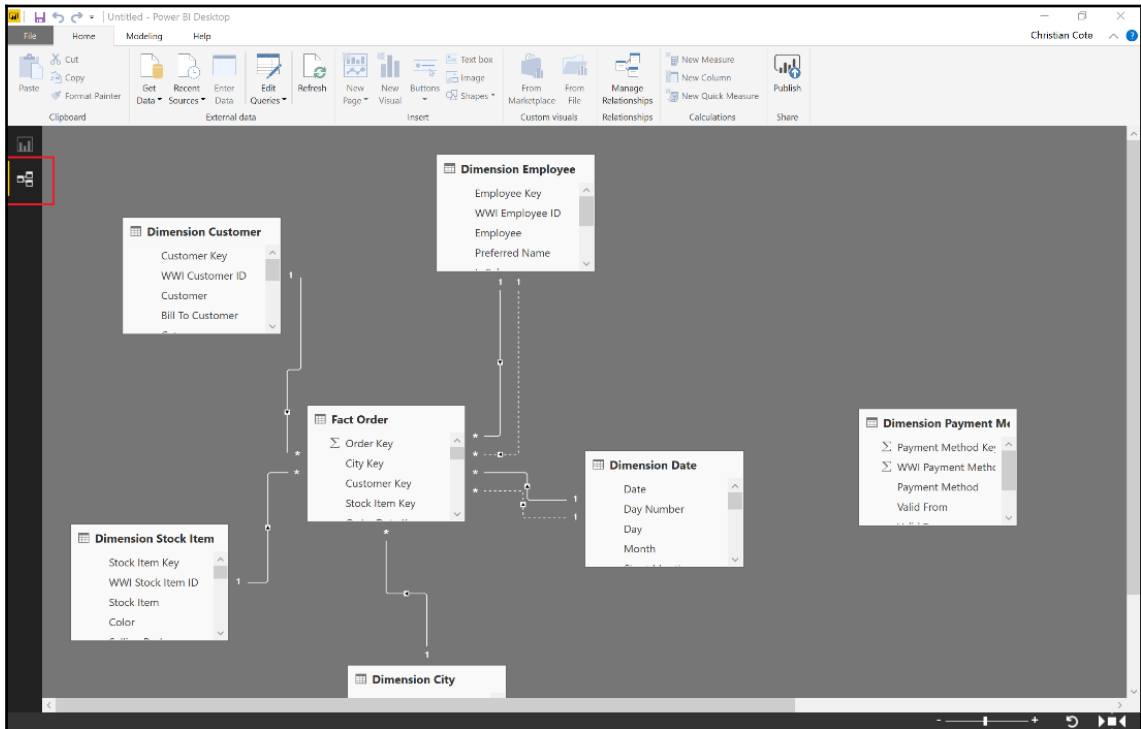
# Encryption Support



We were unable to connect to the data source using an encrypted connection. To access this data source using an unencrypted connection, click OK.

OK

Cancel



Untitled - Power Query Editor

File Home Transform Add Column View Help

Close & Apply Close New Source Recent Sources Enter Data Data source settings Manage Parameters Refresh Preview Properties Advanced Editor Manage Query

Queries [9]

This preview may be up to 2 days old. Refresh

	Date	Day Number	Day
1	1/1/2013	1	1
	1/2/2013	2	2
	1/3/2013	3	3
	1/4/2013	4	4
	1/5/2013	5	5
	1/6/2013	6	6
	1/7/2013	7	7
	1/8/2013	8	8
	1/9/2013	9	9
	1/10/2013	10	10
	1/11/2013	11	11
	1/12/2013	12	12
	1/13/2013	13	13
	1/14/2013	14	14
	1/15/2013	15	15
16	1/16/2013	16	16

Copy  
 Paste  
 Delete  
 Rename  
 Enable load  
 Include in report refresh  
 Duplicate  
 Reference  
 Move To Group  
 Move Up  
 Move Down  
 Create Function...  
 Convert To Parameter  
 Advanced Editor  
 Properties...

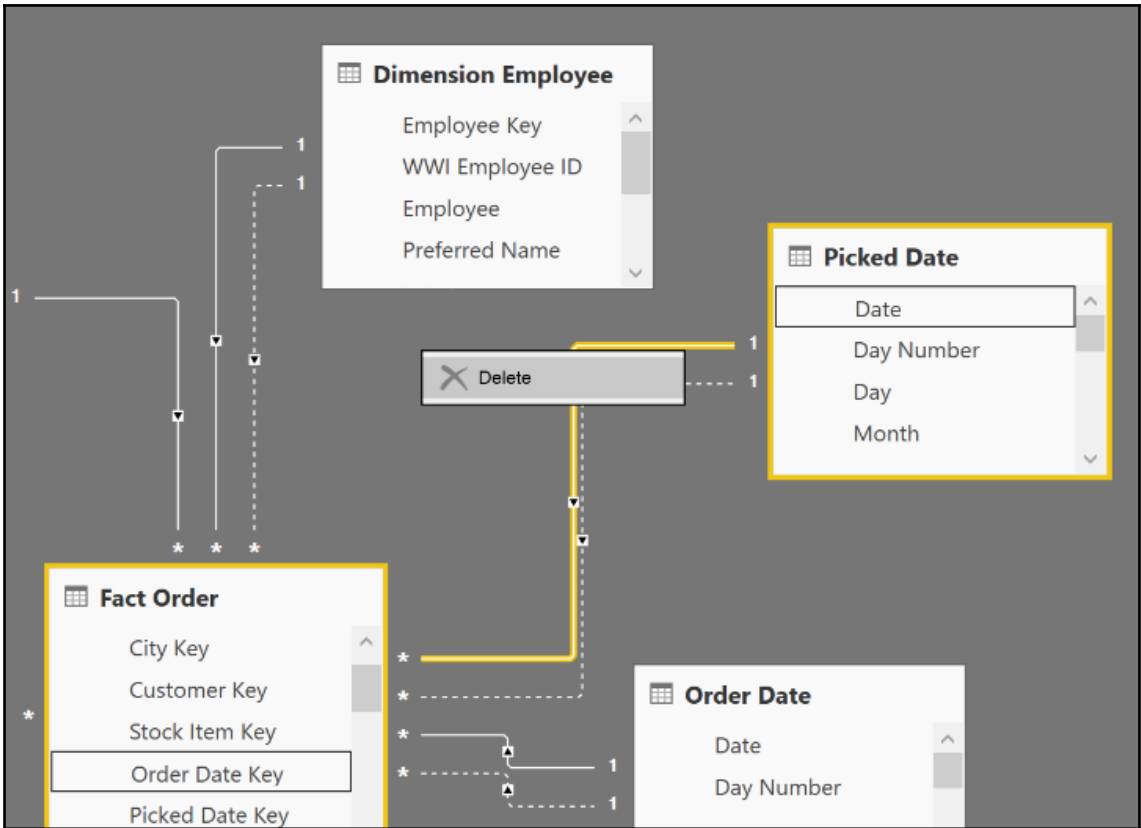
**Queries [10]** <

- Dimension City
- Order Date
- Dimension Customer
- Dimension Employee
- Dimension Payment Met...
- Dimension Stock Item
- Dimension Supplier
- Fact Order
- Dimension Transaction T...
- Order Date (2)**

**i** This preview may be up to 2 days

	Date	1 <sup>2</sup> 3 Day Number
1	1/1/2013	
2	1/2/2013	
3	1/3/2013	
4	1/4/2013	
5	1/5/2013	
6	1/6/2013	
7	1/7/2013	
8	1/8/2013	
9	1/9/2013	
10	1/10/2013	
	1/11/2013	

- Copy
- Paste
- Delete
- Rename**
- Enable load
- Include in report refresh



## Edit relationship



Select tables and columns that are related.

Fact Order

Order Key	City Key	Customer Key	Stock Item Key	Order Date Key	Picked Date Key	Sales
142	56757	0	202	Tuesday, January 1, 2013	Wednesday, January 2, 2013	
81	50119	0	193	Tuesday, January 1, 2013	Tuesday, January 1, 2013	
12	50969	0	180	Tuesday, January 1, 2013	Tuesday, January 1, 2013	

Picked Date

Date	Day Number	Day	Month	Short Month	Calendar Month Number	Calendar M
Tuesday, January 1, 2013	1	1	January	Jan	1	CY2013-Jan
Wednesday, January 2, 2013	2	2	January	Jan	1	CY2013-Jan
Thursday, January 3, 2013	3	3	January	Jan	1	CY2013-Jan

Cardinality

Many to one (\*:1)

Cross filter direction

Single

Make this relationship active

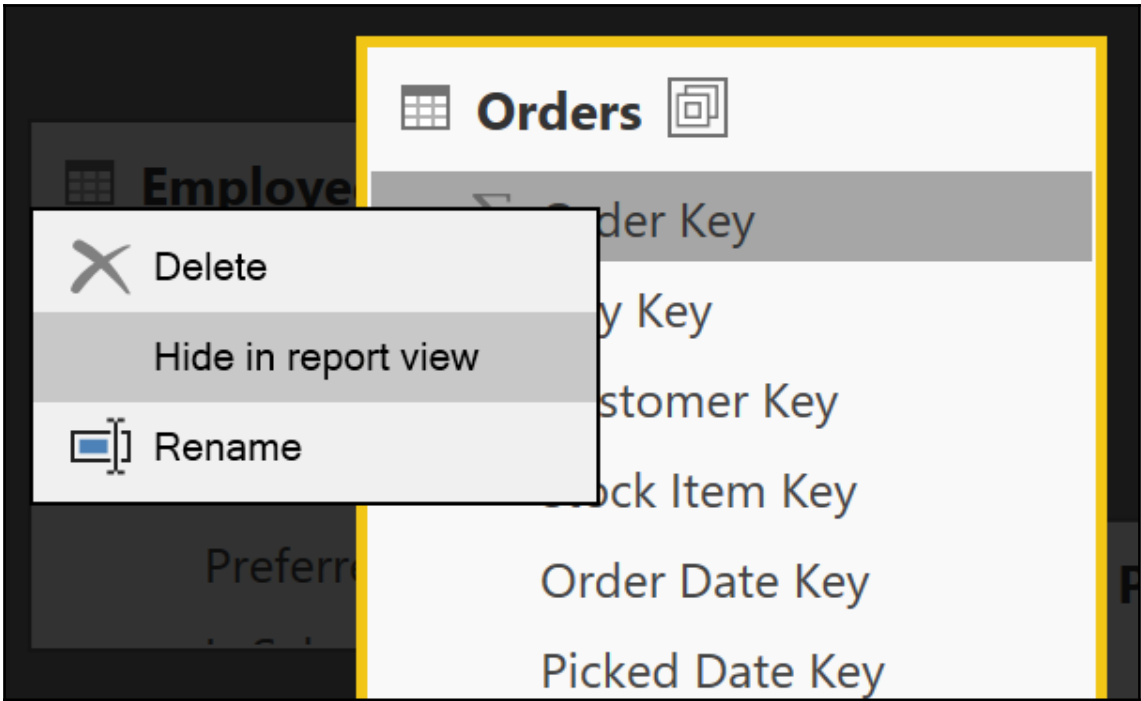
Apply security filter in both directions

Assume referential integrity

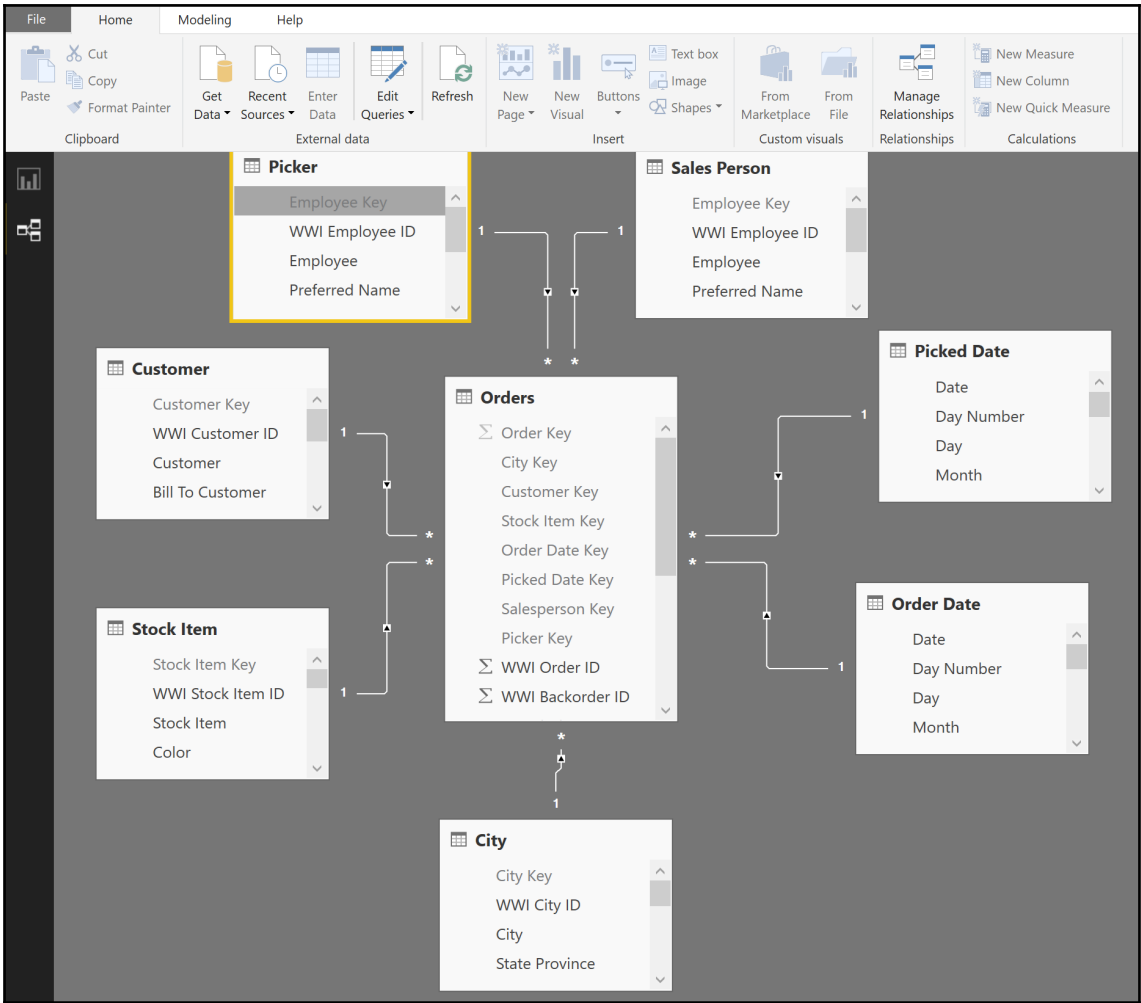
[Learn more](#)

OK

Cancel








---

## VISUALIZATIONS >



Values

Drag data fields here

## FILTERS

Page level filters

Drag data fields here

## FIELDS >

Search

- ▶ City
- ▶ Customer
- ▶ Order Date
- ▶ Orders
- ▶ Picked Date
- ▶ Picker
- ▶ Sales Person
- ▶ Stock Item

---



# World Wide Importers

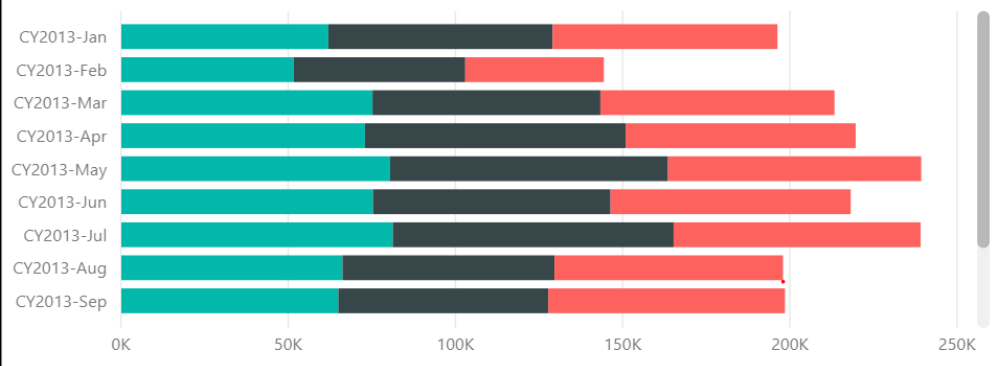
Orders by year

Calendar Year

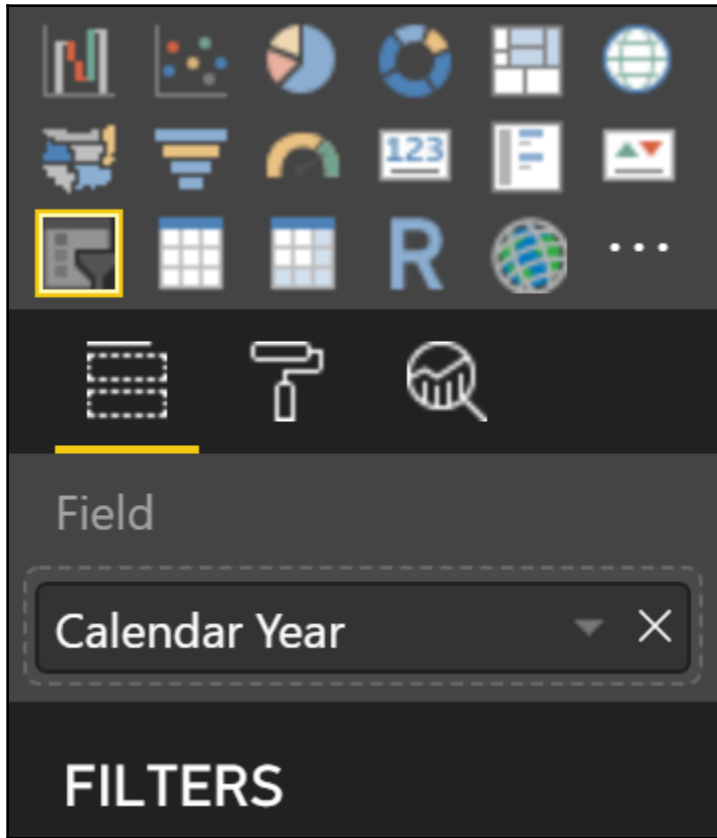
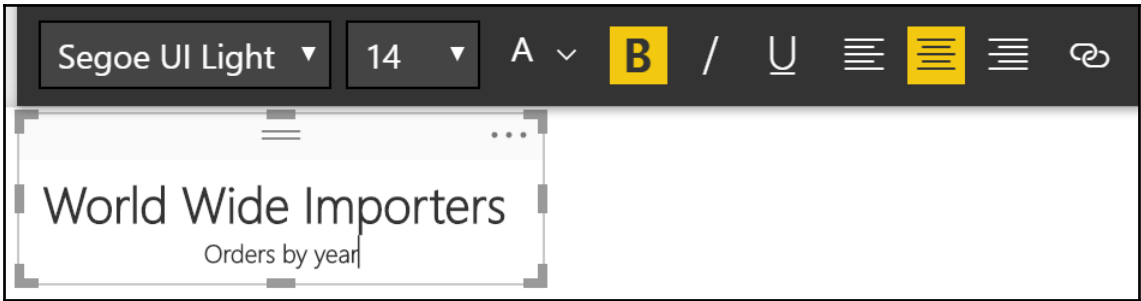
2013

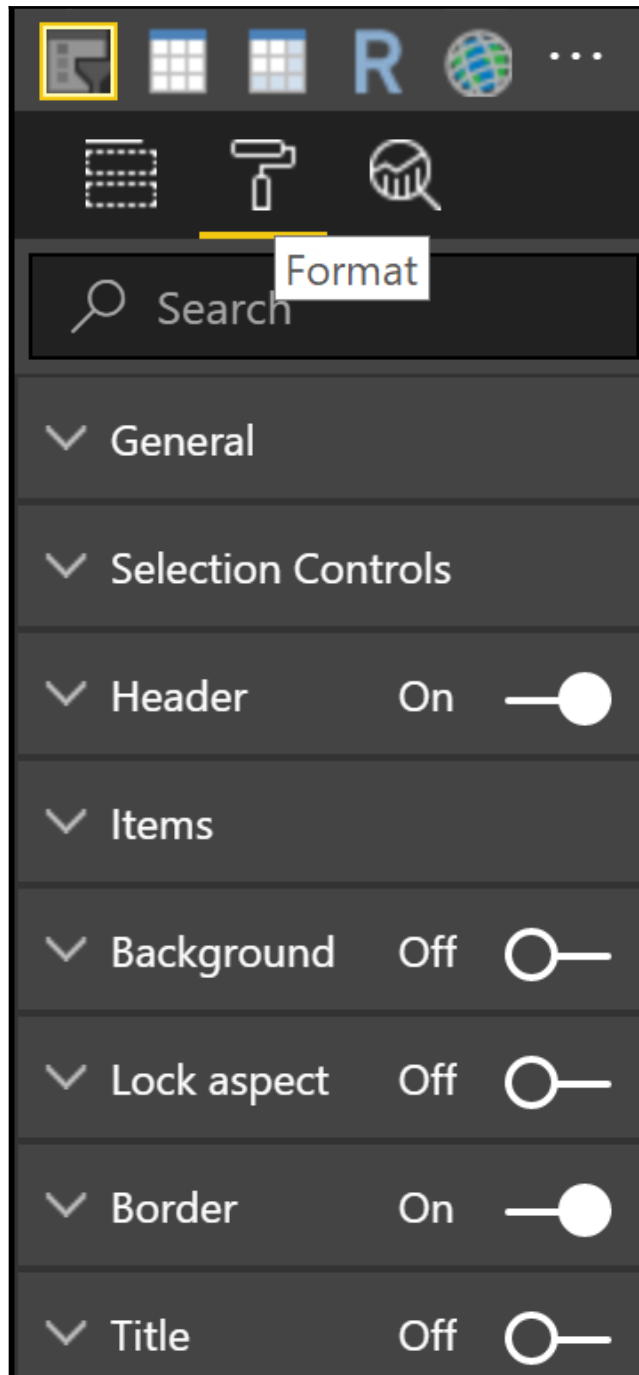
Quantity by Calendar Month Label and Buying Group

Buying Group ● N/A ● Tailspin Toys ● Wingtip Toys







Stock Item	Quantity
Black and orange fragile despatch tape 48mmx75m	58068
Black and orange fragile despatch tape 48mmx100m	51948
Shipping carton (Brown) 500x310x310mm	45625
Black and orange this way up despatch tape 48mmx100m	42840
3 kg Courier post bag (White) 300x190x95mm	42400
Shipping carton (Brown) 305x305x305mm	41925
Shipping carton (Brown) 229x229x229mm	41650
Clear packaging tape 48mmx75m	41366
Shipping carton (Brown) 457x457x457mm	40750
Shipping carton (Brown) 356x356x279mm	40450
Shipping carton (Brown) 413x285x187mm	39825
Express post box 5kg (White) 350x280x130mm	39800
Black and yellow heavy despatch tape 48mmx75m	39336
<b>Total</b>	<b>2473653</b>





### VISUALIZATIONS >



Axis

Calendar Month Label ▼ ×

Date ▼ ×




Legend

Buying Group ▼ ×

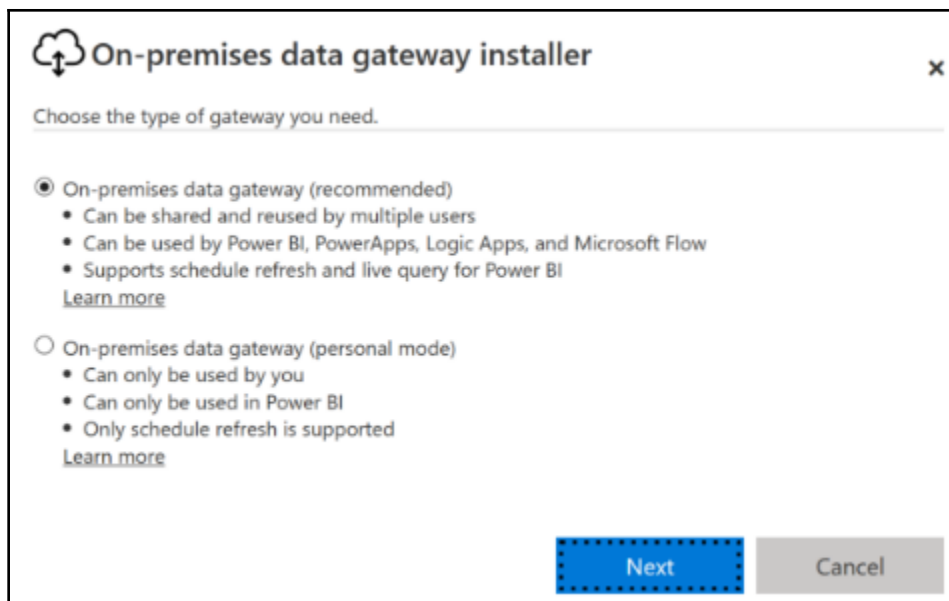
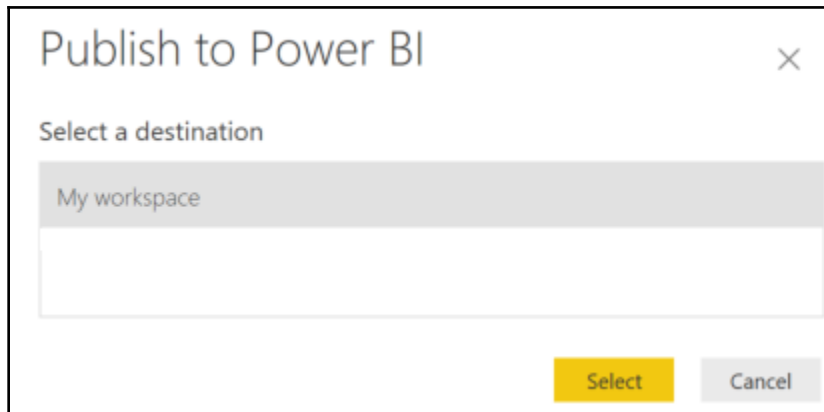
Value


Quantity ▼ ×

### FIELDS >

- ▶  City
- ▶  Customer
- ▶  Order Date

- Calendar Mon...
- Calendar Mon...
- Calendar Year
- Calendar Year ...
- Date
- Day
- Day Number
- Fiscal Month L...
- Fiscal Month ...
- Fiscal Year
- Fiscal Year Lab...
- ISO Week Nu...



 **On-premises data gateway** ? ×

You are signed in as ChristianCote@ia-techconsulting.com and are ready to register the gateway.

New on-premises data gateway name

Add to an existing gateway cluster

Recovery key (8 character minimum)

ⓘ This key is needed to restore the gateway and can't be changed. Record it in a safe place.

Confirm recovery key

[Learn more about gateway clusters](#)

We'll use this region to connect the gateway to cloud services: North Central US [Change Region](#)



**On-premises data gateway**

**Status**

✓ The gateway ADFV12Book is online and ready to be used.

Gateway version number: 14.16.6697.1 (May 2018)

Help us improve the on-premises data gateway by sending usage information to Microsoft.  
[Read the privacy statement online](#)

**Logic Apps, Azure Analysis Services** [Create a gateway in Azure](#)  
North Central US

**PowerApps, Microsoft Flow** ✓ Ready  
Default environment

**Power BI** ✓ Ready  
Default environment

Close

ADD DATA SOURCE

GATEWAY CLUSTERS

▼ ADFV12Book

New data source

Test all connections

### Data Source Settings

Data Source Name

Data Source Type

Server

Database

Authentication Method

The credentials are encrypted using the key stored on-premises on the gateway server. [Learn more](#)

Username

Password

> Advanced settings

Add Discard

ADD DATA SOURCE

GATEWAY CLUSTERS

- ADFV12Book
  - WideWorldImportersDW

Test all connections

### Data Source Settings

Users

✓ Connection Successful

🕒 Next Step: Go to the [Users tab](#) above and add users to this Data Source

Data Source Name: WideWorldImportersDW

Data Source Type: SQL Server

Server: ADFV2Book

Database: WideWorldImportersDW

Authentication Method: Basic

The credentials are encrypted using the key stored on-premises on the gateway server. [Learn more](#)

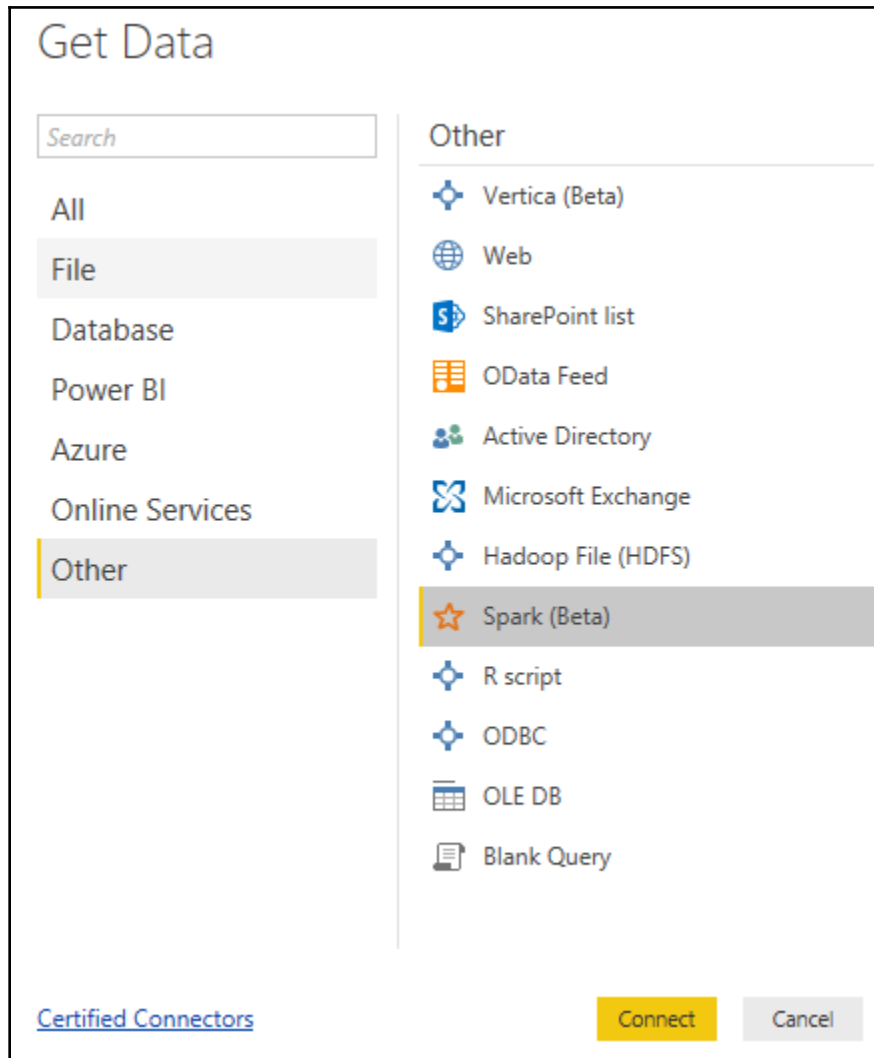
Username: [REDACTED]

Password: [REDACTED]

> Advanced settings

Apply Discard





---

Spark ×

Server

Protocol

Data Connectivity mode ⓘ

Import

DirectQuery

Clusters / ADFV2Book

# ADFV2Book

[Edit](#) [Start](#) [Clone](#)

[Configuration](#) [Notebooks \(0\)](#) [Libraries \(0\)](#) [Event Log](#) [Spark UI](#) [Driver Logs](#) [Spark Cluster UI - Master](#)

Cluster Type  
Serverless Pool (beta, R/Python/SQL) **Standard** [Learn more about Serverless Pools](#)

Databricks Runtime Version  
4.0 (includes Apache Spark 2.3.0, Scala 2.11)

Python Version  
2

Driver Type  
Standard\_DS3\_v2 14.0 GB Memory, 4 Cores, 0.75 DBU

Worker Type  
Standard\_DS3\_v2 14.0 GB Memory, 4 Cores, 0.75 DBU

Min Workers: 2 Max Workers: 8  Enable Autoscaling

Auto Termination  
 Terminate after 120 minutes of inactivity

[Spark](#) [Tags](#) [Logging](#) **[JDBC/ODBC](#)** [Permissions](#)

Server Hostname  
eastus.azuredatabricks.net

Port  
443

Protocol  
HTTPS

HTTP Path  
sql/protocolv1/o/5935092853513601/0519-233033-mimes124 (unique)  
sql/protocolv1/o/5935092853513601/adfv2book (alias, not guaranteed unique)

JDBC URL  
jdbc:hive2://eastus.azuredatabricks.net:443/default;transportMode=http;ssl=true;h  
ttpPath=sql/protocolv1/o/5935092853513601/0519-233033-mimes124

JDBC URL ?

`jdbc:hive2://eastus.azuredatabricks.net:443/`

`sql/protocolv1/o/5935092853513601/0519-233033-mimes124`

Windows

Username / Password

Spark

☆ <https://eastus.azuredatabricks.net:443/sql/protocolv1/o/5935092853513601/0519-233033-mimes124>

User name

token

Password

.....

Back Connect Cancel

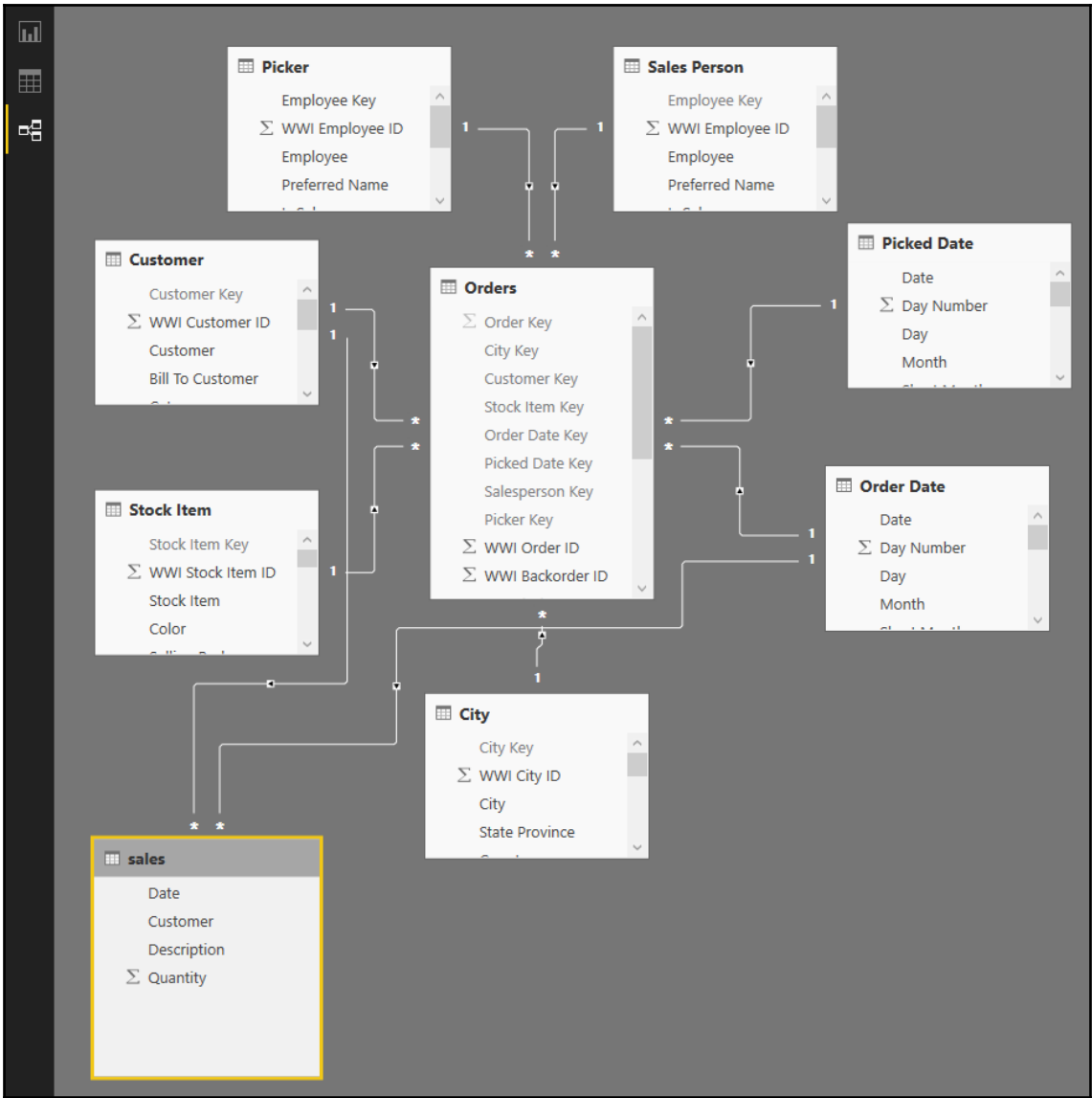
## Switching to import mode

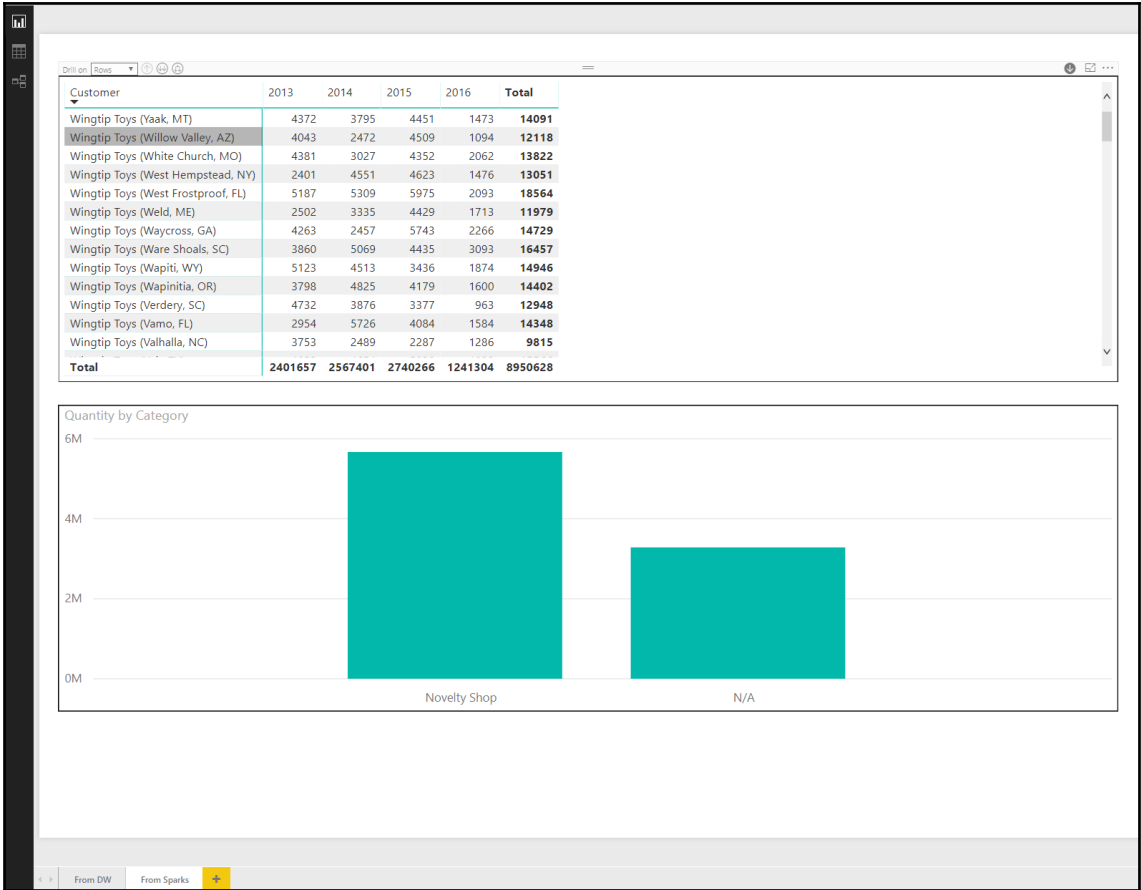
The data source you are trying to connect to is different than the existing DirectQuery data source. To continue, all queries must be switched to import mode, which may result in a large amount of data being loaded.

Switch

Cancel







### VISUALIZATIONS >

Rows

- Customer ▾ ×
- Bill To Customer ▾ ×
- Primary Contact ▾ ×

Columns

- Calendar Year ▾ ×
- Calendar Month Label ▾ ×
- Date ▾ ×

Values

- Quantity ▾ ×

### FILTERS

Visual level filters

- Bill To Customer (All)
- Calendar Month Label (All)
- Calendar Year (All)
- Customer (All)
- Date (All)
- Primary Contact (All)
- Quantity (All)

Page level filters

Drag data fields here

Report level filters

Drag data fields here

### DRILLTHROUGH

Keep all ... On

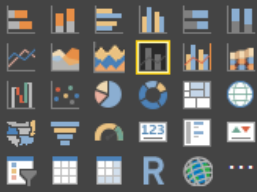
Drag drillthrough fields here

### FIELDS >

Search

- ▶ City
- ▶ Customer
  - Bill To Custom...
  - Buying Group
  - Category
  - Customer
  - Lineage Key
  - Postal Code
  - Primary Contact
  - Valid From
  - Valid To
  - WWI Custome...
- ▶ Order Date
  - Calendar Mon...
  - Calendar Mon...
  - Calendar Year
  - ▶ Calendar Year ...
    - Calendar Year
    - Calendar Mo...
    - Calendar Year ...
    - Date
    - Day
    - Day Number
    - Fiscal Month L...
    - Fiscal Month ...
    - Fiscal Year
    - Fiscal Year Lab...
    - ISO Week Nu...
    - Month
    - Short Month
- ▶ Orders
- ▶ Picked Date
- ▶ Picker
- ▶ sales
  - Customer
  - Date
  - Description
  - Quantity
- ▶ Sales Person
- ▶ Stock Item

### VISUALIZATIONS >



Shared axis

Category X

---

Column series

Drag data fields here

Column values

Quantity X

---

Line values

Drag data fields here

Tooltips

Drag data fields here

### FILTERS

Visual level filters

Category (All)

Quantity (All)

---

Page level filters

Drag data fields here

Report level filters

Drag data fields here

### FIELDS >

- ▶ City
- ▶ Customer
  - Bill To Custom...
  - Buying Group
  - Category
  - Customer
  - Lineage Key
  - Postal Code
  - Primary Contact
  - Valid From
  - Valid To
  - Σ WWI Custome...
- ◀ Order Date
  - Calendar Mon...
  - Σ Calendar Mon...
  - Σ Calendar Year
  - ◀  Calendar Year ...
  - Σ Calendar Year
  - Σ Calendar Mo...
  - Calendar Year ...
  - Date
  - Day
  - Σ Day Number
  - Fiscal Month L...
  - Σ Fiscal Month ...
  - Σ Fiscal Year
  - Fiscal Year Lab...
  - Σ ISO Week Nu...
  - Month
  - Short Month
- ▶ Orders
- ▶ Picked Date
- ▶ Picker
- ◀ sales

### DRILLTHROUGH


Keep all ... On

---

## Publishing to Power BI ✕

• Publishing 'ADFV2Book.pbix' to Power BI

---

 **Did you know?** You can create a portrait view of your report tailored for mobile phones, on the **View** tab select **Phone Layout**. [Learn more](#)

Cancel


## Publishing to Power BI ✕

✓ Success!

[Open 'ADFV2Book.pbix' in Power BI](#)

[Get Quick Insights](#)

---

 **Did you know?** You can create a portrait view of your report tailored for mobile phones, on the **View** tab select **Phone Layout**. [Learn more](#)

Got it

