

# Contrôle continu Microsoft Azure

Noms des participants :

- TCHATCHOUA NJANJO LIONEL
- NDIAYE SERIGNE MBACKE
- FOMENA TSATSOP JOEL VALDES

## Architecture du projet

1. **Ingestion** : Stockage du fichier CSV brut dans **Azure Blob Storage**.
2. **Stockage structuré** : Création d'une **Azure SQL Database**.
3. **ETL & Transformation (Bonus)** : Utilisation d'**Azure Data Factory (Data Flow)** pour nettoyer les données avant insertion.
4. **Analyse** : Requêtes SQL pour extraire des indicateurs.

## Étape 1 : Préparation du Stockage

*Objectif : Créez l'espace pour recevoir le fichier Kaggle.*

1. Connectez-vous au **Portail Azure**.

The screenshot shows the Microsoft Azure portal interface. At the top, there's a navigation bar with 'Microsoft Azure', a search bar, and user information. Below it is the 'Education | Vue d'ensemble' dashboard. The dashboard has several sections: 'Prise en main' (Getting Started) with 'Vue d'ensemble' selected, 'Détails de l'offre étudiant' (Student offer details) showing credits available (100 \$US sur 100 \$US), 'Solutions populaires' (Popular solutions) with links like 'Déployez un conteneur Docker', 'Créez votre première application Node.js', 'Créer et entraîner un modèle Machine Learning', and 'Créez et déployez votre premier site web'; 'Services gratuits' (Free services) with links like 'Machines virtuelles Azure – Windows', 'Stockage Blob Azure', 'Vision par ordinateur', and 'Azure App Service'; 'Logiciel gratuit' (Free software) with links to 'SQL Server 2019 Developer' and 'Agents for Visual Studio 2019 (version 16.0) test Agent'; 'Parcours d'apprentissage' (Learning paths) with 'Scientifique des données' and 'Ingénieur IA'; and a 'Ressources' (Resources) section with links to 'Guide de démarrage pour les développeurs Azure' and 'Calculatrice de prix'. A sidebar on the left lists 'Vue d'ensemble', 'Ressources pédagogiques', and 'Besoin d'aide ?'. At the bottom, there's a note: 'Ajouter ou supprimer des favoris en appuyant sur'.

2. Recherchez et sélectionnez **Comptes de stockage** (Storage accounts).
3. Cliquez sur **+ Créer**.
  - **Groupe de ressources** : Créez un nouveau (ex: ProjetAzureRG).

- **Nom du compte :** stockagevideogames2024 (doit être unique).
- **Région :** (France Central ou West Europe).
- **Redondance :** Choisir **LRS** (Locally-redundant storage) pour minimiser les coûts.

#### 4. Cliquez sur **Revoir + créer**, puis **Créer**.

The screenshot shows the Azure Deployment blade for the storage account 'stockagevideogames2024\_1766878490477'. It displays a success message: 'Votre déploiement a été effectué'. Below this, it shows the deployment details: Name: stockagevideogames2024\_1766878490477, Abonnement: Azure for Students, Groupe de ressources: ProjetAzureRG. The deployment started at 28/12/2025 00:42:41 and has a correlation ID: 3e9a12be-0ae4-4df8-b832-8b3003a70cc2. A table titled 'Détails du déploiement' lists three resources: 'stockagevideogames2024/default' (Type: Microsoft.Storage/storageAccounts, Status: OK), 'stockagevideogames2024/default' (Type: Microsoft.Storage/storageAccounts, Status: OK), and 'stockagevideogames2024' (Type: Microsoft.Storage/storageAccounts, Status: OK). To the right, there are links for 'Cost Management', 'Microsoft Defender', and 'Tutoriels Microsoft'. At the bottom, there are buttons for 'Envoyer des commentaires' and 'Partagez votre expérience avec le déploiement'.

  

The screenshot shows the Azure Storage blade for the storage account 'stockagevideogames2024'. The left sidebar includes options like Journal d'activité, Étiquettes, Diagnostiquer et résoudre les problèmes, Contrôle d'accès (IAM), Migration des données, Événements, Navigateur de stockage, Storage Mover, Solutions de partenaire, Visualiseur de ressources, Stockage des données, Sécurité + réseau, Gestion des données, Paramètres, and Supervision. The main pane displays the storage account properties under the 'Bases' tab. Key settings shown include: Groupe de ressources: ProjetAzureRG, Emplacement: francerégional, Abonnement (déplace): Azure for Students, ID d'abonnement: b033a59c-3ce6-45d5-a95c-a07e46d2c0da, État du disque: Disponible, and Étiquettes (modifier): Ajouter des étiquettes. The 'Propriétés' tab is selected. Other tabs include Monitoring, Fonctionnalités (7), Suggestions (0), Tutoriels, and Outils et Kits de développement logiciel (SDK). The 'Vue JSON' link is also visible.

5. Une fois la ressource créée, allez dans le menu de gauche **Stockage des données > Conteneurs**.
6. Cliquez sur **+ Conteneur**, nommez-le source-data et validez.
7. Entrez dans le conteneur source-data, cliquez sur **Charger (Upload)** et sélectionnez votre fichier vgsales.csv.

The screenshot shows the Azure Storage Explorer interface. The left sidebar has a tree view with 'source-data' selected. The main area shows a table of blobs. One row is selected, showing details: Name is 'vgsales.csv', Last modified is '28/12/2025 00:46:53', Access level is 'Chaud (défaut)', Blob type is 'Objet blob de ...', Size is '1.29 MiB', and Status is 'Disponible'. There are buttons for 'Ajouter un répertoire', 'Charger', 'Modifier le niveau d'accès', 'Actualiser', 'Supprimer', 'Copier', 'Coller', 'Renommez', 'Acquérir le bail', and 'Résilier le bail'.

## Étape 2 : Création de la Base de Données

*Objectif : Préparer la table SQL pour recevoir les données propres.*

1. Dans le portail, recherchez **Bases de données SQL** (SQL Databases).
2. Cliquez sur **+ Créer**.
  - **Nom de la base** : VideoGamesDB.
  - **Serveur** : Cliquez sur "Créer un nouveau". Nommez-le (ex: srv-games-2024).
  - **Authentification** : Choisissez "Utiliser l'authentification SQL". Créez un login (ex: adminuser) et un mot de passe fort.
  - **Niveau de calcul** : Cliquez sur "Configurer la base de données", choisissez **De base (Basic)** ou "Serverless" pour limiter les coûts (très important pour un projet étudiant).
3. Dans l'onglet **Mise en réseau** (Networking) :
  - **Règle de pare-feu** : Cochez **OUI** à l'option "Autoriser les services et les ressources Azure à accéder à ce serveur" (Indispensable pour que Data Factory puisse écrire dedans).
4. Cliquez sur **Créer**.

**Créer une base de données SQL** ...

Microsoft

Région	France Central
Nom de la base de données	VideoGamesDB
Serveur	(nouveau) srv-games-2024
Méthode d'authentification	Authentification SQL
Connexion d'administrateur du serveur	adminuser
Calcul + stockage	Usage général - Serverless: Série standard (Gen5), 1 cœur virtuel, Stockage de 32 Go, non redondant dans une zone
Redondance du stockage de sauvegarde	Stockage de sauvegarde redondant localement

**Mise en réseau**

Autoriser les services et les ressources Azure à accéder à ce serveur	Oui
Ajouter l'adresse IP actuelle du client 89.92.180.62	Oui
Point de terminaison privé	Aucun
Version TLS minimale	1.2
Stratégie de connexion	Default

**Sécurité**

Identité	Non activé
Cryptage transparent des données (niveau serveur)	Clé gérée par le service sélectionnée

**Créer une base de données SQL** ...

Microsoft

Identité	Non activé
Cryptage transparent des données (niveau serveur)	Clé gérée par le service sélectionnée
Clé gérée par le client au niveau de la base de données	Non configuré
Identité gérée attribuée à l'utilisateur au niveau de la base de données	Non configuré
Advanced data security	Pas maintenant
Toujours crypté avec des enclaves sécurisées	Non configuré
Registre SQL (base de données)	Désactivé
Stockage des synthèses	Désactivé

**Paramètres supplémentaires**

Utiliser des données existantes	Blank
Classement	SQL_Latin1_General_CI_AS
Fenêtre de maintenance	Valeur système par défaut (17 h 00 - 8 h 00)

**Balises****Microsoft.SQLDatabase.newDatabaseNewServer\_eafc565f39dd45b5aa7c8 | Vue d'ensemble** ...

Déploiement

Rechercher Supprimer Annuler Redéployer Télécharger Actualiser

**Votre déploiement a été effectué**

Nom du déploiement : Microsoft.SQLDatabase.newDatabaseNewServer\_e... Heure de début : 28/12/2025 01:09:37  
Abonnement : Azure for Students ID de corrélation : 8522ea05-12ce-4f5d-8e3c-4f1524d376be  
Groupe de ressources : ProjetAzureRG

**Détails du déploiement**

Ressource	Type	Statut	Détails de l'opération
srv-games-2024/VideoGamesDB	Microsoft.Sql/servers/databases	Created	Détails de l'opération
srv-games-2024/ClientIp-2025-12-28-1-9-30	Microsoft.Sql/servers/firewallrule	OK	Détails de l'opération
srv-games-2024/Default	Microsoft.Sql/servers/connection	OK	Détails de l'opération
srv-games-2024/AllowAllWindowsAzureips	Microsoft.Sql/servers/firewallrule	OK	Détails de l'opération
srv-games-2024	Microsoft.Sql/servers	Created	Détails de l'opération

**Étapes suivantes**

Accéder à la ressource

**Gestion des coûts**  
Recevez des notifications pour vous aider à respecter votre budget et à éviter des frais imprévus sur votre facture.  
[Configurer les alertes de coût >](#)

**Microsoft Defender pour le cloud**  
Sécuriser vos applications et votre infrastructure  
[Accédez à Microsoft Defender pour le cloud >](#)

**Tutoriels Microsoft gratuits**  
[Commencer l'apprentissage aujourd'hui >](#)

**Travailler avec un expert**  
Les experts Azure sont des partenaires fournisseurs de services qui peuvent vous

5. Une fois créée, allez sur la ressource, menu de gauche **Éditeur de requêtes (aperçu)**. Connectez-vous avec vos identifiants.

6. Collez et exécutez le code suivant pour créer la table cible :

SQL

CREATE TABLE JeuxVideo (

```

Rank INT,
Name NVARCHAR(200),
Platform VARCHAR(50),
Year VARCHAR(10),
Genre VARCHAR(50),
Publisher NVARCHAR(100),
NA_Sales DECIMAL(10, 2),
EU_Sales DECIMAL(10, 2),
JP_Sales DECIMAL(10, 2),
Other_Sales DECIMAL(10, 2),
Global_Sales DECIMAL(10, 2)
);

```

```

5   Year VARCHAR(10),
6   Genre VARCHAR(50),
7   Publisher NVARCHAR(100),
8   NA_Sales DECIMAL(10, 2),
9   EU_Sales DECIMAL(10, 2),
10  JP_Sales DECIMAL(10, 2),
11  Other_Sales DECIMAL(10, 2),
12  Global_Sales DECIMAL(10, 2)
13 );

```

### Étape 3 : ETL avec Nettoyage

*Objectif : Utiliser "Mapping Data Flow" (flux de données de mappage) pour nettoyer les données (supprimer les années "N/A") au lieu d'une simple copie.*

1. Recherchez **Fabriques de données** (Data Factories) > + **Créer**. Nom : adf-video games-2024.

The screenshot shows the Azure Data Factory Studio interface. At the top, there's a navigation bar with a search bar, a 'Supprimer' (Delete) button, and tabs for 'Résumer les propriétés de cette fabrique de données (V2)' and 'Afficher les métriques du runtime d'intégration pour cette fabrique de données (V2)'. Below the navigation is a sidebar with links like 'Rechercher', 'Vue d'ensemble', 'Journal d'activité', 'Contrôle d'accès (IAM)', 'Étiquettes', 'Diagnostiquer et résoudre les problèmes', 'Visualiseur de ressources', 'Paramètres', 'Démarrer', 'Supervision', 'Automatisation', and 'Aide'. The main content area has a title 'Bases' with a collapsible section showing details: Groupe de res... (déplacer) : ProjetAzureRG, Statut : Succeeded, Emplacement : France Central, Abonnement (déplacer) : Azure for Students, ID d'abonnement : b033a59c-3ce6-45d5-a95c-a07e46d2c0da. To the right, it says Type : Fabrique de données (V2) and Démarrer : Démarrage rapide. A 'Vue JSON' link is at the top right. In the center, there's a blue cube icon with three dots, followed by the text 'Azure Data Factory Studio' and a 'Lancer le studio' button. Below this are four cards: 'Démarrages rapides' (Quick Starts) with a cloud icon, 'Tutoriels' (Tutorials) with a laptop icon, 'Galerie des modèles' (Model Gallery) with a document icon, and 'Modules d'entraînement' (Training Modules) with a person icon.

2. Cliquez sur **Lancer Studio** (Launch Studio).

**3. Créez les Services Liés (Connections) :**

- Allez dans l'onglet **Gérer** (la valise) > **Services liés** > **+ Nouveau**.
- Créez une connexion vers **Azure Blob Storage** (pointez vers le compte créé à l'étape 1).

## Nouveau service lié

 Stockage Blob Azure [En savoir plus](#)

Se connecter via un runtime d'intégration \* ⓘ

AutoResolveIntegrationRuntime

Type d'authentification

Clé de compte

**Chaîne de connexion**

Azure Key Vault

Méthode de sélection de compte ⓘ

From Azure subscription  Enter manually

Abonnement Azure ⓘ

Azure for Students (b033a59c-3ce6-45d5-a95c-a07e46d2c0da)

Nom du compte de stockage \*

stockagevideogames2024



Propriétés de connexion supplémentaires

+ Nouveau

Tester la connexion ⓘ

Au service lié  Au chemin de fichier

Annotations

+ Nouveau

Créer

Précédent

 Tester la connexion

Annuler

- Créez une connexion vers **Azure SQL Database** (pointez vers la DB créée à l'étape 2).

## Nouveau service lié

 Azure SQL Database [En savoir plus](#)

Nom \*

Description

Se connecter via un runtime d'intégration \* ⓘ

AutoResolveIntegrationRuntime



Version

2.0 (Recommended)  1.0

 Importer à partir d'une chaîne de connexion

Méthode de sélection de compte ⓘ

From Azure subscription  Enter manually

Abonnement Azure

Azure for Students (b033a59c-3ce6-45d5-a95c-a07e46d2c0da)



Nom de serveur \*

srv-games-2024



Nom de la base de données \*

VideoGamesDB



## Nouveau service lié

 Azure SQL Database [En savoir plus](#) 

Type d'authentification \*

SQL authentication

Nom d'utilisateur \*

adminuser

Ajouter du contenu dynamique [Alt+Maj+D]

Mot de passe

Azure Key Vault

Mot de passe \*

.....

Ajouter du contenu dynamique [Alt+Maj+D]

Always Encrypted 

Encrypt 

Mandatory

Trust server certificate 

Ajouter du contenu dynamique [Alt+Maj+D]

Host name in certificate

Propriétés de connexion supplémentaires

+ Nouveau

### 4. Créer les Jeux de données (Datasets) :

- Onglet **Auteur** (Crayon) > **Datasets** > + Nouveau.
- **Source** : Azure Blob Storage > CSV. Sélectionnez le fichier vgsales.csv.  
Cochez "First row as header".
- **Sink** : Azure SQL Database. Sélectionnez la table dbo.JeuxVideo.

The image consists of two vertically stacked screenshots of the Azure Data Factory interface. Both screenshots show the 'Ressources de fabrique' (Fabric Resources) pane on the left and a detailed configuration pane on the right.

**Screenshot 1: Source CSV Configuration**

- Properties:**
  - Nom:** SourceCSV
  - Description:** [Empty]
  - Annotations:** [Empty]
- Configuration:**
  - Service lié:** AzureBlobStorage1
  - Chemin d'accès au fichier:** source-data / Annuaire / vgsales.csv
  - Type de compression:** Aucune compression
  - Séparateur de colonne:** Comma (,)
  - Délimiteur de ligne:** Default (\r\n or \n\r)
  - Encodage:** Par défaut(UTF-8)
  - Guillemet:** Double quote (")

**Screenshot 2: Sink SQL Configuration**

- Properties:**
  - Nom:** SinkSQL
  - Description:** [Empty]
  - Annotations:** [Empty]
- Configuration:**
  - Service lié:** AzureSqlDatabase1
  - Tableau:** dbo.JeuxVideo
  - Actualiseur:** Actualiseur
  - Aperçu des données:** Entrez manuellement

## 5. Créez le Data Flow :

- Allez dans **Flux de données** (Data flows) > + Nouveau flux de données.
- Activez le bouton "**Déboguer le flux de données**" en haut pour voir les données en direct.
- Ajouter la Source :** Cliquez sur "Add Source", nommez-la SourceCSV, choisissez le dataset CSV.
- Ajouter une transformation (Le Bonus) :** Cliquez sur le petit + à droite de la source et choisissez **Filtrer** (Filter).
  - Nommez-le **FiltrerAnneesManquantes**.
  - Dans l'éditeur d'expression, tapez : `Year != 'N/A'`
  - Explication : Cela supprime les lignes où l'année est différent de "N/A", rendant la donnée plus propre.*

- Ajouter le Récepteur (Sink) : Cliquez sur le + après le filtre et choisissez Récepteur (Sink).

- Choisissez le dataset SQL.

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the 'Ressources de fabrique' (Fabric Resources) sidebar lists various components like Pipelines, Capture des changements de données, Jeux de données, and Flux de données. The main workspace displays a data flow pipeline named 'flux de données1'. This pipeline starts with a 'SourceCSV' activity (Import from CSV) which reads data from a CSV file. This is followed by a 'FiltreAnnées' (Filter by Year) activity, which uses expressions to filter rows based on the 'Year' column. Finally, the data is loaded into a 'VersSQL' (Load to SQL) activity. The 'Aperçu des données' (Data Preview) tab on the right shows a preview of the first five rows of the dataset, which include columns like Rank, Name, Platform, Year, Genre, and Publisher. The pipeline is currently in debug mode, indicated by the green checkmarks in the status bar.

## 6. Créez le Pipeline :

- Allez dans **Pipelines** > + Nouveau pipeline.
- Glissez-déposez l'activité **Flux de données** dans le panneau central.
- Sélectionnez le flux que vous venez de créer.

## 7. Cliquez sur Déboguer (Debug). Attendez que les coches vertes apparaissent.

The screenshot shows the Microsoft Azure Data Factory interface. The 'Ressources de fabrique' (Fabric Resources) sidebar lists a pipeline named 'pipeline1'. The main workspace shows the details of this pipeline. The 'Activités' (Activities) section indicates that a 'Flux de données' (Data flow) activity has been added. The pipeline status is shown as 'Opération réussie' (Operation successful). The pipeline details pane on the right shows the name 'pipeline1' and a description field. The pipeline has been triggered 26 times, with the most recent execution starting at 12/28/2025, 3:37:09 AM.

## Étape 4 : Analyse des Données

## Analyse 1 : Vérification du nettoyage

Vérifier qu'il n'y a plus de "N/A" dans les années :

Requête 1 ×

Annuler la requête  Exporter les données comme  Afficher seulement l'éditeur

```
1  SELECT COUNT(*) as Lignes_Incorrectes
2  FROM JeuxVideo
3  WHERE Year = 'N/A';
4
```

Résultats Messages

Recherche pour filtrer des éléments...

Lignes_Incorrectes
0

## Analyse 2 : Question métier (Top 5 Éditeurs)

Quels sont les éditeurs qui ont vendu le plus de jeux (en millions d'unités) ?

Requête 1 ×

Annuler la requête  Exporter les données comme  Afficher seulement l'éditeur

```
1  SELECT TOP 5 Publisher, SUM(Global_Sales) as Total_Ventes_Millions
2  FROM JeuxVideo
3  GROUP BY Publisher
4  ORDER BY Total_Ventes_Millions DESC;
```

Résultats Messages

Publisher	Total_Ventes_Millions
Nintendo	1784.43
Electronic Arts	1093.39
Activision	721.41
Sony Computer Entertainment	607.28
Ubisoft	473.54

✔ Requête réussie | 0s