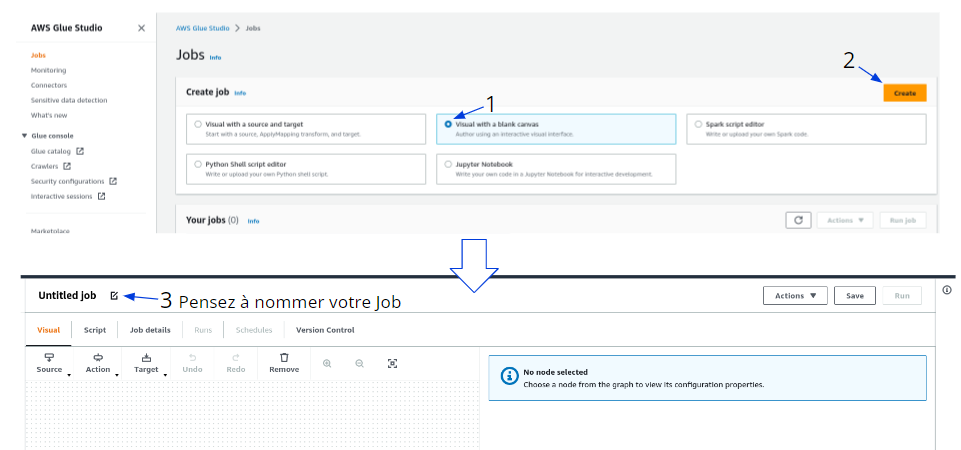
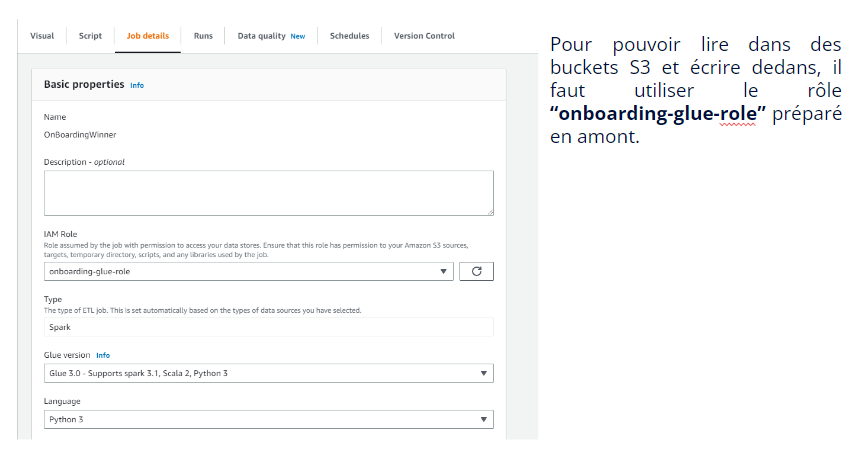
Consignes Glue

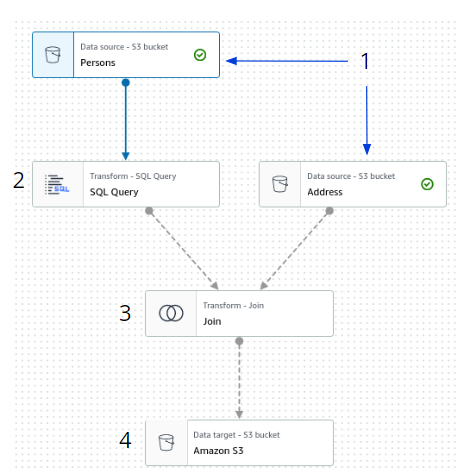
# Création d’un nouveau job



# Utilisez le rôle IAM



# Paramétrage du job



1. Lire les fichiers depuis S3 avec les configurations suivantes :
   * **Data Format** : CSV
   * **Delimiter** : Comma (,)
   * **Quote character** : Double quote (“)
   * **First line of source file contains column headers** : True
   * Cliquer le bouton “**infer schema**”
2. Décoder et mettre en qualité les données qui proviennent du fichier suspects.csv en utilisant la requête suivante :  
    **select**

**cast(unbase64(trim(prenom)) as string) as prenom**

**, cast(unhex(trim(nom)) as string) as nom**

**, trim(id\_adresse) as id\_adresse**

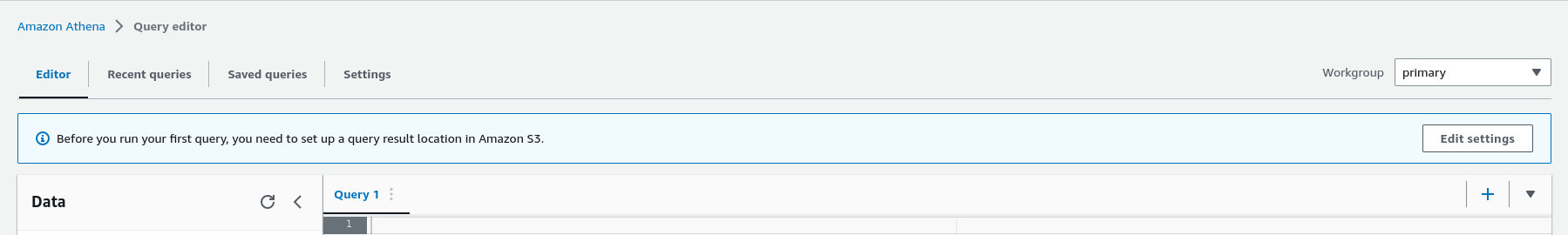
**from myDataSource**

1. Joindre les deux tables tel que id\_adresse = id
2. Ecrire le résultat dans S3 et demander à Glue de créer une table Athena
   * **S3 Target Location** : insérer un chemin vers un préfix vide (le chemin doit finir par un “/”)
   * **Create a table in the Data Catalog and on subsequent runs, update the schema and add new partitions** : True
   * **Database** : default
   * **Table name** : <nom de l’équipe>
3. Dans l’onglet **Job details** sélectionner le rôle onboarding-glue-role
4. Sauvegarder le job et exécuter

# Athena - Pour visionner votre résultat

Nous allons utiliser le service Athena qui permet de requêter des fichiers plats stockés dans S3 avec du SQL.

Mais avant d’écrire votre requête, vous aurez besoin de configurer un chemin vers S3 pour le stockage. Cliquez sur le bouton **Edit Settings** du bandeau bleu :



Renseigner le chemin vers votre bucket d’équipe :

